Chapter 1

Introduction

1.1 Background

An important pillar of modern medicine is the measurement of biomedical signals such as cardiac electrical activity, heart sounds, respiratory flows and sounds, brain and other neural activities. Almost all biomedical signals are contaminated with noise artifacts due to sensing methods or environmental noise. In addition, extracting important information regarding any pathological conditions from measured physiological signals demands careful analysis of the signal by an expert.

Modern analog and digital signal processing techniques have contributed significantly in making the analysis of physiological signals easier and more accurate. For example, these techniques help remove different types of noise, identify inflection points and combine multiple signals. They also detect and classify changes in the signals due to pathological events and conditions and transform the signal to extract hidden information not available in the original signal. The primary goal of processing biomedical signals is to identify any pathological condition of the patient, and monitor the changes in the condition over a course of treatment or with the procedure.

Recent studies have shown a significant development in the field of non-linear dynamics. These studies have shown that simple non-linear systems can display highly complex behaviour [29]. In recent years, there has been an increasing interest in applying non-linear dynamics theory in studying biological systems. Various complex biosignals such as the electromyogram (EMG), electroencephalogram (EEG) and electrocardiogram (ECG) have been studied and tested in the past few decades [23, 28, 34, 45, 47]. The results of these tests showed evidence that the dynamics underlying these signals is non-linear and indicated the possibility of the presence of deterministic chaos.

1.2 Aims and Objectives

The ECG provides vital information about the electrical activity of the heart over time. Analysis of the ECG using feature extraction can yield important information for the diagnosis of cardiac arrhythmias by clinicians. Classical techniques have been used for this purpose, such as the analysis of ECG signals for classification using the autocorrelation function, time-frequency analysis, wavelet transforms and adaptive filtering.

The aim of this thesis was to understand the nature of these complex biosignals (the ECG in the present case) and develop an algorithm that can be used for analytical purposes. To achieve this aim, various theories and published papers by different authors relevant to the subject have been reviewed, and algorithms have been developed and implemented using MATLAB$^{®}$ software.

1.3 Thesis Overview

Chapter 2 gives an overview of cardiac anatomy, and provides details about the conduction system of the heart. Information about the ECG and its distinct waveforms is presented and various heart arrhythmias are described.

Chapter 3 very briefly discusses biosignals. The chapter also gives detailed information about the various theories that offer a better understanding of the nature of biosignals and how these concepts can be used for the interrogation of biosignals using non-linear dynamical analysis techniques. The theories discussed in this chapter include non-linear theory, dynamic systems, chaos theory, fractals, fuzzy logic, and artificial neural networks

Chapter 4 is a literature review of existing and previous work in the analysis of biosignals carried out by various authors in this field. The chapter very briefly discusses the methodology and findings of their work.

Chapter 5 describes algorithm development and implementation using MATLAB$^{®}$ software. It describes briefly the logic used and the step-by-step procedure of the development of the algorithms. The following algorithms were developed (a) for displaying the signal, (b) QRS detection using linear thresholding technique and using discrete wavelet transforms, (c) to calculate correlation between

two signals, (d) to calculate Kolmogorov complexity of the ECG signal and (e) to calculate Lyapunov exponent of the ECG signal.

Chapter 6 discusses the results obtained by testing various signals obtained from the internet.

Chapter 7 concludes the thesis with future work possibilities.

Chapter 2

Cardiology and the Electrocardiogram

2.1 The Heart

The heart is the hardest working muscle in the human body. It is situated to the left of the middle of the thorax underneath the breastbone (sternum). The adult human heart weighs about 250-350 gms and is about the size of two fists held side-by-side. At an average rate of 80 beats per minute, the heart beats about 115,000 times in one day. During an average lifetime, the human heart will beat more than 3 billion times. On average the heart pumps 7,600 litres of blood around the body every day.

2.1.1 Anatomy of the Heart

The essential function of the heart is to pump blood to various parts of the body. Figure 2.1 shows the anatomical features of the heart and the pathways of blood flow within it. The heart consists of four chambers: right and left atria and right and left ventricles. The two atria act as collecting reservoirs for blood returning to the heart while the two ventricles act as pumps to eject the blood to the body. As in any pumping system, the heart comes complete with valves to prevent the back flow of blood.
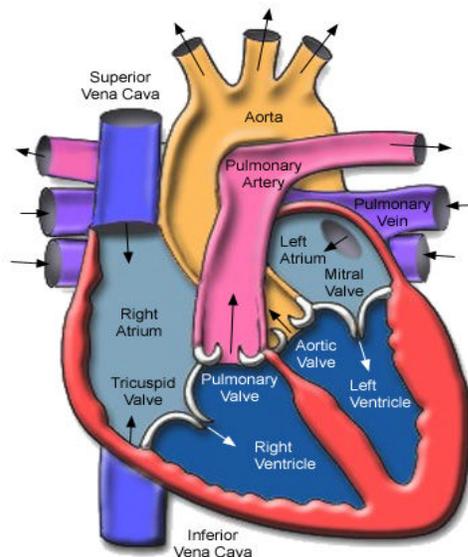

Figure 2.1 Anatomy of the heart.[1]

The right atrium receives blood from the superior vena cava and inferior vena cava, which returns venous blood to the heart from the body. The right atrium is a highly distensible chamber so that it can accommodate the venous return and maintain a low pressure (0-3 mmHg). Blood flows from the right atrium, across the tricuspid valve, and into the right ventricle. The wall of the right ventricle is not as thick as the left ventricle, and anatomically, it wraps itself around the major part of the left ventricle. The outflow tract of the right ventricle is the pulmonary artery which is separated from the ventricle by the semi-lunar pulmonary valve. Blood returns to the heart from the lungs via four pulmonary veins that enter the left atrium. The left atrium, like the right, is highly compliant, but less compliant than the right atrium. Therefore, the left atrial pressure is higher than the right atrial pressure (6-10 mmHg). Blood flows from the left atrium, across the mitral valve, and into the left ventricle. The left ventricle has a very thick muscular wall so that it can generate high pressure during contraction. Blood from the left ventricle is pumped across the aortic valve and into the aorta.

The tricuspid and mitral valves (atrioventricular valves) have fibrous strands of connective tissue known as chordae tendineae attached to their leaflets and to papillary muscles located on the respective ventricular walls. The papillary muscles contract during ventricular contraction and generate tension on the valve leaflets via the chordae tendineae to prevent the atrioventricular valves from bulging back into the atria. The semi-lunar valves do not have such analogous attachments.

2.1.2 Heartbeat Cycle

The heartbeat cycle consists of two parts the diastole and systole. Figure 2.2 illustrates what happens during diastole and systole.
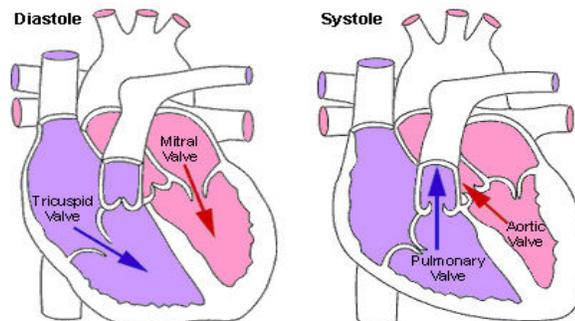

Figure 2.2 Diastole and Systole.[1]

Diastole occurs when the heart is relaxed and not contracting. During diastole the blood fills each atrium and then starts filling each ventricle by pushing blood through the tricuspid and mitral valves.

During systole, the electrical impulses conducted by specialized conducting fibres trigger the heart walls to contract. The left and right atria contract at nearly the same time pumping the remaining blood into ventricles. Systole continues pushing the oxygen-rich blood from the left ventricle to other parts of the body and de-oxygenated blood from the right ventricle to the lungs. The tricuspid and mitral valves close to prevent the back flow of blood.

2.1.3 Cardiac Conduction System and the ECG

When vertebrate muscles are excited, an electrical signal called an "action potential" is produced and spreads to the rest of the muscle cell, causing an increase in the level of calcium ions inside the cell. The calcium ions bind and interact with molecules associated with the cell's contractile machinery, the end result being a mechanical contraction. Even though the heart is a specialized muscle, this fundamental principle still applies.

2.1.3.1 Conduction System

An electrical stimulus is generated by the sinoatrial node, or SA node, which is a small mass of specialized tissue located high on the right atrium close to where the superior vena cava enters the right atrium. The SA node generates an electrical stimulus periodically at about 60-100 times per minute under normal conditions. This electrical stimulus travels across the right and left atria and causes atrial contraction.

The electrical impulse travels from the SA node to the atrioventricular (AV) node, located on the interatrial septum close to the tricuspid valve (it takes about 0.03 seconds for the impulse to travel from the SA node to the AV node). The impulse then continues down the conduction pathways via the Bundle of His into the ventricles. The conduction through the AV node is slow providing a delay. The delay at the AV node allows sufficient time for all of the blood in the atria to fill their respective ventricles.

The Bundle of His is located in the proximal intraventricular septum. From the AV node, the impulse travels rapidly through the Bundle of His or the

atrioventricular bundle. This impulse travels to the right and left bundle branches. Theses bundle branches extend to the right and left sides of the interventricular septum and the apex of the heart. The bundles give rise to thin filaments known as Purkinje fibres. The Bundle of His, bundle branches and the Purkinje fibres rapidly transmit impulses throughout the ventricular myocardium so that both ventricles contract at the same time. The AV node together with the Bundle of His forms the AV junctional tissue which has an intrinsic rate of 40-60 beats per minute. If the SA nodes are injured, AV junctional tissue can take over control of heart rate and rhythm.

As the ventricles contract, blood is forced out through the ventricular valves into the pulmonary trunk and the ascending aorta. After the ventricles complete their contraction phase, they relax and the SA node initiates another impulse to start another cardiac cycle.
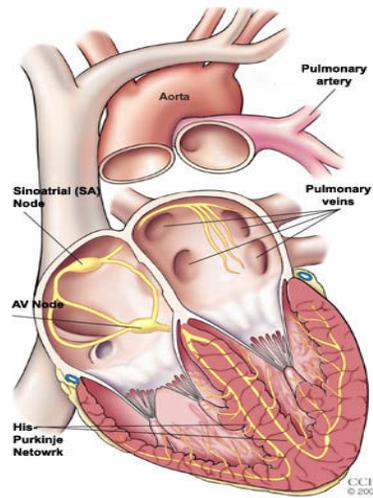

Figure 2.3 Conduction system of the heart.[6]

Any of the electrical tissues in the heart have the ability to be a pacemaker. However, the SA node generates an electric impulse faster than any other tissue so it is normally in control. If the SA node should fail, the other parts of the electrical system can take over, although usually at a slower rate.

2.2 The Electrocardiogram (ECG)

As the heart undergoes depolarization and repolarization, the electrical currents that are generated spread not only within the heart, but also throughout the body. This electrical activity generated by the heart can be measured by an array of electrodes placed on the body's surface. The recorded tracing is called an electrocardiogram (ECG, or EKG). An idealised schematic of a "typical" ECG with its component waveforms is shown below in Figure 2.4.
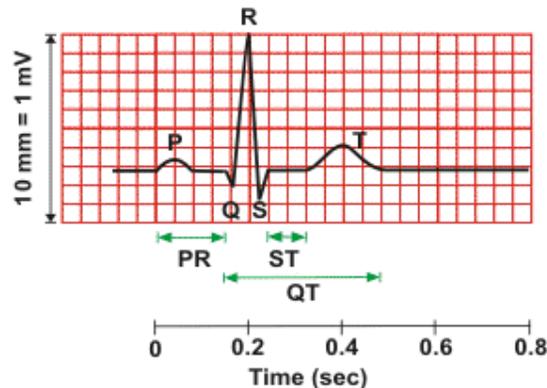

Figure 2.4 Typical electrocardiogram.[35]

The different waves that comprise the ECG represent the sequence of depolarization and repolarization of the atria and ventricles.

P Wave

The P wave represents the wave of depolarization that spreads from the SA node throughout the atria, and is usually 0.08 to 0.1 seconds in duration. The zero voltage, period after the P wave represents the time in for which the impulse travels within the AV node and the Bundle of His.

- The relationship between P waves and the QRS complexes helps distinguish various cardiac arrhythmias.
- The shape and duration of the P waves may indicate atrial enlargement.

QRS Complex

The QRS complex represents depolarisation of the ventricles. The QRS complex is larger than the P wave. This is because the ventricles contain more mass than the atria. The QRS complex appears spiked rather than a rounded smooth curve because

of the increase in conduction velocity in the Bundle of His and Purkinje fibres. A normal QRS complex is 0.06 to 0.10 seconds in duration. The shape of the QRS complex in Figure 2.4 is an idealised waveform and its shape may change depending on the nature of the electrodes used and their positioning. This shape may even change if there is any irregularity in the conduction of electrical impulses through the ventricles. Also, not every QRS complex contains a Q wave, an R wave, and an S wave. Any combination of these waves can be referred to as a QRS complex. The shape, duration and amplitude of the QRS complex are very important in diagnosing cardiac arrhythmias, conduction abnormalities and other heart disease states.

T Wave

The T wave represents ventricular repolarization and is longer in duration than depolarization.

U Wave

Sometimes a small positive U wave may be seen following the T wave. Figure 2.4 does not show this U wave as it is not always seen. The U wave is very small and represents the repolarization of the papillary muscles or Purkinje fibres.

PR Interval

The period of time from the onset of the P wave to the beginning of the QRS complex is termed the PR interval. It is usually 0.12 to 0.20 seconds in duration. This interval represents the time between the onset of atrial depolarization and the onset of ventricular depolarization. If the PR interval is prolonged, it may indicate a first-degree heart block.

ST Segment

The ST segment connects the QRS complex and the T wave and is normally 0.08 to 0.12 sec in duration. It represents the brief time for which the ventricles are in a depolarized state. The ST segment is important in the diagnosis of ventricular ischemia or hypoxia, because under those conditions it can become either depressed or elevated.

QT Interval

The QT interval is measured from the beginning of the QRS complex to the end of the T wave. It represents the time of depolarization and repolarization of both ventricles. The duration of the QT interval ranges from 0.2 to 0.4 seconds depending

on heart rate. A normal QT interval is usually about 0.4 seconds. Prolonged QT intervals can be diagnostic for susceptibility to certain types of tachyarrhythmias.

2.3 Arrhythmias

The electrical impulse generated at the SA node normally governs the rhythm of the atria and ventricles. The normal rhythm of the heart is very regular, with minimal fluctuations and is designed to ensure maximal efficiency and optimal performance. An arrhythmia (also termed a dysrhythmia) is an abnormal rhythm of the heart. It is a condition in which the heart rhythm becomes irregular, too fast (tachycardia) or too slow (bradycardia), or the frequency of the atrial and ventricular beats are different. Arrhythmias are very common and affect more than 700,000 people in the UK [9] and about 14 million people in the USA [35].

Arrhythmias can cause problems with contractions of the heart chambers by:

- not allowing the ventricles to fill with an adequate amount of blood because the electrical signal is causing the heart to pump too fast.
- not allowing a sufficient amount of blood to be pumped out to the body because the electrical signal is causing the heart to pump too slowly or too irregularly.

Everyone experiences some variation in their heartbeat at certain times and may occasionally feel palpitations (awareness of a different heart beat). Some arrhythmias are life threatening medical emergencies that can cause cardiac arrest and sudden death, whereas some are quite small and normal. Examples of normal changes in the heart rate include the following:

- Sinus Arrhythmia: A condition in which the heart rate varies with breathing. Sinus arrhythmia is commonly found in children; adults may often have it as well. This is condition where there may be no symptoms or problems.
- Sinus Tachycardia: A condition in which the heart beats faster i.e. the heart rate is more than 100 beats per minute. This is a normal response in situations when the body is under stress from exercise, strong emotions, fever, or dehydration. The heart regains its normal heart rate once the stress is removed.

- Sinus Bradycardia: A condition in which the heart beats slower i.e. the heart rate is less than 60 beats per minute. This is a normal response when the body is exposed to the cold or has a low body temperature.

- Ectopic Beats: These are extra heart beats and are normal. This condition can be made worse by consumption of caffeine. Ectopic beats are normal and are generally not dangerous and do not damage the heart.

Some arrhythmias that are fatal and life threatening, and cause cardiac arrest or death, are described below.

- Atrial Fibrillation: This is a condition in which the electrical signals emerge from the atria at a very fast and erratic rate. As a result, the ventricles contract in an irregular manner. Because the impulses travel through the atria in an irregular fashion, it results in loss of coordinated atrial contraction. This irregular movement causes turbulent flow of blood in the heart which can lead to the formation of blood clots. These blood clots can move to other parts of the body like the brain, causing a stroke. A typical ECG with atrial fibrillation is shown in Figure 2.5.

Figure 2.5 Atrial fibrillation.[3]

- Atrial Flutter: This is a condition in which the electrical signals come from the atria at a fast but even rate. As a result the ventricles contract faster and increase the heart rate. This can be seen very well in an ECG as a "sawtooth" pattern of flutter waves between each QRS complex as shown in Figure 2.6.
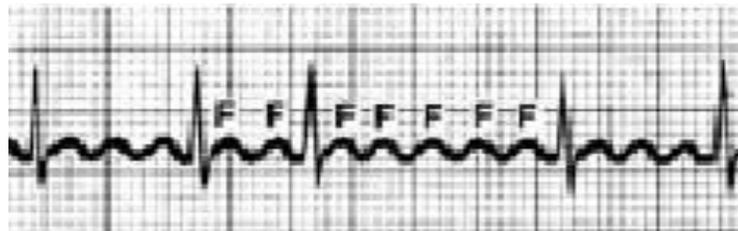
Figure 2.6 Atrial flutter.[3]

- Atrial Tachycardia: It is also known as supraventricular tachycardia (SVT), and occurs when an area in the atrium other than the SA node also produces electrical impulses. This overrides the pacemaker with regular impulses in rapid succession and results in an increased heart rate with 140 to 240 beats per minute. Atrial tachycardia can begin suddenly causing the ventricles to contract rapidly. An attack from atrial tachycardia can last for a few minutes and urgent treatment is required to stop the attack. A typical ECG with atrial tachycardia is shown in Figure 2.7.

Figure 2.7 Atrial tachycardia.[3]

- Ventricular Tachycardia: This is a condition in which an electrical signal is sent from the ventricles at a very fast but often regular rate. The rapid contraction of the ventricles prevents the heart from filling adequately with blood and hence less blood is pumped through the body. Ventricular tachycardia usually causes a sudden collapse. It often happens after a heart attack and if immediate medical attention is not given may result in death. A defibrillator is normally used to retain the normal rhythm of the heart. Figure 2.8 below shows a typical ECG reading of ventricular tachycardia.
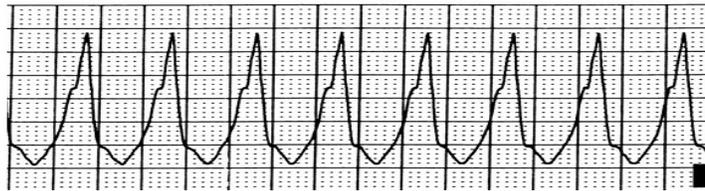
Figure 2.8 Ventricular tachycardia.[3]

- Ventricular Fibrillation: This is a condition in which many electrical signals are sent from the ventricles at a very fast and erratic rate and occurs when the heart stops beating and just flutters. The ventricles are unable to fill with blood and pump. As a result, blood is not circulated to the brain and other

parts of the body and there is a complete loss of pulse and consciousness. A person with ventricular fibrillation needs immediate medical attention and must be treated with cardiopulmonary resuscitation (CPR) and defibrillator as soon as possible. The ECG in this case shows random, apparently unrelated waves. Usually, there is no recognizable QRS complex as shown in Figure 2.9.
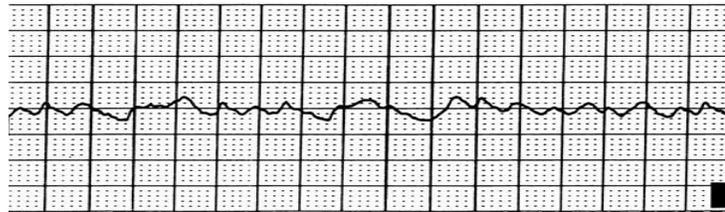

Figure 2.9 Ventricular fibrillation.[3]

Chapter 3

Biosignals

3.1 Biosignals

Biosignals are physiological signals that can be measured and monitored from biological beings. Biosignals not only refer to electrical signals but also to non-electrical signals. Electrical biosignals are electric currents produced by the summation of potential differences across a specialized tissue, organ or cell system like the nervous system. Some of the well-known electrical biosignals include the electrocardiogram (ECG), electroencephalogram (EEG), electromyogram (EMG), galvanic skin response (GSR), and magnetoencephalogram (MEG). Non-electrical biosignals that can be recorded and monitored from a biological being include mechanical signals (e.g. mechanomyogram (MMG)), acoustic signals (e.g. phonetic and non-phonetic utterances, breathing) and visual signals (e.g. movements).

3.2 Theory

3.2.1 Non-linear Theory

Everywhere in our daily life we encounter non-linear phenomena. In fact, it is well known that real life is non-linear [20]. Non-linear problems are difficult to solve and are often linearized based on certain assumptions. These assumptions may lead to loss of important information in the system. Linear methods that have been used in the field of biomedical signal analysis and classification are now slowly starting to show their limitations. It seems that these techniques are not always powerful enough to analyze signals that originate from very complex non-linear living systems.

Non-linear systems do not follow the principle of superposition i.e. non-linear systems do not adhere to the properties of additivity and homogeneity like linear systems. In other words the behaviour of a non-linear system does not depend on the summation of its constituent parts or their multiples. For example a linear function $F(x)$ adheres to the properties of additivity (eqn. 3.1) and homogeneity (eqn. 3.2) in the following equations.

$$F(a+b) = F(a) + F(b) \tag{3.1}$$

$$F(ab) = aF(b) \tag{3.2}$$

Almost all interesting physical systems in existence are non-linear systems. Linear physical systems are very rare in nature. The differential equation of the motion of a pendulum is non-linear. Also the equation: $y = ax(1-x)$, referred to as the logistic equation in population biology, is an example of a non-linear system. The x^2 term of this equation gives rise to a parabola [28]. The linearity of a system allows us to make certain assumptions and approximations for simplification of the system to draw results. However, in non-linear systems these assumptions cannot be made, as non-linear systems cannot be described as additive or homogeneous. Another major complication is that non-linear systems are composed of multiple sub-units that cannot be understood by analyzing these components individually, because these sub-units of a non-linear network interact i.e. they are coupled. An excellent example is the "cross-talk" of pacemaker cells in the heart. Their non-linear coupling generates behaviours that defy explanation using traditional (linear) models as shown in Figure 3.1. As a result, they may exhibit behaviour that is characteristic of non-linear systems, such as self-sustained, periodic waves as in ventricular tachycardia, or abrupt changes in cases of seizure and, possibly, chaos [28].

Universality is a unique property of non-linear dynamics, established from previous experiments that the non-linear systems which might appear to be very different in their specific details are often seen to exhibit common patterns of response. Yet another unique, important and universal class of non-linear transition is called the bifurcation. This explains the situation where a very small variation in any parameter of the system will result in an abrupt change in behaviour of the system. The sudden appearance of regular oscillations that alternate between two values is a universal class of bifurcation. The most common example for this is the beat-to-beat alternations in QRS axis and amplitude seen in cardiac tamponade. This has more to do with the back and forth swinging motion of the heart within the pericardial effusion. Some other examples of alternans in cardiac physiology are the ST-T alternans which may precede ventricular fibrillation and pluses alternans during heart failure. As a result of these reasons, non-linear systems are very difficult to predict and almost impossible to model.
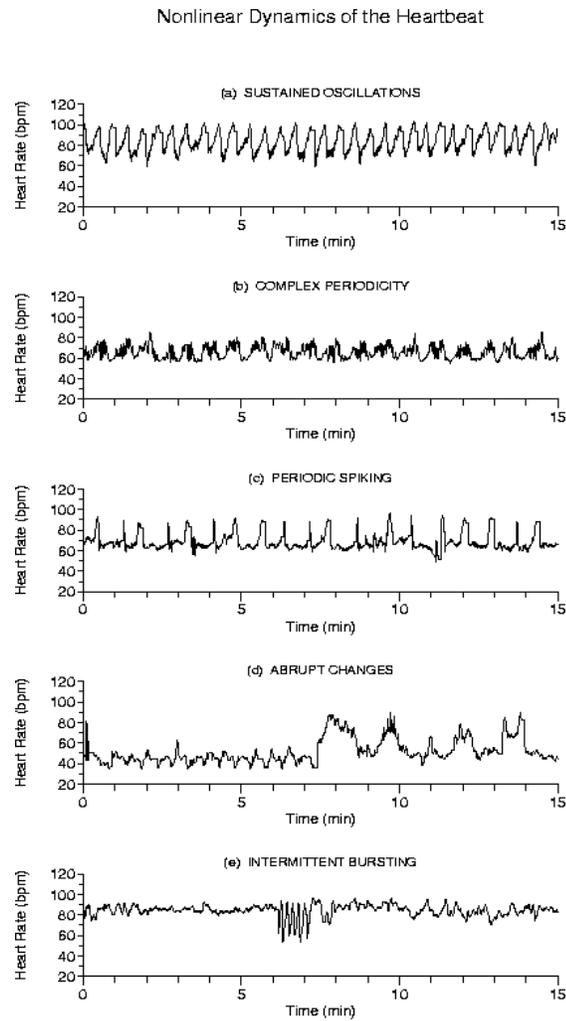
Figure 3.1 Non-linear dynamics of heartbeat.[10]

Figure 3.1 shows various examples of non-linear dynamics of heartbeat. Panels (a)-(c) are from subjects with obstructive sleep apnea syndrome. Panels (d) and (e) are from healthy subjects at high altitude (~15,000 ft).

3.2.2 Dynamic Systems

A dynamic system can be described as the formalisation in mathematical terms of any rule which defines the time dependence of a point's position in its ambient space. For example, we have different phenomena like water flowing in a pipe, or the number of tadpoles every summer in a pond, which have mathematical models to

characterise them. These models are used for financial and economic forecasting, weather forecasting, medical diagnosis, and many other applications.

Application of these systems falls into the following three major categories:

1. Predictive (also referred to as generative), in which the objective is to predict future states of the system from observations of past and present states of the system.

2. Diagnostic, in which the objective is to infer what possible past states of the system might have led to the present state of the system (or observations leading up to the present state).

3. Neither predictive nor diagnostic but rather to provide a theory for the physical phenomena.

These three categories correspond roughly as a need to predict, explain, and understand physical phenomena, respectively.

The state of a dynamic system can be defined by a set of points or real numbers in a state space [14]. Small change in the state of the system under study, leads to changes in these set of points. A rule that describes the future states of a dynamic system from the current state is known as the evolution rule of the dynamic system. This rule is deterministic in nature i.e. for a given time; only one future state follows the current state. Hence, determining the state for all future times requires iterating the relation many times, advancing a small step each time [14].

For example, the state of a system at time 't' is an instantaneous description of the system. This information is sufficient to predict future states of the system, without the knowledge of states prior to 't'. The evolution of the system from the time 't' forms a sequence or continuous trajectory. This sequence can be considered as the space of possible system states in future proceeding time 't'. This space of possible system states is called the state space of the dynamic system. If this rule is well chosen, then the way these points move around the state space move matches the way the measured values of variables change over time. Thus, dynamic system paves a way for the analysis and prediction of chaotic signals [14].

3.2.3 Chaos Theory and Fractals

Chaos is traditionally thought of as being confusion, hysteria and turmoil. Chaos, however, in the sense of chaos theory, is the idea that the final outcome of any event can be extremely sensitive upon some initial conditions. Amazingly, brief organized patterns can be found within chaotic systems. Chaotic systems have three main properties of sensitivity, mixing and periodicity. As chaos theory is still a relatively new field of research, its properties may be apt to change in the near future [3].

Fractals are the models generated by mathematical equations which result in chaotic systems. Fractals are very artistic, complex, and intricate. They also have important properties which include (a) a fine structure, (b) defined by a recursive process, (c) too irregular to be described by traditional geometry, (d) self-similarity, and (e) fractal dimension [20, 28].

In biology, chaotic systems can be used to show the rhythms of heartbeats, walking strides, and even the biological changes of aging. Fractals can be used to model the structures of nerve networks, circulatory systems, lungs and even DNA [28, 43].

3.2.4 Chaos Theory

"Chaos theory is the qualitative study of unstable aperiodic behaviour in deterministic non-linear dynamic systems."[20]. A dynamic system is defined as simplified model of the time-varying behaviour of an actual system, and aperiodic behaviour, is the behaviour that occurs when no variable describing the state of the system undergoes a regular repetition of values. Aperiodic behaviour never repeats and it continues to manifest the effects of any small perturbation thus making it impossible to predict any future state of the system. Human history is an excellent example of aperiodic behaviour, where the events of rise and fall of civilizations are observed and analysed and it can be noticed that no events ever repeat exactly.

An interesting fact about chaos theory is that unstable aperiodic behaviour is found in mathematically simple systems, which display behaviour so complex and unpredictable that they can be considered as random in nature. This effect is known as the butterfly effect. Ian Stewart, FRS, a professor of mathematics at the University of Warwick, describes this phenomenon as follows:

"The flapping of a single butterfly's wing today produces a tiny change in the state of the atmosphere. Over a period of time, what the atmosphere actually does diverges from what it would have done. So, in a month's time, a tornado that would have devastated the Indonesian coast doesn't happen. Or maybe one that wasn't going to happen does."[46].

The phenomenon of butterfly effect common to chaos theory is sensitive dependent to initial conditions. A long term behaviour of a system radically changes because of small changes in initial conditions. These small changes are generally considered as experimental noise, or background noise, or an inaccuracy of the equipment, which are impossible to avoid even under most ideal conditions of an isolated lab. For example the difference between the final results of a non-linear chaotic system, with initial values of 2 and 2.000001 will be very large. In fact the results will be totally different for the two values [46].

Chaos was first discovered in 1960 by a meteorologist, named Edward Lorenz. To understand chaos, Lorenz performed a simple experiment using a waterwheel, which is today known as the Lorenzian Waterwheel Experiment. Lorenz took a waterwheel with eight buckets spaced evenly around its rim with a small hole at the bottom of each bucket. The buckets were mounted on swivels, so that the buckets would always point upwards. The entire system was placed below a waterspout. A slow, constant stream of water flowed from the waterspout into the buckets and the waterwheel began to spin at a constant rate. Lorenz then increased the flow of water and an interesting phenomenon occurred. The increased velocity of the water caused a chaotic motion of the waterwheel. At increased velocity, the waterwheel revolved in one direction as before but then with a sudden jerk it started moving in the opposite direction. The synchronized filling and emptying of the buckets no longer existed and the whole system was chaotic in nature. Lorenz observed this phenomenon for hours and recorded the position and contents of the buckets. Lorenz observed that there was never an instance when the waterwheel was in the same position twice and the waterwheel continued its chaotic behaviour without ever repeating any of its previous conditions [20]. The graph of the waterwheel is known as the Lorenz attractor, which is a chaotic map and is noted for its butterfly shape, and the equations that govern the Lorenz attractor are

$$\frac{dx}{dt} = \sigma(y - x) \tag{3.3}$$

$$\frac{dy}{dt} = x(\rho - z) - y \tag{3.4}$$

$$\frac{dz}{dt} = xy - \beta z \tag{3.5}$$

where σ is called the Prandtl number and ρ is called the Rayleigh number. The Figure 3.2 below shows the 2-D evolution of the trajectories in the Lorenz attractor starting at different initial points.
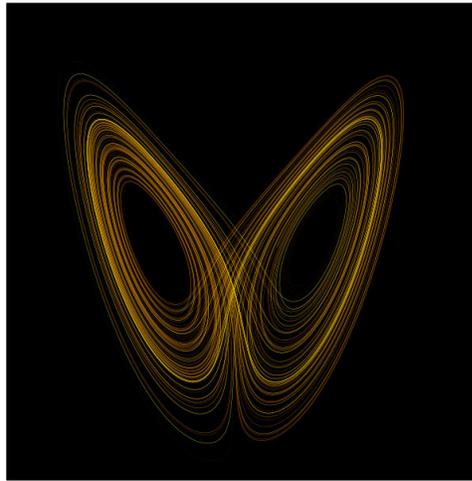

Figure 3.2 Lorenz attractor.[3]

Lorenz had obviously made an immense breakthrough in not only chaos theory, but life. "*Lorenz had proved that complex, dynamical systems show order, but they never repeat. Since our world is classified as a dynamical, complex system, our lives, our weather, and our experiences will never repeat; however, they should form patterns.*" [20]

The human heart has a chaotic pattern. The time between beats does not remain constant, for example it depends on how much activity a person is doing. In certain conditions, the heartbeat speeds up, whereas in other conditions the heartbeat will be erratic and is called a chaotic heartbeat. The analysis of a heartbeat can help medical researchers find ways to put an abnormal heartbeat back into a steady state, instead of uncontrolled chaos [27].

Goldberger, working at Harvard Medical School, analyzed normal and diseased hearts. He found that normal hearts usually stayed within a healthy range, but within that range their behaviour was chaotic. Furthermore he observed that healthy hearts showed more variability in their beating than sick ones [27]. Babloyantz and Destexhe compared epileptic and normal brain EEGs and found that both signals showed chaotic dynamics, but a normal brain was much more chaotic than when under an epileptic seizure [15].

3.2.5 Fractals

Fractal is a term coined by Mandelbrot to refer to irregular fragments or shapes. The term fractal is a geometric concept related to, but not synonymous with chaos [38]. It can be defined as a rough or fragmented geometric shape, made up of small subunits which resemble the whole. Traditional geometric forms are smooth and regular and have integer dimensions (1, 2, and 3, for line, surface, and volume respectively), while fractals are highly irregular and have non-integer or fractional dimensions. A fractal line unlike a smooth Euclidean line is wrinkly and irregular. Magnification of these wrinkles with the low-power lens of a microscope reveals smaller wrinkles on the larger ones. Further magnification shows yet smaller wrinkles, and so on.

Fractals exhibit two important properties

1. Self-similarity: i.e. they appear similar at all levels of magnification. A fractal shape will look almost, or even exactly, the same no matter what size it is viewed at. This repetitive pattern gives fractals their aesthetic nature.
2. Non-integer dimensions: This means that they are entirely different from the graphs of lines and conic sections that belong to Euclidean geometry.

A large variety of natural objects show the property of self-similarity, which includes branching of trees, coral formations, wrinkly coastlines, and ragged mountain ranges. Also a number of cardiopulmonary structures have a fractal-like appearance. Fractal anatomies include the arterial and venous trees, tracheobronchial tree and His-Purkinje network. Figure 3.3 below shows some of the existing fractals in nature and the human body.

The self-similarity of cardiopulmonary structures serves a common physiologic function of fast and efficient transport over a complex and spatially distributed

system. For example the fractal tracheo-bronchial tree (as shown in Figure 3.3) provides an enormous surface area for exchange of gases at the vascular-alveolar interface, efficiently coupling pulmonary and cardiac function. Peskin CS et al. have shown how fractal organization of connective tissue in the aortic valve leaflets relates to the efficient distribution of mechanical forces [43].


(a)

(b)

(c)

(d)
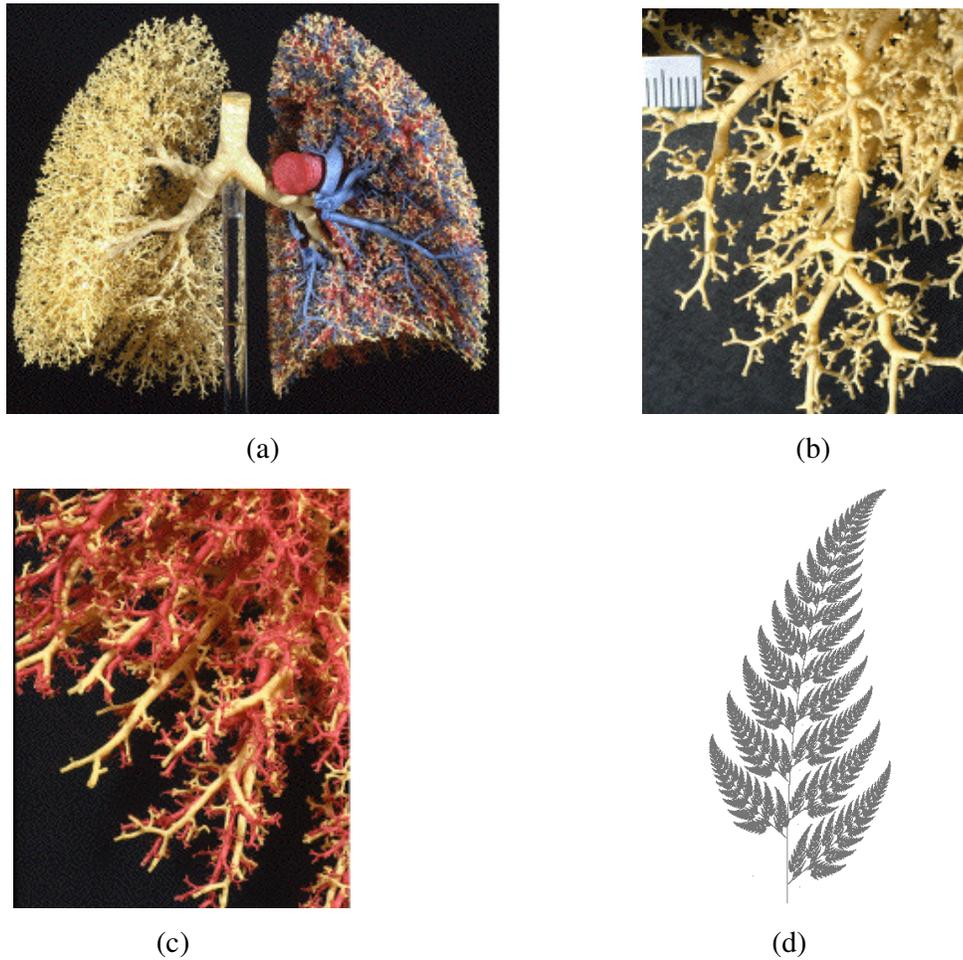
Figure 3.3 Fractals (a) lungs as whole (b) the branching bronchiole tubes (c) the bronchiole tubes and the branching arteries (d) fern leaf.[3, 52]

Fractals are irregular in shape but not all irregular structures or erratic time series representations are fractals [38]. A distinctive type of long-range order is a key feature of the class of fractals observed in biology. This property generates correlations that extend over many scales of space or time for example heart rate at a

particular time, is related not, just to immediately preceding values, but to fluctuations in the remote past. Certain pathologies are marked by a breakdown of this long-range organization property, producing an uncorrelated randomness similar to "white noise." An example is the erratic ventricular response in atrial fibrillation over relatively short time scales as shown in Figure 3.4 below [28].
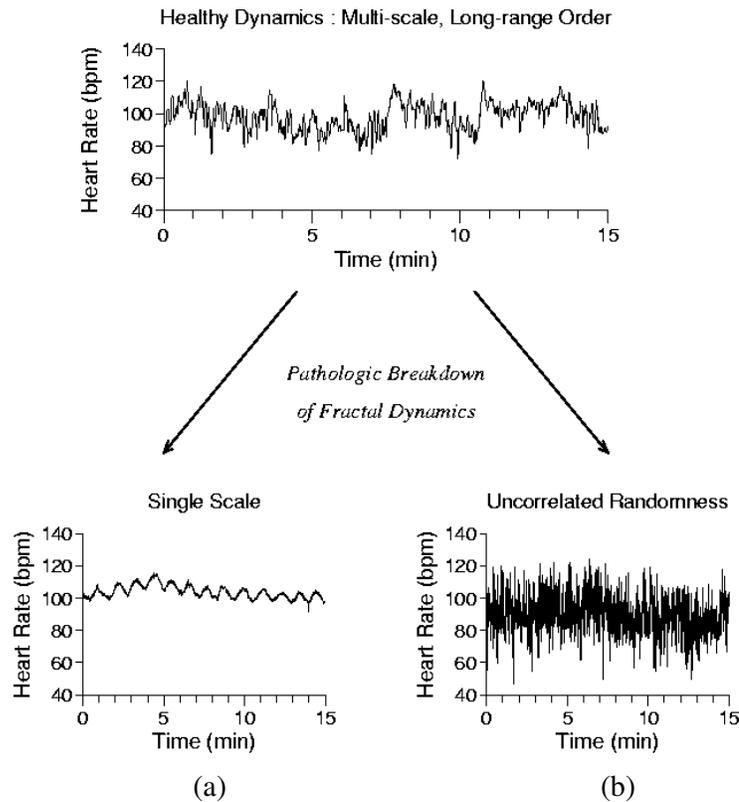
Figure 3.4 Pathological breakdown of fractal dynamics.[28]

Figure 3.4 shows what happens as the heart goes out of its normal state. Figure 3.4 (a) shows a subject with heart failure (the graph shows highly periodic values with little variation). Figure 3.4 (b) shows a subject with atrial fibrillation (the graph shows very erratic heart rate jumping from high value to low value without any specific pattern).

The list of applications of chaos theory and fractals is inexhaustible. A lot of breakthroughs have been made in the area of chaos and fractals. In order to explain more complex systems of nature, work must continue. As quoted by Stephen Hawking *"Understanding chaos is understanding life as we know it"*[20]

3.2.6 Fuzzy Logic

3.2.6.1 Introduction

Fuzzy logic was introduced in 1965 by Zadeh [51], and provides knowledge representation suitable for biological and medical problems because of its ability to work with imprecise information as well as with randomness.

Let us take an example of glycemia level of patients. Considering the threshold of 1.4g/l, all levels almost equal to threshold labelled 'normal' and those greater than the threshold labelled 'abnormal'. Boolean logic assumes that every fact is either entirely true or false (can never be both) i.e. the transition from one level to the other is very sudden because the level of 1.39g/l is considered as normal and the level of 1.41g/l is considered as abnormal. This is overcome by fuzzy logic by establishing a progressive passage from normal levels to abnormal levels. Fuzzy logic uses a membership function, which describes the degree of truth with real values between zero and one (inclusive). It can be achieved by considering the level is normal up to 1.3g/l and as it progresses from 1.3l/g to 1.5l/g the less normal it becomes and finally it can be considered really abnormal when greater than 1.5g/l. This can be seen in the Figure 3.5 below. Fuzzy logic uses a membership function which indicates the extent to which the value belongs to the acceptable range [14].
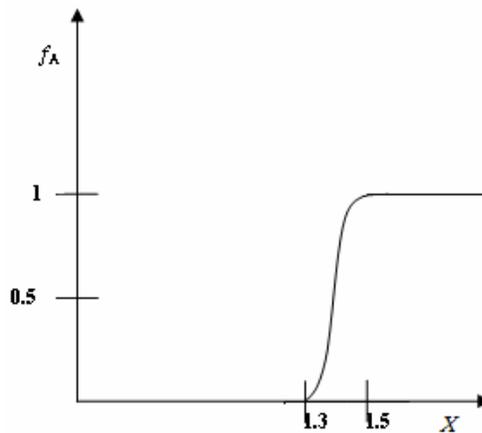
Figure 3.5 Fuzzy set *A* representing the category "abnormal" of the glycemia rate.[14]

3.2.6.2 Features

The uniqueness of fuzzy logic is the concept of partial membership of the element in a class or category i.e. an element in the set can be defined as 'somewhat abnormal'. This concept gives rise to a system of representing knowledge in a progressive manner from one level (normal) to the next level (abnormal). In simple words it can be stated as 'the more the value increases within the given limits, the more abnormal the level' [14]. This concept of knowledge representation is of immense use while modelling biological phenomenon, in which there is no strict boundary between adjoining situations.

This type of representation is very remarkable because it can be adjusted to the environment. For example, if the observed patients are elderly, the membership function of the class of abnormal glycemia levels needs to be adjusted accordingly. Another advantage of this approach is that one can set up an interface between numerical values and symbolic ones that can be expressed in natural language. For example, a patient with a glycemia level of 1.7g/l is associated with the symbolic value 'abnormal.'

3.2.6.3 Application

Thus, fuzzy logic is useful for representing imprecise knowledge with ill-defined boundaries. Fuzzy logic relating to control systems has many commercial applications. Some of these applications are listed below

- Automobile and other vehicle subsystems, such as ABS and cruise control (e.g. Tokyo monorail)
- Air conditioners
- Cameras
- Dishwashers
- Elevators
- Washing machines and other home appliances
- Video game artificial intelligence

Fuzzy logic also has many applications in the field of medicine and biology. Fuzzy techniques are used for diagnosis support systems, which take into account the

clinical indications which are difficult to describe precisely. Fuzzy logic also has extensive applications in the field of medical image processing, such as edge detection and segmentation [14].

3.2.7 Artificial Neural Networks (ANNs)

3.2.7.1 Introduction

Recently, there has been a growing interest in artificial neural networks (ANNs) also known as connectionist models, parallel distributed processing models and neuromorphic systems. ANNs simulate the biological nervous system and their development has been inspired by the ability of nervous system to deal with problems and the potential for parallel implementation. Years of study have been invested in the development of ANNs in the hope of achieving brain-like performance in the field of signal processing and pattern recognition. ANNs represent a promising new generation of information processing systems with the speed and precision of computers and analytical processing similar to the human brain [14].

3.2.7.2 Structure

Artificial neural networks are simple mapping functions that map the values of one or more inputs to single or multiple outputs. They are very similar to any ordinary transfer function except that they provide a greater degree of flexibility in mapping. For example, it is not necessary to have the knowledge of the function that governs the input to an ANN for mapping it. Also ANNs adapt to changes in the input and output i.e. if the inputs change the elements of the ANNs can be adjusted to continue to map the new inputs to the same output and vice versa. This property of ANNs is very promising for their application in signal processing and pattern recognition [14].

Structurally ANNs are similar to the human neural system i.e. they consist of identical computational cells that are interconnected in a fashion similar to the synaptic connection of neurons. These computational cells are also called processing elements (PEs). Figure 3.6 below shows a simple PE with 'n' inputs similar to the biological synaptic connections. The adjustable weights w's are shown that affect the impact of each input on the PE's output.
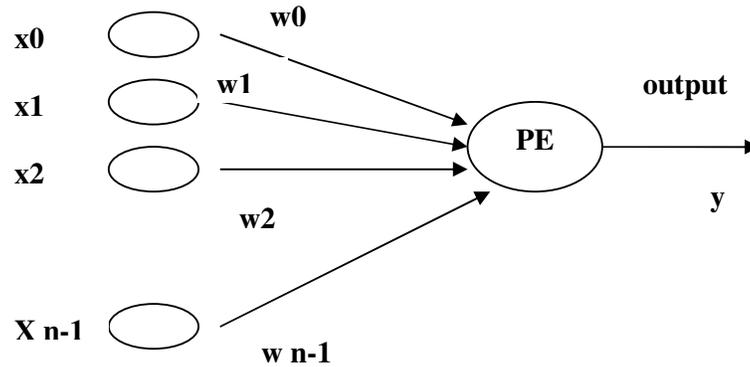
Figure 3.6 Simple artificial neuron consisting of a single PE.[14]

Particularly, the PE maps the inputs x_1 through x_{n-1} to the output y. The mapping consists of two steps. First, the input to the PE is summed to obtain α as shown in the equation below

$$\alpha = \sum_{i=0}^{n-1} w_i x_i \tag{3.6}$$

where x_i and w_i represent the magnitude of the inputs and their associated weight respectively for $i = 1,2,\ldots, n$-1. The PE output Y is computed as a function of α and an offset or threshold value θ as

$$Y = f(\alpha - \theta) \tag{3.7}$$

where f is called an activation function which is usually non-linear.

Thus, a PE unit can have n inputs and one output and a network of identical PEs that are interconnected in a series fashion encompassing an input and an output layer form an artificial neural network. A simple three-layer network of such ANNs form a network called a multilayer perceptron (MLP). In general a MLP has an input layer of source nodes, an output layer of neurons (computation nodes), and one or more "hidden" layers of neurons [14]. The equation for obtaining the output of the network shown in Figure 3.6 is as follows

$$Y = f(\sum_{i=0}^{n-1} w_i x_i - \theta) \tag{3.8}$$

3.2.7.3 Features

ANNs have a general ability to learn which make them potentially able to solve pattern recognition problems that are difficult. Thus, ANNs can be better classifiers than the existing conventional statistical classifiers which depend on the probability density function of the input data. Some of the important features of ANNs are:

- Adaptation
- Pursuing multiple hypotheses in parallel
- Fault tolerance
- Processing degraded or incomplete data
- Performing transformations

3.2.7.4 Application

ANNs have an immense potential to provide information about systems that are intractable and non-linear. This ability is demonstrated by their application in processing and analysis of bio-potentials, medical images, speech and auditory processing and so on. ANNs are more effective when they are used in conjunction with other signal and image processing techniques. Following is the list of a few applications of ANNs [14]:

- Detection and classification of biomedical signals
- Biomedical signal enhancement
- Biomedical signal compression

Chapter 4

Literature Review

4.1 Introduction

The traditional methods of analysing and classifying biomedical signals are the linear methods, which have many limitations. These techniques are not sufficient to analyze signals originating from complex and non-linear living systems [23]. Recently few methods have been developed from the theory of non-linear dynamics for analysis of time series which represent the signals measured from non-linear systems [11, 23, 32, 42]. However, the tools used for non-linear biosignal analysis are different from the ones used for linear biosignal analysis. Even the parameters that describe non-linear biosignals do not give a very clear interpretation like those of linear systems [23].

Non-linear analysis of biosignals (ECG in case of this project) usually deals with reconstruction of their state-space (time-delay plots), correlation dimension estimates [42], measures of the Lyapunov exponent [42], Kolmogorov entropy [42], fractal dimensions of the reconstructed phase space, fuzzy logic representation of knowledge [32], pattern recognition by artificial neural networks [31, 39, 41, 49]. Other techniques such as frequency domain features, time-frequency analysis, wavelet transform, adaptive filtering sequential hypothesis testing, as well as morphological features, have been used to transform the qualitative diagnostic criteria into a more objective quantitative signal feature.

4.2 Lyapunov Exponent

The Lyapunov exponent is an important feature of the chaotic system as it helps in quantifying its sensitivity. It also describes how small changes in the state of a system grow at a fairly exponential rate and then they go on to dominate the system completely. Lyapunov exponents are defined as the long time average exponential rates of divergence of nearby states.

Consider two close points at step n, x_n and $x_n + dx_n$. At the next time step they will have diverged, namely to x_{n+1} and $x_{n+1} + dx_{n+1}$. It is this average rate of divergence (or convergence) that the Lyapunov exponent captures. Another way to

think about the Lyapunov exponent is as the rate at which information about the initial conditions is lost [18].

$$\lambda = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \ln(\frac{dx_{n+1}}{dx_n}) \qquad (4.1)$$

If a system has at least one positive exponent, the system is said to be chaotic. The larger the positive exponent the more chaotic the system becomes. The Lyapunov exponent can be represented by λ_1, λ_2, λ_3 to λ_n corresponding to the most rapidly expanding and contracting principal axes in that order [42]. The magnitude of the Lyapunov exponent is a measure of the sensitivity to initial conditions, which is the most important characteristic of a chaotic system. The absolute value of the exponent indicates the degree of stability. If the Lyapunov exponent is less than zero then the system attracts to a fixed point or stable periodic orbit. If the Lyapunov exponent is zero then the system is neutrally stable.

Kolmogorov complexity is related to the sum of positive Lyapunov exponents is also a measure of chaotic nature of a signal. Kolmogorov complexity is maximized by the random strings present in the signal thus giving us the randomness of the signal or the data set. The fundamental theorem of Kolmogorov complexity for any two random signals x and y is stated as below

$$K(xy) \leq K(x) + K(y) + c \qquad (4.2)$$

Where $K(x)$ and $K(y)$ are the complexities of the respective signals, c is a constant and $K(xy)$ is the joint complexity of the concatenation of the signals [36].

4.3 Correlation Dimension

Another important parameter for the recognition of dynamic diseases from the ECG is the correlation dimension.

In non-linear dynamics, the correlation dimension of a signal can provide a lower bound for the actual dimension of the attractor, the degrees of freedom of the system, which is roughly equal to the number of state variables needed to describe the system. The correlation dimension estimates of cardiac rhythms, using various algorithms that have been proposed, such as the Takens' estimator [48], Judd's algorithm and the Grassberger-Procaccia method [29], have given an estimate of correlation dimension of 2.1 for sinus rhythm that increases to reach more than 7.8

during ventricular fibrillation. This increase corresponds to increase in dimensionality and complexity as also the number of variables governing the dynamics of the system.

Owis et al. have addressed the problem of characterizing the non-linear dynamics of the ECG signal and its variation with different arrhythmia types. They computed correlation dimension, one of the important chaotic system parameter, for a large number of ECG signals. To study the dynamics of the system they first reconstructed the state space trajectory using delay time embedding theorem to create a larger dimensional geometric object by embedding into a larger m dimensional embedding space. A suitably large m value was used so that the orbits of the system do not cross each other and was tested using the false nearest neighbour (FNN) algorithm [11].

The dimension D of an object represents the exponent that scales the bulk b of an object with linear distance r (i.e., $b \propto r^D$) [42]. Owis et al. used the Grassberger-Procaccia algorithm which uses a correlation integral $C(r)$ to represent the bulk b, which is defined as the average number of neighbours each point has within a given distance r. The correlation dimension (D_2) is defined as the slope of the liner region of the plot of $\log(C(r))$ versus $\log(r)$ for small values of r. The authors then used a second order regression for the whole curve to improve the implementation. For the part that appeared linear by vision, a linear regression was obtained. With this more consistent values for D_2 were obtained. An automatic algorithm was then developed by the authors to determine the linear region in order to eliminate the need for human interaction. The algorithm computed the second derivative of the $\log(C(r))$ against $\log(r)$ curve and searched for the longest plateau with values below certain threshold. In case there were more than one linear region which were found to have the same length, the one that yields the maximum D_2 value was chosen. The statistical analysis of the calculated D_2 values confirmed that normal ECG signals can be differentiated from abnormal. However, they were not successful in classifying between different abnormal signals.

Fojt et al. (1998) used the heart rate variability (HRV) signal for the analysis. They tried to make the HRV signal description in a state space simpler, understandable by applying the singular value decomposition method (SVD) [26].

This method helps in transforming the original state space into another state space whose coordinates are defined by the singular vectors of an $m \times n$ data matrix.

$$A = \left[x(\tau(i-1) + j), i = 1,.........,m, j = 1,.........,n \right] \qquad (4.3)$$

Where $x(\ )$ are values of the original samples of the analyzed signal, τ is a value of optimum Takens delay, n is the number of samples involved in decomposition and m is an embedding dimension. Since the information value of the new coordinates is dependent on the corresponding singular eigenvalues, the new coordinates can be arranged or selected thus eliminating the less important components and making the state space representation easier to survey.

Fojt et al. estimated the embedding dimension by using the method proposed by Grassberger and Procaccia [29] which also enabled them to calculate the correlation dimension for the HRV signal. The authors determined from their experiment that the correlation dimension of signals taken from the groups of patients of cardiac dysfunction, showed that the normal ECG signal had the smallest value of the correlation dimension from 2.1 to 2.8 with a mean of 2.46, standard deviation of 0.37. They also determined that the correlation dimension increases with the increase of pathology which was in range of 3.2 to 4.6, mean 3.81 and standard deviation 0.56. With regard to the HRV data, exactly opposite results were obtained, the highest values of the correlation dimension were obtained with signals from healthy people ranging from 5.3 to 7.6 and those values decreased significantly with pathology to 3.7 to 4.9.

4.4 Fuzzy Clustering

State recognition and event prediction are the two important tasks in biomedical signal processing. Some of the examples are tachycardia detection from ECG signals, epileptic seizure prediction from an electroencephalogram signal and prediction of vehicle driver losing control from both signals. Fuzzy clustering is used to examine a set of ordered measurements of the system behaviour and recognise temporal patterns that exist, which can be used to forecast an event or a transition between two different states of the biological system [14].

Biomedical signal processing using fuzzy clustering generally includes the following four steps:

1. rearrangement of the time series into temporal patterns for the clustering procedure,

2. dynamic state recognition and event detection by unsupervised fuzzy clustering,

3. system modelling using the non-continuous temporal pattern of each cluster, and

4. time series prediction by means of similar past temporal patterns from the same cluster of the last temporal pattern [14].

Fuzzy clustering algorithms are widely used for various applications, in which grouping of overlapping elements is necessary, for example (a) complex biomedical signals like EEG and ECG signals and (b) in medical images such as MRI and PET images [17]. Unsupervised fuzzy clustering is one of the most common and widely used methods for determining structure in a given set of data [25]. Weigend et al.(1992) combined the fuzzy clustering techniques with other common methods developing a deterministic versus stochastic algorithm for time series prediction of a system. Hata Y. et al. had published a paper describing an automated medical diagnosis system (AMDS), which determines a normal degree for a medical sign or disease based on fuzzy logic. Their whole concept of AMDS is based on the theory of hierarchical definability proposed by Zadeh.

4.5 ANN

A number of researchers have used artificial neural networks (ANNs) and a variety of features to detect and classify ECG beats. Watrous and Towell [49] used a multilayer perceptron (MLP) with back-propagation learning algorithm to discriminate between normal and ventricular beats. Marques et al. [39] used ANNs to discriminate between normal rhythms, ventricular hypertrophies, and myocardial infracts by feeding different ANNs with diverse set of features as input. Nandal and Bosan [41] used an MLP for discriminating between normal, ventricular and fusion rhythms.

Hu et al. [31] developed an adaptive MLP for classification of ECG beats. They classified the beats into two groups (normal and abnormal) with an average recognition accuracy of 90%. As per MIT/BIH database annotation they further tried to classify beats into 13 different groups and were successful with an average recognition accuracy of 65%. In order to improve this they developed a hierarchical system of MLP networks, which first classified the beat into normal or abnormal and then classified it into the specific beat type thus improving the recognition accuracy to 84.5%.

Barro et al. [16] developed a new ANN which considered the morphological features of heartbeats (detected on a multi-channel ECG signal) for its classification. They concentrated on the adaptive classification behaviour of the ANN with the capacity to self-organize itself dynamically in response to the characteristics of the ECG signal.

Kalayci and Ozdamar [34] developed a three-layer MLP feed-forward neural networks system using the back propagation learning algorithm to detect EEG spikes. They used wavelet transforms to extract features from spike and non- spike wave of EEG to train and test their ANN classifiers.

Xue et al. [50] developed an ANN whitening filter to model the low frequency components of the ECG signal, which were basically non-linear and non-stationary in nature. The filtered signal, mainly containing the higher frequency QRS energy, was then passed through a linear, matched filter to detect the location of QRS complex. The authors observed that the ANN whitening filter was very effective at removing the non-stationary noise of ECG signal.

Hamilton et al. [45] developed a compression algorithm using an autoassociative neural network for ECG signals. The first step was to detect the QRS complex and then compress it using the autoassociative ANN. The autoassociative ANN they used had similar input and output layers. A multilayer perceptron with back propagation learning algorithm was also used. This consisted of hidden layer with a reduced number of nodes to produce compression. As the beats in a given segment had similar gross morphology, the authors stored an average waveform and compressed only the difference, thus achieving optimum gain of the ANN's compression capability.

4.6 Wavelet Transform

The wavelet analysis methods are the most popular methods of processing biomedical signals. These methods have been representing the temporal characteristics of a biomedical signal by the help of its spectral components in the frequency domain. Wavelet analysis is quite important in the analysis of the non stationary signals whose spectral features are changing over a period of time now and it is also the main alternative for the analysis of such non stationary signals. Biological signals are also analysed by the method of wavelet analysis as the statistical characteristics of these signals are non stationary [13].

Wavelet-based QRS detection algorithms are mostly based on Mallat's approach for singularity detection using wavelet coefficients maxima [37]. Vladimir Johneff used a different approach by applying complex valued wavelets for QRS detection by computing the wavelet decomposition at one scale only, thus reducing the calculations [33].

The non-sustained ventricular tachycardia (NSVT) is defined as the three or more consecutive ventricular beats (VPBs) with 120 beats per minute or more and which lasts about 30 seconds [9]. This is usually asymptotic in patients with cardiac diseases. In patients with cardiac diseases the NSVT can act like a prognostic indicator of an increased risk of mortality and sudden cardiac death (SCD) [7]. Szi-Wen Chen brings to light that NSVT as a risk factor for SCD, is difficult to manage as there hasn't been a "cause and effect" relationship established between the two. However, the presence of NSVT is usually associated with an increased occurrence of subsequent sustained Ventricular Tachycardia (VT) and it is being considered as the precursor of SCD. A long term predictor of SCD and arrhythmic death after myocardial infraction has been recognised as a decreased HRV or an increased sympathovagal balance (SB) is quantified by spectral power ratio LF/HF [19].

Szi-Wen Chen performed a WT- based HRV analysis. The numerical results which have been derived by Szi-Wen Chen, from a database consisting of NSVT with coronary artery disease, ischemic and normal episodes indicated that the short time SB estimates encountered significantly sharp increases at few time instants within a period, just prior to NSVT. However the same phenomenon was not observed from the normal or the ischemic case. Szi-Wen Chen proposed that this

method could be the beginning for future WT based analysis of HRV which can be aimed at the prediction of NSVT episodes, which can be valuable for clinical use.

Addison et al. have used the energy based method of interrogating the ECG in ventricular fibrillation (VF) using high resolution, log scale continuous wavelet plots. The underlying structures within the VF waveform are made visible in the wavelet time-scale half space using the energy based method. The authors attempted to meet three objectives (a) to understand the pathophysiological processes occurring in sudden cardiac death, (b) to analyse the VF waveform to predict the usefulness of this therapy and (c) the use of alternative therapies like various electrical or pharmacological therapies to improve resuscitation success rates.

Wavelet transforms aid a signal to be decomposed such that frequency characteristics and the location of certain features in a time series may be highlighted simultaneously. Wavelet transform procedure helps in overcoming the shortcoming of the Fourier analysis procedure [12]. The shortcoming of the Fourier analysis is that the spectrum contains only the globally averaged information which leads to location specific features in the signal being lost. The wavelet transform for continuous time signal $x(t)$ is defined as

$$T(a,b) = \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} g\left[\frac{t-b}{a}\right] x(t) dt \qquad (4.4)$$

Where $g((t-b)/a)$ is the analysing wavelet function [23].

The transform coefficients $T(a,b)$ were found for both specific location ($t=b$) and for specific wavelet periods (which are a function of a). The plot of $T(a,b)$ against a and b in either a surface or contour plot is known as scalogram [12]. By definition, the analyzing wavelet has zero mean. The wavelet transform intrinsically removed the signal mean from the transform space. Therefore, the low frequency components associated with baseline drift did not occur in salient regions of the scalogram as deduced by Addison et al.

The wavelet–based noise reduction method is supposed to have a near optimal non-linear noise reduction. Due to this the wavelet transform in particular is considered as appropriate for the analysis of transient, non stationary signals containing high frequency components [21].

Sun et al. proposed the use of multiscale-based non-linear descriptor (Hurst index), to characterize the ECG episode, to differentiate ventricular tachycardia and ventricular fibrillation from normal sinus rhythm in the descriptor domain. Wavelet transform is an efficient tool for the multiscale analysis, which is a very useful framework for signal processing tasks [47]. "*The Hurst index is defined in the multiscale domain as a feature to quantify the non-linear dynamical behaviour (such as, self-similarity, roughness, and irregularity) of the ECG signal for detecting the life-threatening ventricular arrhythmia*" [47]. The Hurst index, *H*, is a single scalar parameter which describes the fractal Brownian motion model. Fractal Brownian motion model is a very useful model for non-stationary stochastic self-similar processes which have long term dependence over wide ranges of frequencies [22].

Sun et al. tested their new method on MIT-BIH malignant ventricular arrhythmia database and studied the relationship between the ECG episode length and the corresponding recognition performance. The performance of their experiment was very good with an accuracy rate as high as 100% for recognition of ventricular tachycardia and ventricular fibrillation from normal sinus rhythm. They also observed that the accuracy varied from 82.24% to 100% when differentiating between ventricular tachycardia and ventricular fibrillation. They noted that the increase in Hurst index is more from normal sinus rhythm to ventricular tachycardia as compared to the increase from ventricular tachycardia to ventricular fibrillation for a particular episode length.

4.7 Conclusions

To date, many novel methods for decomposition, analysis and display of the ECG signal using principles of non-linear dynamics, have been researched and presented by many researchers [12]. These methods have enabled us to display and measure the hidden characteristics of the ECG waveform.

Goldberger et al. believe that the practical applications of non-linear dynamics are likely within few years, mostly in the form of bedside physiological monitoring equipments. A number of indices derived from chaos theory and fractals have shown promising results in forecasting critical conditions by detecting subtle changes in heart rate oscillations and detecting certain pathologies by breakdown in fractal

scaling. In addition to the diagnostic applications, the study non-linear dynamics of biosignals can also have prospects in therapeutic interventions.

Garfinkel et al. found that a properly timed external stimuli is very effective in controlling certain mathematical or physical systems with complex dynamics. For example a chaotic system can be made more regular or periodic (chaos control) and periodic systems can be made chaotic (chaos anti-control). Garfinkel et al. proposed that certain arrhythmias may be controlled by development of chaos control algorithms that can electrically pace the heart back to normal sinus rhythm at the times of ventricular fibrillation.

Addison et al. believe that wavelet decomposition of ECG signal can be used to remove the cardiopulmonary resuscitation (CPR) artifact through temporal filtering carried out using a modulus maxima technique. This method would allow the assessment of the state of heart during resuscitation without stopping CPR. Addison et al. propose that detection of specific features within ventricular fibrillation waveform may lead to the modifications of the defibrillation technique.

However, the realization of the above mentioned applications requires the application of advanced post-processing techniques, which may include artificial neural networks (adaptive wavelet networks), entropy techniques and Bayesian statistics. The fundamental principle underlying these new applications and interpretations is the importance of analyzing various biosignals such as heart rate using non-linearity principles and not simply relying on averaged values obtained by linear techniques. '*Dynamical analysis often show that there is hidden information in physiological time series and that certain fluctuations previously considered "noise" actually represent important signals* ' [28].

Chapter 5

MATLAB Software Development and Implementation

5.1 Data

The aim of this thesis was to develop algorithms with the aid of MATLAB® software that can be used to analyse biosignals (ECGs in the present case). The first step before algorithm development was to obtain a sufficient variety of ECG signals to work on. These ECG signals were obtained from the Physionet ECG database [10]. These are discrete-time ECG signals in Excel sheets sampled at a certain frequency with the data in the form of Time vs. Voltage.

5.2 Software Development

Basic knowledge of object oriented programming as well as logical program flow was essential to get through with this stage. The code writing stage proceeded as follows

1. Visual representation of signals is the basic step to analyze the data further. The analog representation of the discrete time, sampled data is created using the plot function of MATLAB. The program 'Display_signal.m' is used for analog representation of the signals.

2. The 'Display_signal.m' program was then further modified to detect the R peaks in the signal. The program 'peak_detector.m' detects the positive and negative peaks in the signal. Following are the key steps used in the program to detect peaks.

    I. Firstly, it calculates and stores the size of the input signal. The size of the input vector is calculated using function $size(x)$, which returns the sizes of each dimension of array 'x'.

    II. Then it starts checking for the peaks above a set threshold value and finally displays the result. The *max* function is used to detect and store the peaks and index values of the input signal.

3. The program 'ecg_correlation.m' was developed which is used to calculate the percentage correlation between two signals. This program first reads the two signals and stores them. Then it checks the length of the two signals and

modifies the signals to the same length by padding zeroes at its end if one of the signals is shorter in length. The program then calculated the mean autocorrelation and the mean correlation coefficient using the predefined function in MATLAB. The program used basic beats for comparison of the two signals. Finally, it displayed the percentage correlation of the two signals.

4. The 'ecg_correlation.m' code was further modified and a new code 'correlation_refset.m' was developed. This program first reads all the signals in the folder and stores them as reference signals and creates data for further modification. This program uses the same logic of correlation. Just by giving the position of the signals in the reference set as input the program calculates the correlation of the two signals and displays it as output.

5. The new code 'find_refset.m' accepts a new unknown signal and indicates which category this unknown signal belongs. This program works similar to the previous program of 'correlation_refset.m'. However, now, when a new unknown signal is supplied as an input, the program calculates the correlation of this unknown signal with all the reference signals and then displays the reference signal with the highest correlation.

6. The above algorithms developed used linear analysis techniques. In order to compare linear techniques with non-linear techniques, another algorithm was developed for peak detection and heart rate using the discrete wavelet transform. The code developed is 'heart_rate_DWT.m'. The program first accepts the signal as input and then smooths the data using a moving average filter. The Matlab command used is 'smooth'. A moving average filter is mathematically represented by the following equation

$$y(n) = \frac{1}{M} \sum_{k=0}^{M-1} x(n-k) \qquad (5.1)$$

Further the signal is de-noised using wavelet transform and this filtered signal is further used for peak detection and heart beat rate computation.

7. Now, concentrating on the non-linear aspect of the signal processing, a new code was developed 'Kolmogorov_complexity.m' which calculates the complexity of the ECG signal. The Kolmogorov complexity $K(x)$ is computed using the entropy $H(p)$ of the weight of ones in the data set. $K(x)$ is defined in

equation (5.2) below where $x\#1$ is the number of 1 bits and $x\#0$ is the number of 0 bits in the data set whose complexity is to be determined. Entropy $H(p)$ is defined in equation (5.3) [36].

$$K(x) \approx l(x)H(\frac{x\#1}{x\#1+x\#0})+\log_2(l(x)) \qquad (5.2)$$

$$H(p) = -p\log_2 p - (1-p)\log_2(1-p) \qquad (5.3)$$

The program gives the complexity measure of the signal as an output which can be used to determine the randomness and chaotic behaviour of the signal.

8. Another important non-linear descriptor is the Lyapunov exponent. An algorithm 'Lyapunov.m' was developed which calculates the average Lyapunov exponent of the given ECG signal. The algorithm gives us state space representation of the signal and Fourier coefficient distribution of the signal along with average Lyapunov exponent.

    *   State Space Representation: This is a graphical device suited to detect hidden dynamical patterns and non-linearities in some data. To evaluate the non-linear characteristics of a signal, the single time series should be embedded in a state space. The ECG signal is a record of discrete time series consisting of N samples with a sampling interval, Ts. To embed the sequence in an M-dimensional state space, the time-delay reconstruction is used with a constant set of delays, $Td >= Ts$. The time series of N-samples is observed as a set of $N_m = (N-M+1)$ vectors, each consisting of M components corresponding to the delayed data samples by,

$$X_t = \left\{ x_t, x_{t-Td}, x_{t-2Td}, \ldots\ldots\ldots x_{t-(m-1)Td} \right\} \qquad (5.4)$$

    Described above is the Takens method of delay [25]. This representation adapted from the chaos theory helps retain the dynamical properties of the system.

    *   Fourier coefficient distribution: the following equation calculates the discrete fourier transform (DFT) of input vector x with N elements.

$$X(k) = \sum_{n=1}^{N} x(n).\exp\left[-j.2.\pi.(k-1).\frac{n-1}{N} \right], \quad 1 \le k \le N \qquad (5.5)$$

- Lyapunov exponent: The numerical calculation of Lyapunov exponent is as follows

    1. Start with any initial condition in the basin of attraction ($n/4$ in present case)

    2. Iterate until the orbit is on the attractor

    3. Select (almost any) nearby point (separated by d_0)

    4. Advance both orbits one iteration and calculate new separation d_1

    5. Evaluate $\log |d_1/d_0|$ in any convenient base

    6. Readjust one orbit so its separation is d_0 in same direction as d_1

    7. Repeat steps 4-6 many times and calculate average of step 5

    8. The Lyapunov exponent is $\lambda_1 = \log |d_1/d_0|$

    9. The average Lyapunov exponent is $\lambda_1 = (\log |d_1/d_0|) / n$

The results of these programs are shown and discussed in detail in the next chapters.

Chapter 6

Results and Analysis

The aim of this thesis was to understand the complex nature of an ECG signal and to develop algorithms that can be used for analytical purposes. To achieve this aim, five algorithms were written in MATLAB® software and tested on the ECG signals obtained from the Physionet website [10]. The results of the algorithms are shown and discussed in this chapter.

6.1 Results of Algorithm 1 (Time-Series Representation of ECG Signal)

The first algorithm is used for the time-series representation of the signal. Following are the results of some of the signals

(a)   (b)
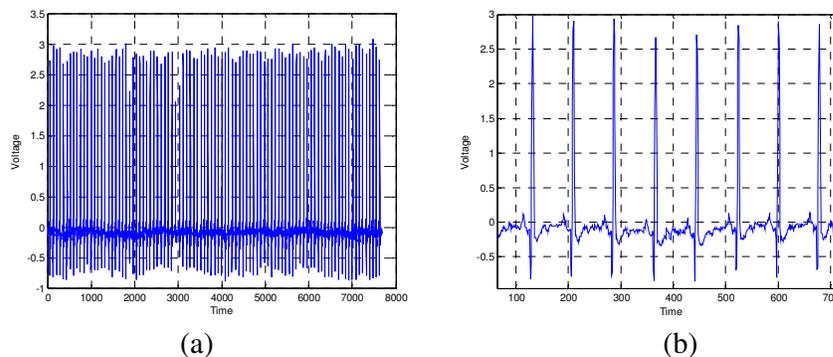
Figure 6.1 Normal sinus rhythm. (a) Entire signal (b) Zoomed-in section of the signal.

Figure 6.1 above shows the time-series representation of a normal ECG. Various features of the signal can be clearly distinguished in the zoomed-in part of the Figure 6.1 (b).

<center>(a)   (b)</center>
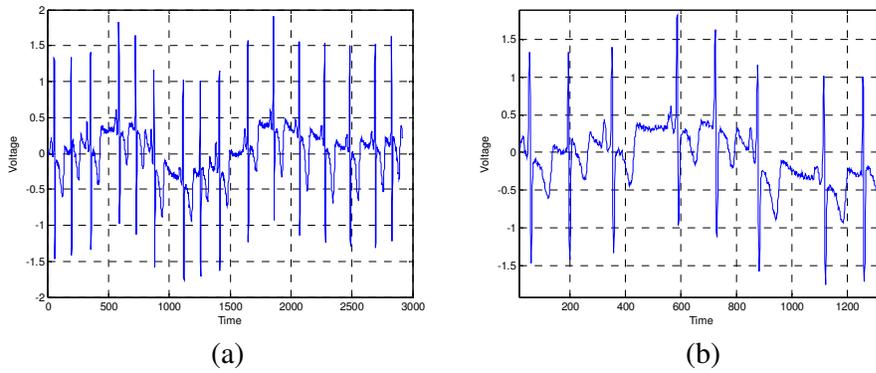
Figure 6.2 Atrial tachycardia.

Figure 6.2 above shows the time series representation of atrial tachycardia. The zoomed in section of Figure 6.2 (b) shows the occurrence of an abnormal P wave before each QRS complex.

<center>(a)   (b)</center>
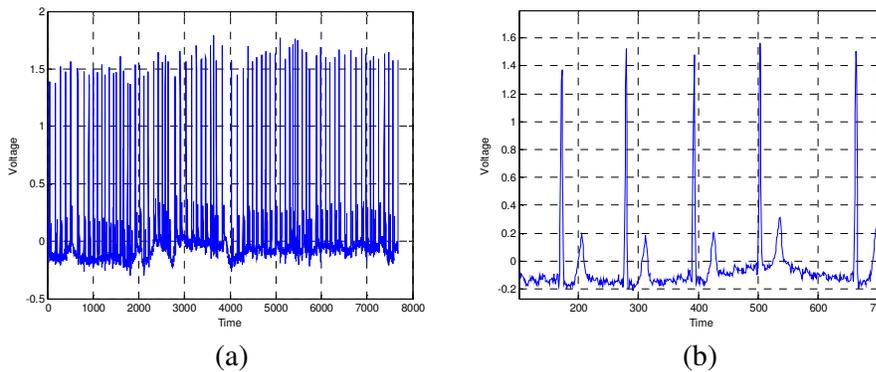
Figure 6.3 Atrial fibrillation.

Figure 6.3 shows the time series representation of atrial fibrillation.The zoomed in section of Figure 6.3 (b) shows the absence of a P-wave before each QRS complex.

(a)   (b)
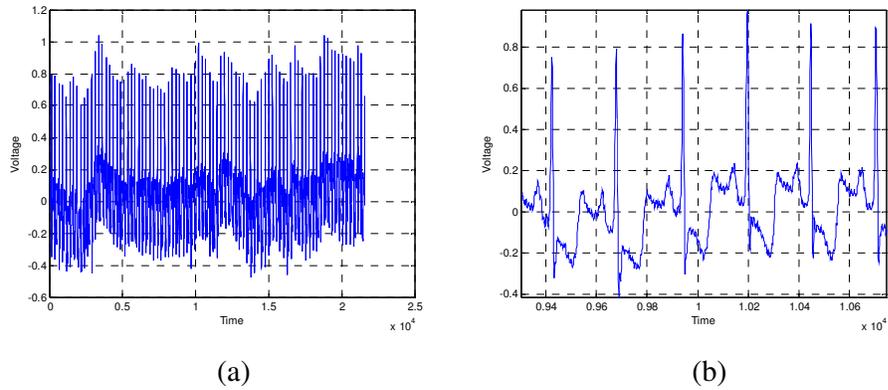
Figure 6.4 Ventricular tachycardia.

Figure 6.4 shows the time series representation of ventricular tachycardia.

The zoomed in section of Figure 6.4 (b) shows the occurrence of abnormal QRS complexes.

(a)   (b)
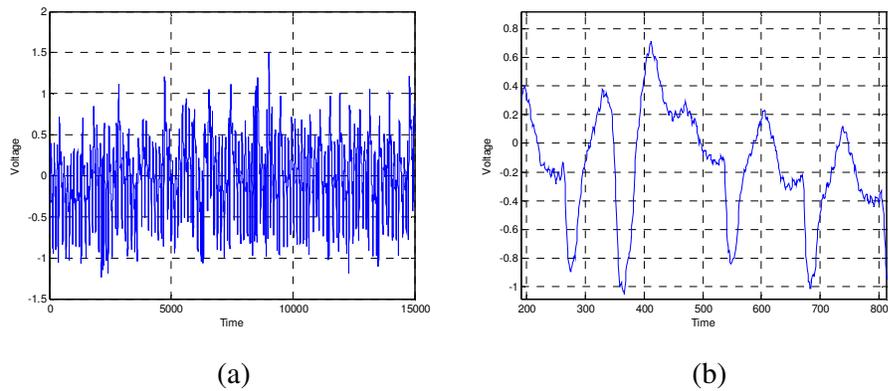
Figure 6.5 Ventricular fibrillation.

Figure 6.5 shows the time series representation of ventricular fibrillation.

The zoomed in section of Figure 6.5 (b) shows the occurrence of a totally random rhythm. Also the P wave is absent and the QRS complex cannot be defined.

6.2 Results of Algorithm 2 (Peak Detection using Threshold Method)

The second algorithm is a peak detection algorithm. It detects the R-peak of the given ECG signal thus giving the position of QRS complex which is of diagnostic importance.

(a)   (b)
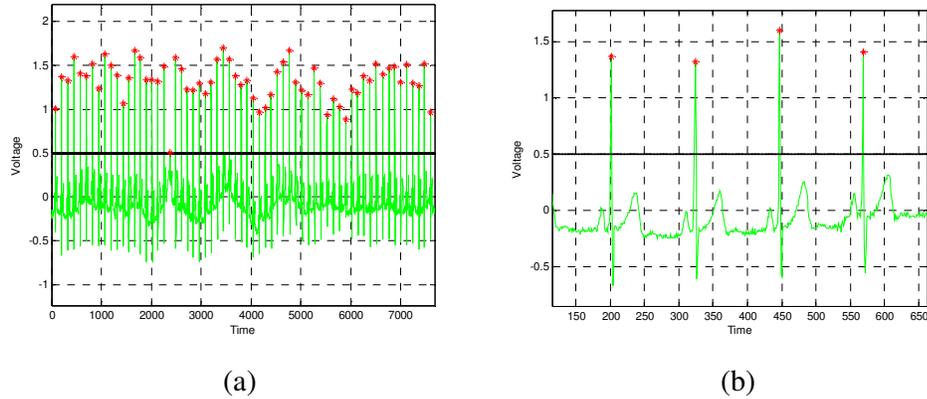
Figure 6.6 Normal sinus rhythm.

Figure 6.6 (a) and (b) show the position of the R-peaks of QRS complexes in a normal ECG signal.

(a)   (b)
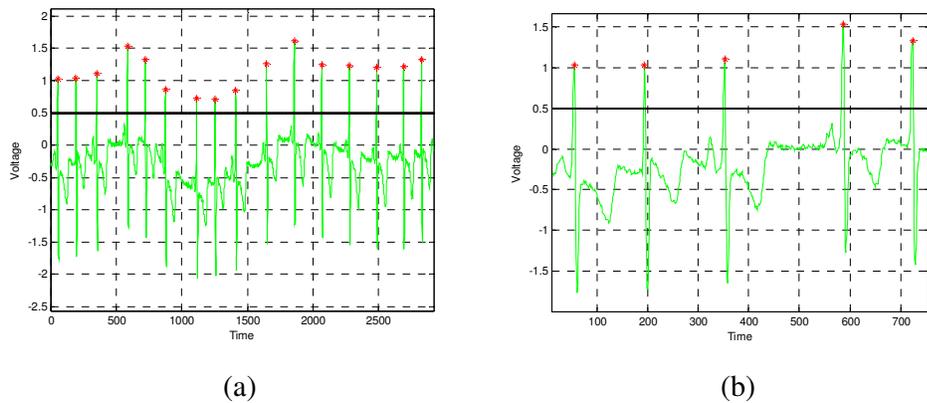
Figure 6.7 Atrial tachycardia.

Figure 6.7 (a) and (b) show the position of the R-peaks of QRS complexes in atrial tachycardia.

(a) (b)
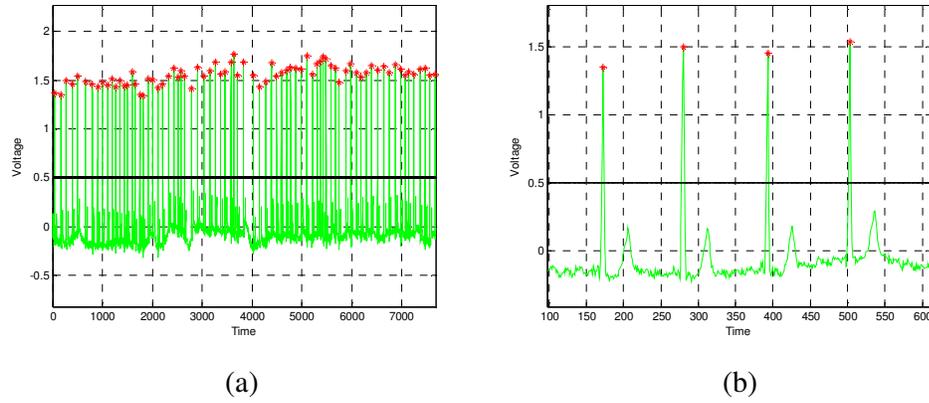
Figure 6.8 Atrial fibrillation

Figure 6.8 (a) and (b) show the position of R-peaks of QRS complexes in atrial fibrillation.

(a) (b)
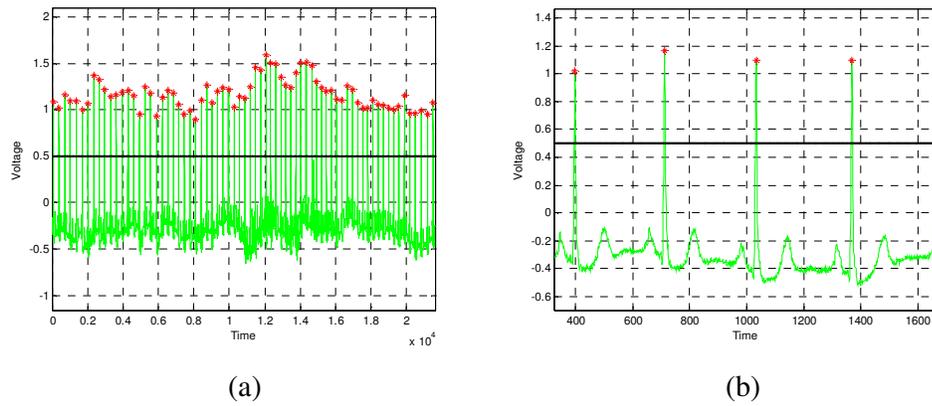
Figure 6.9 Ventricular tachycardia.

Figure 6.9 (a) and (b) show the position of R-peaks of QRS complexes in ventricular tachycardia.
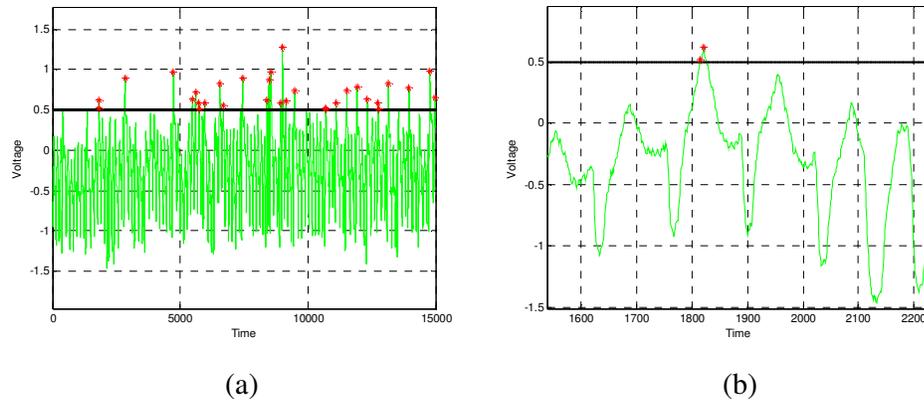
(a)   (b)

Figure 6.10 Ventricular fibrillation.

As seen in Figure 6.10 (a) and (b) the R-peaks cannot be determined properly as the ventricular fibrillation signal is random and chaotic in nature.

6.3 Results of Algorithm 3 (Correlation of Unknown Signal with Known Signals)

The third algorithm calculates the correlation between two signals. This algorithm is used to find the correlation of unknown ECG signals with the existing set of signals. The set of signals includes normal sinus, atrial tachycardia, atrial fibrillation, ventricular tachycardia and ventricular fibrillation. The average percentage of the correlation for every set of signals is calculated. This helps in classifying the unknown signal into a particular category as shown in the Table 6.1 below depending on the maximum percentage of correlation.

Percentage correlation with defined set of signals					
	Normal Sinus (ns)	**Atrial Tachycardia (at)**	**Atrial Fibrillation (atfb)**	**Ventricular Tachycardia (vt)**	**Ventricular Fibrillation (vf)**
Unknown signal 1 (unknw1)	50.62%	77.50%	9.42%	13.06%	15.96%
Unknown signal 2 (unknw2)	33.25%	14.57%	41.32%	3.74%	4.31%
Unknown signal 3 (unknw3)	68.12%	38.41%	13.55%	9.85%	11.35%
Unknown signal 4 (unknw4)	29.31%	19.02%	12.94%	34.35%	27.80%
Unknown signal 5 (unknw5)	17.33%	33.58%	16.85%	30.23%	36.71%

Table 6.1 Percentage correlation matrix.

The above Table 6.1 shows the percentage correlation matrix of unknown ECG signals with the predefined sets of ECG signals.

Table 6.1 shows that each unknown signal can be identified on the basis of the highest value obtained under the five mentioned signal types. For instance, unknown signal 1 gave the highest correlation percentage under atrial tachycardia. At 77.50% it can be said that this unknown signal can be classified as atrial tachycardia and the only one closest to this output is the normal sinus at 50.62%. There is a huge

difference of 20% which shows that atrial tachycardia is the unknown signal. On the other hand, unknown signal 5 gives the highest value for ventricular fibrillation at 36.71% which is very closely followed by atrial tachycardia at 33.58%. Since this signal gives outputs for two categories without a substantial difference in percentage I believe that if there were more reference signals to be tested with the unknown signal, it could have been identified by obtaining a higher output. This could have helped in obtaining a considerable difference from the next highest percentage correlation output

6.4 Results of Algorithm 4 (Correlation of Signals in Data Set)

The algorithm 4 is similar to algorithm 3. This algorithm gave the correlation of the signals in the given set of signals

a) belonging to same set (for example correlation between two atrial tachycardia signals)

b) belonging to different sets (for example correlation between atrial fibrillation and ventricular tachycardia)
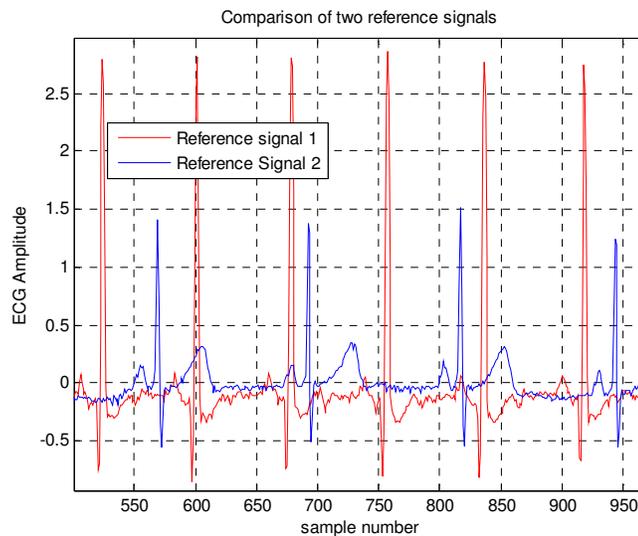
Figure 6.11 Comparison of two normal ECG signals.

Figure 6.11 above shows the comparison of two normal ECG signals. It can be observed that the two signals are morphologically similar and have distinct ECG features. The correlation between the two signals is 0.8615.
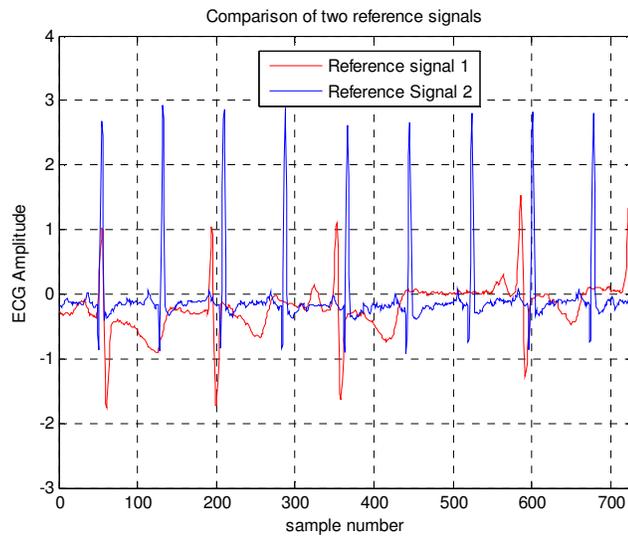
Figure 6.12 Comparison between normal ECG and atrial tachycardia.

Figure 6.12 above shows the comparison between normal ECG and atrial tachycardia signals. It can be observed that the two signals are morphologically different with an abnormal P wave before each QRS complex in case of atrial tachycardia signal (reference signal 1). The calculated correlation between the two signals is 0.2725.
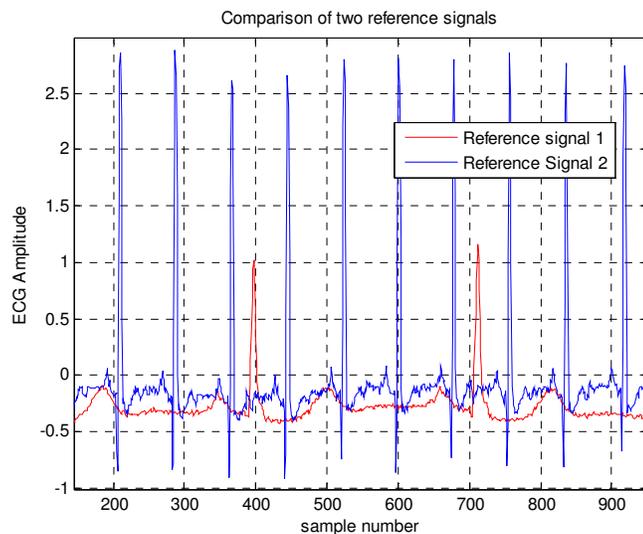
Figure 6.13 Comparison between normal ECG and ventricular tachycardia.

Figure 6.13 shows the comparison between normal ECG and ventricular tachycardia signals. It can be seen that the ventricular tachycardia signal (reference

signal 1) has an abnormal QRS complex and the correlation between the two signals is 0.1072.
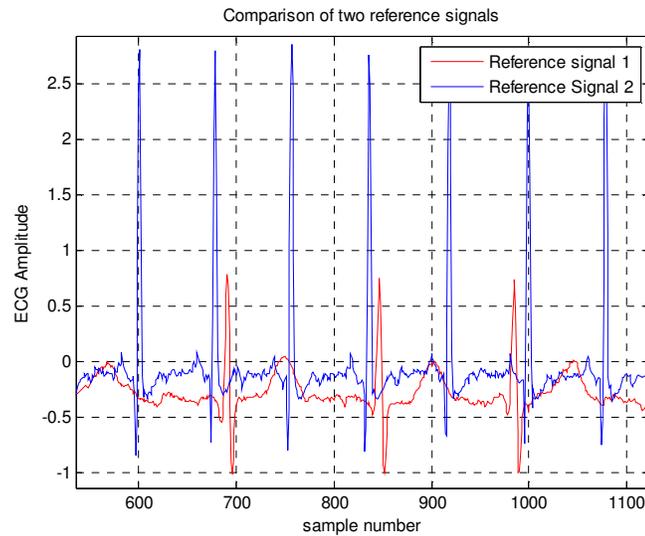
Figure 6.14 Comparison between normal ECG and atrial fibrillation.

Figure 6.14 shows the comparison between normal ECG and atrial fibrillation signals. The reference signal 1 in the above figure shows the absence of P wave in atrial fibrillation signal (reference signal 1). The calculated correlation between the two signals is 0.1873.
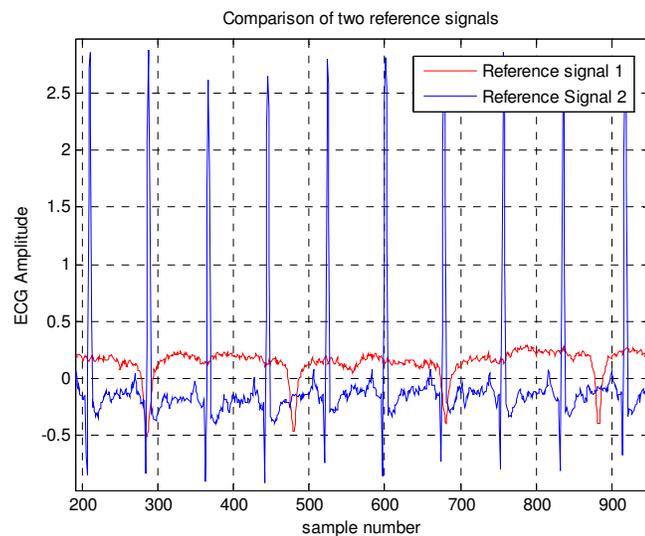
Figure 6.15 Comparison between normal ECG and ventricular fibrillation.

Figure 6.15 shows the comparison between normal ECG and ventricular fibrillation signals. It can be seen that the ventricular fibrillation signal (reference signal 1) is very random with no morphological similarity with the normal ECG signal. The correlation between the two signals is -0.1958.

6.5 Results of Algorithm 5(Identifying Unknown Signals into Various Category)

Algorithm 5 calculates the correlation of the given unknown signal with all reference signals and gives the name of the signal which has maximum correlation with it. The Table 6.2 below shows the results of the algorithm.

	Reference signal with highest correlation
Unknown signal 1 (unknw1)	at2 (atrial tachycardia)
Unknown signal 2 (unknw2)	atfb1 (atrial fibrillation)
Unknown signal 3 (unknw3)	ns1 (normal sinus)
Unknown signal 4 (unknw4)	vt1 (ventricular tachycardia)
Unknown signal 5 (unknw5)	vf1 (ventricular fibrillation)

Table 6.2 Classification of unknown signals.

6.6 Results of Algorithm 6 (Peak Detection and Heart Beat Rate using DWT)

Algorithm 6 is a peak detection algorithm using the discrete wavelet transform. The signal is pre-processed using moving average filter. The smooth signal is then further processed using discrete wavelet transform to detect peaks. Later the algorithm calculates and displays the heart beat rate as shown below.
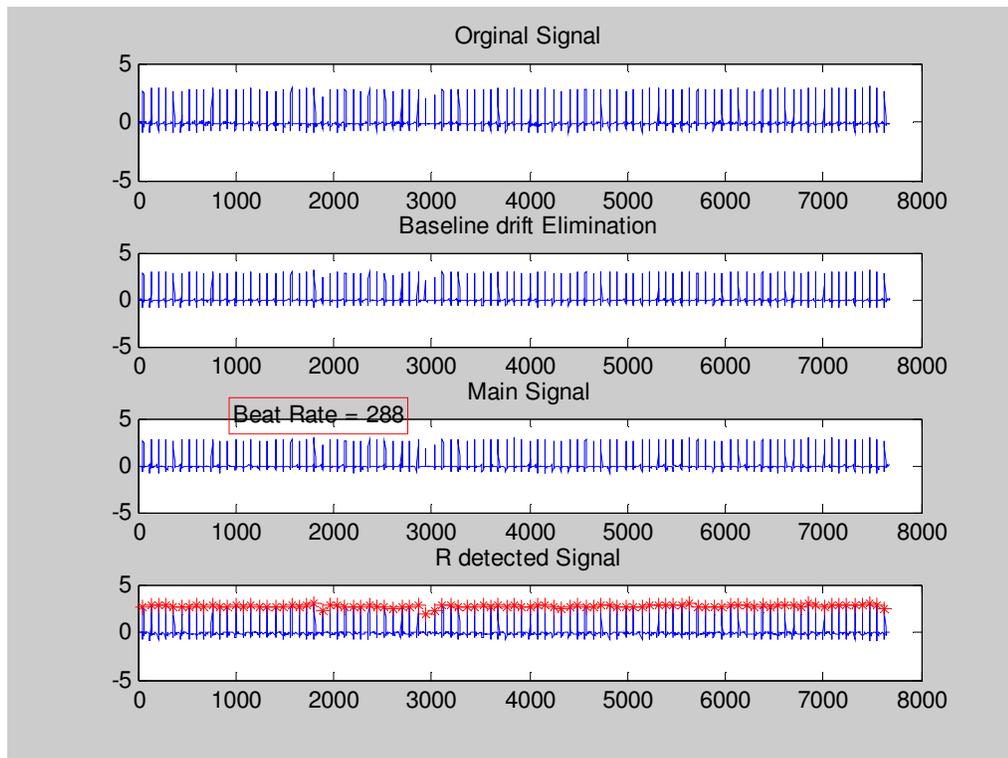
Figure 6.16 Heart beat rate of normal ECG.
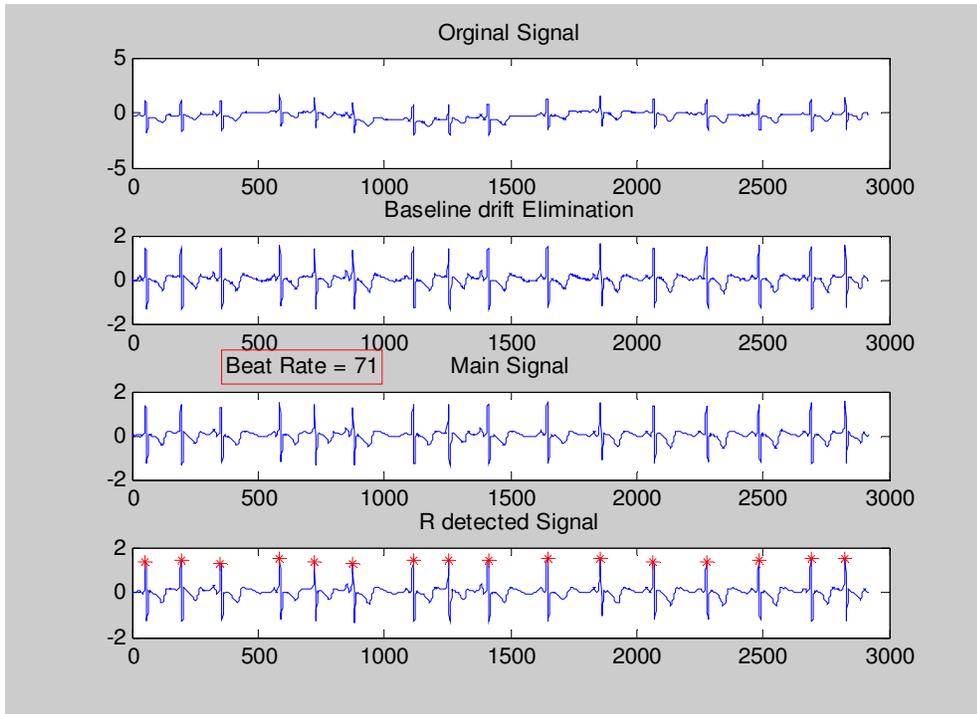
Figure 6.17 Heart beat rate of atrial tachycardia.
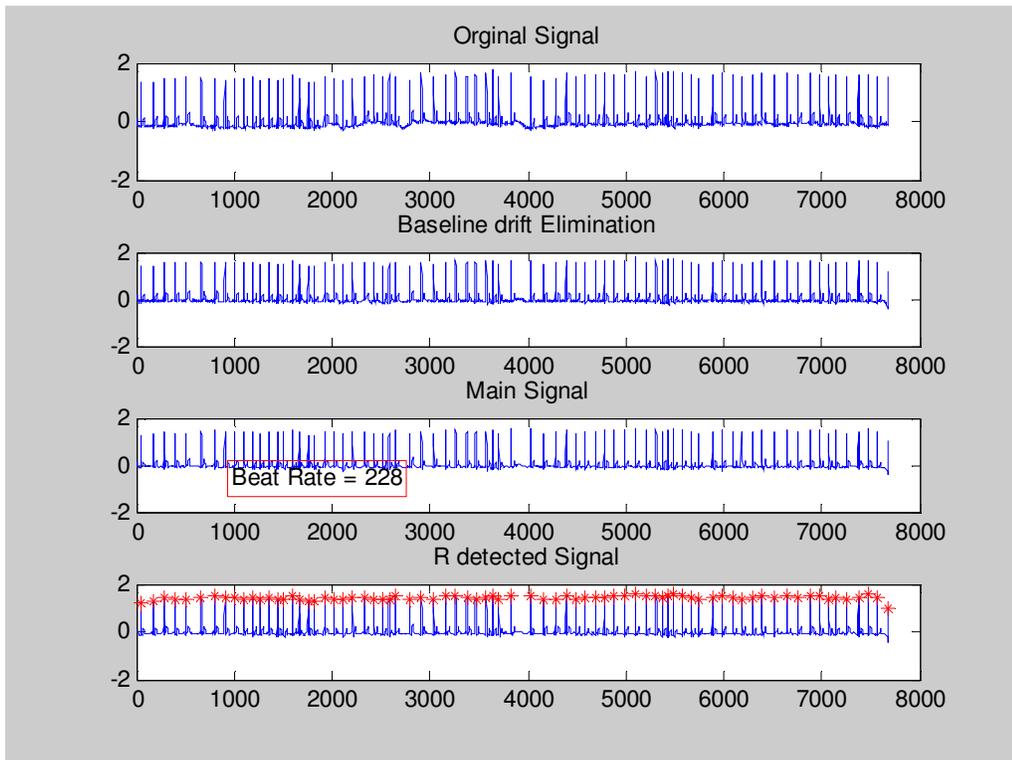
Figure 6.18 Heart beat rate of atrial fibrillation.
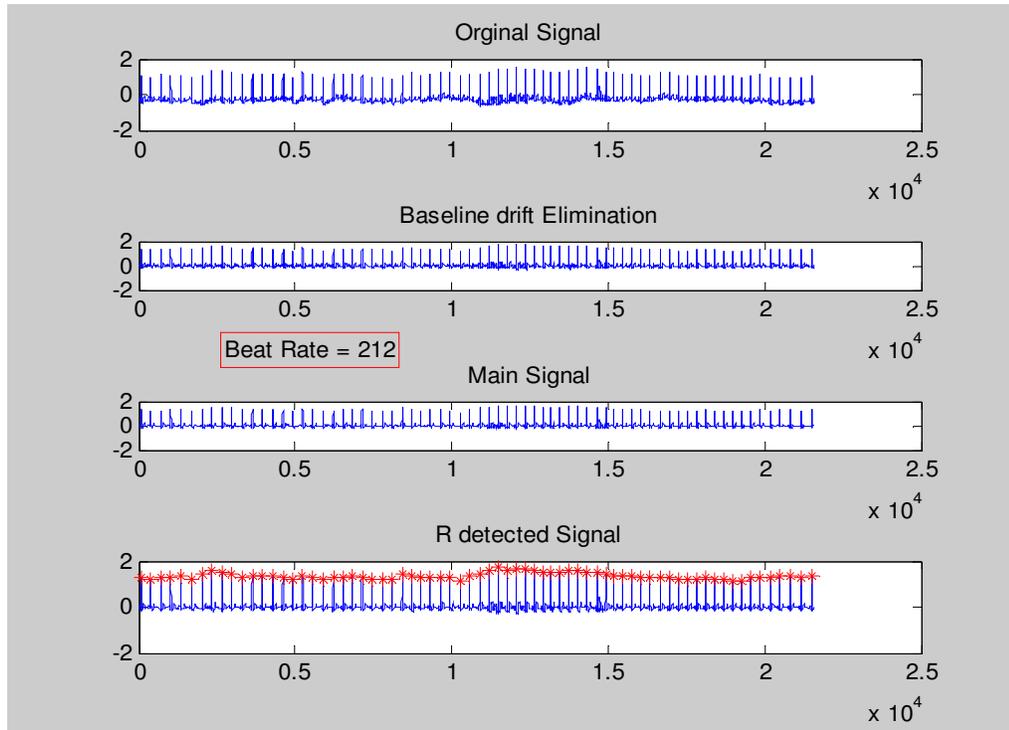
Figure 6.19 Heart beat rate of ventricular tachycardia.
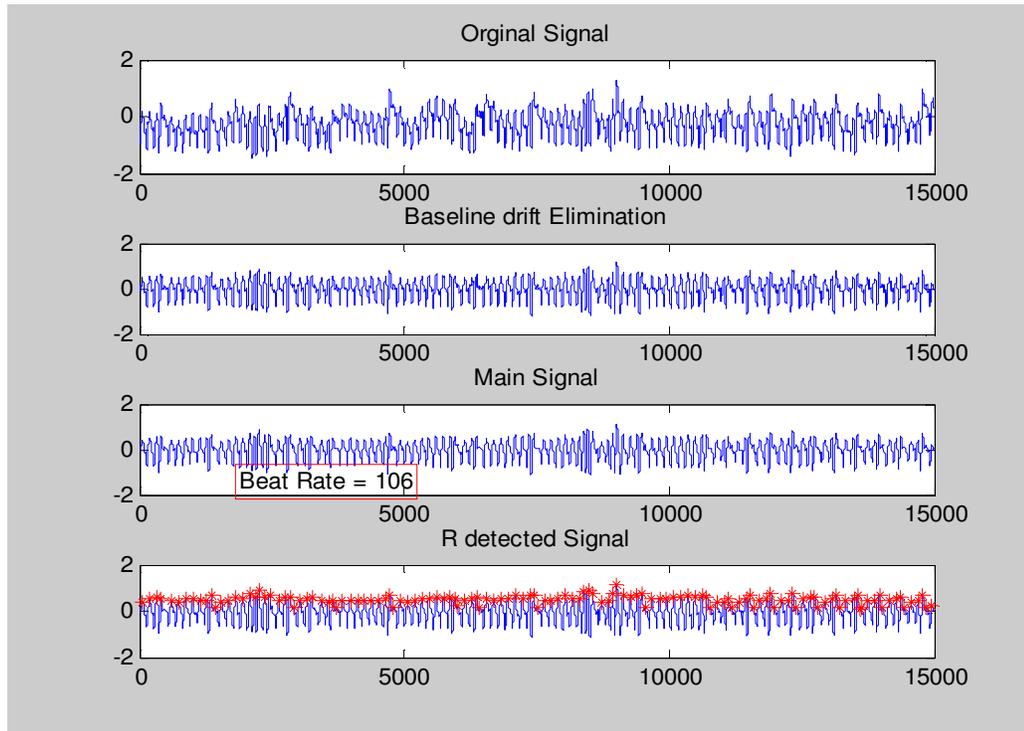
Figure 6.20 Heart beat rate of ventricular fibrillation.

6.7 Results of Algorithm 7 (Kolmogorov Complexity of ECG Signal)

Algorithm 7 calculates Kolmogorov complexity of the signal. Kolmogorov complexity gives us the measure of randomness and chaotic nature of the signal.

Signal	Kolmogorov Complexity
ns1	12.90
ns6	12.89
at1	11.51
at2	11.509
atfb1	12.9069
atfb5	13.86
vt1	14.39
vt4	14.39
vf1	13.8727
vf8	13.8727

Table 6.3 Kolmogorov complexity of signals.

The above Table 6.3 shows the complexity measure of signals belonging to different category. It can be observed that the complexity measure is highest for ventricular tachycardia signals and is lowest for atrial fibrillation signals. This shows that the state of ventricular tachycardia is very random in nature which might lead to life threatening condition of ventricular fibrillation resulting in heart attack.

6.8 Results of Algorithm 8 (State-Space Representation and Lyapunov Exponent)

This algorithm gives us the state space representation of the ECG signals. It also gives us the Fourier coefficient distribution of the signals. Finally the algorithm gives us the average Lyapunov exponent of the given signal.

The figures below show the state space representation of the signals
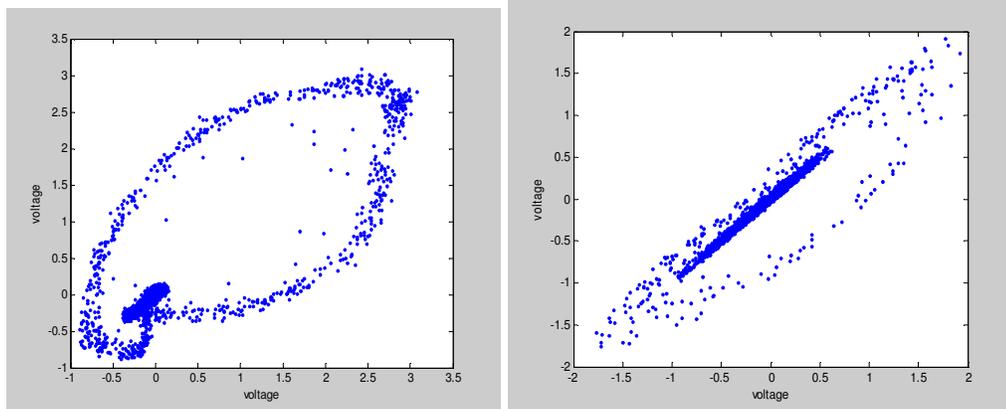
Figure 6.21 Normal sinus rhythm vs. atrial tachycardia.
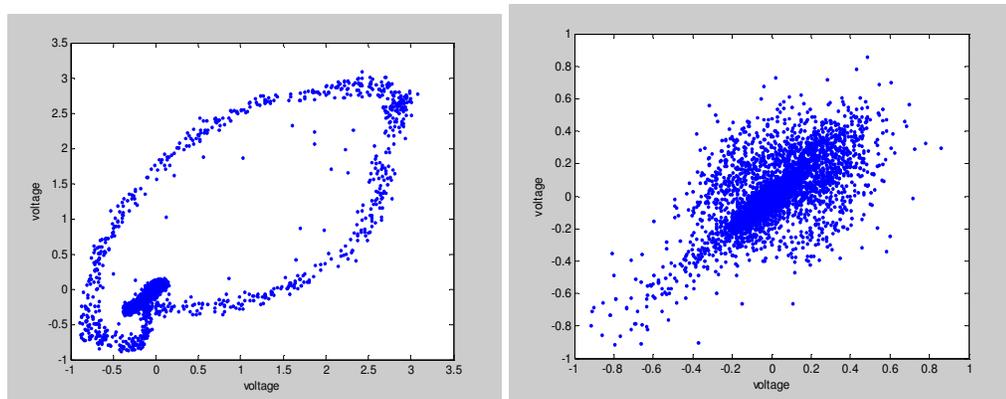
Figure 6.22 normal sinus rhythm vs. atrial fibrillation.
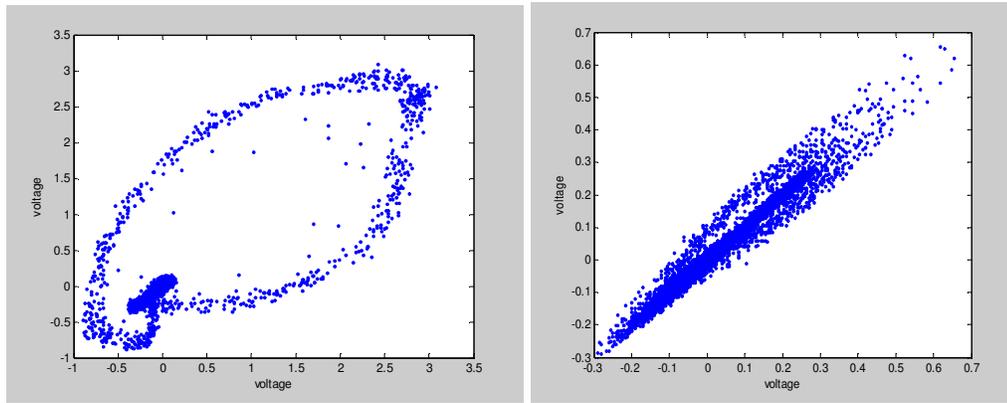
Figure 6.23 Normal sinus rhythm vs. ventricular tachycardia.
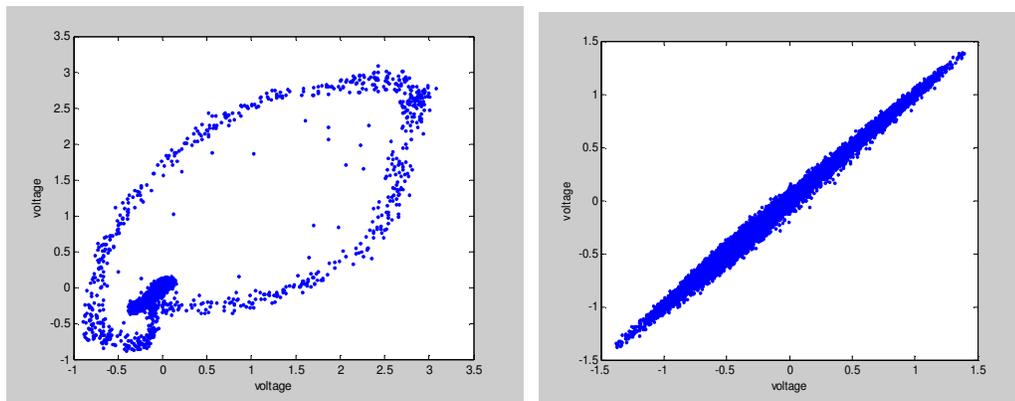
Figure 6.24 Normal sinus rhythm vs. ventricular fibrillation.

Figures below show the distribution of Fourier coefficients of the ECG signals. The graphical representation of Fourier coefficient helps in examining the cyclic structure in the frequency domain.
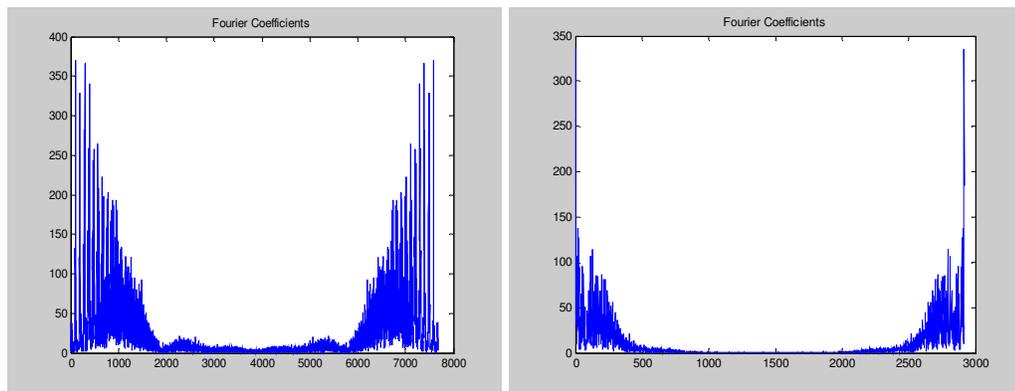
Figure 6.25 Normal sinus rhythm vs. atrial tachycardia.
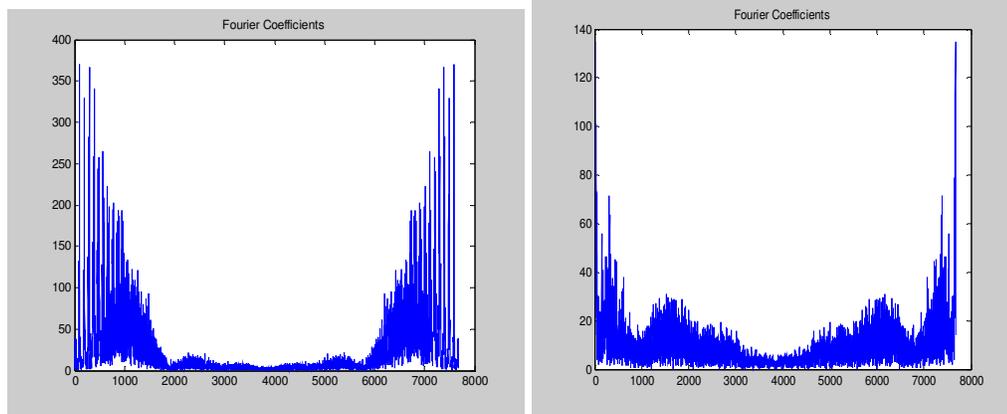
Figure 6.26 Normal sinus rhythm vs. atrial fibrillation.
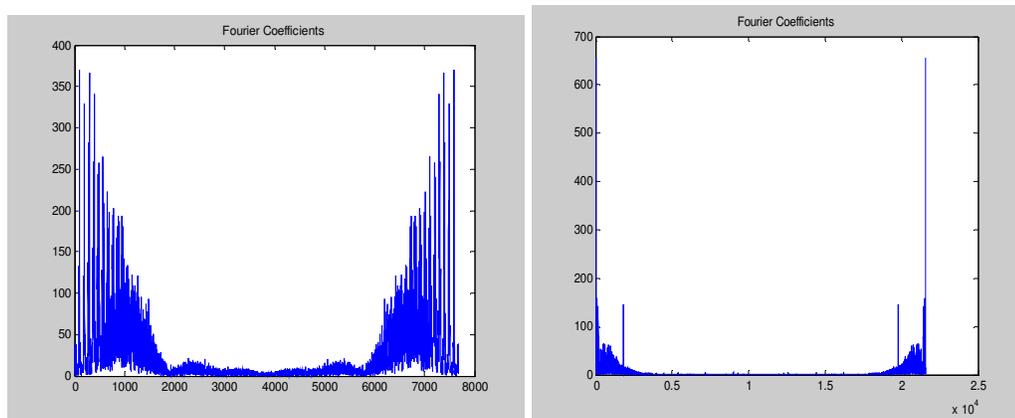
Figure 6.27 Normal sinus rhythm vs. ventricular tachycardia.
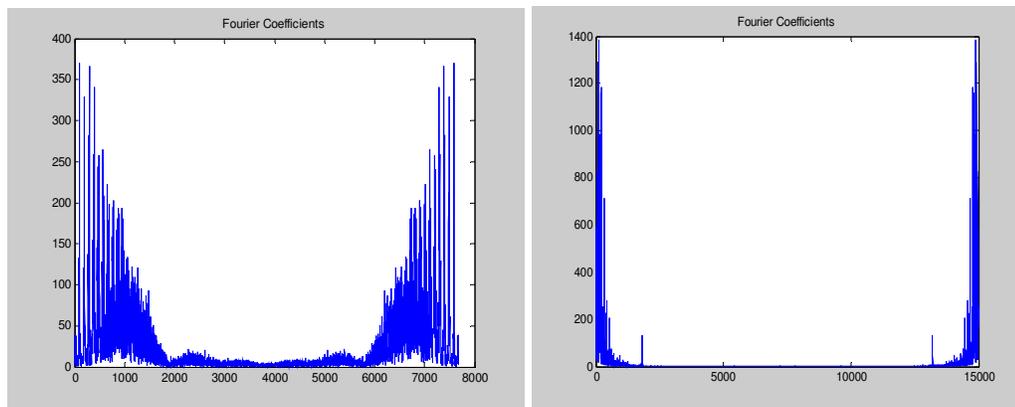
Figure 6.28 Normal sinus rhythm vs. ventricular fibrillation.

The Table 6.4 below gives us the average Lyapunov exponent of ECG signals. The table shows that normal sinus rhythm is less chaotic and the chaotic behaviour increases with increase in pathology.

Signal	Average Lyapunov Exponent
Ns1	0.0095
Ns5	0.0013
Ns8	0
Ns9	0
At1	0.0264
At2	0.0313
atfb3	0.0737
atfb4	0.0100
atfb6	0.0634
Vt1	0.0199
Vt3	0.1235
Vt4	0.0648
Vf2	0.1681
Vf3	0.1799
Vf8	0.1114
vf11	0.1507

Table 6.4 Lyapunov exponent of signals.

6.9 Comparison of Linear and Non-linear Techniques

This section compares the results of Algorithm 2 and Algorithm 6. Algorithm 2 is the peak detection algorithm which uses the linear technique of thresholding, and Algorithm 6 uses discrete wavelet transforms for detecting the peaks. Below are the results of both algorithms with common input signals.

(a)

(b)

(c)

(d)

(e)

(f)

(g)   (h)

(i)   (j)
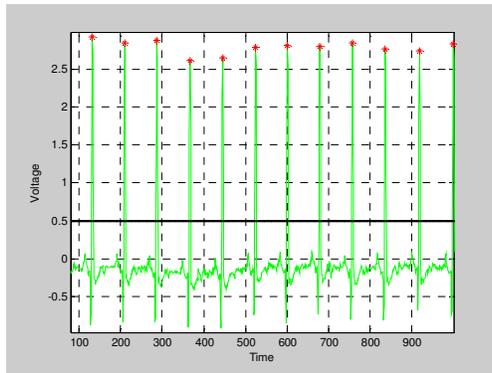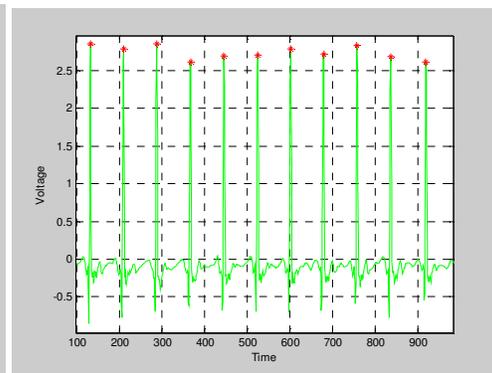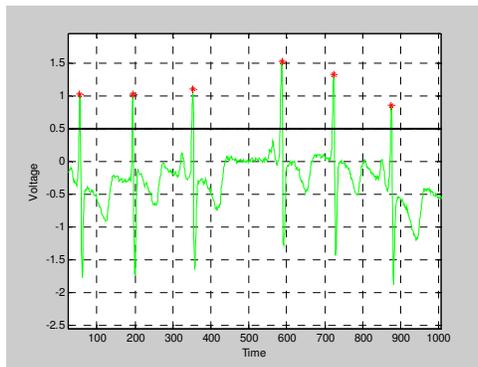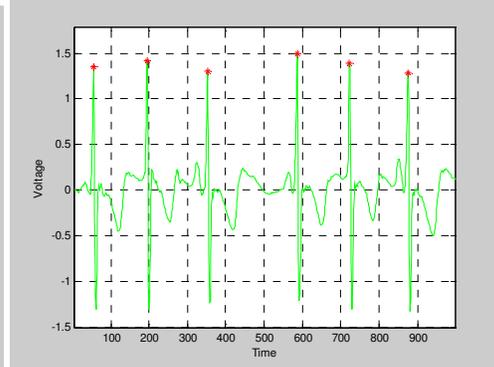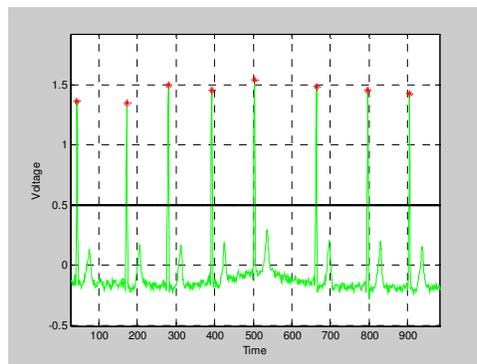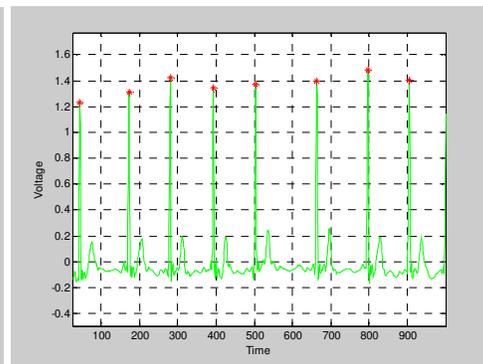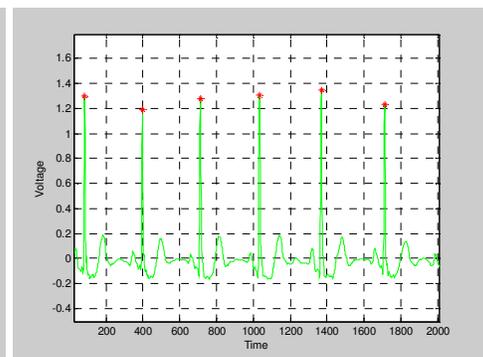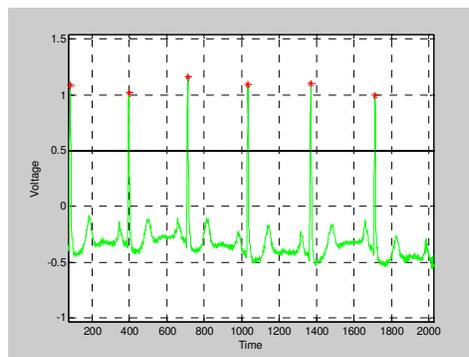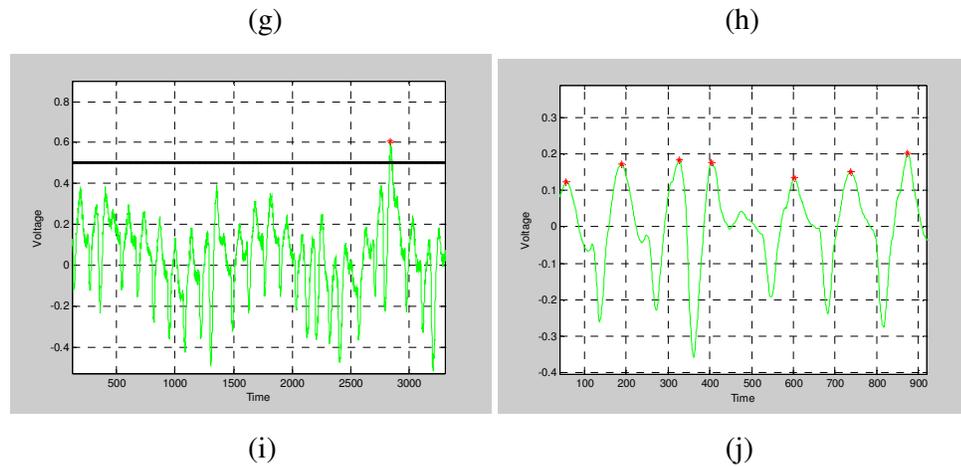
Figure 6.29 Comparison of results for peak detection using thresholding and discrete wavelet transform.

Figure 6.29 above shows the results of two algorithms that were developed to detect the peaks of ECG signal. Figures 6.29 (a), (c), (e), (g), (i) are the results of algorithm 2 which implemented the thresholding technique for peak detection. Figures 6.29 (b), (d), (f), (h), (j) are the results of algorithm 6, which used the discrete wavelet transform for processing and peak detection of the ECG signal. Figures (a) and (b), (c) and (d), (e) and (f), (g) and (h) and (i) and (j) are results of normal ECG, atrial tachycardia, atrial fibrillation, ventricular tachycardia and ventricular fibrillation respectively.

The results show a noticeable difference between the two techniques. Following are some of the advantages of algorithm 6 over algorithm 2.

1. The base line drift is eliminated.
2. The signal is more smooth and filtered.
3. The peaks are very well detected (peaks of ventricular fibrillation are detected).
4. The time for calculation is less.

Chapter 7

Conclusions and Future work

7.1 Conclusions

As discussed in previous chapters, the aim of this project was to understand the nature of complex biosignals (the ECG in this project) and then use the various tools available for non-linear signal processing for the analysis of these signals. This aim was achieved by implementing algorithms using MATLAB$^{®}$ software.

The algorithms developed can be used for (a) the time series representation of the signals, (b) the peak detection, (c) the classification of signals, (d) calculating heart beat rate using DWT, (e) calculating Kolmogorov complexity and (f) calculating Lyapunov exponent. The algorithms used for the time series representation, state space representation and the peak detection gave very excellent results. This can be observed from the results of algorithm 1, algorithm 2 and algorithm 8. The peak detection algorithm could detect peaks of all signals except for the ventricular fibrillation signals. This is because the ventricular fibrillation signal is very random and chaotic in nature as shown in Figure 6.5 and Figure 6.9. However the peak detection was easy and fast using non-linear approach of DWT. This can be observed from Figure 6.20. Figures 6.16 to 6.20 show the effectiveness of non-linear techniques over linear techniques.

The algorithms 3, 4 and 5 use the logic of correlation for classifying the signals. Table 6.2 shows the classification of unknown signals into different ECG signal types. The unknown signal 1 and unknown signal 3 are very well classified as atrial tachycardia and normal sinus respectively. The classification of the other unknown signals is not very good. This is because of limited number of samples in reference set. The accuracy and classification can be improved by increasing the number of sample ECG signals in the reference set there by improving the knowledge database. Even with limited samples in reference set, the algorithm3 and algorithm 5 had same results there by confirming that the classification of unknown signals was correct.

Algorithm 4 gave us the correlation of the signals present in the reference set. It can be seen from the results of Algorithm 4, that it can efficiently differentiate between the signals belonging to different types.

The drawback of these algorithms is that they take more time for execution ranging from 30 seconds to 300 seconds. Hence these algorithms are no good for real time applications. Even with the time lag it can be concluded that the overall performance of the algorithms was good and with more time and efficient programming the performance of these algorithms can be further improved.

Algorithms 6, 7 and 8 used the non-linear approach of analysis towards the ECG signals. The technique of discrete wavelet transform was used to detect peaks in Algorithm 6. This approach of peak detection was very efficient. The algorithm gave very accurate and fast results. Also the execution time was very less as compared to the linear approach for peak detection.

Algorithm 7 calculated the Kolmogorov complexity of the ECG signals. Kolmogorov complexity is the measure of randomness. Table 6.3 summarizes the results of this algorithm. The results show that atrial tachycardia signals have lowest complexity (11.51) whereas the ventricular tachycardia signals have highest complexity (14.39) followed very closely by ventricular fibrillation (13.87), atrial fibrillation (12.91) and normal sinus (12.90). This shows that the behaviour of the heart is more random when in ventricular tachycardia state.

Algorithm 8 gave the state space representation and fourier coefficient distribution of the ECG signal. It also gave the average Lyapunov exponent values which are recorded in table 6.4. The results show that the Lyapunov exponent of normal sinus rhythm is the smallest and is largest for ventricular fibrillation. The difference between the Lyapunov exponents of normal ECG and ventricular fibrillation ECG is distinguishable. Thus it can be considered as an excellent tool for differentiating between normal and abnormal ECG's. However classification of different arrhythmias is difficult. According to the results we can conclude that the normal ECG is less chaotic than ventricular fibrillation ECG.

The project includes two algorithms for finding the peaks of the given ECG signal. Algorithm 2 used the linear technique while Algorithm 6 used a non-linear technique for finding peaks. It was observed that the number of programming lines

of code required for implementation of algorithm 2 were more then those required for algorithm 6. The implementation of code was much easier using the DWT. Also the time required for executing the program was less for algorithm 6 as compared to algorithm 2. Thus the overall performance of non-linear approach was much better than linear approach.

Non-linear approach of analysis of biosignals required study of complex theories and more advanced programming to obtain answers and for better understanding of various phenomenons. Though difficult and complex, non-linear dynamic analysis assures promising results for analysis of biosignals.

7.2 Future Work

This project is just the first step towards achieving a bigger and better goal in this vast field of non-linear analysis of ECG signals. The amount of data and information that can be extracted from these signals is very immense. This information that these signals carry is of great importance in understanding the very complex nature of the human body. In order to improve the work done in this project following are some of the points that need to be addressed for future work

- Signal pre-processing techniques such as filtering for removing the noise artifacts should be used before the actual analysis of the signals. Following this procedure would give better results.

- The algorithms should be tested more on real ECG signals and should be modified to make the algorithm more robust.

- Increasing the knowledge data base will ensure more accurate results.

- The algorithms that calculate Kolmogorov complexity and Lyapunov exponent should be further modified by using Jacobian approach to give more accurate results.

- The algorithms should be modified to calculate correlation dimension, which can give us more information about the chaotic and fractal nature of the ECG signals.

- Wavelet transform based interrogation of the ECG signals should be the next step implemented towards the analysis and classification.

- Fuzzy logic theory can be used for analysis of ECG signal and algorithms should be developed using this concept. There by giving a more statistical approach towards the analysis of signals.

- The future works should involve analysis of biosignals using various non-linear techniques as the results obtained are more accurate.

References

[1]   "Texas   Heart   Institute   at   St.   Luke's   Episcopal   Hospital," http://www.texasheartinstitute.org/.

[2]   "Oregon Health and Science University," http://www.ohsuhealth.com.

[3]   "Wikipedia: The Free Encyclopedia," http://en.wikipedia.org.

[4]   "University   of   Virginia   Health   System," http://www.healthsystem.virginia.edu/.

[5]   "Howard Hughes Medical Institute," http://www.hhmi.org.

[6]   "Cleveland Clinic," http://www.clevelandclinic.org/.

[7]   "St.   Luke's-Roosevelt   Hospital   Center   'The   Arrhythmia   Service'," http://www.arrhythmia.org.

[8]   "American Heart Association," http://www.americanheart.org.

[9]   "The British United Provident Association," http://www.bupa.co.uk.

[10]   "PhysioNet," http://www.physionet.org/.

[11]   H. D. I. Abarbanel, R. Brown, J. J. Sidorowich *et al.*, "The analysis of observed chaotic data in physical systems," *Reviews of Modern Physics,* vol. 65, no. 4, pp. 1331-1392, 1993.

[12]   P. S. Addison, J. N. Watson, G. R. Clegg *et al.*, "Evaluating arrhythmias in ECG signals using wavelet transforms," *IEEE Engineering in Medicine and Biology Magazine,* vol. 19, no. 5, pp. 104-109, 2000.

[13]   M. Akay, " Introduction: Wavelet transforms in biomedical engineering," *Annals of Biomedical Engineering,* vol. 23, no. 5, pp. 529-530, 1995.

[14]   M. Akay, *Nonlinear Dynamical Analysis of Biosignals*, New York: IEEE Press, 2000.

[15]   A. Babloyantz, and A. Destexhe, "Low-Dimensional Chaos in an Instance of Epilepsy " *National Academy of Sciences,* vol. 83, pp. 3513-3517, 1986.

[16]   S. Barro, M. Fernandez-Delgado, J. A. Vila-Sobrino *et al.*, "Classifying multichannel ECG patterns with an adaptive neural network," *IEEE Engineering in Medicine and Biology Magazine,* vol. 17, no. 1, pp. 45-55, 1998.

[17]   J. C. Bezdek, and S. K. Pal, *Fuzzy Models for Pattern Recognition*, New York: IEEE Press, 1992.

[18]   P. Bourke. "Random Attractors Found using Lyapunov Exponents," http://local.wasp.uwa.edu.au/~pbourke/fractals/lyapunov/.

[19]   S.-W. Chen, "A wavelet-based heart rate variability analysis for the study of nonsustained ventricular tachycardia," *IEEE Transactions on Biomedical engineering,* vol. 49, no. 7, pp. 736-742, 2002.

[20]   M. J. Donahue. "An Introduction to Mathematical Chaos Theory and Fractal Geometry," http://www.duke.edu/~mjd/chaos/chaos.html.

[21]   D. L. Donoho, and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika,* vol. 81, no. 3, pp. 425-455, 1994.

[22]   P. Flandrin, "Wavelet analysis and synthesis of fractal Brownian motion," *ieee Transactions on Information Theory,* vol. 38, no. 2, pp. 910-917, 1992.

[23]   O. Fojt, and J. Holcik, "Applying nonlinear dynamics to ECG signal processing," *Engineering in Medicine and Biology Magazine, IEEE,* vol. 17, no. 2, pp. 96-101, 1998.

[24]   A. Garfinkel, M. Spano, W. Ditto *et al.*, "Controlling cardiac chaos," *Science,* vol. 257, no. 5074, pp. 1230-1235, 1992.

[25]   I. Gath, and A. B. Geva, "Unsupervised Optimal Fuzzy Clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 11, no. 7, pp. 773-780, 1989.

[26]   J. F. Gibson, J. D. Farmer, M. Casdagli *et al.*, "An Analytical Approach to Practical State Space Reconstruction," *Physica D,* vol. 57, no. 1-2, pp. 1-30, 1992.

[27]   L. Glass, A. L. Goldberger, M. Courtemanche *et al.*, "Nonlinear Dynamics, Chaos and Complex Cardiac Arrhythmias," *Mathematical and Physical Sciences,* vol. 413, no. 1844, pp. 9-26, 1987.

[28]   A. L. Goldberger, L. A. N. Amaral, L. Glass *et al.*, "Nonlinear Dynamical, Fractals, and Chaos Theory: Implications for Neuroautonomic Heart Rate Control in Health and Disease," *Circulation,* vol. 101, pp. 215-220, 1999.

[29]   P. Grassberger, and I. Procaccia, "Measuring the Strangeness of Strange Attractors," *Physica D,* vol. 9, no. 1-2, pp. 189-208, 1983.

[30]   Y. Hata, O. Ishikawa, S. Kobashi *et al.*, "Automated Medical Diagnosis System (AMDS) with Normal Degree based on Fuzzy Logic," *IASTED Biomedical Engineering Proceeding,* vol. 417, no. 42, 2004.

[31]   Y. H. Hu, W. J. Tompkins, J. L. Urrusti *et al.*, "Applications of artificial neural networks for ECG signal detection and classification," *Journal of Electrocardiology,* vol. 26, pp. 66-73, 1993.

[32]   T. Jaditz, "Time Series Prediction: Forecasting the Future and Understanding the Past," *Journal of Economic Behavior & Organization,* vol. 26, no. 2, pp. 302-305, 1995.

[33]   V. Johneff, "Complex valued wavelet analysis for QRS detection in ECG signals," *Bioinformatics, Images, and Wavelets*, pp. 134-136, 2004.

[34]   T. Kalayci, and O. Ozdamar, "Wavelet preprocessing for automated neural network detection of EEG spikes," *IEEE Engineering in Medicine and Biology Magazine,* vol. 14, no. 2, pp. 160-166, 1995.

[35]   R. E. Klabunde. "Cardiovascular Physiology Concepts," http://www.cvphysiology.com/.

[36]   A. Kulkarni, and S. Bush, "Detecting Distributed Denial-of-Service Attacks using Kolmogorov Complexity Metrics," *Journal of Network and Systems Management,* vol. 14, no. 1, pp. 69-80, 2006.

[37]   S. Mallat, "Zero-crossing of a wavelet transform," *IEEE Transactions on Information Theory,* vol. 37, no. 4, pp. 1019-1033, 1991.

[38]   B. B. Mandelbrot, *The Fractal Geometry of Nature*, New York: W.H. Freeman and Company, 1982.

[39]   J. P. Marques-de-Sa, A. P. Goncalves, F. O. Ferreira *et al.*, "Comparison of artificial neural networks based ECG classifiers using different feature types," *Computers in Cardiology*, pp. 545-547, 1994.

[40]   F. H. Martini, *Fundamentals of Anatomy & Physiology*, 7 ed., p.^pp. .698,699,700,701, Canada: Pearson Education International, 2005.

[41]   J. Nadal, and M. d. C. Bossan, "Classification of cardiac arrhythmias based on principal component analysis and feed-forward neural networks," *Computers in Cardiology*, pp. 341-344, 1993.

[42]   M. I. Owis, A. H. Abou-Zied, A.-B. M. Youssef *et al.*, "Study of feature based on nonlinear dynamical modeling in ECG arrhythmia detection and classification," *IEEE Transactions on Biomedical engineering,* vol. 49, no. 7, pp. 733-736, 2002.

[43]   C. S. Peskin, and D. M. McQueen, "Mechanical equilibrium determines the fractal fiber architecture of aortic heart valve leaflets," *American Journal of Physiology (Heart and Circulatory Physiology),* vol. 266, no. 1, pp. 319-328, 1994.

[44]   V. Pichot, J.-M. Gaspoz, S. Molliex *et al.*, "Wavelet transform to quantify heart rate variability and to assess its instantaneous changes," *Journal of Electrocardiology,* vol. 86, no. 3, pp. 1081-1091, 1999.

[45]   W. A. Sandham, D. J. Hamilton, and D. C. Thomson, " ANN compression of morphologically similar ECG complexes," *Medical and Biological Engineering and Computing,* vol. 33, no. 6, pp. 841-843, 1995.

[46]   I. Stewart, *Does God Play Dice? The Mathematics of Chaos*, Middlesex: Penguin Books Ltd, 1989.

[47]   Y. Sun, K. L. Chan, and S. M. Krishnan, "Life-threatening ventricular arrhythmia recognition by nonlinear descriptor," *BioMedical Engineering OnLine,* vol. 4, no. 6, 2005.

[48]   F. Takens, "Detecting strange attractors in turbulence," *Dynamical Systems and Turbulence*, pp. 366-381, Berlin: Springer, 1981.

[49]   R. Watrous, and G. Towell, "A patient-adaptive neural network ECG patient monitoring algorithm," *Computers in Cardiology*, pp. 229-232, 1995.

[50]   Q. Xue, Y. H. Hu, and W. J. Tompkins, "Neural-network-based adaptive matched filtering for QRS detection," *IEEE Transactions on Biomedical engineering,* vol. 39, no. 4, pp. 317-329, 1992.

[51]   L. A. Zadeh, "Fuzzy Sets " *Information and Control,* vol. 8, pp. 338-353, 1965.

[52]   "Virtual High School Global Consortium," http://www.govhs.org/vhsweb/Gallery.nsf/Files/Fractals+&+Chaos+In+Biolo gy+/$file/fractals_and_chaos_in_biology.ppt.

Bibliography

1. M. Akay, *Nonlinear Dynamical Analysis of Biosignals*, New York: IEEE Press, 2000.

2. J. C. Bezdek, and S. K. Pal, *Fuzzy Models for Pattern Recognition*, New York: IEEE Press, 1992.

3. R. E. Klabunde, *Cardiovascular Physiology Concepts*: Lippincott Williams & Wilkins, 2004.

4. B. B. Mandelbrot, *The Fractal Geometry of Nature*, New York: W.H. Freeman and Company, 1982.

5. F. H. Martini, *Fundamentals of Anatomy & Physiology*, 7 ed., Canada: Pearson Education International, 2005.

6. I. Stewart, *Does God Play Dice? The Mathematics of Chaos*, Middlesex: Penguin Books Ltd, 1989.

Appendix I: List of Figures

Appendix II: List of Tables

Appendix III: Contribution to the Project

The primary goal of this project was to analyse biosignals, to identify pathological condition of the patient and monitor the changes in the condition over a course of treatment or with the procedure. Also, it was important to define and identify the use of such a study with relevance to the current needs.

The very first task was to review various biosignals and the non-linear processing techniques associated with their analysis. After reviewing a lot of published papers ECG signal was selected for analysis. The ECG signal was thoroughly studied to understand the physiological, technical and dynamical aspects of the signal.

The next task involved reviewing the theories involved in non-linear dynamic analysis of signals. While understanding these theories certain parameters like correlation dimension, Lyapunov exponent and Kolmogorov entropy were studied for their application in non-linear analysis of biosignals. This task followed with the completion of literature review.

The next impediment in the project was learning of MATLAB software. Getting familiarised with the syntax and commands of MATLAB was an excellent learning experience. The HELP command in MATLAB and the mathworks website were the two very helpful sources for understanding the syntax and commands used in MATLAB programming. This stage involved the task of developing logical and error-free algorithms and associated codes for performing specific calculations. A gradual approach was used towards the development of algorithms. The algorithms were then tested with the ECG signals acquired from web source. The algorithms were tested for robustness by using different kinds of ECG signals as input.

Finally, eight algorithms were developed and implemented for

1. Time series representation of ECG signals
2. Peak detection of ECG signals using threshold technique
3. Finding percentage correlation of unknown ECG signal with known ECG signal sets.
4. Finding correlation between any two ECG signals
5. Classifying unknown ECG signals.
6. Calculating heart beat rate and peak detection using DWT.

7. Calculating Kolmogorov complexity of a signal

8. Calculating Lyapunov exponent of the signal

The results obtained are fairly accurate as discussed in chapter 6.

Appendix IV: Contents of Compact Disc

The accompanying compact disc (CD) contains the following:

1. A folder '**Algorithm 1**' consisting of

    a. a code for the time series representation of the signals

    b. word file (Instructions.doc) which gives instructions for executing the code

2. A folder '**Algorithm 2**' consisting of

    a. code for peak detection of the signals

    b. word file (Instructions.doc) which gives instructions for executing the code

3. A folder '**Algorithm 3**' consisting of

    a. code for finding the percentage correlation between two signals

    b. word file (Instructions.doc) which gives instructions for executing the code

4. A folder '**Algorithm 4**' consisting of

    a. code for finding correlation of two signals within the reference set

    b. word file (Instructions.doc) which gives instructions for executing the code

5. A folder '**Algorithm 5**' consisting of

    a. code for classifying the unknown signals

    b. word file (Instructions.doc) which gives instructions for executing the code

6. A folder '**Algorithm 6**' consisting of

    a. code for peak detection and calculating heart beat rate using DWT

    b. word files (Instructions.doc) which gives instructions for executing the   code

7. A folder '**Algorithm 7**' consisting of

    a. code for calculating Kolmogorov complexity

    b. word file (Instructions.doc) which gives instructions for executing the code

8. A folder '**Algorithm 8**' consisting of

    a. code for calculating Lyapunov exponent

    b. word file (Instructions.doc) which gives instructions for executing the code

9. A folder '**cardiacsignals**' consisting of ECG data files in an excel sheet format. The data files have following notations

    - atfb: atrial fibrillation

    - at: atrial tachchycardia

    - ns: normal sinus

    - s: sinus

    - vt: ventricular tachychycardia

    - vf: ventricular fibrillation

10. A folder '**unknownsignals**' consisting of unknown ECG data files in an excel sheet format.

11. A folder 'Thesis' consisting of

    a. abstract_patil.pdf (initial abstract submitted)

    b. MSc_research_project_gantt_chart.mpp (gantt chart)

    c. Patil.ppt (initial presentation on project development)

    d. Final abstract_anup patil.pdf (final abstract submitted)

    e. patil_thesis.pdf (final thesis)