A Feasibility Study of the Use of Artificial Neural Networks in the Diagnosis and Treatment of Schizophrenia

Sabina Choudhury

A Thesis

in

The Department

of

Educational Technology

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Arts (Educational Technology) at
Concordia University
Montreal, Quebec, Canada

April, 2004

Your file Votre référence
ISBN: 0-612-91012-1
Our file Notre référence
ISBN: 0-612-91012-1

Abstract

A Feasibility Study of the Use of Artificial Neural Networks in the Diagnosis and Treatment of Schizophrenia

Sabina Choudhury

A training problem exists in the field of psychiatry since the traditional methods of varying effectiveness in diagnosing and selecting treatment options are passed onto new medical residents, learning from established professionals. The lack of consensus within the psychiatric community towards diagnosis and treatment selection probably impacts the quality of patient care. Difficulties in the diagnosis processes may lead to inaccurate diagnoses and delays in administering timely and effective treatment alternatives. To address this educational need, it is proposed that the development of a multi-layered neural network application with back-propagation facilities that would be taught to categorise the various forms of schizophrenia and the various treatment options should be created as a tool to assist medical residents in the diagnosis and treatment selection for schizophrenia. An overview of neural networks, psychiatric disorders, the steps in the diagnostic process and current assessment tools for schizophrenia is provided. The feasibility of developing a neural network as a diagnostic and treatment selection tool for schizophrenia is assessed, taking into account current neural network's limitations. It is concluded that the lack of consensus amongst psychiatric professionals towards current diagnosis and treatment selection process will result in data inadequate to train such a neural network, thereby resulting in potentially false results. The long-term methods to overcome these limitations and some alternatives towards providing a short-term computer-based educational solution are presented.

Acknowledgements

Given the intensive writing process that is often associated with a thesis, I am finding it quite difficult to express in my own words things that should be said and people who need to be acknowledged. Therefore I will do so, using the inspirational words of my favourite poet, Kahlil Gibran (The Prophet, 1923).

"When you work, you fulfil a part of earth's furthest dream, assigned to you when that dream was born" and "[work] is to charge all things you fashion with a breath of your own spirit".
This was a topic close to my heart and spirit and truly feels like a dream fulfilled.

****

To my parents: "*You are the bows from which your children as living arrows are sent forth"*. Thank you for launching that arrow and providing the foundation that permits me to find my own path in this world.

To Ruby, Josh and Sandra: *"When he is silent your heart ceases not to listen to his heart: For without words, in friendship, all thoughts, all desires, all expectations are born and shared, with joy that is un-acclaimed".* Your unconditional support and words of wisdom carried me through some difficult moments. Thank you for knowing when to listen, even when no words were spoken.

All of my friends (you know who you are): "*Your friends are your needs answered"*. To all of you who supported me and put up with my never-ending bouts of self-doubt. Each of you fills a need that makes me a complete person.

Professor Gary Boyd: *"If he is indeed wise he does not bid you enter the house of his wisdom, but rather leads you to the threshold of your own mind".* This epitomises the wealth of knowledge that I have gained by opening my mind to all that you exposed me.

Anne Brown: "*You who travel with the wind, what weather vane shall direct your course?"* I think I speak for all students when I thank you for guidance in all activities required during the academic process. In addition, thank you for keeping me sane and providing the voice of reason when I thought that I would never finish.

Dr Vida: "*The vision of one man lends not its wings to another man"*. Thank you seeing my vision and for not trying to talk me out of what was a challenging topic. You assisted me even with the limitations.

****

Some final words: "*You would know in words that which you have always known in thought" and "The treasure of your infinite depths would be revealed to your eyes".*
This process has challenged and pushed me and has revealed an insight into myself that has forever shaped my soul.

Table of Contents

List of Figures

List of Tables

I. Training Problem

Chapter 1: Training Problem Overview

Much of psychiatric diagnosis is based upon standards and practices that have been in place for a long period of time. Diagnostic tools such as standardised tests, Diagnostic and Statistical Manual of Mental Disorders (DSM), psychiatric experiences etc. form the foundation by which diagnoses are made. Although psychiatric treatment options are closely regulated, much of the choices rest upon trial and error and the expertise of psychiatric professionals.

However, even with the strict diagnostic rules and regulated treatment options established within the psychiatric community, there are many patients who "fall through the cracks" and do not receive adequate care. In addition, many patients who do not necessarily meet the diagnostic guidelines established may receive less than adequate treatment because of problematic diagnosis. In addition, chemical treatments are not always conducive towards adequate patient functioning.

The problems are not necessarily products of poor training / teaching or inadequate tools to assess and determine treatment choices. Rather, there may be certain undefined factors that have yet to be considered during the current diagnostic process and treatment selection. Perhaps there are patterns that have yet to be considered and are therefore excluded during this process. Are there undetermined factors that should be considered that could lead to better treatment choices in an effort to reduce the current trial and error method of treatment selection?

The method of training medical residents to diagnose and select treatment options rests with those in the psychiatric community (psychiatrists, nurses, social workers etc.). The already established and "proven" tools are passed onto the new group of medical psychiatric residents. As such, treatments are still chosen on a "best-effort" basis or trial and error method. While new psychiatric medication attempts to minimise any potential side effects and offer a healthier way of living, many patients are still prescribed a variety of medication until a "best fit" is achieved. This method is considered to be problematic within the psychiatric community as it is not necessarily conducive to the goals of achieving and maintaining a healthier lifestyle for the patients. In addition, this method is not cost effective as this revolving cycle can result in patient recidivism, medication waste, and excess work hours by the psychiatric professionals.

Given these factors, there is a definite training problem that needs to be addressed. There is a real difficulty in training residents to diagnose psychiatric disorders optimally in order to determine the best and cost efficient treatment solution. Psychiatrists often make uncertain diagnoses and treatment selections for psychiatric patients. They are also the ones primarily responsible for the training of new psychiatric medical residents. As illustrated above and will be described later, there is a need to enhance the diagnostic process and treatment selection. How can we minimise the difficulty of between diagnosing psychiatric patients early and provide the best treatment alternative? Part of this solution rests with providing tools that would improve human performance (of the medical residents) during the diagnostic process and treatment selection.

Chapter 2: General Overview

The world of psychiatry has evolved over the years. In all areas from diagnostic terminology to treatment methods, there has been a shift in this domain over the years. Ancient societies used to attribute mental illness to supernatural forces and as such, used exorcisms or even drilling holes in the skulls of afflicted individuals to "drive out" the evil spirits. The Greek and the Roman civilization saw the first attempts to classify mental ailments. Hippocrates observed cases of mental disturbance and attempted to determine possible organic or biological causes of mental illnesses. However, no attempts were made to cure the disorders. With the advent of possible biological causes, there was still a common belief the in the evil spirit aetiology of mental illness, as evidenced by the witch-hunts. The eighteenth and nineteenth centuries saw the creation of asylums to house the mentally ill, although usually with less than hygienic conditions.

The late nineteenth century saw the beginning of new ways to study and classify mental illness. It also saw the re-introduction of a belief in biological elements as a probable cause. In addition, Freud introduced psychoanalysis to the field of psychiatry, where the root cause of mental illness was believed to be a product of "unconscious conflict", most of which could be resolved through the use of hypnosis and talk. The introduction of various psychological perspectives (cognitive, behavioural, humanistic, etc.), have provided psychologists and psychiatrists with different approaches to treat the mentally ill.

Scientific advancements have led researchers, as opposed to clinicians, to concentrate on the biological or neurological elements of psychiatric disorders. Disorders such as schizophrenia or depression have been found to be partially the result of chemical imbalances in the brain (schizophrenia has been linked to the lack of dopamine receptors in the brain). These have led to the development of a class of psychiatric drugs that will target or correct for these imbalances.

Cybernetics can be used to "structure" the field of psychiatry. Norbert Wiener has defined Cybernetics as the "science of control and communication in the animal and machine". More specifically, it is a field that outlines the organization of relating systems in a manner that will generate the desired results. The methodology of systems science allows for the understanding of relational elements of various components within a system. This relational view is important when considering the field of psychiatry and the implications for patient care. The changes in aetiology that psychiatry has undergone over the years have led to various treatment options, from psychoanalysis to the now popular pharmacological variations. Regardless of the changes over time, there are specific processes usually involved when the initial assessment is made, to the eventual diagnosis and finally the ultimate treatment option(s). This sequence of processes, and all its relevant components can be understood using a systems analysis.

Although biological science appears to now be the focal point for diagnosing and treating psychiatric disorders, it is important that the social and psychological issues (different for each individual) not be ignored when helping the psychiatric patient. Based upon Wiener's definition, there is a global systemic component to the diagnosis, treatment and management of psychiatric disorders. The major goal of this "Patient System" is to ensure the well being of patients (both in and out patients). In order to accomplish this goal, it requires the work and cooperation of other sub-systems. In addition, all must work within the confines of governmental regulations, limited budgets, and demands of other interest groups (patient advocate groups etc.). For the purposes of this document, the following sub-systems have been determined to be key to the diagnosis and treatment. Figure 1.1 provides an overview of the major sub-systems deemed critical for the purposes of this document.

a) *Medical*: this represents the doctors who assess, diagnose and determine the psychiatric treatment and the medical residents who train with the experienced medical professionals. The psychiatrist makes the initial, official diagnosis. This is often done after clinical interviews, an assessment of past psychiatric history, family background, the administration of tests (psychometric and/or biological) and other methods.

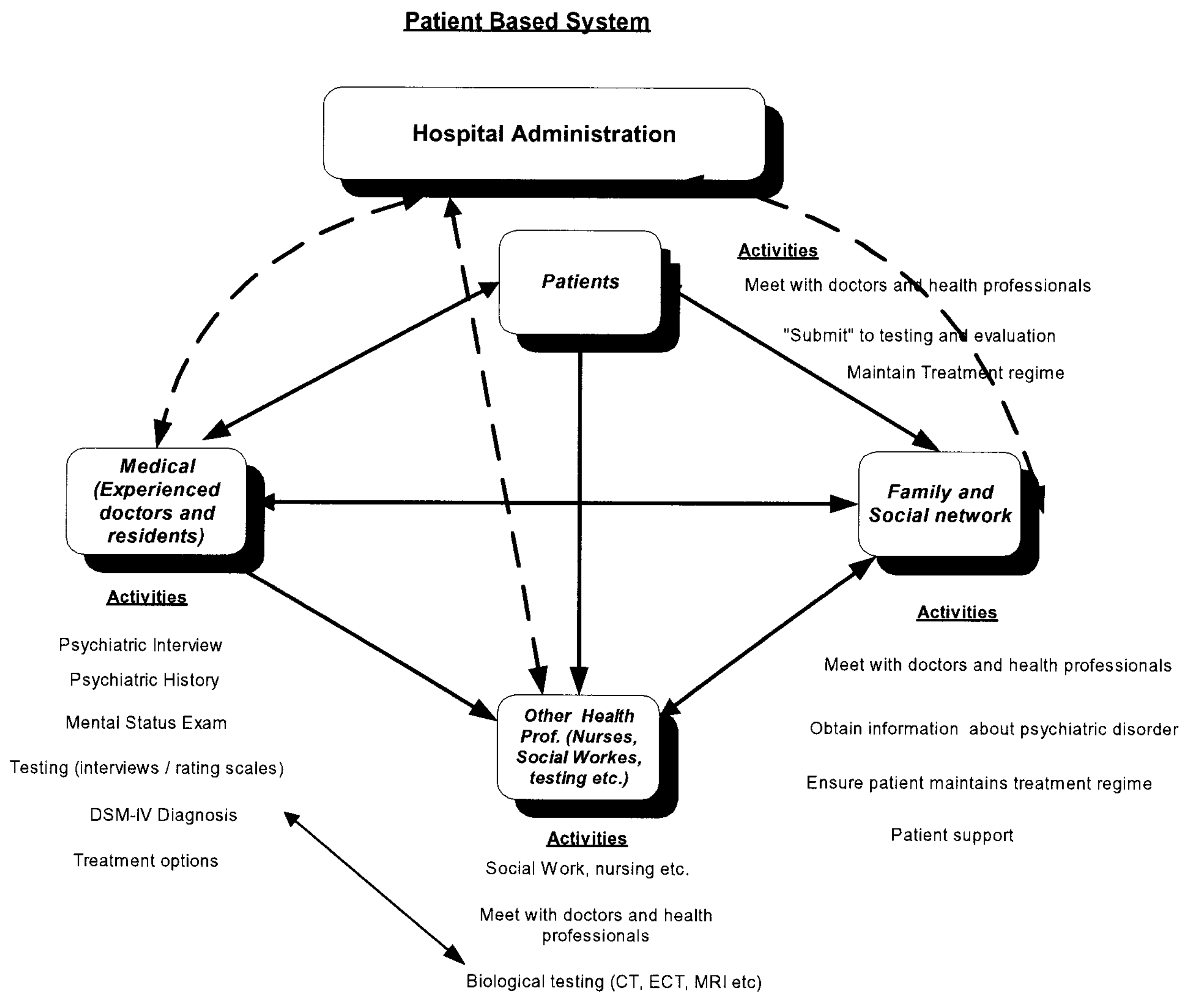
b) *Patients*: these are the psychiatric patients who are interviewed, tested, diagnosed, and who are recipients of the treatment prescribed by the physicians.

c) *Family and Social Support*: this represents the support structure that is often deemed critical to maintaining the well being of the psychiatric patient. They provide the emotional assistance, medication reminders to patients, ensure appointments are kept, etc.

d) *Other*: this represents the other participants required to ensure the well being of the patient. These include nurses, social workers, and technicians (those that administer the ECT, MRI, CT etc.). Interviews with social workers and other health care professionals is often performed in order to ensure that the patient is not only managed from a medical or psychiatric perspective, but that adequate measures are in place to ensure that the patient is a functioning member within society. This is important, primarily for those patients admitted into psychiatric wards.

All of these sub-systems have individual responsibilities and interact within and among one another.

Figure I: 2.1: Patient sub-system



The following table is a summary of the responsibilities and goals of the key sub-systems defined for the purposes of this document.

Table I – 3.1 Sub-system responsibilities and goals

Responsibilities	Goals
Sub-system I: Medical	
Psychiatric interview	• To establish rapport with the patient. • To informally explore the patient's past history • To obtain overview of patient functioning and potential diagnosis and treatment.
Psychiatric history	• To assist in diagnosis and treatment options (i.e. obtain onset of illness, precipitating factors etc.).
Mental Status Exam	• To assess patient functioning and ability to process information.
Testing (interviews and rating scales)	• To provide quantitative or qualitative information to assist in the diagnostic process.
Provide DSM-IV diagnosis	• To determine appropriate course of treatment action.
Treatment	• To ensure patient functioning and well-being
Sub-system II: Patients	
Meet with doctors and other health care professionals	• To obtain diagnosis • To obtain treatment options • To ensure patient well-being • To satisfy and maintain well-being of family and social support • To ensure maintenance of treatment regime
Testing and Evaluation	• To obtain diagnosis • To obtain treatment options
Maintain treatment regime	• To ensure well-being of patient • To ensure well-being of family, social support and society
Sub-system III: Family and Social Network	
Meet with doctors and other health care professionals	• To obtain diagnosis • To obtain treatment options • To obtain social support (e.g. support groups, therapy etc.) • To ensure patient maintains treatment
Obtain information about illness	• To educate themselves on the prognosis • To ensure patient maintains treatment regime • To obtain support from services and other families
Ensure patient maintains treatment regime	• To ensure well-being of patient • To ensure well-being of family, social support and society
Patient support	• To ensure well-being of patient
Sub-system: Other Health professionals	
Nursing, social work etc.	• To ensure patient integration into society.
Meet with doctors and other health care professionals	• To obtain diagnosis • To obtain treatment options • To ensure patient well-being • To satisfy and maintain well-being of family and social support • To ensure maintenance of treatment regime
Biological testing (MRI, CT etc)	• To provide and reinforce diagnostic process.

Chapter 3: Thesis Overview

It is the author's appraisal that the current methods of assessment and diagnosis are too time consuming and inaccurate, often amounting to little more than trial and error diagnosis and treatment. In order to determine the proper methods of treatment, it is essential that a correct diagnosis be made for the patient in question. With psychiatric disorders, there is often a lack of consensus amongst health care professionals regarding the primary diagnosis. Even with standardized criteria (e.g. the Diagnostic and Statistical Manual – DSM) and the various tools to assess psychiatric disorders, the inconsistencies in diagnosis and ultimately in treatment, still exists. This can hinder the process of teaching the methods of diagnosis, assessment and treatment, to potential health care professionals, keen to ensure the well being of their patients.

The treatment approach is determined from the diagnosis. As such, it is important that the proper diagnosis be made. There are various methods used. Many experienced psychiatrists rely on their own evaluation methods (past psychiatric episodes, family history, symptoms, etc.). There are a multitude of standardized tools available, ranging from structured interviews to psychological tests. The final diagnosis is often made in accordance to criteria specified in the DSM for legal and financial reasons and as a method to standardise the diagnostic process. However, as most of these tools have not been subjected to reliability and validity studies, they are primarily considered as research tools and still depend on the individual, professional, interpretation. The results obtained do not always point exactly to a specific disorder category. They therefore can lead to different professional diagnostic opinions.

As discussed previously, much of the treatment begins with the administration of psychiatric drugs to control the symptoms. However, it is often uncertain as to how each individual will respond to the drugs prescribed. This, coupled with the side effects often associated with the drugs, makes the process of determining adequate treatment difficult and as such, results in the trial and error approach to determine the best medication suited for each individual.

As illustrated above, diagnosing and treating mental disorders is not a straightforward task. There are many elements and systems involved, each of which could vary from individual to individual. Today, not all depressed patients respond equally to the same medication, even with similarities in symptoms. There are also variations within particular psychiatric disorders. With all of these variants, it would be ideal to have a tool that would account for many of these variances and thus enable professionals to properly diagnose and treat each individual case. There are many elements that go into the diagnosis (tests, professional opinions, diagnostic criteria etc.). As such, a tool that would incorporate all of these variables should greatly assist in the diagnosis and determination of appropriate treatment in a timely and cost-efficient manner, without jeopardising patient care. A tool such as this would and should also incorporate the characteristics of the individual and his/her environment and not simply produce a generalized "one size fits all" type of output.

It is proposed that creating an *artificial neural network* (ANN), which could incorporate the vast amounts of information that would be required to accurately diagnose and determine specific treatment options tailored to each individual. This would be worthwhile, at least for educating physicians. Artificial neural networks are expert systems that mimic human brain functioning. They were originally popular in the 60s, but lost favour due to the complexity and expanse of developing such systems. With the creation of sophisticated computer tools, artificial neural networks have since regained prominence and are currently used in many areas that require forecasting (stock market) and diagnoses (medicine). The main advantage of using artificial neural networks over standard computerised tools is their ability to learn adaptively. Similar to human brain functioning, artificial neural networks have the capabilities of learning to reconfigure themselves over time, for better categorising patterns and their recognition abilities. In addition, unlike standard computers which process information in a serial format, artificial neural networks work with parallel processes and as such, can take in a multiplicity of information at once in order to produce valuable output. Many other systems have been built to simulate parallel processing, by incorporating many serial processes, but most still lack sufficient ability to learn and adjust outputs based upon "prior experience".

The ultimate purpose of this thesis will be to analyse whether neural networks can really help to overcome the inadequacies of varying diagnoses and therefore varying treatment options often made by different psychiatrists in order to provide an adequate teaching tool for new medical psychiatric residents. It is assumed that such a tool would minimise diagnostic differences and may ensure that the optimal medication be administered at first to avoid the trial and error method of drug treatment. It is proposed that such a tool would be better suited for such a task because of its ability to learn. A neural network could potentially take a multitude of psychiatrists' diagnostic and treatment data and simultaneously process the information in order to provide the optimal diagnosis and treatment for the specific patient without so much of the variances common amongst practising psychiatrists today.

The thesis will attempt to specify the components required to create an artificial neural network to diagnose, assess and provide treatment alternatives for an individual with a psychiatric disorder. It is the author's contention that such a tool, designed to incorporate most facets of the diagnostic and treatment conditions will be able to offer an accurate and speedy basis for diagnosis and target specific alternatives for each individual as opposed to providing a generalised diagnosis with a generalised treatment. This tool will not initially be developed for practicing psychiatrists, as many have the experience and their own criteria to diagnose, assess and provide treatment alternatives. In addition, the legal, liability ramifications would be problematic. This will be a teaching tool that will enable psychiatric medical residents to validate their results obtained in the traditional clinical assessment and supplement what they learn. However, it is the author's dream that such a tool could potentially be used as rapid and accurate diagnostic and treatment measure that could eliminate the need for multiple assessment tools and reduce the variability that exists amongst the mental health professionals. This tool must provide an alternative method that would enhance or supplement the current teaching tools available for medical residents. In addition, creating an accurate tool may decrease the amount of recidivism currently experienced in the mental health field.

As the field of psychiatry is vast, the thesis will focus on the feasibility of developing an algorithm that will evaluate, assess, diagnose and treat various forms of schizophrenia (paranoid, catatonic, etc.). The development of an artificial neural network is a complex task. The broad and complex nature of psychiatry sometimes lends itself to a lack of diagnostic and treatment consensus amongst experts in this field. As such, focusing on one element of psychiatry (schizophrenia) will permit a feasible assessment of the use of such a tool. Here, the assumption is that schizophrenia is a specific medical disorder rather than a broad category of complexity and interacting source of pain and malfunction. It is therefore reasonable to use the DSM-IV criteria as a source to define schizophrenia.

Discussing the possibility of such a tool has originally led to some reservations. This is not due to the fear that an artificial neural network would potentially replace the health care professionals, as this would never happen, but rather as a question as to its added usefulness, given the realm of experience with psychiatrists, the depth of assessment tools and DSM-IV decision trees available. It was a psychiatrist's (Vida, S. , 2002) question about the added value, which prompted the decision to create a patient based tool that would essentially provide an "A to Z" blueprint for each patient, from the diagnosis to the treatment and treatment outcome.

The proposal is divided into various sections, each considered essential towards the development of such a tool.

Section I – Training Problem: Described above.

Section II - Literature review: This section contains the key areas of research deemed important towards the development of the algorithm. It is comprised of the following chapters.

a) *Chapter 1* provides a description of artificial neural networks, given that they are the focus for this thesis. A comparison with biological neural networks is presented, in addition to the components, theories, algorithms and various types of artificial neural networks. This section is quite extensive, as the selection of the types of networks is crucial to form the building blocks towards a full diagnostic and treatment tool.

b) *Chapter 2* provides an overview of psychiatric disorders with a description of schizophrenia. The DSM-IV is the current standard for determining psychiatric diagnoses. The DSM-IV criteria will form one element that is involved in the development of the diagnostic algorithm. The various treatments available to schizophrenics will also be discussed.

c) *Chapter 3* discusses the diagnostic process commonly used amongst the psychiatric community. These steps can be used as an illustration of the time involved to assess the patient as well as the difficulties that can sometimes ensue when providing a diagnosis. In addition, they can also be used as the potential criteria (inputs) required in the creation of an artificial neural network. In order to validate that an artificial neural network would be beneficial as a teaching and potentially a diagnostic tool, a review of difficulties often associated with diagnosis and treatment will also be discussed.

d) *Chapter 4* reviews the various assessment tools used to diagnose schizophrenia. Their criteria can be used as inputs in the development of an algorithm. For example, historical information, psychiatric history etc. can potentially be "fed into" an artificial neural network in order for it to establish the proper diagnosis, treatment and predict treatment outcomes.

Section III - The Educational Technology Problem: As a means to consolidate these chapters, this section will discuss the limitations of the current system, from the diagnostic to the treatment process. It will assimilate the ideas in the literature review in order to justify the development of an artificial neural network as a tool for medical residents. It will illustrate that the time and effort required for a medical resident to learn the current methods of evaluation, diagnosis and treatment options can be better enhanced by the development of such a tool in order better serve the schizophrenic population.

Section IV – Hypothetical Solution: This section will argue for the usefulness of an artificial neural network to diagnose, assess and treat schizophrenia. The type of artificial neural network deemed suitable for such a problem will be introduced as well as some of the components required to build the algorithm.

Section V: Conclusion: This section will discuss the feasibility of a neural network to accomplish the training problem and other possible solutions.

Section VI: Recommendations: This section will review the best course of action that could be conducted NOW in order to obtain an educational solution to the training problem.

II. Literature Review

Chapter 1: Neural Networks

A. Introduction to Artificial Neural Networks

Artificial neural networks (ANNs) are mathematical models of information processing systems that try to replicate the functioning of biological neural networks (Fausett, 1994). They are computer simulations of human brain performance (Moallemi, 1991). They are named after the cells in the human brain that perform intelligent operations. Artificial neural networks are formed from hundreds or thousands of simulated neurons connected together in much the same way as the brain's neurons (http://www.calsci.com/whatare.html, 2000).

Artificial neural networks are designed to imitate biological neural networks in their parallel, interconnected architectures and thus, parallel processing. This parallel processing differs from conventional systems whereby instructions are processed sequentially (i.e. one instruction at a time). Conventional systems can be created with multiple serial processors, to give it the impression of parallel activity, but information is still decimated one input at a time. In addition, it is very difficult for programmers to write programs for parallel machines, as they must consider many possible outcomes occurring simultaneously as opposed to one solution leading to another. As a result of this interconnected and parallel activity, artificial neural networks, have the ability to learn and can process various pieces of information at the same time.

Artificial neural networks are trained by repeatedly presenting examples to the network (http://www.calsci.com/whatare.html, 2000). Each example includes both inputs (information required to make a decision) and outputs (the resulting decision, prediction, or response). If the generated output does not match the target output, the internal connections (weights) are adjusted. This process is continued until the network reaches the specified level of accuracy. Once the network is trained and tested, a new input can be presented to the network in order to produce a response or prediction.

It is the ability to learn that distinguishes artificial neural networks from conventional models, and thus, makes them attractive tools for assessment, classification and recognition (California Scientific Software, 1990, Moallemi, 1991). In addition, neural networks are more robust than serial processors, in that each part participates in the performance of any given task. Should part of the neural network fail, the system will still be able to function, albeit with reduced accuracy.

The development of neural networks is not dependant on particular hardware or programs. It is the architecture and the learning (training) algorithm that determines how the neural network will perform.

B. Brain Functioning

Neural networks try to simulate the brain's ability to learn (Neural networks with Java, 2001). As mentioned previously, information is passed between neurons along the dendrites. If stimulated, appropriately, an output is generated to all connected neurons in order to elicit a response. There are various theories as to how learning occurs. From a biological perspective, the first time information is received, the connection between neurons is changed. These connections remain firm, such that when similar input information is received in the future, these same neural cells will be stimulated and will adapt in accordance to the new information. This is where the learning occurs in biological networks. If related information is not sent to the network for a certain period of time, the already established network will not be stimulated, and will become weak. This is analogous to humans forgetting what had once been taught or vague recollections. These networks are changing in order to respond to the various inputs.

C. Characteristics of Artificial Neural Networks

Neural networks are characterized by the following (Fausett, 1994):

- architecture: the pattern of connections between neurons
- training / learning algorithm: the method of determining the weights on the connections
- activation functions: the internal state of the neuron.

Neural nets consist of the following components:

- *Nodes (units):* Neural networks consist of layers of nodes or units connected to one another. These are analogous to neurons in biological networks. *Input neurons* receive information and send it to the *inner (hidden)* layer of neurons. The *output neurons* produce the network's response to the input. Connections are how neurons communicate with one another. They can be inhibitory or excitatory. Inhibitory connections prevent the firing of neurons, while excitatory connections enhance firing. It is the connections that determine how neural networks operate. For example, some models can be constructed with a feedback mechanism whereby connections can go from the output of one layer to the input of a previous layer. The architectures of artificial neural networks will be discussed in a future section.

- *Weights:* Weights reside on each connection at the input of a neuron. Weights are analogous to synapses in biological networks and represent the information required to solve the problem (Fausett, 1994). It is these weights that control the strength of the incoming signal and multiply the output of the sending neuron (Elman et al., 1998). A neuron will calculate the output by finding the weighted sum of all of the inputs. At any point in time, the neuron adds up the weighted inputs to produce an activation value.

- *Activation values*: An *activation function* dictates the neuron's response to the signals based upon the weighted signals. It signifies what the node actually does. After the inputs are summed, the net input of the neuron is combined with the previous state in order to produce a new activation value. Inhibitory signals, as mentioned before, prevent the firing of neurons, and as such, will have a negative value that reduces the activation value. This activation is passed through an output or transfer function, which will produce the actual output. The most common activation function is a sigmoid function (Malmgrem, 2000) as their computational and mathematical properties are relatively straightforward. In addition, sigmoid functions are seen in biological neurons.

- *Transfer functions:* Transfer functions produce the output and are considered the response function (Elman et al., 1998). There are various types of transfer functions. One type is a *threshold function*, which is an all or nothing function. If the activation is greater than a fixed amount, the neuron output will be a 1. Conversely, if the activation is less than the fixed amount, the output will be a 0. Another type of transfer function is a *saturation function*, where by once a maximum has been reached, additional excitation will have no effect. A *sigmoid function* has a high and low saturation limit and range in between. The sigmoid function will be 0 if the activation value is a large negative number or 1 when the activation value is a large positive number. Regardless of the type of function, a neuron will not have any impact on other neurons unless the activation value has a minimum value.

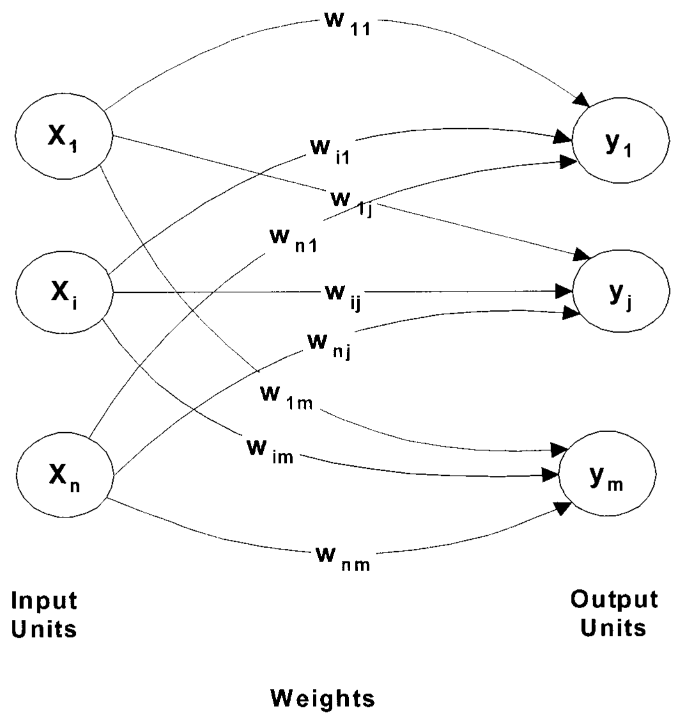
D. Operation of Neural Networks

I. Architecture

The architecture represents the pattern of connections between neurons. Neural nets are often classified as single or multi-layered (Fausett, 1994). Input layers are not counted when determining the number of layers because they perform no form of computation. It is the number of layers of weighted connections that determine whether a net is single or multi-layered.

As illustrated by Figure II.1.1, *single layer networks* have one layer of connection weights (Fausett, 1994).

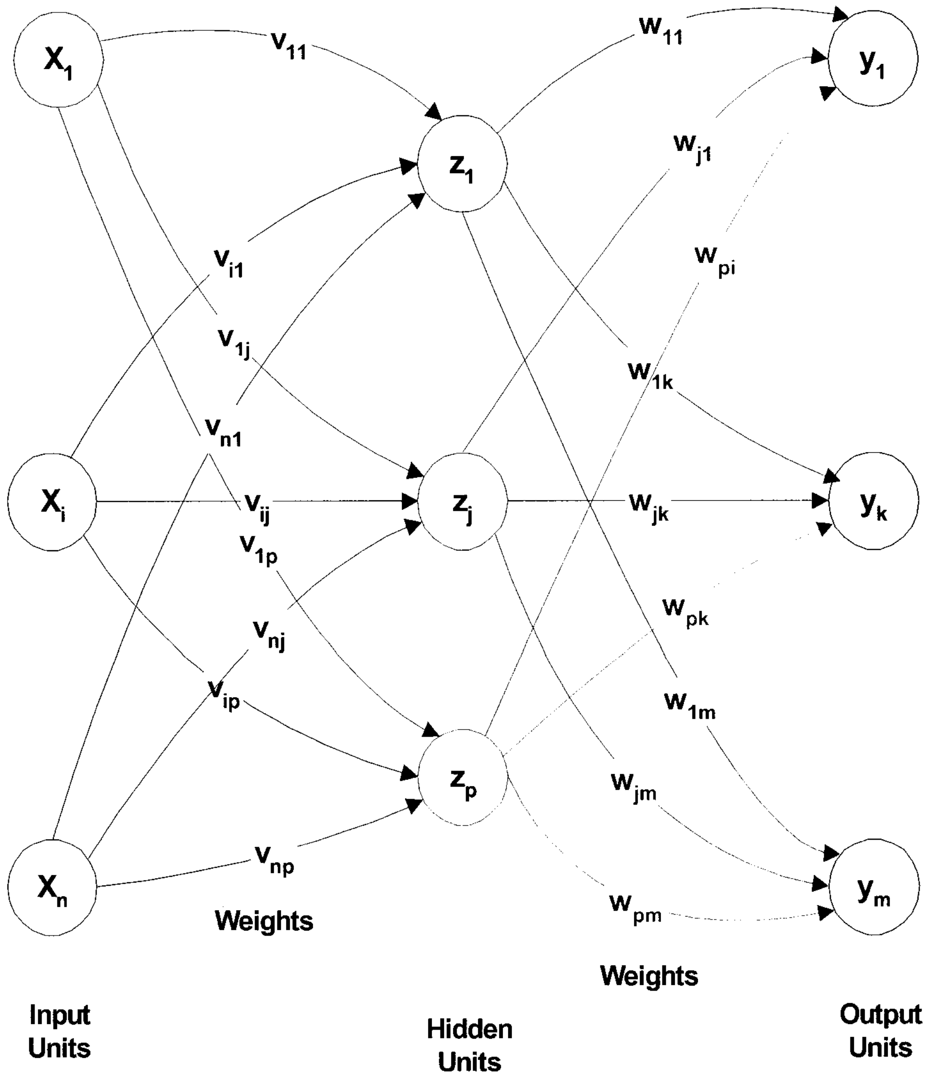
Figure II :1.1: Single layer network



Here, the input units are connected to output units, but are not connected to other input units. In addition, the output units are also not connected to other output units and the weights for one output do not influence the weight for another output.

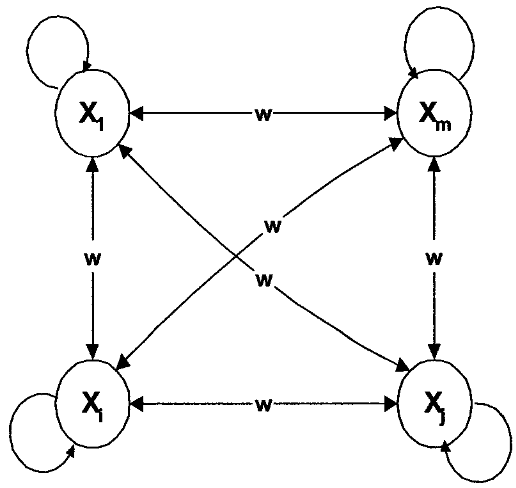
With *multi-layer nets*, (Figure II. 1.2), there are one or more layer of nodes (i.e. hidden units) between the input and output units (Fausett, 1994). These nets can solve complicated problems that single layer networks cannot. However, training the network is more difficult.

Figure II 1.2: A multi-layered neural network



With *competitive layers*, the interconnections between are not shown in the architecture. Here, a group of neurons compete for the right to become active (Fausett, 1994). Figure II. 1.3 illustrates a simple competitive neural network.

Figure II:1.3: Competitive neural network



II. Learning (setting the Weights)

Setting the weights is the method by which neural networks are trained and distinguishes neural networks from other tools (Fausett, 1994). These weights are made up of units, which change based upon the inputs. It is the value of these weights that determine how artificial neural networks learn. To summarize, when information is received by the input neuron, it is passed to connected neurons, provided that the appropriate threshold has been reached. The value of the weights should be increased if information is to be sent to the output neuron or decreased should it not to be sent to the output neuron. The weights are adjusted accordingly until a desired value has been attained, such that a given input will result in the desired output.

In order to design artificial neural networks, there must be a learning algorithm. These learning rules are algorithms to determine the appropriate weight (connection) changes required for the artificial neural network to learn (Malmgren, 2000). The learning algorithm defines the neural network's ultimate learning capabilities when presented with novel or unfamiliar inputs. Malmgren defines these weight changes as analogous to the brain's long-term memory, whereas the activation of a unit is analogous to the short-term memory.

The most common training methods are supervised, unsupervised and self-supervised training (Fausett, 1994).

Supervised learning: Here, the teacher or creator of the network provides the inputs and the desired outputs for training examples. Training is performed by presenting a series of training vectors or input patterns. Each input is associated with a desired output (Malgrem, 2000, Fausett, 1994). Weights are adjusted based upon a learning algorithm (e.g. Hebb rule or delta rule). Although used in various artificial neural network models, the fact that they require a teacher makes them restricted in terms of their applications (Elman et al., 1998). There have been attempts to overcome these restrictions.

One method of supervised training is "auto-association" (Elman et al., 1998). Here, a network is given an input and is trained to reproduce the same input pattern as the output. In order to do this, the network looks for the lowest representation of the inputs (i.e. the easiest common element amongst inputs). The teacher becomes the input.

Another form of supervised learning is the "simple recurrent network" (Elman, 1990 as cited by Elman et al., 1998). This network has predictive elements. They contain connections from the hidden units to what Elman describes as context units. The context units store the hidden unit activations for one step and feed them back to the hidden units on the next step. This ensures that the hidden units have a record of prior activation, thus enabling them to perform tasks over time. In essence, the target output is the next input and therefore, requires no teacher. This is called self-supervised learning.

Another form of supervised learning / training is called "reinforcement learning" (Elman et al., 1998). Here, the network is not instructed if the output is right or wrong, but rather is given a scalar value (i.e., the network is getting closer or farther away from the correct output, not an absolute value). The problem with this form of learning is determining the error, and makes this form of learning slower than the other forms of supervised learning.

Unsupervised learning: this form of learning uses no external teacher and relies only on local information (Stergiou and Siganos, 1996). They use similar input vectors without the use of training data (Fausett, 1994). No target output is specified. The net modifies the weights so that the most similar inputs are assigned the same outputs. There are various types of unsupervised learning (competitive learning, feature mapping, and multiple mapping. They often generate novel and interesting categorisations of the data.

III. Activation Functions

An *activation function* dictates the neuron's response to the signals based upon the weighted signals. Artificial neural networks sums the weighted signals and applies an output or activation function. Summing the weighted inputs is often the identity function [e.g. (f(x) = x for all x]. Mathematically, the most common activation functions, primarily amongst multi-layer networks are represented by non-linear functions (Fausett, 1994).

Fausett identifies these common methods of converting the net input:

- *Binary step function*: Single layer nets use this function to convert a continuous net input to an output, which is binary (1 or 0) or bipolar (1 or –1).
- *Sigmoid functions*: these are s-shaped curves often used for neural nets trained by back-propagation.
- *Logistic function (binary sigmoid)*: this is a sigmoid function with a range from 0 to 1. These are used as the activation function for neural networks where the desired output values are either binary or in the interval 0 to 1.
- *Bipolar sigmoid*: this function permits a range of values appropriate for any given problem. The most common range is –1 to 1.

E. Types of Simple Neural Networks for Pattern Classification

Many of these simple networks were designed for pattern classification, where each input belongs or does not belong to a particular category (Fausett, 1994) and that the set of training patterns for a particular category is known. Here, the output unit is 1 or 0 (or –1). An example Fausett uses is with EKG, whereby the output is "normal" or "abnormal". The artificial neural network, when presented with a novel input, would produce an output response of normal or abnormal, based upon the information or patterns for which it had been previously trained.

I. Architecture of simple networks

The architecture is made up of a layer of input units and a single output unit. They also consist of a bias or threshold value (Fausett, 1994). A bias acts as a weight on a connection where the activation is always 1. Increasing the bias increases the net input. Including a bias, the activation function is as follows:

$$f(net) = \{1 \text{ if net } \square\ 0 \text{ and } -1 \text{ if net } \square\ 0 \text{ and net} = b + \square\ x_i w_i$$

In order to illicit a positive response from the output unit, the net input (i.e. b+ $x_1w_1 + x_2w_2$) must be greater than zero.

Another feature of simple nets is a concept called linear separability. Any input that belongs to a class where the desired output is "yes" would be represented by an output class of 1. A "no" response would conversely fall into the output class of –1 (Fausett, 1994). A response of yes or no would have a step function as its activation function. Represented graphically, all yes responses would lie on one side of a decision boundary and no responses on the other. This is called linear separability, for which Minsky and Papert (1988, as cited by Fausett, 1994), showed to be the only type of problems that simple neural networks can solve.

The three most common simple networks (Hebb, Perceptron and ADALINE) are illustrated below. As these form the foundation of other neural networks, they will be discussed in greater detail.

II. Hebb Net

The most attractive feature of neural networks is their ability to learn (California Scientific Software, 1990; Stader, 1992; Elman et al., 1998). Learning occurs when strengths of the connections are modified during the training process. As these authors state, most neural networks use some form of Hebbian learning to adjust the strengths of the connections. Hebb proposes that learning is driven by correlations and that the connections between units become stronger if the units' activities are correlated.

With respect to biological systems, Hebb proposed that biological associative memory lies in the synaptic connections (weights) between nerve cells (nodes). When pairs of neurons are activated simultaneously, the connection becomes stronger and reinforced. Hebb's theory has its advantages to modelling theory in that correlations are easy to determine when training neural networks. The limitations occur when one wants to go beyond simple pair-wise correlations. Other learning methods are therefore required.

The following represents the algorithm to train a Hebb net (Fausett, 1994):

1. Initialise the weights: $w_i = 0$ ($i = 1$ to n)

2. Set the activations for the input units: $x_i = s_i$, (i = 1 to n)
3. Set the activation for the output unit: y = t
4. Adjust the weights: w_i (new) = w_i (old) + $x_i y$, (i = 1 to n) or $\square w = xy$
5. Adjust the bias: b (new) = b(old) + y

As this simple pattern is used as a basis for more complex networks discussed later, an example is presented. The following is an example of the training process for the AND function and binary inputs and targets (1 or 0):

Input	Target	Weight Changes	Weights
(x_1, x_2, 1)		($\square w_1$ $\square w_2$, $\square b$)	(w_1 w_2, b)
(1, 1, 1)	1	(1, 1, 1)	(1, 1, 1)
(1, 0, 1)	0	(0, 0, 0)	(1, 1, 1)
(0, 1, 1)	0	(0, 0, 0)	(1, 1, 1)
(0, 0, 1)	0	(0, 0, 0)	(1, 1, 1)

Step 1: Set the weights and weight changes:

Weight changes = input x target ($\square w_1 = x_1 t$, $\square w_2 = x_2 t$, $\square b = t$) *e.g. for* $\square w_1 = 1, 1, 1$

New weights are the sun of previous weights. For target 1, weights are 1,1,1 since previous weights before learning were 0.

Note: Since the target is 1, for a binary system, this output would be considered "on".

Step 2: Present the remaining training inputs. Since the target is 0, no learning occurs and prevents the net from learning ay pattern when the target is "off".

Substituting the above 0 for bipolar targets (-1), one would perform the same computations. However, given the bipolar nature of 1 and –1, modification of a weight when the input unit and target values are both "on" at the same time and "off" at the same time means that all units will learn, whether there is an error.

III. Perceptron

This method was devised in order to overcome the limitations with pair-wise correlations associated with Hebbian learning. The perceptron was one of the earlier artificial neural networks and was designed for supervised pattern classification (Malgrem, 2000). Each perceptron has a threshold value and every time an input pattern is received, a net weighted value is calculated and is compared to a threshold (Neusciences.com). This procedure is simple in that there are no hidden units and only one output unit, which is connected to all inputs (Stader, 1992). For a binary system, if the threshold is greater than the net input (i.e. the net input is less than the threshold), a zero value is given to the output, otherwise a 1 is given. An error value is produced if the output is not equal to the desired or target output. Weights are adjusted until the desired output is achieved.

The goal in this supervised form of learning is to minimize the error. It specifies the manner by which inputs will generate the appropriate outputs and appropriate, defined weights will be adjusted automatically (Elman et al., 1998). At the end of the network training, the weights will yield the correct output for any input.

Weights are updated using a simple algorithm (Stader, 1992) that is a variant of the Hebbian learning rule (California Scientific Software, 1990). If the output is correct, the weights are not adjusted. If the output is not correct, the weights are adjusted appropriately.

Although the perceptron was developed as a means to address the limitations of Hebb's rule, limitations still exist. It works for simple two-layer networks of inputs and outputs (Elman et al., 1998; Stader, 1992) and can only generalise based upon physical similarities or similarities of input pattern. If the perceptron does not have a similar input-output pattern when presented with a novel input, it will not generate the desired output. There have been attempts to expand upon the theory (e.g. the Least Mean Square solution, Gaussian classifier, etc.), but these limitations enabled researchers to focus on other areas of neural network research.

IV. ADALINE

The ADALINE usually uses bipolar activations for its input and target output (Fausett, 1994). It is a single unit that receives input from several units. The weights connecting the two are adjustable and there is only one output unit. The ADALINE can be trained using the delta (or the Least Mean Squared) rule. This rule states that the weight change □w_{ij} is dependant on the difference between the actual output (x_j) and the desired output (t), (Malgrem, 2000).

$$\square w_{ij} = k\, x_i\, (t - x_j)$$

During training, the activation function is its net input. Therefore, the learning rule minimizes the mean squared error between the activation and target value and thereby allows the network to continue learning on all training patterns, even after the correct output value has been reached.

V. Pattern Association

When referring to the human mind or memory, pattern association refers to process of forming associations between patterns that are similar. When presented an input pattern that is similar, but not exactly the same as a previous stimulus, the same response pattern can be generated, even if the new input is not the same as the previous input, but similar. For example, initially presented with a particular word of a particular height or size on a written page, will generally elicit the same response when presented in italics or smaller size when viewed on a different book, paper etc.

One of the limitations of conventional serial processing computers is the inability to carry out pattern association. In conventional systems, information is stored and retrieved in storage addresses whereas neural networks have information distributed throughout the systems in the weights.

Associative memory neural networks are single-layer nets where the weights are set up such that it can store a set of pattern associations (Fausett, 1994). Each association is an input-output vector pair (e.g. s:t). *Auto-associative* memory is a network where each vector t is the same as the vector s. Here, an auto-associative neural network can be used to determine whether an input vector is known or unknown. *Hetero-associative* networks are where the t's are not the same as input s with which they are associated. In these networks, they learn and are able to recall the desired response for given inputs, which are identical as well as those inputs, which are similar, but not identical to the training input.

Training algorithms for pattern association can use the Hebb rule (as illustrated above) or the Delta rule. One limitation about neural networks that use pattern associations is when there could be more than one possible output as its correct response to the input. This is solved by the concept of competition, which is discussed in the next section.

One example of pattern associations is *Hopfield nets*. These are unstructured networks (Stader, 1992) in that they are not organized into layers. They are made of simple neurons (California Scientific Software, 1990). The connections between neurons are symmetric. The advantage of such networks is in their abilities to store memories or patterns similar to the human brain (Nethttp://brain.web-us.com/brain/neur_hopfield.html#hop). If partial input is presented to the network, it has the capabilities of producing a full output. If there is a connection from input a to output b, then there is a connection from b to a as well (Stader, 1992). The connections can be excitatory or inhibitory and are trained by self-supervised learning. They can recognise patterns by matching new inputs with previously stored (learned) patterns (California Scientific Software). When an input is presented to the network, the output will be the closest response, based upon the stored patterns that the network has learned.

Similar to the previous training methods, Hopfield networks must be presented a series of training data. Patterns are presented repeatedly. The actual output is compared with the target output and the weights or connections are adjusted accordingly, until the error between the actual and target output is minimised.

Once the training has occurred, the network will be able to recall the training patterns, and will therefore be able to produce appropriate outputs with new input information. As mentioned previously, if a new input is presented for which the Hopfield does not have a corresponding output, a similar output will be produced (similar to memories).

F. Neural Networks and Competition

To overcome the limitations within pattern associations, an additional structure can be included such that the network is forced to make a response. This structure is called competition (Fausett, 1994). These networks combine competition with some type of learning to adjust the weights. The most common form of competition is the "*winner-takes-all*", where only one neuron in the competing group will have a non-zero value. Here, the neuron with the largest activation will remain "on". If target values are used for the input training patterns, the training is supervised. In unsupervised training, the network will look for patterns in the input data.

The *Mexican Hat network* uses a concept called "on-centre-off-surround" (Fausett, 1994). Here, each neuron is connected to excitatory links (positive weights) to some "cooperative neighbours" that are in close proximity to these links. Each neuron is also connected to inhibitory links (negative weights) to a number of "competitive neighbours" (neurons that are far away). Therefore, the positive reinforcement from nearby units and negative reinforcement from those farther away increases the activation of the excitatory weights.

Kohonen's self-organizing feature maps are networks that are also based upon competition. However, unlike the Mexican Hat, these utilise unsupervised training. They are patterned after the topographic maps within the brain, where each area specialises in different items (Stader, 1992). They use simple adaptive neurons that receive signals from an event space (California Scientific Software, 1990). They simulate the brain whereby aspects of sensory input are represented in two-dimensional areas (Stader, 1992). The brain receives a wide array of information. In order to simplify these images, the brain must be able to generalise or classify the information. Units are connected so that there are excitatory connections between nearby units and inhibitory connections between distant units. Here, all units will respond to an input. The network is trained with an unsupervised winner take all strategy. The unit with the highest response is located and the weights are adjusted accordingly. The neighbouring units will also have their weights updated.

In the Kohonen network, all units respond randomly to the input. They use simple adaptive neurons that receive signals from an event space (California Scientific Software, 1990), which contains a wide array of data. As mentioned previously, to process information, the network requires a method to simplify images and relate them to previously stored or learned information sets. This is done by generalisation or classification.

I. Backpropagation Neural Networks

This is a form of supervised learning that requires feed forward neural networks (www.neusciences.com). All networks discussed previously were single layered networks. Back-propagation networks are multi-layered. It is a gradient descent method that minimises the total squared error of the output (Fausett, 1994). Here, inputs are mapped to a given set of target outputs. Similar to other neural networks, the aim is to train the networks to respond correctly to patterns that are trained and the ability to yield outputs or responses to input that are similar.

Multi-layered neural networks have input neurons, hidden layers and output units. The network is trained with a series of input patterns (the feedforward component) that passes to the hidden layers, which pass to the output units to produce a given response. Any error is adjusted backwards and the weights are adjusted accordingly.

Given that this is the first discussion of multi-layered networks, the following represents the steps required to train with a network with a back-prop algorithm (www. Neusciences.com):

1. *Randomly choose a pair from the training data.* The input pattern of the pair is given to the input layer of the network. Each signal of the pattern is assigned to one neuron on this layer.

2. *The network passes these signals to the neurons on the hidden layer.* For each neuron on the hidden layer, a net input value is calculated:

   $NetL_{pi} = a * O(L\text{-}1)_{pj} * W_{ij}$

   P = index of a pair of patterns
   $NetL_{pi} =$ net input of neuron I on layer L
   $O(L\text{-}1)_{pj} =$ output of neuron j on layer L
   $W_{ij} =$ weight of the connection

3. After all neurons in the pattern have received a net input, an *activation value (output) is calculated.* Each hidden unit computes its activation and sends the signal to the output unit. As mentioned previously, the most common activation function is the sigmoid function.
4. *These outputs are passed to the next layer* and the same process of computing net inputs and activations are performed until the output layer of the neural network is reached. The output values of the neurons on the output layer are taken as one pattern or signals, which therefore represents the actual output.
5. The *actual output is compared to the target output.*
6. An *error value is computed*, which is the difference between the actual and target output, squared.

If the error is zero, no changes are made to the network. If the error is not zero, changes are made to the weights of the connections to reduce the error.

The method of error reduction is where the name back-propagation arises. The error is sent backwards through the network at each layer. The error is distributed at the output unit back to the previous hidden layers. The error is not propagated back to the input layers, but is used to update the weights between the hidden layers and the input layers.

Once the networks have been trained, it is usually validated by using a set of input/output patterns similar to the training set used.

G. Applications of Neural Networks

There have been quite a number of neural network applications cited in the literature. Given that the focus of this paper is geared towards health care and the medical profession, this section will focus on applications within health care.

Artificial neural networks have been applied as predictive and diagnostic tools. Lee and Park (2001) have demonstrated its use to classify HIV versus AIDS patients. They used a three-layer, back-propagation architectural model. The input layer (1^{st} layer) contains 15 nodes and is multiplied by computer generated random numbers to create the hidden layers (layer 2). The net activation of the hidden layer, become input values of the output layer, which calculates HIV status or AIDS status.

Neural networks have also been used as a means to predict the risk of lymph node spread (Bowser, 2000). He states that an artificial neural network is 98% accurate in classifying localised prostate cancer patients at being low risk for lymph node spread. Although not intended to replace traditional methods of clinical assessment or diagnosis, he states its use as a clinical tool to measure the percentage of spread (or "no spread").

Artificial Neural Networks have also been used to detect abnormal fetal heart patterns (Lee, Ulbricht, Dorffner, 1999). Lee, Ulbricht and Dorffner used a recurrent, artificial neural network with hidden layers and compared it with conventional methods currently used (cardiotocography CTG), which are considered labour intensive in that they require intensive physician monitoring, are subject to user variability in terms of results and unable to adopt or learn from new or unusual circumstances. They state that their artificial neural network was able to learn and adapt to various circumstances to detect fetal abnormality significantly better and faster than the CTG.

Sordo (2002), in a document prepared by Open Clinical outlines some of the commercial uses of artificial neural networks. Papnet is a commercial neural network that assists in the reading of Pap smears (cells taken from the cervix to assess the presence of pre-cancerous or cancerous cells. Early detection allows for the successful treatment of cancer, however, detection is reliant on human's analyses under a microscope. To minimise human errors, Papnet assistance has resulted in more accurate screening process, thereby resulting in earlier detection of cancerous cells.

Another clinical diagnostic tool described by Sordo is Entropy Maximization Network (EMN), which predicts metastases in breast cancer patients. Characteristics that are fed into this artificial neural network are the age of the patient at the time of the diagnosis, mitotic count (the number of hyperchromatic nuclei in the invasive tumour), size of the primary tumour nuclei, etc.

To illustrate the depth of use of artificial neural networks within health care, Sordo also provides examples for image analysis and interpretation. As artificial neural networks can be constructed to recognise patterns, their use as imaging tools within the health care is strong for X-rays, CT scans etc. Many imaging results are dependant on the subjectivity of the person analysing the results, albeit with many years of experience. Artificial neural networks can minimise this subjectivity. Aizenberg et al (2001, as cited by Sordo, 2002) used neural networks to improve resolution in brain tomographies for the detection of microcalcification in mammographies.

Burke et al (1997), has used artificial neural networks to improve the accuracy of predicting cancer survival rates. They have demonstrated that artificial neural networks were more accurate in predicting the five-year survival rates for breast and colerectal cancer than the traditional TNM staging system. They used a multi-layer artificial neural network (NevProp) with back-propagation training. Their data was obtained from the American Colleges of Surgeon (ACS) tumour registries in the United States. The artificial neural networks were significantly better predictors of the five-year cancer survival than the standards TNM staging system. This is because the network was able to include prognostic factors not captured within the TNM such as demographics, menopausal characteristics, pre-surgery, post-surgery, etc. which makes the artificial neural networks powerful predictors.

Artificial neural networks have started to gain prominence within psychiatric diagnoses. Hatzakis et al attempt to classify five main psychiatric diseases using PSE – CATEGO categories (Present State Examination). They think that artificial neural networks are better able to diagnose psychiatric illnesses when compared with the classic method of decision trees or expert systems. However, neural networks are not widely used since it has been difficult to demonstrate their advantages over the traditional techniques.

Although it has been illustrated that artificial neural networks are powerful diagnostic predictors within healthcare, their use and benefits must be appropriate for the actual application. (Malgrem, 2000), describes two problems common with artificial neural networks - “not coding the data in the best way” and using “too powerful networks”. Not coding the data accurately will obviously result in incorrect or misrepresented outputs. Using too powerful networks can fail if there is limited data available, which may therefore result in the network’s inability to generalise.

H. Expert Systems

Expert Systems are a form of artificial intelligence that allows modelling of information at complex levels (http://www.ghg.net/clips/ExpertSystems.html, 1997). These systems are developed to emulate human logic and expertise. They provide expert quality advice, diagnoses and recommendations given real world problems. Rule-based programming (i.e. If-Then logic) is the most common method of developing expert systems. The *if* portion of the rule indicates the facts required to solve the problem and matches these facts against patterns. The *then* portion of the rule is the set of actions required.

I. Architecture

A *User Interface* is the method by which the user interacts with the system (http://www.cee.hw.ac.uk/~alison/ai3notes/subsection2_5_2_1.html). It could be through the use of menus or natural language. An *interface engine* is used to reason by the expert knowledge (i.e. the if-then logic) and the specific data required to solve the problem. The data includes data obtained from the user and partial conclusions. Almost all expert systems also have an *explanation subsystem*, which allows the program to explain its reasoning to the user. Some systems also have a *knowledge base editor,* which help the expert or knowledge engineer to easily update and check the knowledge base. Expert systems have the capability of separating domain specific knowledge from more general-purpose reasoning and representation techniques.

Chapter 2: Psychiatric Disorders

A. Diagnosing Mental Illnesses – Overview

Although there exist many tools to diagnose psychiatric disorders, professionals often debate as to the correct diagnosis as well as the most effective treatment for the psychiatric population. Psychiatrists are responsible for providing a final diagnosis for patients. Much of their diagnosis is based upon interviews, medical histories, etc. Structured diagnostic interviews are superior in their diagnostic accuracy than non-structured methods such as the simple clinical interview.

There have been various studies developed to improve the accuracy of diagnosis. Basco, Bostic, Davies et al (2000), studied the effect of adding a more structured dimension to improve diagnostic accuracy for outpatients with severe mental illnesses. By adding extra layers to the diagnostic process, their study revealed an improvement in the accuracy of psychiatric assessment. They state that the inclusion of psychiatric nurses and other non-physicians involved in mental health improves the thoroughness and accuracy of psychiatric diagnosis. These added elements, are defined as the Gold Standard evaluation. Elements involved in the diagnostic process included interviews using the Structured Clinical Interview for the DSM-III-R (SCID) and other diagnostic measures, the medical records which contained information such as progress reports, physicians' orders, hospital admission etc. In addition to assessing the diagnosis at each process of the study, the time required to diagnose the disorder at each stage was also calculated. Their study revealed that adding layers to the diagnostic process improved the accuracy. This demonstrates that psychiatrists can arrive at accurate diagnoses with the assistance of SCID and medical record results, provided that competent and available staff is used. The average time for psychiatrists to arrive at a diagnosis was also less than the above, mentioned gold-standard method, given that other personnel were able to provide the initial screening with the SCID and medical records.

The above has shown that while using a multi-layered approach can improve psychiatric diagnosis, it can still be time consuming. In the Basco, Bostic, Davies et al study (2000), they recorded the length of time of each diagnosis during this four-layered approach. The time to administer the SCID and medical records was averaged at one hour and 44 minutes. The chart review took an average of 42 minutes. The average time for the follow-up interview leading to the gold standard diagnosis was 40 minutes. While this may improve the accuracy of diagnosis, it is still time consuming and relies heavily on the availability of personnel at each diagnostic level. Basco et al even states that the 40-minute gold standard may still be insufficient for a large part of the psychiatric population.

Often, when the initial diagnosis is inaccurate, changes in diagnosis and treatment can ensue. In the Basco, Bostic, Davies et al study, the original patient's psychiatrists not involved in the research had asked to see the results of the study. In 50% of the cases, feedback had led to a change in patient care. Had proper tools been used initially, a correct diagnosis could have been made, and patients would not necessarily be subjected to many forms of treatment or care. Correct treatment could have been provided at the onset of diagnostic discovery, provided one believes in the notion of proper diagnosis and correct treatment.

There are a variety of diagnostic tools available, many of which will be discussed in a future section. However, even with the multitude of tools available, determining psychiatric diagnoses is often not a straightforward process. There are many factors that vary amongst the psychiatric population that makes diagnosis a difficult process. As illustrated above, using a multi-layered process can be time consuming and may result in the delay of psychiatric treatment, which could ultimately impede the patient recovery. The development of a tool to expedite the diagnostic process could ensure that treatment be delivered in a timely fashion so as to ensure the well being of the patient and society as a whole.

The following section will discuss the current standard for diagnosing mental illnesses.

B. The Diagnostic and Statistical Manual of Mental Disorders (DSM)

The current standard for diagnosing mental illness is the Diagnostic and Statistical Manual of Mental Disorders, revision IV (DMS-IV), developed by the American Psychiatric Association. The DSM was originally developed to classify mental disorders in order to collect statistical information (DSM-IV, 1994). The original version (DSM-I) was developed in 1952, with subsequent revisions made to include medical nomenclature and enhance its use as a clinical, educational, research and theoretical tool. The DSM-IV is the current diagnostic tool used for professionals diagnosing psychiatric ailments. It is used by professionals in a wide variety of disciplines and occupations. It should be noted that a DSM-IV diagnosis is usually applied to an individual's current condition and not previous diagnoses.

Other classification tools exist such as the International Classification of Diseases (ICD), published by the World Health Organization. As much of the literature reviewed in this thesis makes reference to the DSM, more emphasis will be devoted to the review of this manual.

I. Definition of Mental Disorders

The DSM-IV provides a classification of mental disorders, but they admit to there being no strong definition that encompasses the term "mental disorder". As the DSM-IV states, "each is a useful of indicator for a mental disorder, but none is equivalent to the concept, and different situations call for different definitions".

As such, the DSM-IV has maintained the definition used in the DSM-III and the DSM-III-R. In the DSM-IV, each mental disorder is classified as a "clinically significant behavioural of psychological syndrome or pattern that occurs in an individual and is associated with present distress or disability or with a significantly increased risk of suffering death, pain, disability or an important loss of freedom". It also must be considered as "a manifestation of a behavioural, psychological or biological dysfunction in the individual".

II. Summary of the DSM-IV Organization

Each disorder is coded in accordance to the ICD-9-CM. The coding is important for medical record keeping and facilitates data collection, retrieval and compilation of statistical information (DSM-IV, 1994). They are also required in order to report information to third parties, governmental bodies, WHO etc.

In addition, they include the severity of the illness, which is broken down into *Mild, Moderate, Severe, In Partial remission, In full remission* and *Prior history*. Each section (mental disorder) is divided into the diagnostic features, sub-types, recording procedures (i.e. reporting the name of the disorder), associated features and disorders, culture, age and gender features, prevalence, course, familial pattern and differential diagnosis.

In addition to providing specifications of specific diagnoses for particular psychiatric disorders, the DSM-IV also has a multi-axial assessment method to aid in the planning of treatment and to predict outcomes. The DSM-IV includes 5 axes, the first two of which contain the high-level categories of psychiatric disorders documented in the manual. Rather than provide a detailed listing of the components that make up each axis, a summary will be provided. Detailed descriptions for further interest and research can be obtained from the DSM-IV.

- *Axis I: Clinical Disorders and Other Conditions that may be a focus of Clinical Attention*: this axis is for reporting the various disorders or conditions classified in the DSM-IV, except for the personality disorders and mental retardation. These include schizophrenia and other psychotic disorders, eating disorders, etc.
- *Axis II: Personality Disorders and Mental Retardation:* this is separated from the Axis I categories to ensure that special attention be given to their presence as they may be otherwise overlooked. This includes paranoid personality disorder, schizoid personality disorder, narcissistic personality disorder, etc.
- *Axis III: General Medical Conditions:* this is for reporting medical conditions that may be relevant to understanding and treating the psychiatric condition. Some may include complications of pregnancy, childbirth and the peurperium, diseases of the blood and blood-forming organs etc.
- *Axis IV: Psychosocial and Environmental Problems:* these are problems that may affect the diagnosis, treatment and prognosis of the disorders classified in Axis I or Axis II. These may include a negative life event (death), inadequate social support, environmental difficulties etc.
- *Axis V: Global Assessment of Functioning*: this section is for the clinician's overall assessment of the patient. This is used to plan treatment and to predict the possible outcome. As the focus of this thesis also concentrates on treatment, this section will be essential.

C. The DSM-IV Criteria for Schizophrenia and Other Psychotic Disorders

As the focus of this thesis is on Schizophrenia, the DSM-IV criteria for this disorder will be discussed. For information on other psychiatric disorders, reference can be made to the DSM-IV.

Schizophrenia is defined as a disturbance that lasts for at least 6 months and includes at least one month of "active-phase" symptoms (delusions, hallucinations etc.). Other disorders included in this section of the DSM-IV include schizophreniform disorder, schizoaffective disorders and others that are characterized by duration of symptoms or of differing aetiology.

I. Diagnostic criteria for Schizophrenia

The criteria below, has been taken directly from the DSM-IV (1994).

Table II: 2.1: DSM-IV criteria for Schizophrenia

Criteria	Description / Symptoms
a) *Characteristic symptoms*	At least two or more of the following symptoms must be present during a one-month period.
Positive Symptoms	• *Delusions*: abnormality in content and thought and a misrepresentation of perceptions or experiences (examples are persecutory, somatic, religious or grandiose beliefs). • *Hallucinations:* abnormality in perceptions, which can take the form as visual, olfactory, auditory and tactile. Auditory hallucinations are the most common (i.e. the person hears voices). • *Disorganized speech*: the speech patterns are such that the patient may jump from one topic to another, may provide unrelated answers to questions, or incoherent speech (word salad). • *Grossly disorganized or catatonic behaviour*: this can range from childlike silliness to unpredictable agitation, thereby leading to difficulties in performing daily activities (cooking meals, maintaining hygiene etc.). The catatonic behaviour includes a marked decrease in movement, reactivity to the environment and the extreme catatonic stupor (unawareness of the environment).
Negative Symptoms	• *Affective flattening:* this is represented as a loss of emotional expression, reactivity and feeling. It includes poor eye contact, and reduced body language. • *Alogia*: poverty of speech. It is commonly represented by brief, laconic and empty replies. • *Avolition*: the inability to initiate and persist in goal-directed activities. • *Anhedonia*: the difficulty in experiencing pleasure or interest in activities. • *Dysphoric mood (attention)*: the sub-scale contains measures of social inattentiveness, and inattentiveness during mental status testing.
b) ***Social / occupational dysfunction***	One or more major areas of functioning (work, interpersonal relations, hygiene) are below the level achieved prior to the onset of schizophrenia.
c) ***Duration***	Continuous signs of the disturbance should persist for at least 6 months. The 6-month period must include at least one month of symptoms (or less if successfully treated). These may include prodromal or residual periods, which can be mild forms of the positive symptoms specified in Criterion A. Negative symptoms can also be seen during the prodromal or residual phase of the disorder.
d) ***Schizo-affective and Mood Disorder exclusion***	These two features with psychotic features must be ruled out because there are no major depressive, manic or mixed episodes that have occurred with the active phase symptoms and if mood episodes have occurred during active phase symptoms, the total duration is brief.
e) ***Substance / general medical condition exclusion***	The disturbance is not related to physiological effects of a substance (drug abuse, medication etc.) or a general medical condition.
f) ***Relationship to a Pervasive Developmental Disorder***	If there is a history of autism or another Pervasive Developmental disorder, the additional diagnosis of schizophrenia is made only if prominent delusions or hallucinations are also present for at least a month.

II. Schizophrenia Sub-types

Sub-types are determined by the presence of dominant symptoms during the time of evaluation (DSM-IV, 1994). The following are the DSM-IV summaries of the criteria, which must be met in order to meet one of these subtypes.

Table II: 2.2: Schizophrenic Sub-Types

Sub-Type	Description
Paranoid Type	A. Preoccupation with one or more delusions or frequent auditory hallucinations. B. None of the following is prominent: disorganized speech, or catatonic behaviour or flat or inappropriate affect.
Disorganized Type	A. All of the following are prominent: (1) disorganized speech (2) disorganized behaviour (3) flat or inappropriate affect B. The criteria are not met for Catatonic Type
Catatonic Type	The patient must have at least two of the following characteristics: (1) motoric immobility as evidenced by catalepsy or stupor (2) excessive motor activity (that is apparently purposeless an not influenced by external stimuli) (3) extreme negativism (an apparently motiveless resistance to all instructions or maintenance of rigid posture) or mutism. (4) peculiarities of voluntary movement as evidenced by posturing, stereotyped movements, prominent mannerisms, or prominent grimacing. (5) echolalia or echopraxia
Undifferentiated type	Symptoms are met for Criterion A, but are not met for the Paranoid, Disorganized or Catatonic type.
Residual Type	When there is at least one episode of schizophrenia, but without the prominent positive psychotic symptoms. A. Absence of prominent delusions, hallucinations, disorganized speech and grossly disorganized or catatonic behaviour. B. There is continuing evidence of schizophrenia (the presence of negative symptoms or two or symptoms listed in Criterion A).

III. Causes / other features of Schizophrenia

Although there is a wide range of theories available, the causes of schizophrenia are still unknown. The onset of schizophrenia occurs between the late teens and mid-30s. Women tend to have late diagnoses, more prominent mood symptoms and better prognosis. The course of schizophrenia is variable and as such, it is difficult to determine the prognosis for the individual.

The DSM-IV has reported various laboratory findings that point to abnormalities in brain functioning and development as well as subtle neurological abnormalities. Based upon these findings, one would assume the importance of performing biological or neurological tests (discussed later) in order to accurately diagnose a patient with Schizophrenia. In addition, the increased activity of dopamine receptors found in schizophrenics has been at the forefront in the development of anti-psychotic medications used to alleviate the positive symptoms (hallucinations, delusions etc.).

The DSM-IV states that the most common structural abnormalities are enlargement of the ventricular system, and sulci in the cortex. There have also been studies that have found decreased temporal and hippocampal size, increased size of the basal ganglia, and decreased cerebral size. Some schizophrenics have been found to have abnormal cerebral blood flow or glucose utilization. Neuro-physiological studies have found schizophrenics to have slower reaction times, abnormalities in eye tracking or impairments in sensory gating.

While increased dopamine receptors are responsible for many of the positive symptoms of schizophrenia, these are not considered to be the enduring or lasting effects that often debilitate these individuals (Coyle, 1997). The deficits in social functioning have resulted in brain changes that correlate with a reduction in volume of the cerebral cortex, which is responsible for higher mental functioning.

The DSM-IV also states that some schizophrenics are sometimes physically awkward and may display neurological soft signs (left-right confusion, poor coordination etc.). Many motor abnormalities may be associated with treatments using anti-psychotic medication.

Studies also suggest a genetic component to schizophrenia. First-degree biological relatives with schizophrenia have 10 times the risk of that in the general population. There is still uncertainty if there exists a genetic predisposition to schizophrenia or whether interaction with environmental influences could cause the onset.

Chapter 3: Steps in the Diagnostic Process

MacKinnon and Yudovsky (1991), provide an overview of the evaluation process, and will be used as the basis for this section. When referring to a health care professional, reference will be made to the psychiatrist or clinician.

A. The Psychiatric Interview

This is the process where the psychiatrist can first establish rapport with the patient. The psychiatrist or clinician can introduce the concepts of confidentiality, informally explore the patient's past history and assess the anxiety or pain level of the patient. If additional meetings are required, the psychiatrist can advise the patient of such and can introduce various treatment options. Interviews with relatives can also occur in this phase.

B. The Clinical Exam

This contains the psychiatric history and the mental status examination.

I. The Psychiatric History / Evaluation

This is important, as psychiatrists rely upon historical information in order to provide diagnosis and treatment options. This should also include a medical history and physical examination. For psychotic patients, it is recommended that more structure be in place in order to elicit responses. The components required in the psychiatric history are as follows:

- *Preliminary Identification:* this is information written by the psychiatrist regarding the patient's name, age, marital status, sex, occupation, language, race, nationality, religion and residence.
- *Chief complaint*: the problem for which the patient seeks help.
- *History of Present Illness:* this area should include *onset* of the illness, and an evaluation of the patient's highest level of functioning. It should also include the *precipitating* factors that led to the current onset and the *impact* of the illness to the patient and the family.
- *Psychiatric Review of Systems*: a medical review of the patient's medical health from the perspective of the psychiatrist. This could include sleep patterns, weight changes, sexual functioning etc.
- *Previous Psychiatric Illnesses*: prior episodes of emotional and mental disturbances are described.
- *Personal History:* this is information about a patient's past life. Here, the psychiatrist can go through the stages of the patient's life to get a thorough and complete understanding of the patient's history. The psychiatrist will also inquire about the patient's history during the adult life (marital, social, sexual, etc.).

II. The Mental Status Examination (MSE)

This involves an assessment of how well the patient is able to function and process information (Choca and Van Denburg, 1996). They and MacKinnon and Yudovsky describe the following elements that should be included in the MSE.

- *Appearance, Attitude and Behaviour*: this represents the patient's appearance, hygiene, motor behaviour, attitude towards examiner, speech patterns (rapid, slow speech, emotional etc.)
- *Thought Process and Content*: thought process includes the manner of approach, logic, degree of clarity, whereas thought content refers to apparent level of reality, appropriateness of thought and insight. Thought content will also include preoccupations, obsessions, phobias, delusions, etc.
- *Perception:* hallucinations, delusions etc.
- *Mood, Affect and Emotional Regulation*
- *Consciousness*: this includes alertness and wakefulness.
- *Orientation:* is the patient aware of who and where he/she is?
- *Memory, planning and mental control*
- *Impulse Control and Frustration Tolerance:* the patient's ability to control aggressive, hostile, fearful, guilty etc. expressions.
- *Judgment and Insight*
- *Anxiety level*: the level of tension, and nervousness.

C. Testing

There are a variety of test types available to aid in the diagnosis and determination of treatment of psychiatric disorders. The following will focus on those used for psychotic disorders.

I. Biological testing

Electroencephalogram (EEG)

This involves amplifying and measuring the electrical activity in the brain (MacKinnon and Yudovsky, 1991). The electrical activity is evaluated according to the frequency, amplitude and type of brain wave tracing. There are no specific EEG abnormalities in schizophrenics, but abnormal readings are 2 to 3 times that found in the normal population. Schizophrenics have been found to have an increased presence of delta activity (delta activities constitute frequencies less than 4 Hz), and reduction in REM sleep.

Computerized Tomography (CT)

This is formally known as the CAT scan. It is a method whereby variations in the bone density, cerebrospinal fluid, blood, vessels and grey and white matter can be assessed by X-ray, computer processing and photographic material (MacKinnon and Yudofsky, 1991). Lesions and various brain changes can be visualized. It has become a strong source of recognizing psychiatric disorders caused by brain disease. Schizophrenics have shown to have enlarged ventricles compared with the general population. In addition, the size of the anterior hippocampus is smaller for schizophrenics. Weinberger (1987, as cited by MacKinnon and Yudofsky, 1991) has stated that these changes are not the secondary results of medication.

Magnetic Resonance Imaging (MRI)

This is an advanced technique of acquiring brain images (and other images of the body). With respect to brain images, the patient's head is placed in an electromagnetic field, after which the magnetic field is removed. Sensors, computers and monitors organize the electrical changes into images of the brain (MacKinnon and Yudofsky, 1991). They state that white and grey matter in the brain can be differentiated and can indicate the presence of lesions better than most imaging techniques. Although the use of MRI is only starting to receive prominence in the field of psychiatry, it is anticipated that imaging of neurotransmitter defects in schizophrenia can be detected.

Positron Emission Tomography (PET)

This technique makes use of the brain's use of glucose during brain activity. A radioactive glucose substance (F-FDG) is injected into a patient's vein. The patient is placed in a device that can measure the presence and path of F-FDG in the brain (MacKinnon and Yudofsky, 1991). With respect to schizophrenic patients, PET scans have revealed reduced metabolism in the anterior cortex relative to the posterior cortex. There have also been studies that have found lower metabolic ratios in the central grey matter on the left side. The data shows that schizophrenics have abnormalities in the cortical and sub-cortical levels.

II. Psychological Testing

There are a number of methods used to identify and classify psychotic disorders. They can be grouped in the following manner (Peck and Shapiro, 1990).

- *Psychiatric interviews:* these are standardized interview tools by which the clinician can ask questions to the patient. Some of the most common tools are the Diagnostic Interview Schedule (DIS), the Schedule for Affective Disorders (SADS) and the Structured Clinical Interview (SCID).

- *Psychiatric Rating Scale:* as mentioned previously, there are a variety of sub-types associated with schizophrenia, which respond to different treatments, therefore, it is necessary to determine under which type of schizophrenia the patient falls. The most common tools are the Scale for the Assessment of Negative Symptoms (SANS), the Scale for the Assessment of Positive Symptoms (SAPS), and the Brief Psychiatric Rating Scale (BPRS).

- *Psychological Tests*: these tests include self-reporting questionnaires, projective tests and tests that assess perception. The most common tools are the Minnesota Muliphasic Personality Inventory (MMPI), the Wechsler Adult Intelligence Scale (WAIS), the Rorschach and the Thematic Apperception Test (TAT).

MacKinnon and Yudofsky (1991) state that the use of psychological testing has declined over the years. This is primarily a result of the rating scales that attempt to quantify mental illness and the fact that psychological tests can be time consuming. They do not however, discount the use of psychological tests, as they are successful in pinpointing psychopathology sometimes missed using the rating scales or interviews. In addition, the diagnostic tools or rating scales often measure limitations of intellectual functioning. Providing a battery of tests that include the WAIS-R or MMPI-2 can measure the intellectual capacities of the patient.

It should be noted that although there are formal psychometric tools, many psychiatrists rely on their own scales and methods of assessment, as many institutions do not always have the luxury of trained psychometrists. In Canada, much of this often results from funding cutbacks in the health care system, thereby causing staff shortages. Institutions that do have psychometrists to administer psychological tests are also backlogged with a number of requests and consequently, timely and quick assessments are not always possible.

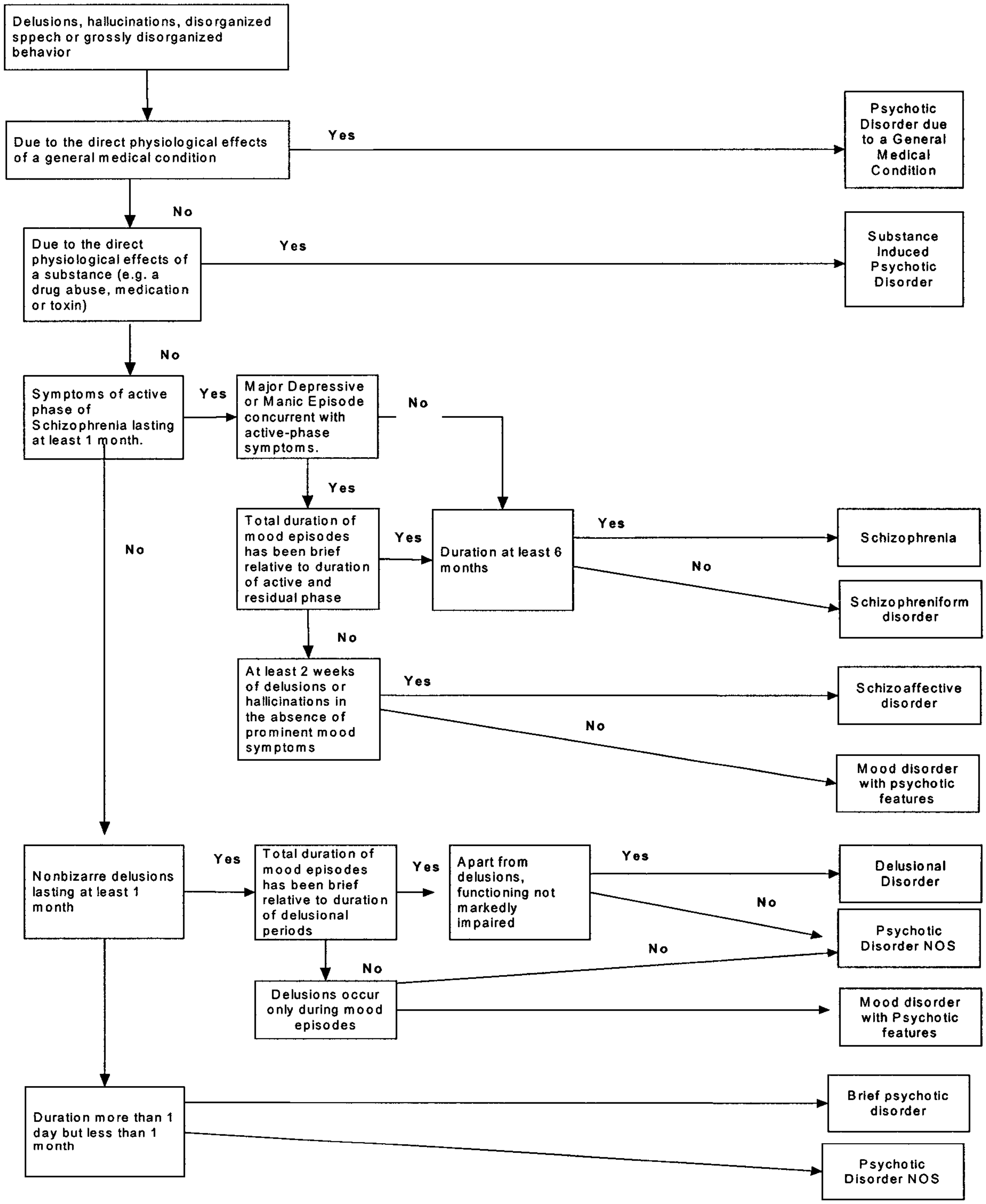
Since much of what is assessed in these tests will form the basis of the artificial neural network input criteria, Chapter 4 discusses some of psychological assessment tools currently used in greater detail.

D. DSM-IV Diagnosis

As specified in Chapter 2, the DSM-IV provides a multi-axial system of assessment, which can assist in the diagnosis process (DSM-IV, 1994). This enables a comprehensive and systematic evaluation of, medical condition, psychosocial functioning and general functioning that otherwise may be overlooked if there was a focus on only one parameter.

In addition to the multi-axial system, the DSM-IV also provides decision trees that can assist the professional in the organization and hierarchal nature of the manual in order to determine accurate diagnosis. The following represents the DSM-IV decision tree for Schizophrenia and its related disorders:

Figure II. 3.1:A Diagnostic Algorithm



E. Treatment

There are a variety of treatment options available for schizophrenics, many in the form of anti-psychotic medication. However, much evidence has indicated the need to provide support to both patients and family as a means to provide a holistic and balanced treatment regiment as well as to treat patients immediately.

Stevens (2001) outlines a balanced method of treating schizophrenia:

a) *Medication*: As there is no cure for schizophrenia, anti-psychotic drugs are the most common method of alleviating symptoms.
b) *Individual and group therapy* may help schizophrenics understand their illness and offer means to cope. In addition, it also offers a setting to improve communication and social skills.
c) *Family education*: this may help families understand and cope with the nature of the illness.
d) *Self-help groups*: these provide support to patients and families.

I. Medication

Anti-psychotic drugs by far are the most effective treatment aimed at alleviating schizophrenic symptoms. McGrath and Emmerson (1999) state that many individuals show a moderate to increased reduction in positive symptoms whereas the negative symptoms are less responsive to treatment. Although successful in alleviating symptoms, there are many side effects that make the use of drugs less appealing to patients and family members. Side effects include tardive dyskinesia, or akthasia (restlessness in the legs and body).

The newer bodies of anti-psychotic drugs have been found to be more effective in treating positive symptoms and are more effective at preventing relapse if the drug usage is maintained (McGrath and Emmerson, 1999). The following represents a summary of McGrath's and Emmerson's and the Harvard Mental Health Letter's (1998) findings of the current anti-psychotic medication, when compared to the previous, traditional anti-psychotics:

Table II: 3.1: Anti-psychotic medication and side effects

Drug (Dose)	Findings / Effects	Side Effects
Clozapine (200-800 mg/day)	• Effective anti-psychotic • Fewer relapses • Greater reduction in symptom • Fewer drop-outs • Greater patient satisfaction • Can relieve deficiencies in memory and attention • Little effect on body movement and can improve tardive dyskinesia	• Sedation • Hyper-salivation • Increased risk of convulsions at higher dosages • Lower blood pressure • Weight gain • agranulocytosis (potentially deadly decrease in the capacity to manufacture white blood cells , therefore patients may require haematological monitoring
Risperidone (2-6 mg/day)	• Effective anti-psychotic • Greater clinical improvement • Little or no effect on positive or negative symptoms • Fewer drop-outs • No risk of agranulocytosis	• Weight gain • Hyperprolactinaemia • Rapid heart beat • Lowered blood pressure • Modest effect on body movement • Raises prolactin levels
Olanzapine (10-20 mg/day)	• Effective anti-psychotic • Lower depression scores • Less sedation • Fewer extra-pyramidal side • effects • Outperformed Haloperidol • Effective for both positive and negative symptoms	• Sedation • Weight gain • Dizziness • Constipation • Little known on long-term effectiveness
Ziprasidone (80-160 mg/day)	• Effective anti-psychotic • Fewer extra-pyramidal side effects	• Sedation • Headaches • Nausea • Constipation
Quetiapine (300-450 mg/day)	• Effective anti-psychotic • Fewer extra-pyramidal side effects	• Dizziness • Dry mouth • Sedation

II. Cognitive and Behavioural Support

The Harvard Medical School (1998) states that while anti-psychotic drugs are effective for alleviating symptoms, they should be coupled with some form of support. This includes providing the patient with an understanding of the illness, the importance of taking drugs on a regular basis, responding to the signs of relapse, support for housing, job searches, social skills, etc. Behavioural techniques, whereby the patients are reinforced for positive behaviours have proven to be effective methods of support. Emotional support is also a new trend. Here, patients are coached on how to monitor thoughts, paranoia, overcome withdrawal tendencies, feelings of humiliation, improve memory functioning, planning and problem solving. Some cognitive approaches such as controlling thought contents have been attempted.

Chapter 4: Assessment Tools

There are a number of assessment tools currently available to enable professionals to make psychiatric diagnoses. Although there are numerous tests available, this section will primarily focus on those given to schizophrenic or other potential psychotic patients, as this is the theme of this thesis. In addition, emphasis will be placed on the rating scales, as these are the tools most frequently used by trained practitioners.

A. Schedule for Affective Disorders and Schizophrenia (SADS)

Purpose: This was developed by Spitzer and Endicott in 1978 (MacKinnon and Yudofsky, 1991). It is an interview assessment scale whereby a clinician conducts clinical interviews and completes a standardized assessment form.

Administration and Summary of Items: The SADS is composed of a 78-page booklet organized into two parts and takes a trained interviewer approximately ½ to 2 hours to complete.

- Part 1: this is designed to provide information about the patient's current condition and psychosocial functioning one week prior to the interview. This part can provide a diagnosis and prognosis as well as to measure symptom change (MacKinnon and Yudofsky, 1991).
- Part 2: this section focuses on historical information about past psychiatric episodes experienced by the patient.

Scoring: The final rating provided does not determine the ultimate diagnosis. Rather, professionals rely on a multitude of other information, such as interviews with the family, case records, etc.

B. Structured Clinical Interview for DSM-IV (SCID)

Purpose: This was developed by Spitzer and Williams in 1985 with a variety of psychometricians, theoreticians, clinicians etc. (MacKinnon and Yudofsky, 1991). It's primary purpose was to provide trained clinicians a tool to make rapid, but valid DSM-III-R (updated to DSM-IV) diagnoses. At the time this test was developed, the DSM-III-R was used as the diagnostic source. It has since been updated to the DSM-IV. It is modelled on the clinic interview and is targeted to clinicians who are trained in the process of clinical interviews, have a basic knowledge of psychopathology and are familiar with the DSM.

Administration and Summary of Items: At the beginning of the interview, the clinician obtains information about the present illness and past episodes. There are a variety of open-ended questions that are developed in order to prompt the patient to describe his / her symptoms. Each axis I diagnostic category is reviewed and the clinician is instructed to follow the questions as written, in order to allow for standardisation.

Scoring: The scoring and interpretation are based upon the criteria used in the DSM-III-R (MacKinnon and Yudofsky, 1991). Each diagnostic criterion is coded as ?, 1,2 or 3.

- Question mark: more information is required.
- Rating 1: refers to whether symptoms described in a criterion are absent or if a statement is false.
- Rating 2: sub-threshold situation whereby the threshold for the criterion is almost, but not met.
- Rating 3: the threshold for the diagnostic criterion in met or exceeded.

C. Scale for the Assessment of Negative Symptoms (SANS) and Scale for the Assessment of Positive Symptoms (SAPS)

Purpose: The SANS, as the name indicates, was developed in 1982 by Anderson (Schutte and Malouff, 1995), to measure the negative symptoms of schizophrenia (affective blunting, alogia, anhedonia, etc.). The SAPS, was developed to measure the positive symptoms associated with schizophrenia, such as hallucinations, delusions etc. These two scales were created because of research findings that suggested the existence of two types of schizophrenia with separate effects, causes and outcomes (Schutte and Malouff, 1995). These two tests are not always used in conjunction, but can be.

Administration and Summary of Items – SANS: For the SANS, there are five scales, each measuring and containing various sub-scales. The following represents the five scales. It rates the negative symptoms listed in the DSM-IV.

1. *Affective flattening or blunting*: loss of emotional expression, reactivity and feeling
2. *Alogia:* impoverished thinking and cognition often associated with schizophrenic patients.
3. *Avolition – Apathy*: lack of energy, drive and interest, not usually accompanied by saddened or depressed affect seen on depression.
4. *Anhedonia:* difficulties in experiencing pleasure or interest in activities.
5. *Attention*: measures of social inattentiveness, and inattentiveness during mental status testing.

Administration and Summary of Items – SAPS: The SAPS contains six scales, measuring a total of 31 items. Schutte and Malouff (1995), describe the following six scales.

1. *Hallucinations:* abnormality in perception, such as auditory hallucinations (voice commenting and voice conversing), somatic or tactile hallucinations, olfactory hallucinations, and visual hallucinations.
2. *Delusions*: abnormality in the content of thought.
3. *Bizarre Behaviour*: schizophrenic's bizarre, unusual or fantastic behaviour (e.g. painting two halves of his body, urinating in a sugar bowl etc.).
4. *Positive formal thought disorder*: fluent speech that communicates poorly.

Scoring: There are three types of scores that can be derived.

a) *Global Score*: the SANS and SAPS contain scales. Within each scale, there is a global rating (i.e. global rating for attention, global rating of alogia etc.) with a score of 0 to 5. This is then summed to provide a "summary global score", with a higher number indicating more severe symptoms.
b) *Symptom Complex Score*: This is the sum for each subscale. Each sub-scale is given a rating of 0 to 5, depending on the severity of symptoms. Summing these individual scores within each subscale leads to a "symptom complex score"
c) *Composite Score*: The sum of the symptom complex score.

D. Brief Psychiatric Rating Scale (BPRS)

Purpose: The Brief Psychiatric Rating Scale (BPRS) as the name indicates, is a brief, but rapid rating scale that provides a global assessment of symptoms for a variety of psychiatric patients (Maruish, 1999). It was developed from clinical evaluations of hundreds of psychiatric patients. Its primary use for schizophrenics allowed this test to gain recognition as a useful clinical research tool.

Administration and Summary of Items: The BPRS can be completed based upon observations in a clinical interview that lasts for approximately 20 to 30 minutes (Maruish, 1999). It is dependant on the clinicians experience and knowledge of the constructs that the scale is rating. The interviewer should spend approximately 3 minutes to develop rapport. This is then followed by 10 minutes of informal non-directive interaction. The last 5 to 10 minutes asks more questions not obtained during the non-directive phase of the interview. It evaluates the following 18 symptoms.

1. *Somatic Concern*: the number and nature of physical complaints or fears or suspicions of bodily illness or malfunction. The frequency and severity of complaints are rated.
2. *Anxiety*: the subjective experience of worry, over-concern, apprehension or fear. Ratings are based upon verbal responses made by the patient.
3. *Emotional Withdrawal:* the ability of the patient to relate in the interpersonal interview situation in order to distinguish from motor retardation and the emotional aspects of withdrawal.

4. *Conceptual Disorganization:* involves the disruption of thought processes, often indicated by confusion, irrelevance, inconsistency, disjointedness, disconnectedness, etc. Ratings are based upon the patient's spontaneous verbal statements and facial expressions.
5. *Guilt feelings:* ratings are based upon the frequency and intensity of reported past feelings of remorse.
6. *Tension:* the physical and motor signs commonly associated with anxiety and do not involve the objective experiences of the patient. The rater should pay attention to signs such as change of posture, sweating, fidgeting, twitching etc.
7. *Mannerisms and Posturing:* includes unusual and bizarre motor behaviour, the severity of which depends on the nature and number. The rating is based upon the unusual nature of mannerisms and not those associated with the tension rating.
8. *Grandiosity:* involves the reported feelings of power, ability importance etc. The ratings should be based upon the current opinions held by the patient and with its foundation in reality.
9. *Depressive Mood:* this should be rated on the basis of expressions, sadness, hopelessness, facial expressions, weeping etc.
10. *Hostility:* the reported feeling of animosity, contempt etc. and ratings are based on the sincerity and affect present during the interview.
11. *Suspiciousness:* when the patient believes he/she has been wronged by others. Ratings are based upon the degree to which the patient tends to project blame.
12. *Hallucinatory Behaviour:* this requires judgment on the part of the rater as to whether or not such behaviour exists.

13. *Motor Retardation:* weakening of motor responses such as speech or voluntary movements.

14. *Uncooperativeness:* signs of hostility and resistance to the interviewer and interview. It is rated based upon the responses to the interview situation (unlike hostility, which is based upon feelings towards others).

15. *Unusual Though Content*: the content of the patient's verbal responses, which are odd, strange, or bizarre. It is the unusualness of responses that are rated, not the degree of organization / disorganization.

16. *Blunted affect:* the reduced emotional tone and apparent lack of normal feelings.

17. *Excitement:* the emotional, mental and psychological aspects of activity.

18. *Disorientation:* the confusion commonly found in psychiatric patients.

Scoring: Each of the 18 items is scored on a 7-pont scale, ranging from not present to extremely severe. The scores are added for each of the items to form a BRPS total score (global severity index) or by forming linear combinations of items based upon factor analytical studies. Maruish states that the global severity index yields minimal qualitative information and should not be used alone. This score is often used as baseline scores in research studies. As such, factor analysis has been applied in order to determine four factors that have emerged in psychiatric samples (thinking disturbance, withdrawal – retardation, hostile- suspiciousness, and anxiety-depression).

E. Strauss – Carpenter Levels of Functioning Scale

Purpose: This test was developed by Strauss and Carpenter in 1975 (Schutte and Malouff, 1995) as a means to assess the functioning of schizophrenics, primarily for the purpose of outcome evaluation.

Administration and Summary of Items: This scale required 9 clinician ratings, each with individual scores. The clinician is asked to describe each item and then rate them, with the associated scale provided. The items are as follows:

1. Duration of non-hospitalisation for psychiatric disorder
2. (A) Frequency of social contacts
2. (B) Quality of social relations
3. (A) Quantity of useful work
3. (B) Quality of useful work
4. Absence of symptoms
5. Ability to meet own basic needs
6. Fullness of life
7. Overall level of function.

Scoring: Sum the responses to the 9 items. The higher the score, the more severe the symptoms.

F. Other Tests

The Wechsler Adult Intelligence Scale (WAIS) was originally developed in 1939 by David Wechsler (Golden, 1990). Originally developed as an alternative to the Stanford-Binet test, its primary purpose was to identify the intellectually challenged and those of higher than average intelligence. A general IQ score is calculated based upon the scores obtained in each section. However, analysing the patterns of scores as opposed to obtaining a general IQ score has enabled professionals to make certain psychiatric diagnoses (Rapaport, Gill and Schafer, 1970, as cited by Golden 1990). The WAIS-R (revised), consists of eleven sub-tests, six of which are verbal and five performance related (Golden, 1990). Administering this test can be time consuming and requires extensive training.

Another test is the Minnesota Multiphasic Personality Inventory (MMPI). The main purpose is to allow professionals to establish a psychiatric diagnosis (Golden, 1990) of personality and other emotional disorders (MacKinnon and Yudofsky, 1991. Scores on the individual scales are associated with various traits and behaviours and allows professionals to form a picture of the individual's behaviours, traits, contact with reality, attitudes etc. This picture and theories of personality disorders enables professionals to establish diagnosis, prognosis, and likely future behaviour (Golden, 1990).

Because of the complexities of the test and the familiarity required, and given that administration and scoring are extensive, time these test will not be used as a component in the development of the Artificial Neural Network.

III. Educational Technology Problem

Given the challenges associated with diagnosis and treatment and the lengthy process from initial diagnosis to treatment selection, there is an educational need to help medical residents make rapid, reliable and valid diagnoses of the various forms of schizophrenia. This would ensure the timely and appropriate delivery of various treatment options. The reasons for the educational need are described below.

As illustrated above, the process of psychiatric evaluation and diagnosis is often a time consuming and tedious process. There are many resources and assessment tools available to enable professionals to arrive at a diagnosis and provide appropriate treatment. When medical residents approach the field of psychiatry, they are taught the principles of medicine in medical school and during their rotations in a clinical setting. Much of their training also comes from trained psychiatrists, many of who have been in the profession for a number of years and thus, have the benefit of experience to guide them towards a correct diagnosis and adequate treatment.

The time it can take for a trained professional to arrive at a diagnosis can be long and can potentially delay the administration of treatment. Psychiatrists require a variety of historical information (medical, familial, etc.), conduct a variety of interviews, etc., before accurate diagnoses are made and treatment options can be discussed. In addition, there is not always a consensus from different psychiatrists with respect to the diagnosis and subsequent treatment option. Although much of the trained professional's task becomes second nature, medical residents usually require a longer time to assess the patient and as such, could benefit from a tool that could assist in their learning. This would therefore ensure that they are learning the tasks and theory necessary to help themselves, and provide the best course of treatment, while not impacting diagnosis and delivery of treatment. As mentioned in a previous section, Basco, Bostic, Davies et al (2000), indicated that adding more layers to the diagnostic process improves the accuracy of diagnosis, however, it is still a more time consuming process. This would therefore delay the administration of treatment.

Although it is the role of teaching hospitals and educational institutions to provide students and residents with the proper tools to function in the health care setting, the teaching should not be at the cost of reducing adequate patient care and well being. Diagnoses will always be required in order to determine the best method of treatment. As mentioned previously, diagnosis can be a time consuming task, which can require many professionals and many tools. In addition, there have been studies that have illustrated that identifying the first episode of schizophrenia is critical in the treatment (Leiberman and Fenton, 2000). As mentioned previously, the first episode of schizophrenia usually occurs in late adolescence or early adulthood. However, it has been determined that there is usually a one to-two year delay between the first psychotic episode and the initiation of treatment. The consequences of such a delay are not quite certain, however, some researchers have suggested that untreated psychosis may result in irreversible brain damage, deterioration and treatment resistant (Lieberman, 1999, as cited by Lieberman and Fenton, 2000). Although there have been studies that have shown otherwise (Hoff et al, 2000), it has been suggested that longitudinal studies will be required to determine the probable brain deterioration when treatment is delayed.

While there is not much evidence of neurological impairment resulting from late detection, it is quite certain that there are social consequences (Lieberman and Fenton, 2000). Maintaining relationships, academics, independent living etc. are affected. In addition, family members suffer the additional strain as caregivers until appropriate diagnosis and treatment have ultimately been provided.

Because of the various permutations and combinations in diagnosing schizophrenia, there are possibilities for minute variations in patient symptoms and characteristics, which can possibly affect and determine appropriate treatment options. In addition, there are a great many treatment options available, which can vary in effectiveness, depending upon the individual in question. These can be chemical or psychological (behaviourist, cognitive etc.). With respect to the chemical treatment, many patients undergo a variety of medication before one is found that will alleviate the symptoms. This trial and error method of treatment is, in the author's view one that is truly detrimental to the well being of the patient and one, which should be avoided, if possible.

With decreased funding allocated for health care, and in particular, for psychiatry, it will be necessary to determine alternate, quick and efficient measures to ensure that diagnosis and evaluation, treatment and support are provided from all caregivers (doctors, nurses, social workers etc.) that will not unduly compromise patient care.

IV. Hypothetical Solution

Chapter 1: Introduction

With the multitude of diagnostic tools and treatment options, it is proposed that the development of a multi-layered neural network application with back-propagation facilities that would be taught to categorise the various forms of schizophrenia and treatment options can be created. The biochemical treatment options best suited for the patient would also be categorised based upon the various side-effects associated with the treatment.

This tool would provide medical residents a strong teaching tool to diagnose and determine the appropriate course of treatment for schizophrenic patients. Given the diverse treatment options currently available and to avoid the long trial and error method to "see which psychiatric medication works best", it is proposed that this neural network would also be a helpful tool for psychiatrists currently in the profession. However, its use amongst psychiatrists would be limited, given the possible legal ramifications.

There are many reasons why such a tool would be beneficial. First of all, the time and expense it takes to teach medical professionals to evaluate the results of psychiatric tests and interview the patients individually and in conjunction with each other is high. It is proposed that a neural network could assimilate the results of all tests and generate an appropriate diagnosis in less time than a human. In addition, a neural network can incorporate the results of these tests, the patient's psychiatric and medical history in addition to those of the family as well as other demographic features. The decreased time to determine a diagnosis would enable the patient to start treatment immediately. In addition, not all trained psychiatrists or clinicians always have the luxury of using the multitude of tools available. As such, pinpointing exact aetiology per patient may be difficult and may be missed, which may result in inappropriate treatment options.

Creating a tool that would allow all areas of assessment, evaluation and diagnosis would permit a stronger patient-centred tool that could ultimately lead to pinpointing an exact course of treatment, thereby eliminating a potential trial and error method or hit-and-miss treatment regiment. In addition, it is hoped that such a tool would minimise the diagnostic and treatment variances that exist between those in the profession.

It is the author's belief that such a tool should be patient centred, primarily where the choice of treatment is concerned. Depending upon the individual inputs (which are determined from the individual patient's characteristics), it is proposed that an artificial neural network such as this one could generate proposals for optimal treatment, based upon the symptoms, patient information and potential treatment side effects inputted. It is believed that this would minimize if not potentially eliminate the need for patients to try various treatment options in order to determine the "best medication fit".

In addition, it is the neural network's ability to process information in parallel that makes such a tool valuable for the diagnosis and treatment. Processing a multitude of data simultaneously is more efficient and faster than the regular human mind and would therefore ensure that treatment commence sooner than otherwise would with traditional techniques.

It is not the intent of this thesis to develop the actual neural network. Rather, it will focus on the components required (inputs, training etc.) and the sources by which these components will be derived (rating scales, DSM, etc.). In order to determine the relevant components, a "Diagnosis and Treatment" system will be developed based upon the literature review presented. This system will look into the flow and link from when the patient enters a health care facility, undergoes psychiatric testing, receives a diagnosis and finally, appropriate treatment. A Patient System has already been described in the introduction; the Diagnosis and Treatment system will extend beyond this and will form the basis of the eventual components required to develop an appropriate artificial neural network.

After this system is outlined, the actual neural network design will be discussed. Using the model provided by the California Scientific Software (Brainmaker), and the components presented in the literature review, a training algorithm will be proposed. The design will include, the architecture, training process, input criteria, etc.

Chapter 2: Neural Network Design

A. Architecture

The artificial neural network will be a **multi-layered, feed-forward** application. A multi-layered design is required because of the many structural components required to evaluate and select the appropriate treatment and the need for hidden layers to connect to the output layer. Input parameters that will be included in the design will be the DSM-IV characteristics for schizophrenia as well as the schizophrenic sub-types. Another region will incorporate the potential side-effects that the patient may experience in order to determine the best biochemical treatment options. The social (family) network and other medical professional networks (nurses, social workers, etc.) will not be considered in the artificial neural network design as it is assumed that these elements are constants in administering proper patient care. The DSM-IV criteria for diagnosing schizophrenia will form part of the input criteria. These are the elements currently involved in the diagnostic and treatment process. A psychiatric resident could take the patient's history, symptoms, etc., feed it into the artificial neural network and potentially arrive at the appropriate diagnosis and treatment alternative. Currently, there is little known about translating and correlating historical patient information and the mental status examination results into a diagnosis. This is one of the potential strengths of a neural network – to amass results and provide potential correlations to otherwise meaningless or unnecessary data.

For this type of neural network, a feed-forward software application is required because of the need for multiple layers. A feed-forward network will take the inputs from a previous layer and feed them to the next hidden layer in order to determine the appropriate output. These hidden layers will calculate the results sent from the input neurons in order to determine the output. After the hidden neurons calculate their results, the output neurons calculate the outputs based upon the weighted sum of the signals sent from the hidden neurons (California Scientific Software). As training sets will be required to determine what the neural network should measure, a feedback network would not be feasible as that is possible for only one layer that feeds information back until it is in a stable state.

The neural network will be an associator, not an evaluator. An associator neural network is one that recognises patterns or correlations, whereas an evaluator evaluates and provides data based upon the training data. It will take input patterns and associate them with an output pattern. As indicated above, it will have many output patterns for each item.

B. Learning Algorithm

The network will use a back-propagation learning algorithm. Here, any error signals (discrepancies between the actual data and the training set) will be fed back to the network altering the weights as it goes along. This will reduce the error and eventually prevent the same error from re-occurring. This network will use non-linear neurons, because many facts will need to be learned. Linear neurons can only learn one fact at a time.

C. Assumptions

Prior to determining what the artificial network should measure and how it should be trained, there are various assumptions or restrictions that should be made, some of which should be evident as the thesis continues.

1. The artificial neural network will diagnose schizophrenia and not alternate psychiatric diagnoses.
2. It is understood that the network will need connections or weights (synapses), but the thesis will not develop the actual calculations required.
3. The connections are the intermediate layers that will store the data. These will compute the weighted sums from the previous layers.
4. The training set should be made up of actual patient data stored in a database or in physicians' files. Should an attempt to create this network be made, this data should be derived from actual schizophrenic patient information. This data would include the patients' medical and psychiatric history, family information, treatment history etc. This data could be derived from patient records available within hospital or psychiatric institutions. Obviously, the appropriate approvals would be necessary in order to conduct such a study. This thesis will not take actual patient data, but will outline the elements that should be extracted if actual training sets were to be developed
5. There may be a correlation between treatment types, various social factors, demographics etc. that has yet to be studied. There also may be unidentified correlations too extensive to be covered within the scope of this thesis. This will

need to be studied further in order to determine the true A to Z patterns of this neural network.

6. There is a correlation between the side effects experienced from family members.
7. It is assumed that the necessary cognitive or behavioural treatments that usually accompany medicinal treatment will continue, and is therefore not factored into any of the diagrams or discussions.
8. It is assumed that the medical community (physicians, nurses, social workers etc.) will perform regular patient follow-ups.

D. How the data will be collected

As mentioned within the literature review, there are many methods to collect information that can result in the diagnosis of schizophrenia and the determination of the appropriate treatment option. Given that the aim of the artificial neural network is to simply expedite the diagnostic and treatment choice, a standardised form should be developed. This will form the basis of the input information, which will be fed into the network in order to produce the output. It is anticipated that the best output will be determined because of the training sets that have been provided to the neural network.

It is proposed that the following questionnaire format be used. This was developed using a combination of the questions used in the Mental Status Exam, DSM-IV diagnostic criteria and documented treatment side – effects. This, in itself, may be a contribution to medical education, as it incorporates much of the components required to diagnose and provide an effective treatment solution.

Table IV: 2.1: Artificial Neural Network Questionnaire

Patient Name:				
Residence:				
Age:	Sex:	Marital Status:		
Occupation:			Income:	
Languages:				

GENERAL INFORMATION		
	Schizophrenic Indicator	Patient Information
Medical Review from Psychiatrist's perspectives		
Sleep Patterns		
Somulant Insomnia	No Yes	
Weight Changes	Yes	
Change in sexual Functioning	Yes	
Change in eating patterns	Yes	
Previous psychiatric illnesses		
Prior schizophrenic or related episode	Yes	
Duration of episode > 6 months	Yes	
Personal History		
Marital Status	Yes or No	
Social Functioning	Yes or No	
Ability to maintain employment	No	
Appearance		
Decline in appearance	Yes	
Decline on personal Hygiene care	Yes	
Abnormal Motor behaviour (e.g. slowed speech, agitation)	Yes	
Negative attitude towards examiner	Yes	
Change in speech patterns (rapid, slow, speech, emotional)	Yes	
Other Medical Conditions		
Diabetes	Yes	
Cardiac problems	Yes	
Parkinson's Disease	Yes	
Parkinsonism	Yes	
Delirium	Yes	
Dementia	Yes	
Liver Disease	Yes	
Mood Disorders	Yes	
Blood problems (e.g. blood dysclasias)	Yes	
Cataracts	Yes	
Endocrine abnormalities	Yes	
Glaucoma	Yes	
Prostatism	Yes	
Constipation	Yes	
FAMILY HISTORY		
Family history of Mental illness	Yes	
Family history of schizophrenia	Yes	
DSM-IV CRITERIA AND DECISION TREE		
Positive Symptoms (At least 2 or more symptoms)		
Delusions	Yes	
Hallucinations	Yes	
Disorganised Speech	Yes	
Disorganised behaviour	Yes	

Are positive symptoms a result of physiological effects from a general medical condition	No	
Are positive symptoms due to the direct physiological effects of a substance (e.g. a drug abuse, medication or toxin)	No	
Symptoms of active phase of Schizophrenia lasting at least 1 month.	Yes	
Major Depressive or Manic Episode concurrent with active-phase symptoms.	Yes or No	
If # 10 is Yes, total duration of mood episodes has been brief relative to duration of active and residual phase	Yes	
If # 10 is No, and # 11 is Yes, continuous duration at least 6 months	Yes	
Negative Symptoms (at least 2 or more symptoms)		
Affective flattening: loss of emotional expression, reactivity.	Yes	
Alogia: poverty of speech (brief, laconic and empty replies)	Yes	
Avolition: inability to initiate and persist in goal-directed activities.	Yes	
Anhedonia: difficulty in experiencing pleasure or interest in activities.	Yes	
Dysphoric mood (attention): measures of social inattentiveness, and inattentiveness during mental status testing.	Yes	
Social / Occupational functioning are below the level achieved prior to the onset of schizophrenia.	Yes	
Does not Include Schizo-affective or Mood Disorders	Yes	
History of Autism or Pervasive Developmental Disorder and presence of hallucinations for at least one month	Yes	
SCHIZOPHRENIC SUB-TYPE		
Paranoid		
Preoccupation with one or more delusions or frequent auditory hallucinations.	Yes	
None of the following is prominent: disorganized speech, or catatonic behaviour or flat or inappropriate affect	Yes	
Disorganized Type		
All of the following are prominent: disorganized speech, disorganized behaviour, flat or inappropriate affect	Yes	
The criteria are not met for Catatonic Type	Yes	
Catatonic Type		
The patient must have at least two of the following characteristics: motoric immobility as evidenced by catalepsy or stupor excessive motor activity (that is apparently purposeless an not influenced by external stimuli) extreme negativism (an apparently motiveless resistance to all instructions or maintenance of rigid posture) or mutism.	Yes	
peculiarities of voluntary movement as evidenced by posturing, stereotyped movements, prominent mannerisms, or prominent grimacing.	Yes	
echolalia or echopraxia	Yes	
Residual Type		
at least one episode of schizophrenia, but without the prominent positive psychotic symptoms.	Yes	
Absence of prominent delusions, hallucinations, disorganized speech and grossly disorganized or catatonic behaviour.	Yes	
There is continuing evidence of schizophrenia (the presence of negative symptoms or two or symptoms listed in Criterion A (positive symptoms)	Yes	
Undifferentiated type		
Symptoms are met for Criterion A (positive and negative symptoms), but are not met for the Paranoid, Disorganized or Catatonic type.	Yes	

PREVIOUS TREATMENT and SIDE EFFECTS		
What was the previous treatment		
Clozapine		
Risperidone		
Olanzapine		
Ziprasidone		
Quetiapine		
Side Effects displayed		
Parkinsonism	Yes	
Dystonia	Yes	
Tardive dyskinesia	Yes	
Urine retention,)	Yes	
Constipation.	Yes	
Visual blurriness	Yes	
Dry mouth	Yes	
Delirium	Yes	
Agronulocytosis	Yes	
Liver disease	Yes	
Cataracts	Yes	
Electrocardiograph abnormalities - prolonged QT intervals	Yes	
A-V conduction delay	Yes	
Arrythmia	Yes	
Diabetes and weight gain	Yes	
Delirium	Yes	
Hypotension	Yes	
Parkinson's Disease	Yes	
Elevated prolactin	Yes	

The schizophrenic indicator represents the frequently seen responses from schizophrenics. The patient indicator is the response or data accumulated from the test administrator.

As illustrated, it is still required that the clinicians pose these questions. The artificial neural network will not eliminate this component. It is, however, these responses that will be fed into the network in order for it to diagnose and offer the best course of treatment. The neural network will automatically weight the relative importance of all the data.

E. What the Neural Network will require as inputs

I. Summary

Prior to determining the actual parameters of the neural network, it is important to define the information that will be presented to the network and what type of information the network should measure. In other words, the data elements (set of information that will be presented to the input neurons or received from the output neurons) need to be defined.

a) *Is the patient Schizophrenic*: The neural network will first need to evaluate if the patient is schizophrenic. The input data will be taken from the components used to develop the DSM-IV decision tree as well as familial, social or genetic indicators (i.e. is there a history of mental illness, etc.). The output will be in a "Yes-No" format – Yes, a diagnosis of schizophrenia is likely or No, the patient does not meet the criteria. This artificial neural network will not be used to determine alternate diagnoses.

In order for the artificial neural network to analyse the data, all components must be converted to numerical values. This will be discussed further during the training data sets.

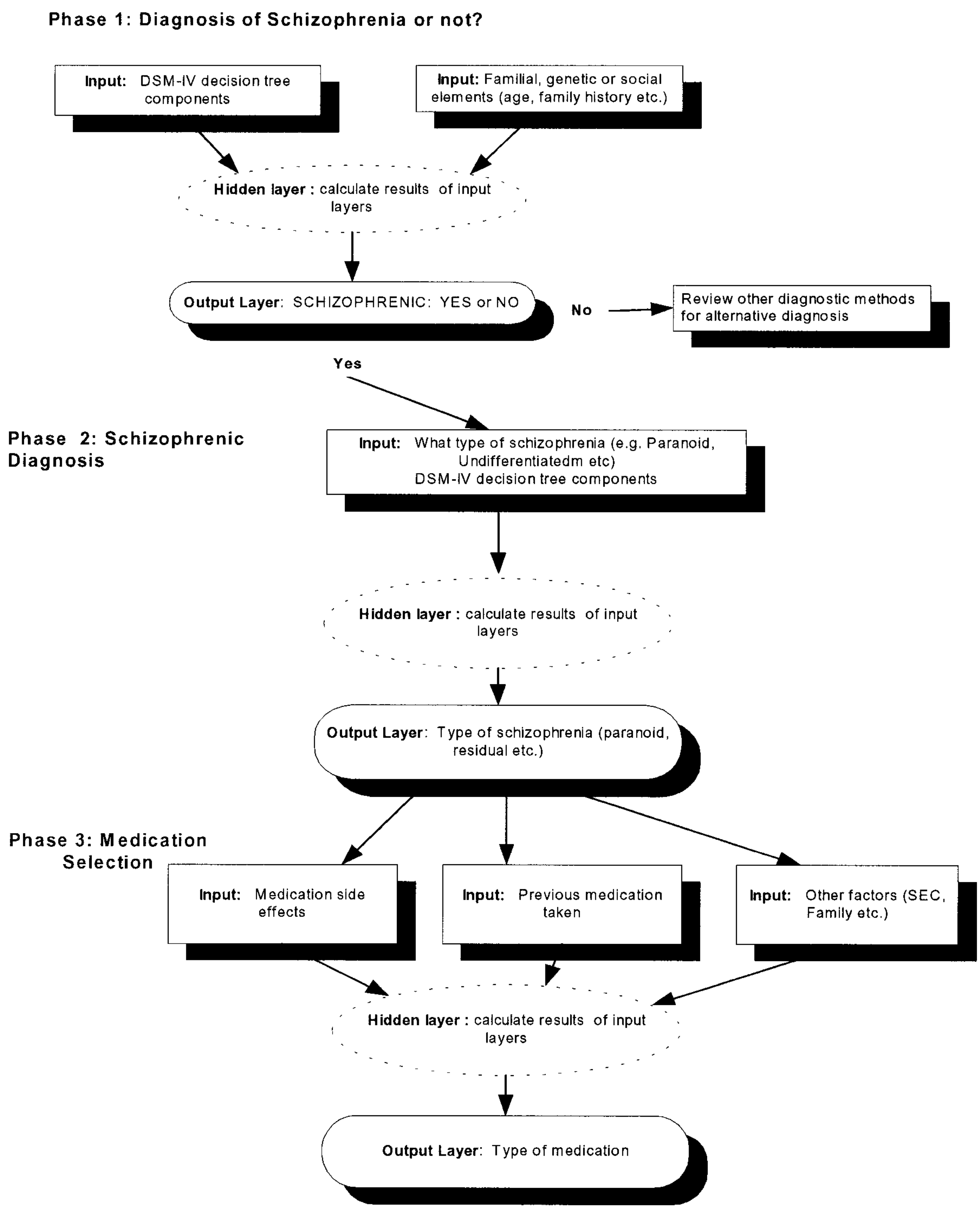
b) *Type of schizophrenia:* If the neural network determines that the patient is schizophrenic (i.e. a YES value), it must determine the schizophrenic sub-type (paranoia, residual, undifferentiated etc.). The input data will be derived from the DSM-IV definitions. Once again, all data must be converted to numerical values.

c) *Treatment options*: Once the diagnosis is determined, the neural network must be able to determine the best biochemical treatment option. This will be based upon all of the various input factors - the current medical history, previous biochemical treatments taken, and most importantly, the side effects. It is the potential for side effects that will have the most weight in determining the best treatment options. It should be noted that different medications have different rates of side effects. Therefore, the neural network should not only measure the incidence of the side effect, but also the frequency of the occurrence. There is literature that measures the frequency of adverse reactions for anti-psychotic medication (Clinical Handbook of Psychotropic Drugs).

The following represents a summary diagram of the components that the artificial neural network should measure. The next sections will break each component further to describe the actual parameters (inputs) that the artificial neural network should contain.

The summary illustrates the neural network in three phases. The phase 1 design is essentially independent of phases 2 and 3 in that it must first determine the diagnosis of schizophrenia or not. Phases 2 and 3 would be developed together in that these results are dependent on the initial diagnosis of Schizophrenia.

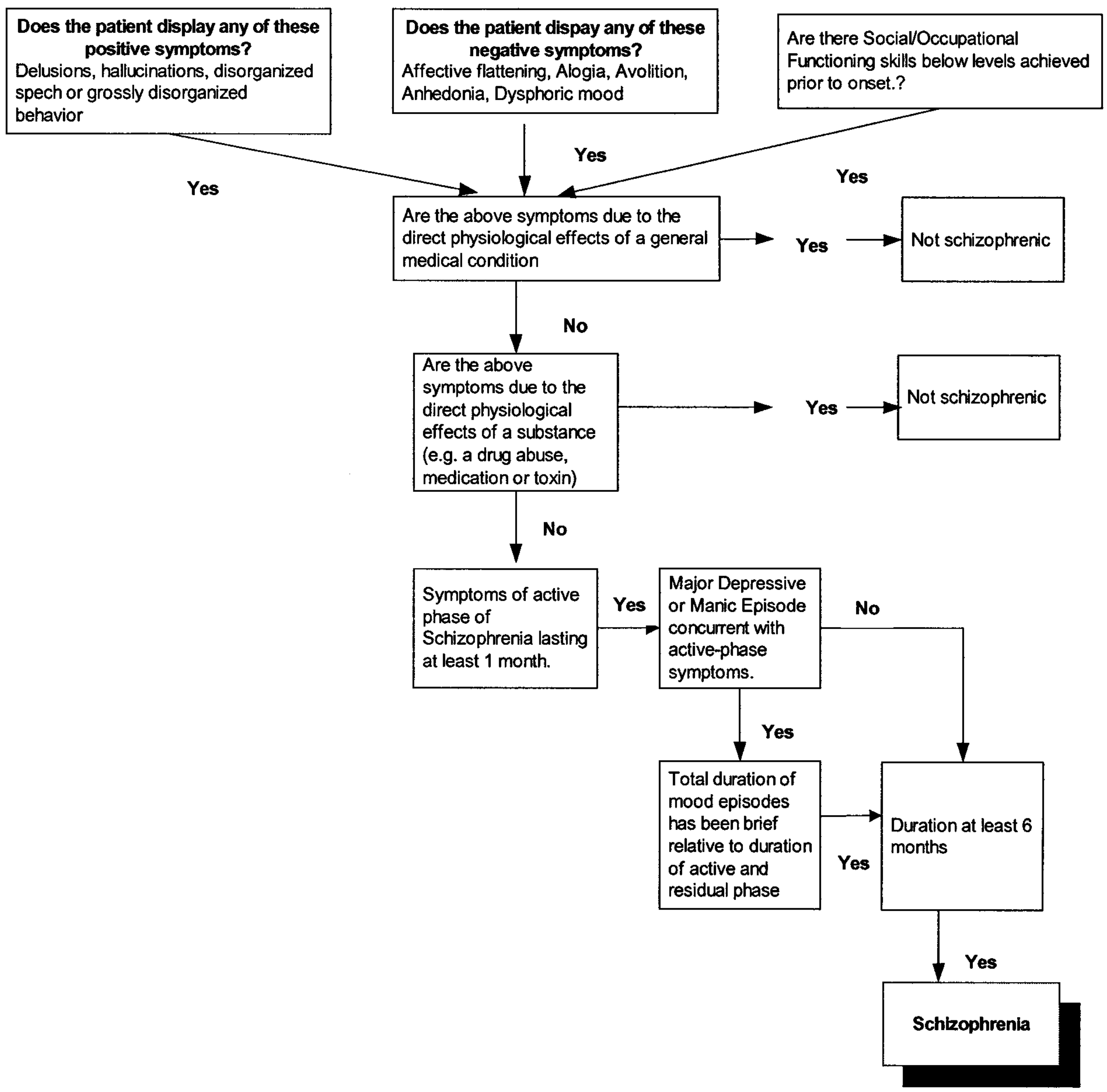
Figure IV: 2.1 Summary: Mode of Operations of Neural Network



II. Phase 1: General diagnosis of Schizophrenia

All data was taken from a combination of the DSM-IV criteria and other sources that enable the accurate diagnosis of schizophrenia. The following represents a flow chart of the data that will be required to form the inputs for the artificial neural network design. There are other components to the DSM-IV decision tree, but these have not been included as it is not the intention for the network to determine alternate diagnoses, but rather a YES (schizophrenia) or NO (not schizophrenic) decision.

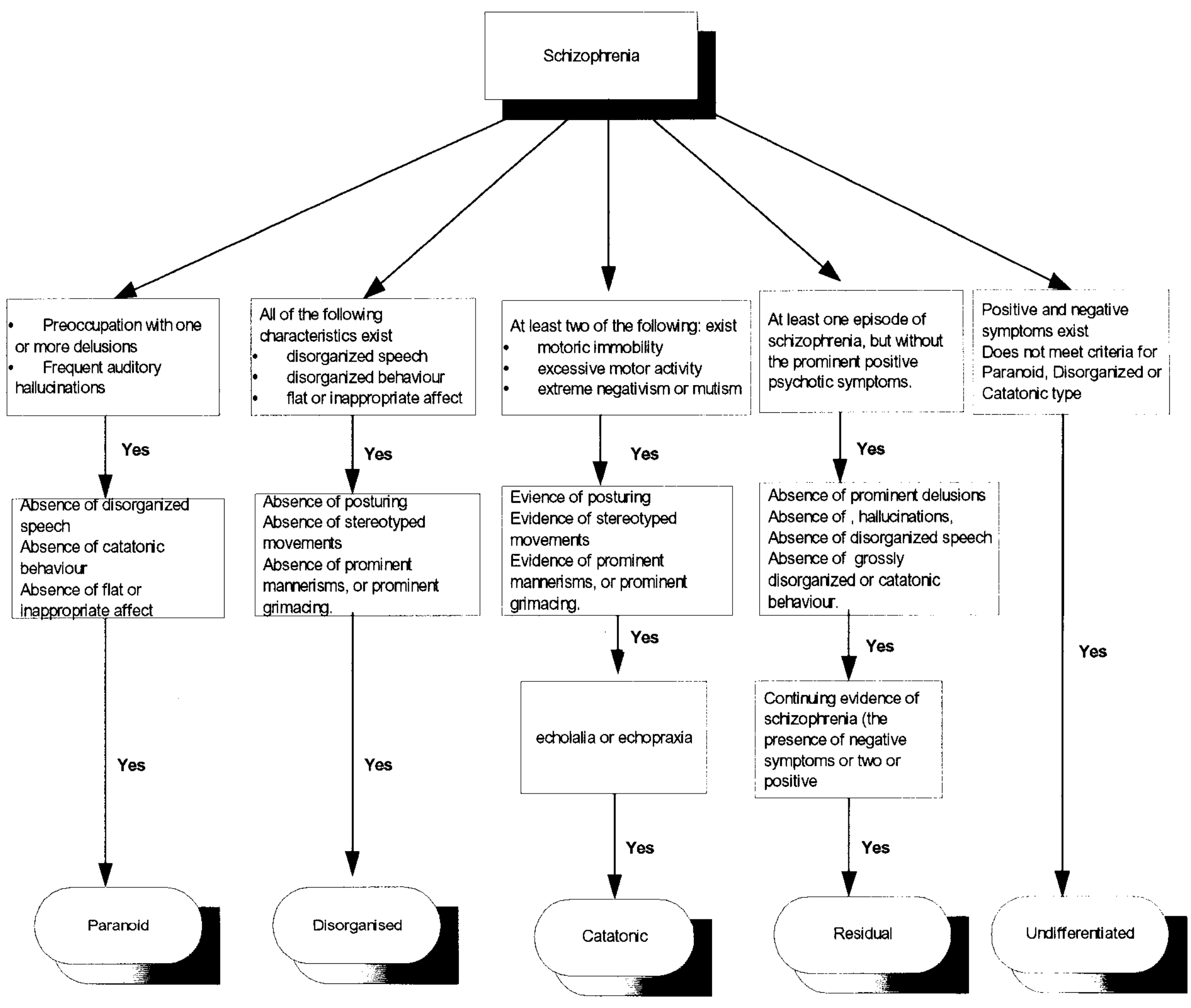
Figure IV: 2.2 DSM-IV Schizophrenia decision tree



III. Phase 2: Diagnosis of the Schizophrenic sub-type

If the network produces a YES output for a diagnosis of schizophrenia, the sub-types must be determined. Although, symptoms may overlap between categories and they are not always mutually exclusive, these are the characteristics described in the DSM-IV and therefore included in the diagnostic process.

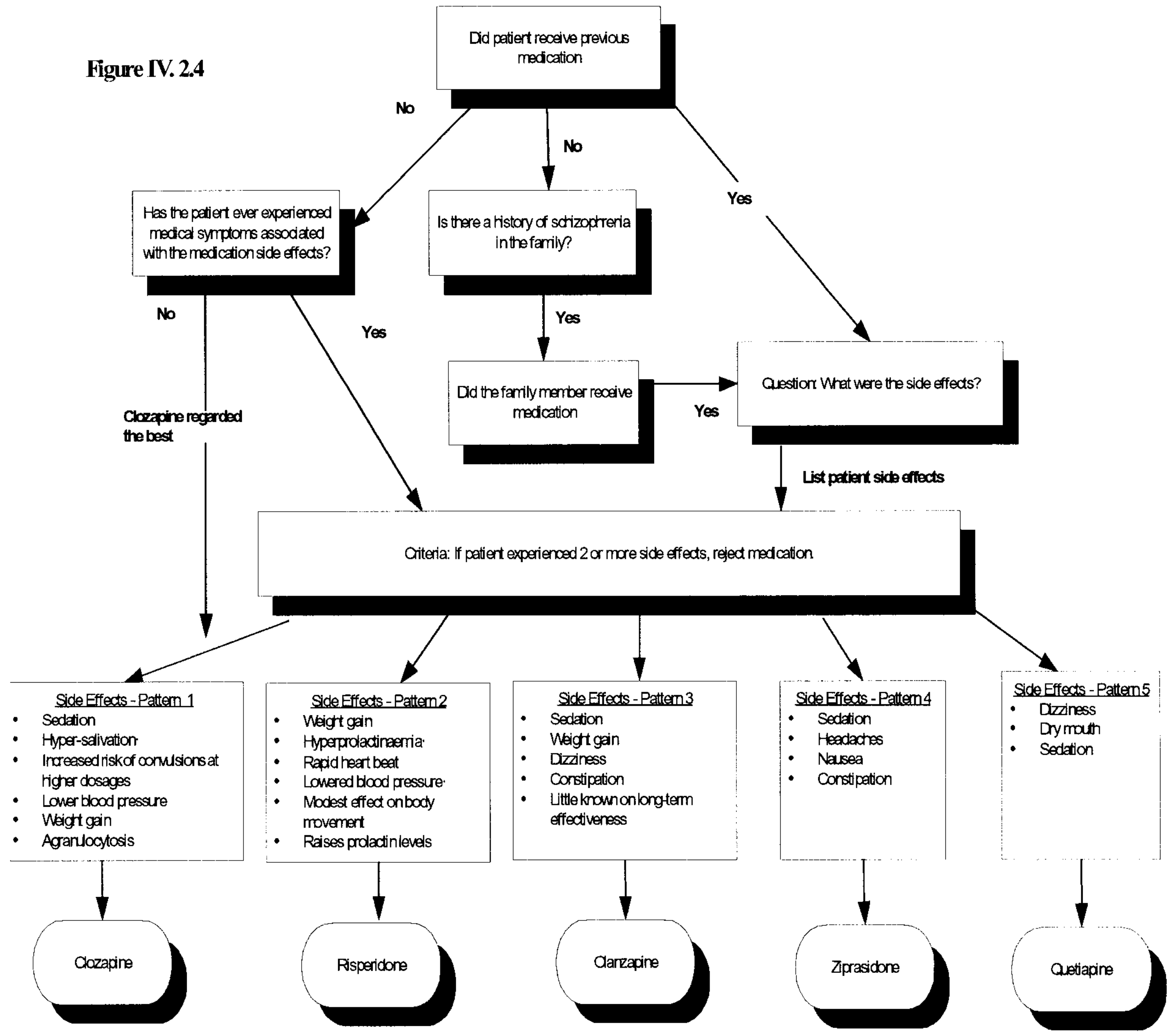
Figure IV: 2.3 Schizophrenic Sub-Types



IV. Step 3: Treatment Options

If the network produces a YES output for schizophrenia and determines the sub-type, the network would need to determine the biochemical treatment. This will be contingent upon various factors (previous treatment options, socio-economic factors etc.), but the primary factor will be the side effects commonly experienced when taking the specific medication in addition to the frequency of side effects. When developing the network, a cut-off must be established to determine the number of "allowable" side effects in order for the network to disregard a particular medication (i.e., if patient had 2 or more side effects, do not choose the medication)

Figure IV: 2.4 Schizophrenia Treatment Choice

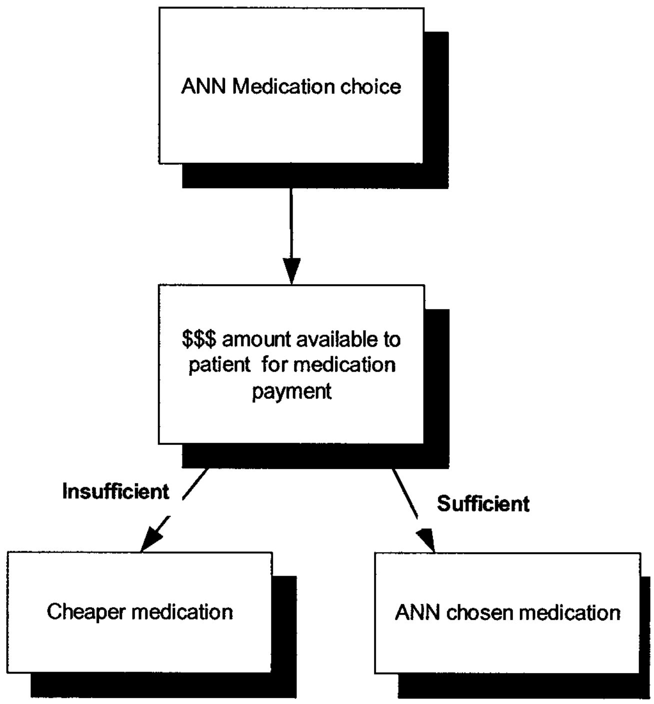


V. Considerations

There are several scenarios that must be considered when using such a tool.

Consideration 1: A realistic consideration that must be factored in is the income level of the patient vis-à-vis the price of the medication. Once the network has determined the appropriate biochemical treatment, the patient may not be able to purchase it due to insufficient income. This has not been factored into the above diagram, but it is a necessary and important consideration. As such, alternative treatments based upon income level or income source must be factored. Although not part of this tool, it should be considered during the actual development. Ideally, the diagram would contain an extension as follows:

Figure IV: 2.5: Medication choice



Consideration 2: If the patient has received medication and has not displayed any side effects, it can be assumed that the patient will continue to receive the medication. Should side effects develop during the course of treatment, all data should be re-entered into the neural network to determine the alternative treatment. All data should be re-entered so the neural network can establish any patterns with familial history, socio-economic factors, previous symptoms, medications, etc. that otherwise may not be easily established by humans. As mentioned previously, the main purpose of the network is to determine if patterns exist in order to determine the best treatment. Given the complexity of such a task, this correlation or pattern analysis has never been done, which is what makes the development of an artificial neural network for such a task quite appealing.

Chapter 3: Training Method

A. Summary

The training process will be supervised where matched sets of inputs and outputs will be provided to the network. There are desired response patterns. These were provided in the flowcharts illustrated in the previous section. In the supervised training, the network will take each input pattern, produce output patterns and compare actual output to desired output. If the actual output is different than the desired output, the network will adjust its internal connection strength in order to reduce the difference between actual and desired output.

A training pattern is presented to the network. An output is determined. If this output is different than the actual output, the network will adjust the weights (equivalent to the synapses). Then it repeats the process for the next training set. Here, a series of training data is continuously presented until all of the facts are correct or within acceptable limits. There should be a large number of training sets in order for the neural network to generalise when presented with "live" information.

Once the training sets have been presented, the network needs to be tested and de-bugged. To test, the neural network must be presented with information that it has not previously seen and by observing the output. The output should match the training sets presented.

B. Network Analysis

As mentioned in the previous section, the network must be able to perform the following functions:

1. Provide a diagnosis of schizophrenic (YES) or not schizophrenic (NO)
2. Provide a diagnosis of the schizophrenic sub-type
3. Provide the optimal treatment option

C. Data Representation

The following represents data requirements for the artificial neural network.

1. *Demographic Information*

Although there is a number of information that psychiatrists use in assessing the patient, the following are considered key elements towards the development of the ANN.

Table IV: 3.1: Demographic data

Input	Categorization	
Onset of Illness		
<20	1	
21-35	2	
36-45	3	
> 45	4	
Medical Review		
Abnormal sleep patterns	Yes = 1	No = 0
Weight changes	Yes = 1	No = 0
Change in sexual functioning	Yes = 1	No = 0
Change in eating habits	Yes = 1	No = 0
Previous Psychiatric episode	Yes = 1	No = 0
Change in appearance	Yes = 1	No = 0
Decline in personal hygiene	Yes = 1	No = 0
Change in motor behaviour	Yes = 1	No = 0
Change in speech patterns	Yes = 1	No = 0
Family history of mental illness	Yes = 1	No = 0
Family history of schizophrenia	Yes = 1	No = 0

2. *Schizophrenic diagnosis*

Table IV: 3.2: Schizophrenic Categorisations

Input	Yes	No
Positive Symptoms 1 – delusions	1	0
Positive Symptom 2 - Hallucinations	1	0
Positive Symptom 3 – Disorganised speech	1	0
Positive Symptom 4 – Disorganised behaviour	1	0
Negative Symptoms 1 - Affective flattening	1	0
Negative Symptom 2 – Alogia	1	0
Negative Symptom 3 - Avolition	1	0
Negative Symptom 4 – Anhedonia	1	0
Negative Symptom 5 – Dysphoric mood	1	0
Low social functioning	1	0
Symptoms not due to medical condition	1	0
Symptoms not due to physiological effects of a substance	1	0
Symptoms of active schizophrenia > 1 month	1	0
Absence of major depressive or manic episode concurrent with active phase	1	0

The **Output** data is Schizophrenia or not Schizophrenia.

There is no range of values, therefore a continuous neuron transfer is not required.

A binary function will suffice.

3. *Schizophrenic Sub-types*

If the diagnosis of schizophrenia has been received, the network must be able to distinguish the type of schizophrenia. The following is the data required for the network:

Table IV: 3.3: Schizophrenic Sub-Type Categorisation

Input	Yes	No	Diagnosis
Preoccupation with one or more delusions	1	0	
Frequent auditory hallucinations	1	0	
Absence of disorganized speech	1	0	
Absence of catatonic behaviour	1	0	
Absence of flat or inappropriate affect	1	0	
Output			***Paranoid***
Disorganized speech	1	0	
Disorganized behaviour	1	0	
Flat or inappropriate affect	1	0	
Absence of posturing	1	0	
Absence of stereotyped movements	1	0	
Absence of prominent mannerisms, or prominent grimacing.	1	0	
Output			***Disorganised***
Motoric immobility	1	0	
Excessive motor activity	1	0	
Extreme negativism or mutism	1	0	
Evidence of posturing	1	0	
Evidence of stereotyped movements	1	0	
Evidence of prominent mannerisms, or prominent grimacing.	1	0	
Echolalia or echopraxia	1	0	
Output			***Catatonic***
At least one episode of schizophrenia, but no prominent positive psychotic symptoms.	1	0	
Absence of prominent delusions	1	0	
Absence of hallucinations	1	0	
Absence of disorganized speech	1	0	
Absence of grossly disorganized or catatonic behaviour.	1	0	
Continuing evidence of schizophrenia	1	0	
Output			***Residual***
Existence of positive and negative symptoms	1	0	
Does not meet criteria for Paranoid, Disorganized or Catatonic type	1	0	
Output			***Undifferentiated***

It should be noted that the neural network's results may, not fit the DSM-IV algorithm flowchart because of the patterns it discovers – patterns not yet discovered by the psychiatric professional.

Medication

Table IV: 3.4: Medication Categorisation

Input	Yes	No
Has patient received previous anti-psychotic medication	1	0
Has the patient ever experienced medical symptoms associated with the medication side effects	1	0
Is there a history of schizophrenia in the family?	1	0
Did Family member receive medication?	1	0
Side effects of patient or family member		
Hyper-salivation	1	0
Increased risk of convulsions at higher dosages	1	0
Lower blood pressure	1	0
Weight gain	1	0
Agranulocytosis	1	0
Hyperprolactinaemia·	1	0
Rapid heart beat	1	0
Lowered blood pressure·	1	0
Modest effect on body movement	1	0
Raises prolactin levels	1	0
Dizziness	1	0
Constipation	1	0
Headaches	1	0
Nausea	1	0
Dry mouth	1	0

The possible outputs are the medication specified in Figure IV.2.4 (section IV, Chapter 2).

It should be noted that based upon the training patterns developed by past patient history, the neural network will be able to determine the most effective medication for the particular patient.

As mentioned previously, not only should the presence of side effects be measured, but the frequency of their occurrence. In their Clinical Handbook of Psychotropic Drugs, Bezchlibnyk-Butler and Jeffries (2002), describe the frequency of adverse reactions to anti-psychotics at therapeutic doses. The frequency is measured in percentages. The following is a summary of their results vis-à-vis the medication most commonly administered. All of their comparisons are based on currently used / approved doses.

Table IV: 3.5: Frequency of Adverse Reactions of Schizophrenic medications

Side Effects	**Clozapine**	**Risperidone**	**Olanzapine**	**Quetiapine**	**Ziprasidone**
Drowsiness, Sedation	> 30	> 10 [a]	> 30	>10	>10
Insomnia, Agitation	> 2	> 10	> 10	>10	>2
Extra-pyramidal Effects					
Parkinsonism	> 2	> 10 [b]	> 2	> 2	> 2
Akathisia	> 10	> 10 [b]	> 10	> 2	> 2
Dystonic reactions	< 2	< 2 [b]	< 2	< 2	> 2
Cardiovascular Effects					
Orthostatic hypotension	> 30	> 30 [a]	> 2	> 10	> 2
Tachycardia	> 30	> 10	> 2	> 2	> 2
ECG Abnormalities	> 30 [c]	< 2	< 2	< 2	> 2 [c]
QTc prolongation (>450 ms)	> 2 [c]	< 2	< 2	< 2	< 2 [c]
Anticholinergic Effects	> 30	> 2	> 10	> 2	> 2
Endocrine Effects					
Sexual dysfunction	< 2	> 10	> 2	< 2	< 2
Galactorrhea	< 2	> 2	< 2	-	-
Weigh gain	> 30	> 30	> 30	> 10	-
Skin Reactions					
Photosensitivity	< 2	> 2	-	-	-
Rashes	> 2	< 2	< 2	< 2	> 2
Pigmentation	-	< 2	-	-	-
Ocular Effects					
Lenticular pigmentation	-	-	-	< 2	-
Pigmentary retinopathy	-	-	-	-	-
Blood dyscrasias	< 2 [d]	< 2	< 2	-	< 2
Hepatic disorder	> 2	< 2	> 2	< 2	-
Seizures	< 2 [e]	< 2	< 2	< 2	-

a) At start of therapy or with rapid dose increase
b) Increased risk with doses above 16 mg daily
c) Higher does pose greater risks
d) Risk < 2% with strict monitering
e) Risks lower with doses below 300 mg

The above represents the frequency of adverse reactions. Therefore the neural network must contain an element that measures this factor.

D. Training Sets

The optimal method of developing training sets is to use a significant number of real patient information. This would be obtained from psychiatric institutions, professionals, databases etc. It should be noted that obtaining such information would be quite extensive and expensive and beyond the scope of this thesis. However, should this project be attempted or examined further, extensive patient data must be obtained. Should any group wish to pursue the development of such a tool, it is important that real patient data be used as part of the training set in order for the neural network to train and develop patterns with realistic data. This will ensure that accurate responses are obtained when the network is presented with "real" patient information away from the training set.

In order to obtain accurate and sufficiently representative, varied data, it is estimated that approximately 100 to 200 patient results be used to develop the training sets. The amount discussed was based upon discussions with a psychiatric professional (2002). In addition, they should correspond to the questionnaire developed previously (Hypothetical solution, Chapter 2) and the data sets, as this will form the basis for future patient assessments. The following represents where the key pieces of information can be obtained for the training sets:

1. *Clinical Exam Information (General Information):* This is the information obtained from the psychiatrist during the initial interview. It contains information regarding chief complaint, history of the present illness, medical information from the perspective of the psychiatrist (e.g. sleep patterns, weight changes, etc.), personal history and previous psychiatric illnesses.
2. *Mental Status examination Results*: This involves an assessment as to how well the patient is able to function and process information. It contains information such as appearance, attitude, behaviour, thought processes, perception, mood, affect, orientation, memory, judgement, anxiety, etc.
3. *Family History:* This would include information about previous family's history of mental illness.
4. *SCID results*: Although there are a variety of assessment tools available, the SCID is the most widely used tool and one that provides DSM-IV criteria within the actual test (e.g. Hallucinations, Delusions, etc.).
5. *DSM-IV diagnosis:* This would include whether or not the patient met the criteria for schizophrenia and if yes, what sub-type (catatonic, residual, paranoid, etc.) can he/she be categorised.
6. It should be noted that there has not been an extensive study that examines the relationship between the patient / family history to the actual treatments best suited to each individual. Usually, a trial and error method of administering

treatment is performed. As such, the type of data and correlations that the neural network could ultimately reveal would be interesting and would warrant future study. The neural networks could potentially accumulate all information that looks at treatment options as part of diagnostic history in order to see if patterns or relationships exist.

7. *Treatment History:* In order for the neural network to ascertain patterns between elements 1 to 6, and various treatments, previous treatments and resultant side effects (wherever applicable) would be required. Including this would allow the network to determine patterns otherwise not evident by humans.

The training will be supervised. Therefore matched sets of inputs and outputs will be provided to the network in order for it to eventually generalise and produce outputs when provided with actual information not within the training set. In order for the neural network to generalise, a large training set will be needed therefore. Although not part of this thesis, the training set should be obtained by patient data maintained within the psychiatric community (hospitals, databases etc.).

The training sets should be comprised of the components mentioned in the previous section. It should be noted that it is not necessary to provide the neural network with rules. It will determine the rules and correlations based upon the training information received.

Chapter 4: Simulation and Validation

As mentioned previously, the scope of this thesis does not include the actual development of the artificial neural network. The data collection required to develop the artificial neural network involves a separate project outside the intent of the thesis. As such, it is important to review the processes required to ensure that this tool will actually perform and produce the desired and stated outputs.

The methodologies discussed in this section have been referenced directly from Simulation Validation: A Confidence Assessment Methodology (Knepell and Arango, 1993). The methodology was chosen primarily because it provides "tools and techniques" that should lead to an efficient, credible assessment of a simulation model". Although the use and success of neural networks in medicine has been demonstrated, the network described in this thesis has not been created and subsequently tested and assessed for its validity. In addition, should development of this network ever be considered, it is recommended that this methodology be followed.

A. Key Players

In order to evaluate the validity, reliability and usefulness of any tool, it is first important to assemble the key players. These are those people responsible for the creation, management, administration, testing and maintenance of the artificial neural network. The following represent the people who should be assembled in order to validate the artificial neural network described in this thesis:

a) ***Assessment Team***: This represents the psychiatrists, computer programmers, medical residents, students, nurses, etc. In essence, they should represent people who were involved with the preparation, planning and evaluation process during the onset of the project. There should be a familiarity with confidence assessment methodology, psychiatric terminology, current psychiatric diagnostic evaluation methods and processes, psychiatric medication, software applications, etc.

b) ***Management:*** This represents the hospital administrators, IT directors, etc. These are the individuals responsible for the day-to-day bureaucracy.

c) ***Developers***: These are the individuals responsible for the design and development (programming) of the artificial neural network. Although Brainmaker by California Scientific Software was chosen partly because they state that limited programming is required to develop a neural network using their tool, the developers role is essential in ensuring the technical integrity of the chosen software, their abilities to understand the technical components of Brainmaker and their input with respect to other potential tools available that may be better suited to produce an accurate schizophrenic diagnosis and treatment selection tool. If possible, it is recommended that an independent group of developers not involved in the actual process be available in order to provide an independent evaluation of the actual product.

d) ***Users***: These represent the actual end users of the software. This would include medical residents, psychiatrists, psychologists, etc. who would ultimately use the product. They are essential players in order to provide an assessment regarding the validity of the tool. Does the artificial neural network produce the correct diagnosis and provide a realistic treatment option?

e) ***Community experts***: These are the subject matter experts who can provide an objective evaluation of the artificial neural network. These would be individuals not represented in the categories specified above. They could be other psychiatrists, students, other staff in the hospitals not represented in this process, family members with a schizophrenic relative, etc.

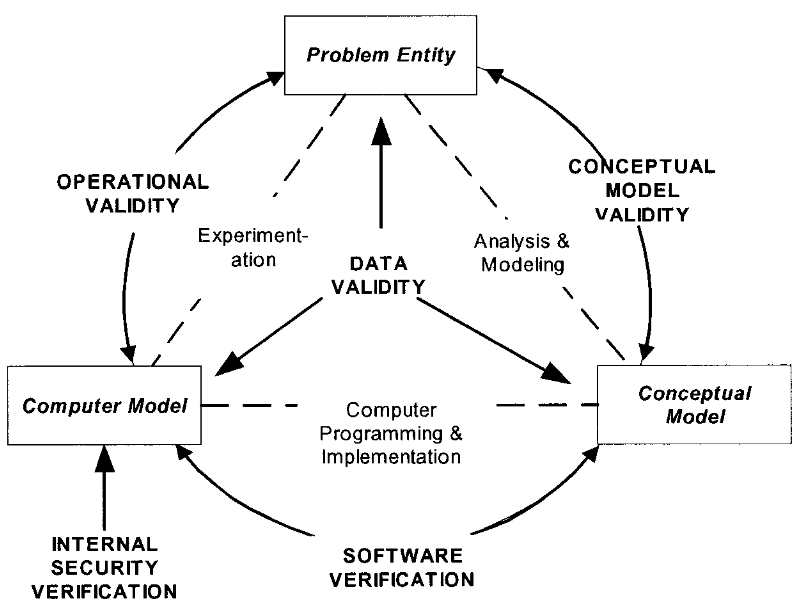
f) ***Review Team:*** These are individuals chosen to participate in intermittent reviews of the artificial neural network. They would potentially be individuals in the categories specified above (psychiatrists, residents, developers, etc.)

B. Model for Simulation Evaluation

The following model represents a framework of a simulation model created by Dr. Robert Sargent from Syracuse University. For the purposes of this thesis, it will be used to incorporate both the actual artificial neural network design as well as how validity testing can be conducted.

The model development process (inner red triangle) represents the Problem Entity, Conceptual Model and Computer Model. These are the elements required to develop the system in question. The remainder of the figure represents the technical processes required to ensure that the implementation processes and integrity are maintained during the development process.

Figure IV: 4.1: Model for Simulation Evaluation



C. Model Development Process

This process summarises the requirements needed to develop the application.

a) ***Problem Entity:*** This is the idea or situation that needs to be modelled. With respect to this thesis, this represents the need to develop a tool that would accurately diagnose and provide treatment options for each patient. There was a need to develop such a tool in order to avoid the trial and error method of administering accurate medication that would enable a patient to function adequately without the trauma and additional side effects associated with various medications. It is anticipated that an artificial neural network would be able to recognise patterns and correlations that would take humans years of research to study.

b) ***Conceptual Model:*** This is the mathematical /logical/verbal representation of the problem. For the purposes of this thesis, they are represented by the diagrams depicted in the previous section. The models have been constructed in order to facilitate the data collection and development processes in addition to outlining the key elements required to diagnose and treat the schizophrenic patient. Training and test data will be required to "feed" the neural network with the appropriate data such that it can generalise when confronted with new patient information. This will be performed and tested by the development team.

c) ***Computer Model:*** This is developed through computer programming. In this thesis, it was determined that Brainmaker would be the application of choice, however, it is the responsibility of the development team to determine if other methods would better suit the needs outlined within the conceptual model. For the development of an artificial neural network, this would include the data collection, training and testing the network and assessing the model's validity against a live patient data set. In neural networks, the key element is the data collection and training. In order to ensure that the network will be able to generalise and represent an adequate data set that would diagnose schizophrenia and provide treatment options, data collection would be a lengthy, tedious and necessary process. This would entail obtaining patient records (approximately 100 to 200 records and pending necessary ethical consent requirements) required to "feed" into the neural network It is estimated that such a venture would take approximately six to twelve months. After the network has been trained, the simulation with "real" patients is necessary, with the necessary validation by subject matter experts, psychiatrists, students, etc. to ensure that the network is producing the correct outputs.

D. Technical Processes

The balance of the diagram represent the technical processes required for simulation evaluation. There are five main processes depicted that attempts to establish the reliability and validity of the model. Although Knepell and Arango discuss this in great detail, they will be summarised here in order to provide a guideline for future consideration.

a) ***Conceptual Model Validation:*** this determines an acceptable level of agreement with the real world. This needs to be analysed in order to "validate its assumptions, theories, fidelity, derivation, logic and interfaces". With respect to this thesis, this would represent the decision to develop an artificial neural network instead of expert systems. Here, a statement of all assumptions would be required in order to ensure that each factor was considered in the final decision. For example, it was the decision to proceed with a neural network because of its ability to generalise and determine patterns that may exist between demographics, previous medication taken, etc. but would take researchers years of work to determine. The team would be required to document all assumptions and theories assessed.

b) ***Software Verification:*** This process determines if the computer program performs what it is supposed to perform. In this case, does the artificial neural network effectively diagnose schizophrenia and provide accurate treatment options. In addition, does the network perform more effectively than other types of programming (e.g. expert systems). By performing software validation, it ensures that the simulation and testing results are credible. To assist in this process, independent developers outside the scope of the initial project team can be brought in as a means to offer an objective and independent viewpoint. They can indicate the validity of the software choice, are appropriate standards being met, etc.

c) ***Operational Validation***: This process assesses whether the application possesses a satisfactory range of accuracy consistent with its intention. This includes the creation of an operational test plan. Knepell and Arango state that the purpose is to baseline the model, stress the model and establish parametric comparisons with previous test results. However, with an artificial neural network, this is done automatically. Training data is required in order for the artificial neural network to learn and generalise when faced with new data. Part of the training process is to use data to test the "generalisabilty" of the network. The artificial neural network produces its own baseline and data to determine if the network has learned what it has been trained to do.

d) ***Data Validation:*** This process ensures that the data used in the model is adequate to measure what it has purported to measure. It requires an analysis of the data and its derivation, and data consistency (are the data sets used consistent across the board). With respect to this thesis, this would involve a review of all schizophrenic patient records obtained for the artificial neural network's training and testing data sets. This review would need be conducted by participants / subject matter experts who are not part of the data collection to ensure reliability and validity of the selection. This committee would first need to ensure that the initial records were documented accurately (for correct diagnosis and treatment selection) and whether or not it is appropriate for the neural network training and test data sets.

e) ***Internal Security Validation:*** These are the security precautions required to ensure system integrity and to minimise external tampering.

The other three elements linking the triangle (experimentation, analysis and modelling and computer programming and implementation) can be represented as the documentation processes. In is important that each phase is documented to ensure a proper audit trail for future projects and revisions to the current application.

E. Assessment Phases

These phases are integral to the both the initial development and testing of the software. Each of these phases is conducive to organising an effective and efficient assessment. These steps should be followed in order to assess the validity and reliability of the artificial neural network, after its development.

a) ***Preparation:*** in this phase, information about the simulation and its use are collected. In addition, the assessment team is organised and the initial plan is developed that outlines timelines, resource planning, and the specific objectives. This is the phase where it can be determined if the assessment can be performed with the constraints available. With respect to the scope of this thesis, development of the actual neural network is not possible due to the limited time and resources available to collect the test and training data. However, as this data is the key to the development of the artificial neural network, this time should be factored in during the preparation phase.

b) ***Planning:*** This phase consists of the detailed review of the artificial neural network and its intended use. The testing and development processes should be documented should future changes to the application be required. The review should contain the strength and weaknesses of the artificial neural network. If necessary, further development may be required in order to ensure that the proper functionalities are in place in order for the application to produce the correct diagnosis and treatment option.

c) ***Application:*** This phase incorporates the actual execution of the steps outlined in the preparation and planning phases. Interim results should be well documented in order to maintain an appropriate audit trail. Here, the actual testing is performed in order to identify problem areas or areas that require refinement.

d) ***Evaluation:*** Here, assessment results are compiled and reviewed. Recommendations for the use, is also established. Here, the assessment team compiles the results. Preliminary results should be documented and discussed and then re-assigned to the appropriate members, if required. A re-evaluation is conducted afterwards and a final recommendation is put forth.

V. Conclusion

Chap 1: Overview

The purpose of this thesis is to analyse the feasibility of artificial neural networks as an effective tool to diagnose and provide treatment alternatives for schizophrenia. A review of the various forms of neural networks, and the necessary components required to create these networks has been done in order to determine the type of network best suited for the proposed project. The review of neural network within health care illustrates the relevance, and importance that neural networks have gained within this field.

The idea to develop a neural network to provide diagnostic and treatment options arose from a need to minimise this variability and to minimise the current trial and error practice of administering medication without knowing the full impact of their effects. Currently, psychiatrists obtain psychiatric and medical histories for each patient. Medical treatments are administered based upon the professional judgements, experiences, medical problems and previous medications taken (where applicable). The time between diagnosis and treatment can be lengthy and can therefore hinder the progress of the particular patient. This tool could potentially provide a solid teaching tool for medical residents entering the field of psychiatry without hindering patient care. In addition, it was anticipated that the network could uncover correlations presently unknown amongst the psychiatric community, and produce patterns that far surpass human's capabilities.

In an effort to provide an analysis of the elements required to create such a network, detailed research had been conducted to understand and document neural networks, the field of psychiatry, the diagnostic processes, the current treatment options available and the side effects most commonly associated with the respective treatments. After a thorough analysis and much discussion with various subject matter experts, it has been concluded that the actual development of such a tool with the components described in the previous sections, would be faulty and would not work. The following section discusses the key issues that would hinder the development of such a network.

Chap 2: Limitations of a Neural Network Design

I. Problems with the Diagnostic Process

The validity and reliability of neural networks rests within the training sets. In order to obtain training data, real patient information is needed. This would be obtained from patient files, psychiatrist's input, etc. As the information within the files would probably vary from patient to patient, there would be a need to standardise the data. The most important "first" data is the diagnosis given to each patient. This is where the first problem with the neural network would occur. The training set is contingent upon accurate and consistent data. However, as is often the case, the diagnostic process can be a variable process, even with the DSM-IV criteria. Two psychiatrists may not agree with the diagnosis given to the particular patient. Given that the training sets would include the diagnosis given by various psychiatrists, who may have differing opinions and methodologies with the diagnostic process, the training data would be faulty and incorrect. In order to ensure that the neural network is providing an accurate diagnosis when presented with a new patient, the training set must be consistent, which would not happen given the divergent views commonly associated with the diagnostic practice.

II. Problems with the Treatment Option

Many of the same issues discussed in the previous section are relevant for the treatment options. Two psychiatrists may not necessarily agree on the treatment option given for the same patient. This variability, once again, would impact the training set and would therefore cause the neural network to produce inconsistent treatment alternatives when presented with a novel patient. It is imperative that consistent patient data be obtained from the patient files. This consistency would be lacking given the variability amongst psychiatrists. One would need to accumulate vast amounts of patient information required for the training sets. However, this would be fruitless if there is limited consistency and agreement within the psychiatric community about the appropriate biochemical treatment.

In addition, the patient files contain limited information with respect to the efficacy of the treatment choice. If a particular medication is ineffective, a change is made and a record may be entered within the patient files. However, there is limited information with respect to why the medication was not working. As this treatment choices outlined within this thesis is contingent upon the network being trained on documented side effects and why various treatments may or may not work, the training data would not be valid.

III. Inconsistent / Missing information within the patient files

This section states the importance of consistency in order to have accurate training data. Training sets would be derived from current patient files maintained within the institution where such a project would occur. There is no consistent method of recording and maintaining patient records. It would be difficult to obtain all the training requirements for one patient – to obtain consistent and similar data amongst patients would be even more difficult.

In addition, as stated above, much of the patient information resides within the psychiatrists' heads – their expertise in dealing with their patients and differences in their methodology eliminates the controlled environment that one would require to accumulate the necessary data required for the training set.

IV. Summary:

As illustrated above, the main reason why development of such a network would not work is because of the quality of the training sets that would be derived based upon the variability of patient information. Psychiatrists may not often agree on the diagnosis and treatment choices for one single patient. This would result in faulty training sets. Therefore, two different inputs can lead to two different outputs. The lack of consistency within the psychiatric field will taint the data used to train the network.

This lack of reliability / consistency or the presence of variability within the diagnostic process and treatment selection is one element that separates this idea from others that have been created within the health care field. Lee and Park (2001) developed an artificial neural network to classify and distinguish the HIV and AIDS patients. As with many fields that attempt to use neural networks, they state, "if attributes of the factors are ill defined and/or ill-structured, finding a solution will be very difficult and complicated". As was illustrated above, the use of artificial neural networks as a diagnostic and treatment tool for Schizophrenia would contain factors that are "ill defined" and "ill structured", given the divergent opinions prevalent within the psychiatric community. Although Lee and Park were working within a field that can have ill structured data, they were able to create a neural network because their data was taken from ACSUS dataset, which contained data from a longitudinal study of persons with HIV/AIDS-related diseases from 10 US cities. Information was gathered on 1949 HIV/AIDS-infected persons (narrowed down to 1,171 cases for their network) in a series of interviews over a total of six time-periods with quarterly follow-up surveys. The interviewers also contacted other medical health professionals to collect information. From the outset, their data set contained long-term, consistent and reliable data. Using the model proposed in this thesis, it is obvious that using the patient data currently in patient charts would not be consistent. In addition, the data used by the Lee and Park neural network contained complete information.

The importance of consistent and reliable data is also shown in Bowser's (2000) study. He created a neural network that was 98% accurate in classifying prostate cancer patients as being at low risk for lymph node spread. The input data used was the clinical stage, biopsy Gleason count and pre-treatment PSA (prostate specific antigen). The output was the presence or absence of lymph node spread. These data sets are measurable, reliable data elements that are easy to obtain and maintain. There is minimal room for error and the data is easy to "feed" in a neural network. However, he expresses caution in that it is still quite difficult to determine the relationships between the input and output variables. In addition, he states that neural networks will not override the judgment of an experienced practitioner, hence some of the scepticism received from this thesis proposal, however, Bowser does see the merits of neural networks for rural or inexperienced physician's with, "prediction based on a large database of patients with similar risk characteristics".

As illustrated, it is quite evident that the benefits of using neural networks as predictors are contingent upon the training sets used. The ability for the neural network to generalise when presented with a novel situation rests with the data with which it has been trained. The training sets should be derived from live data if the network is going to generalise. However, the live data must be consistent and reliable. In addition, the data must be coded accurately.

As a result of the issues and inconsistencies mentioned in the above discussion, it can be concluded that the development of an artificial neural network, at this point, is not necessarily the best long-term educational solution required in order to bridge the gap between the diagnostic process and ultimate treatment selection. However, it does have its merits as a research project. It is the author's perspective that the development of a neural network should not be discounted and is worthy of further research, because it has been illustrated that the network's ability to learn and associate patterns lends itself favourably to such a task illustrated in this thesis. The artificial neural network would potentially allow for a cost efficient solution that would provide diagnostic analysis and treatment selection in a timely fashion that would help patient functioning. In addition, such a tool would definitely account for many indeterminate factors that may not be currently considered by psychiatric professionals because of the vast amount of data and factors that ultimately are too much for individuals to assimilate.

However, the question should be asked – can such a tool still be developed? If so, what conditions would be conducive towards the development of such a tool?

It is the author's contention that a tool to diagnose and provide treatment options for schizophrenic patients can potentially be developed. However, the training sets would need to be developed and created under strict conditions. The possibility that certain factors may not be considered in the diagnosis and training process affects the training that medical residents currently receive from the psychiatric professionals. Given the need for this type of training solution, and given the limitations specified above, it is still important that this tool be considered a viable solution to this training problem. The following section therefore provides a framework and overview of the conditions required to ultimately develop an artificial neural network to diagnose and provide treatment options for schizophrenia.

Chap 3: Activities and Elements required to develop the Network

The following represents the key requirements that would potentially enable the development of an effective artificial neural network to diagnose and provide treatment alternatives for schizophrenia.

1. Participants

This project would require a longitudinal study that would involve the following participants:

Accredited medical institution(s): this would be at least the hospital, community group etc. that actively and consistently has schizophrenic patients.

Psychiatrists: this would include the psychiatrists, and medical residents that provide diagnosis and treatment alternatives to the patients.

Other psychiatric professionals: this would include the support staff, psychometrists, and others associated with the patient's well being.

Patients: the schizophrenic patients who agree to participate in this study.

Review board: this would be made up of subject matter experts who would be in place to independently review the diagnostic and treatment choices made by the attending psychiatrists.

2. Collecting New Patient Medical and General Information

It is recommended that this study should use patients who have never been assessed. Therefore a complete medical review must be performed and documented. This review would also include familial history, general information etc. that would potentially assist in the diagnostic process and treatment selection.

3. Collecting New Patient Diagnostic Data

It has been stated that the current information contained in patient files would be inadequate as training data. In order to capture consistent, patient information that could be used as training sets, new patient data would need to be accumulated in a longitudinal study. As mentioned previously, there is much disagreement amongst the psychiatric community with diagnosing patient information. The first step would be to develop a diagnostic tool that would properly assess if a patient suffers from schizophrenia. This could be a combination of DSM-IV or ICD-10 criteria.

After this diagnostic tool has been created, new patient data would be accumulated. A psychiatric professional would arrive at a diagnosis. This diagnosis would be subsequently subjected to review by an independent board. In order to ensure that reliable and valid data is collected, each member must arrive at the same diagnostic conclusion. This will be documented in the patient files.

4. Biochemical Medical Option

Based upon the diagnosis and history of the patient, the psychiatrist assigned to the patient would make the appropriate treatment choice. This choice would be reviewed by the same independent review panel, who approved the original diagnosis. Once again, this would be documented in the patient records.

The patients would need to be followed continuously in order to document any side effects that may occur. A scale that measures the side effects and frequency of these side effects would need to be created such that a record could be maintained.

Should these side effects be detrimental to the welfare of the patient, a change in medication would be required. Once again, the change would be determined by the patient's psychiatrist and be subjected to review and approval from the independent panel.

These patients would be followed for approximately three to five years. Any subsequent new patient participating in the study would be subjected to the same diagnostic, treatment choice and review.

5. Accumulation of training data

After the patients have been assessed and followed, the collection of training data for the neural network must be made. Accumulation of large amounts of data is required in order to ensure appropriate neural network training. The same tool developed in Section IV, Chapter 2 could be used. The training data would be made up of these patients used in this new study. As each process was reviewed by an independent board, the data collected and subsequently used would be reliable and consistent.

A portion of this training data should be maintained for testing purposes.

After the neural network has been trained and tested, the ultimate test would the presentation of a live patient in order to determine if the network is functioning (i.e. producing the correct diagnosis and treatment selection). The results would be reviewed by the same independent panel.

6. Further Study

As a side note, it would be interesting to see if the above review process aids in the diagnostic and treatment selection process and subsequently standardisation within the psychiatric community. For that, a control group that receives the same type of status quo care would need to be used and compared against the above experimental group.

Chap 4: Other Tools

Given the current limitations of neural networks as a diagnostic and treatment selection tool for schizophrenia, the question arises as to whether alternative artificial intelligence tools can be created. There are a number of artificial intelligence tools that have been studied, however, the options listed below summarise the two most likely alternatives.

The first tool is an Expert System. As mentioned in the literature review, expert systems are computer programs that simulate human intelligence and behaviour. They can range from simple text to logical rules. They differ from neural networks in that expert systems are rule based (e.g. IF-THEN rules). They do not "learn" or associate patterns as do neural networks. The benefits of expert systems are many. Expert systems can maintain a large database of knowledge, which can be updated when necessary. In addition, unlike humans, there is no loss of knowledge. The knowledge and expertise is maintained in this database and can be accessed by those with limited knowledge.

Another promising artificial intelligence tool is the "Multi Agent System" (MAS, 2001). A multi-agent tool is a network of various software agents that interact to solve problems. The advantages of this tool are numerous. They allow the distribution of "computational resources and capabilities across a network of interconnected agents". A centralised system is limited to a smaller network. In addition, multi-agent systems allow for the interconnection of existing systems. They model problems as do teams in the real world, which operate in temporally and spatially distributed fashion. Finally, using multi-agent systems should result in enhanced system performance because they use various software agents to solve the problem. Instead of relying on one system to do "all of the work", one can develop multi agent systems, where each component or agent is responsible for various activities.

VI. Recommendations / Final Note

Given the limitations stated of creating the type of neural network proposed in this thesis, it is obvious that training data used be consistent, reliable and valid. Although psychiatry does have measures in place to ensure correct diagnosis and appropriate medication, discrepancies and disagreements still occur. This resulting process has revealed a lack of standardisation and consensus amongst the psychiatric community. The professionals this author had spoken with during the course of this thesis maintain and administer superior patient care. However, it is quite apparent that the lack of consensus can possibly hinder optimal patient treatment, though this is not necessarily the fault of the practitioners.

The question was often asked – will this tool benefit the field of psychiatry? Why should an artificial neural network be used as opposed to the current diagnostic and treatment selection process? Based upon the discussions above, and primarily due to the limited consensus, it is the author's contention that such a tool may potentially benefit the psychiatric community, in particular, as a teaching tool for medical residents. Currently, it is a tool that is ahead of its time. With respect to the second question - should it be used as an alternative to the current process – not necessarily, unless the institution chooses this tool to be the standard selection tool. Sole use would most definitely be an unlikely option, as most psychiatric professionals would prefer not to rely on an automated tool for what they feel should be a human judgement, especially when the field of psychiatry is involved. In addition, there are possible legal ramifications to introducing this tool as the sole diagnostic and treatment selection method. Rather, the author feels that this tool could be used as a supplement to the current methods and as a tool to teach medical residents, provided that the process above is used in order to standardise the training data.

The question was also asked why a straightforward medical diagnostic tool such as those described in the literature review was not initially proposed. In addition, the question was also asked as to why would one choose the field of psychiatry and attempt to propose a neural network to diagnose and provide treatment alternatives, given the limited standardisation. The answer to the first question should be self-evident. Given the increased attention to neural networks in the medical field, it seemed rather fruitless to propose an idea that many have already proposed. This would therefore defeat the purpose of living the full academic experience. This reasoning begins to answer the second question. There has been limited study of artificial neural networks within the field of psychiatry. Studies have attempted to use neural networks to assess the length of stay in hospitals, but very limited research has been performed with respect to the diagnostic feature and subsequently treatment alternatives. The idea was to present a blueprint for the development of an all-inclusive patient system that would focus on the individual patient and not simply a composite sketch that includes many patients. Currently, the literature does not contain this type of network. It was hoped that the treatment choice determined by the network would obviate the trial and error method of administering medication, and would therefore improve patient care. It was hoped that the network could recognise patterns not yet studied or evident during the medical treatment process.

If it were possible to overcome the limitations, the ultimate recommendation would be the development of a neural network, given its ability to learn and recognise patterns. As mentioned above, it is an idea ahead of its time, and would most likely be used or developed primarily as a research tool. However, this thesis has demonstrated a variety of limitations and effort (financial and limited resources) required to, develop a neural network. Given these constraints, the question arises as to what can be done now to obtain an educational teaching solution to the problem in question.

It is therefore recommend that an Expert System would be the most viable short-term option, with some modifications to the original design. Psychiatric professionals rely on the DSM-IV decision tree to form or consider the patient diagnosis (for schizophrenia and the sub-types). With this standard accepted within the community, it is recommended that this element be removed from the expert system. Rather, the focus should be on the treatment portion. It has already been established that a treatment selection tool would be beneficial as a teaching tool in order to minimise the trial and error decision of prescribing medication. The diagrams in the *Hypothetical Solution* section are algorithms. Therefore, it would be quite easy to create rules of treatment selection based upon known patient or familial side effects (e.g. IF patient received prior medication and IF patient experienced side effect SEDATION and WEIGHT GAIN, THEN NOT CLOZAPINE or THEN RISPERIDONE). To make the system more comprehensive, the rules would also consider the frequency of occurrence.

If it is the intention to have a completely automated diagnostic and treatment tool, it would be very easy to incorporate the diagnostic decision tree into the expert system. Once again, the algorithm would be a simple rule-based IF-THEN system (IF patient has positive symptoms and IF symptoms due to physiological effects of a general medical system, THEN not schizophrenia).

The drawback to creating an expert system that focuses solely on the treatment selection process is that the potential patterns or correlations (demographics, previous medication, family history, etc.) not previously considered (or beyond the scope of the human mind) is lost. As mentioned previously, the choice of neural networks was due to its power to incorporate vast amounts of information and produce patterns based upon the learning rule. There is no need to create a series of rules. While expert systems can incorporate a lot of data, they are still rule based – rules created by humans and therefore potentially biased and with constraints.

If anything, it is the author's contention that this proposal has opened up the possibility of standardising elements within psychiatry. It presented a thorough analysis of the current diagnostic process and treatment options. However, as seen, an attempt to create a tool to measure these outputs would be flawed should current patient data be used. This is due to the variability psychiatric practitioners encounter during their day-today activities. In addition, it is the author's wish that improved standardisation should be an area of future study and consideration, as it would increase the credibility of a field that is often looked at askance in the medical community.

VIII. References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders. (4th ed).* Washington, D.C.: APA.

Basco, M.R.; Bostic, J.Q.; Davies, D.; Rush, A.J.; Witte, B.; Hendrikse, W.; Barnett, V. (2000). Methods to Improve Diagnostic Accuracy in a Community Mental Health Setting. *American Journal of Psychiatry, 157*, 1599-1605.

Bezchlibnyk-Butler, K., Jeffries, J.J. (2002). *Clinical Handbook of Psychotropic Drugs.* Toronto: Hogrefe & Huber Publishers.

Bowser, A. (2000). Neural Network Predicts Risk of Lymph Node Spread. *Urology Times.* Copyright Advanstar Communications.

BrainMaker: How BrainMaker Neural Networks Work (2002) [On-Line]. Available: http://www.calsci.com/whatare.html

Burke, H.B., Goodman, P.H., Rosen, D.B., Henson, D.E., Weinstein, J.N., Harrell, F.E., Marks, J.R., Winchester, D.P., Bostwick, D.G. (1997). Artificial Neural Networks Improve the Accuracy of Cancer Survival Prediction. *American Cancer Society 79 (4)*, 857-862.

California Scientific Software. (1990*). Introduction to Neural Networks.* (3rd ed.). California.

Choca, J.P.; Van Denburg, E.J. (1996). *Manual for Clinical Psychology Trainees.* (2nd Ed.), New York: Brunner/Mazel.

Coyle, J.T. (1997). Glutamate and Schizophrenia's Symptoms. *BrainWork – The Neuroscience Newsletter, 7(3)*, 4-8.

Elman, J.L.; Bates, E.A.; Johnson, M.H; Karmiloff-Smith, A.; Parisi, D.; Plunkett, K. (1998). *Rethinking Innateness: A Connectionist Perspective on Development.* (1st ed.). MIT. Library of Congress Cataloguing-in-Publication Data.

Expert Systems Architecture (1994). [On-line]. Available: http://www.cee.hw.ac.uk/~alison/ai3notes/subsection2_5_2_1.html

Fausett, L.V. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications.* Prentice-Hall.

Golden, C.J. (1977). Validity of the Halstead-Reitan Neuropsychological Battery in a mixed psychiatric and brain-injured population. *Journal of Consulting and Clinical Psychology, 45*, 1043.

Golden, C.J. (1990). *Clinical Interpretation of Objective Psychological Tests.*

Hoff, A.L.; Sakuma, M.; Razi, K,; Heydebrand, G.; Csernansky, J.G.; DeLisi, L.E. (2000). Lack of Association Between Duration of Untreated Illness and Severity of Cognitive and Structural Brain Deficits at the First Episode of Schizophrenia. *American Journal of Psychiatry. 157.* 1824-1828.

Knepell, P.L., Arangno, D.C. (1993). *Simulation and Validation: A Confidence Assessment Methodology.* California: IEEE Computer Society Press.

Lee, C.W.; Park, J-A. (2001). Assessment of HIV/AIDS-related health performance using an artificial neural network. *Information and Management, 38(4)*, 231-238.

Lee, A.; Ulbricht, C.; Dorffner, G. (1999). Application of Artificial Neural Networks for Detection of Abnormal Fetal Heart Rate Pattern: A Comparison with Conventional Algorithms. *Journal of Obstetrics and Gynaecology, 19(5)*, 482-485.

Lieberman, J.A. (1999). Is Schizophrenia a Neuro-degenerative Disorder? A Clinical and Biological Perspective. *Biological Psychiatry, 46*, 729-739.

Lieberman, J.A.; Fenton, W.S. (2000). Editorial. *American Journal of Psychiatry, 157(11)*, 1727-1730.

MacKinnon, R.A.; Yudovsky, S.C. (1991) Principles of Psychiatric Evaluation.

Malmgrem, H. (2000). Artificial Neural Networks in Medicine and Biology. A philosophical introduction. *Opening lecture at the ANNIMAB-1 Conference, May 13-16, 2000.* [On-line]. Available: www.phil.gu.se/Papers.html

Maruish, M.E. (1999). *The Use of Psychological Testing for Treatment Planning and Outcomes Assessment.* (2nd Ed.). Mahwah, N.J. : Lawrence Erlbaum Associates

McGrath, J., Emmerson, W.B. (1999). Treatment of Schizophrenia. *British Medical Journal (319).* 1045-1048.

Memory – The Hopfield Net: Discussion of Memory in Humans. (2002). [On-Line]. http://brain.web-us.com/brain/neur_hopfield.html#hop).

Minsky, M. and Papert, S. (1969). *Perceptrons; an introduction to computational geometry.* Cambridge, Mass.: MIT Press

Moallemi, C. (1991). Classifying Cells for Cancer Diagnosis Using Neural Networks. *IEEE Expert, 6, (6).* 8-12.

Neural networks with Java. (2001). Retrieved February 14, 2001. [On-line]. Available: http://rfhs8012.fh-regensburg.de/~saj39122/jfroel/diplom/e-13-text.html.

Neural Networks. (2001). Retrieved January 30, 2001. [On-line]. Available: http://www.neusciences.com.

Norman, R.P.; Wilensky, H. (1961). Item difficulty of the WAIS information sub-test for a chronic schizophrenic sample. *Journal of Clinical Psychology, 17*, 56.

Peck, D.F.; Shapiro, C.M. (1990). Measuring Human Problems. A Practical Guide.

Pudit, V. *Neural Networks Tutorial.* (2001). Retrieved March 14, 2001. [On-line]. Available: http://dsl.serc.iisc.ernet.in/~vikram/nm_intro.html.

Rapaport, D.; Gill, M.M.; Schafer, R. (1970). *Diagnostic Psychological Testing.* (Rev. ed). New York: International Universities Press.

Riley, G. (1997). What are Expert Systems. [On-line]. Available: http://www.ghg.net/clips/ExpertSystems.html

Schutte, N.S.S; Malouff, J.M. (1995) Sourcebook of Adult Assessment Strategies.

Sordo, M. (2002). Introduction to Neural Networks in Healthcare. *Open Clinical – Knowledge Management for Medical Care.*

Stader, J. (1992). *Applying Neural Networks.* Retrieved March 19, 2001. From: The University of Edinburgh. [On-line]. Available: http://citeseer.nj.nec/stader92.applying.html.

Stergiou, C.; Siganos, D. (1996). *Neural Networks.* Retrieved April 26, 2001. [On-line]. Available: http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.

Stevens, L.M. (2001). Schizophrenia. JAMA Patient Page. *The Journal of the American Medical Association. 286(4).* 494.

The Harvard Mental Health Letter. (1998). *Harvard Medical School. 14(10).*

Vida, Stephen, 2002. Personal Interview. Montreal General Hospital.

Walczak, S.; Scharf, J.E. (2000). Reducing surgical patients costs through use of an artificial neural network to predict transfusion requirements. *Decision Support Systems, 30(2000)* 125-138.

Weinberger, D.R. (1987). Implications of normal brain development for the pathogenesis of schizophrenia. *Archives of General Psychiatry, 44(660).*

Weiner, L.B. (1966). Psychoanalysis in Schizophrenia.

Wechsler, D. (1958). *The Measurement and Appraisal of Adult Intelligence.* (4th ed). Baltimore: Williams & Wilkins.

Zimmerman, I.L.; Woo-Sam, J.M. (1973). *Clinical Interpretation of The Wechsler Adult Intelligence Scale.* New York, Grune & Stratton.