PART I - INTRODUCTION AND SYSTEM DEVELOPMENT

1. INTRODUCTION

1.1. GENES, CHIPS AND INFORMATION

One hundred years ago, to someone walking through the gas-lit streets of Barcelona, the simple thought of sending real-time video streams recorded with a handheld device to a friend in Honolulu over a global wireless communication network would have seemed impossible, unthinkable or, borrowing from the always inspired words of *Arthur C. Clarke*, "utterly undistinguishable from magic". Yet, in the biological blink of four generations, the revolution of information technologies and computer sciences in the second half of the twentieth century has reshaped the world, and our perception of it, beyond recognition.

Today, connecting to a remote server in Alaska while driving through the streets of London can be as commonplace as walking through the virtually recreated streets of 1900's Barcelona. And not only has the prowess of information technologies abated, in a roundabout way, our sense of wonder, but it has also led to the foundation of a society that is completely dependent on these technologies and in which, as the late *Carl Sagan* warned, only a few are aware of their inner workings and underlying concepts.

Nonetheless, there is still much room for bewilderment at the beginning of this new millennium, although it may come from a, seemingly, quite different field. Genetics is perhaps the only scientific discipline that has been able to keep pace with Moore's Law-driven computer sciences in the last fifty years. The advances brought forth by genetics and genetic engineering in the last few years (GM crops, forensic DNA typing or massive screening with DNA-arrays) are only but the *hors d'oeuvre* of what we are bound to see in the next few decades. Children *à la carte*, cut-and-paste replacement organs derived from stem cells and even resurrected dinosaurs are concepts that already spur our imagination. What lies beyond, for the time being, still abides in the fascinating depths of science fiction writers' top hats.

1.1.1. GENETICS AND INFORMATION TECHNOLOGIES

The parallel rise of genetics and computer sciences in the second half of the twentieth century may seem coincidental, but, as we shall see, it is not. Brewing on the revolutionary ideas of *Charles Darwin* [Darwin1859] and the theoretical breakthroughs of *Gregor Mendel* [Mendel1866], passionate debates about the nature of life still abound today. But, although there are clashing arguments on the exact nature and origin of the processes of life, the fact that life is an information process seems now to be established beyond reasonable doubt. This is not a mere analogy, but the only perspective that can cope with the far-reaching implications of Darwinian theory [Dennett1996]. In fact, in the light of physics, life and its off-products can be simply understood as anti-entropic phenomena ([Schrödinger1967], [Davies1998]). It is not surprising, then, to find that we can take benefit of the intermingling of two very different approaches on information processing: human-made Boole's algebra and nature's own genetic code. This merging is already taking place and, clearly, the revolution of genetics in the last decades could not, and should not, be explained without its parallel in human-based information technologies.

Since the late 1970's, computer sciences have assumed Darwinian algorithmics in the form of AI genetic algorithms [Winston1984], while parallel and redundant computing have also drawn upon different aspects of life information technologies ([Avienzis1997], [Bonabeau2000]). More recently, the boundary between artificial and life information systems has shrunk and muddled up, as in the case of DNA-computing approaches [Adleman1994] or live neural network guided robots [Fleming2000]. Likewise, genetics and genetic engineering have extensively borrowed on traditional Boolean algorithms and computers to churn out some sense from the vast amounts of data they manage. This has lead to the birth of bioinformatics and subsidiary fields, like functional genomics and proteomics ([Aldhous1993], [Casari1995]). Moreover, in the last few years, genetic engineering has also started to take profit from the substrate technologies of conventional microelectronics, with the creation of micro-devices that speed up and/or automate genetic engineering analytical techniques [Stipp1997] and bringing about the emerging field of DNA-chips in which this doctoral work lies.

1.1.2. AMPLIFICATION

As anyone that has ever held a conversation over a bad telephone line knows, or as any well-learned computer scientist should know for true from the pioneering work of *Claude Shannon* [Shannon1948], information systems have their toughest enemy in the entropic effect of noise. Many techniques, such as filtering, multi-channel transmission or error correcting codes, can be used to partially circumvent or minimize the problems originated by noise, but they would all be nearly useless were it not for a method to boost the original noisy signal in the first place, a method that has played an essential role in the advance of both computer and genetic sciences by helping to raise signal/noise ratios up to workable levels. To electronic and genetic engineers alike, this technique is known as amplification and, in this broad sense, the scope of this doctoral work comprises the blending of biochemical, computational and technological techniques to produce micro-devices for fast, automated and efficient PCR amplification of DNA samples.

1.2. AN INTRODUCTION TO GENETICS

As its name suggests, this section offers a brief insight into the world of genetics and molecular biology, focusing mainly on the properties and structure of the DNA molecule. The reader already acquainted with these matters is welcome to skip all the oversimplifications and basic assumptions committed hereinafter and proceed directly to the following chapter.

1.2.1. BASIC CONCEPTS

The molecule of life

Many schools of thought differ in their opinion over which is the finest definition of life, a discussion that can be reduced to which is the simplest living form. From selfish transposons to bacteria, the room is open for a definition that covers the fundamental aspects of what we perceive as alive. But, no matter what definition life takes, the fact is that DNA and RNA are the universal information storage systems of all, to date, known living organisms and, certainly, life as we know it would be utterly impossible in their absence. Ribonucleic acids encode information using a four-letter alphabet that is, in turn, redundantly translated into the (conventionally) twenty-letter alphabet of amino acids, the constituents of proteins. The sequence of DNA/RNA that encodes a full protein, together with its previous promoter and regulation regions, is called a gene; and the conglomerate of genes that constitute the full recipe for a living being is called the genome. Proteins are used by living organisms in an enormous number of ways, but mainly as structural components and as enzymes that catalyze complex reactions. One of such reactions, involving a myriad of proteins, protein complexes and specialized RNA sequences, is the miracle of life: DNA/RNA replication.

The DNA molecule

DNA structure

At the beginning of the twentieth century, nucleic acids (defined at the time as those found in an eukaryote cell nucleus) were known to be composed of a sugar, a phosphate and a base [Miescher1871]. It took almost two decades for *Phoebus Levene* to determine the chemical composition of the sugars, ribose [Levene1909] and deoxyribose [Levene1929], and to elucidate the structure of the constituents of nucleic acids: nucleotides.

Nucleotides

Nucleotides are composed of a pentose (ribose or deoxyribose) sugar, up to three phosphate groups and a base (see Figure 1), which may be a purine (adenine and cytosine) or a pyrimidine (cytosine, thymine and uracil).

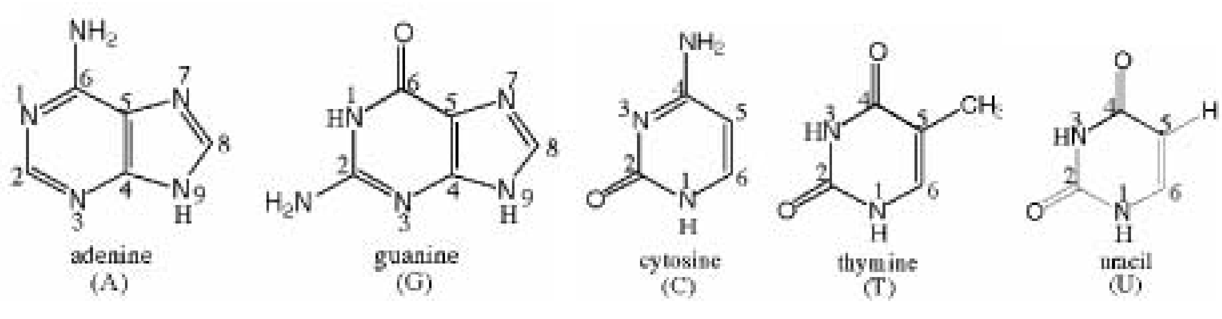

Figure 1 - Nucleic bases.

The pentose sugar (see Figure 2) is a fundamental element in nucleotide formation, since it acts as the backbone of the nucleotide and, hence, of the DNA molecule.

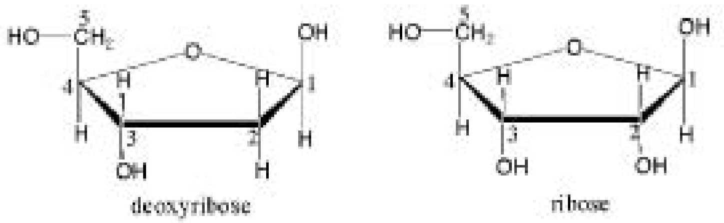

Figure 2 - Deoxyribose and ribose sugar molecules.

In DNA, the purine or pyrimidine base attach themselves to the 1' carbon atom of the deoxyribose, while phosphate groups may attach both to 3' or 5' carbon atoms and in chains of up to three phosphates. The result of this complex grouping is the nucleotide.

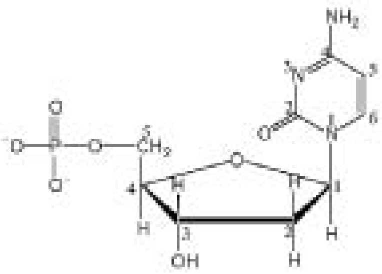

Figure 3 - A nucleotide: deoxy-cytosine mono-phosphate (dCMP).

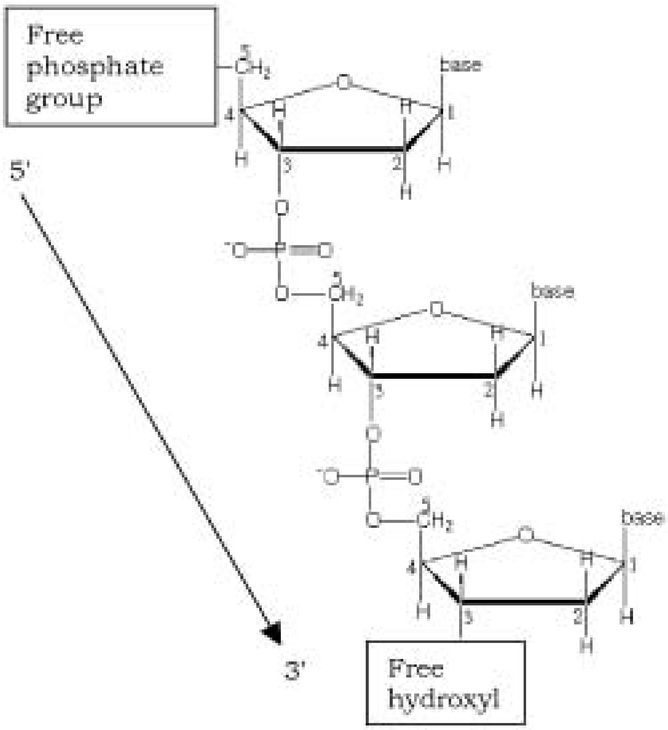

Figure 4 - Single stranded 5'-3' sense DNA sequence.