Chapter 1.1 The plan for this book

The plan for this book is to compare the information and the ways it is coded, stored, and transmitted on the Internet and by living organisms. The way information is handled on the Internet is well defined and documented, while we have a very incomplete picture of information in biology. Thus the Internet will be discussed before we explore similar functions in biology. Chapters on history are at the end of each of the sections under the assumption that you have to know something about each system before the history has much interest.

But what do the Internet and life have in common?

The Internet and living organisms certainly look different. The Internet is a tangled net of wires, optical fibers, terrestrial and satellite microwave links between millions of computers around the world, allowing them to exchange information. Living organisms are generally soft objects that move around the land, sea or air, without obvious connection to each other. They occasionally create progeny, small copies that grow in size to become very similar to the parents. The also die after at characteristic age. However, strong similarities between the Internet and life appear when we examine the information in the two systems.

The purpose of the Internet is to transmit information

The Internet transfers information between computers[1]. You and I use the Internet for that purpose. If you ask the engineers that created the Internet or are now maintaining, improving and expanding it, they will say that the function of the Internet is to transfer information. If a segment of the Internet fails to transmit information, it is changed until it regains that function, or it is removed. All the wires, chips, plugs, power supplies, lasers, and antennae are just hardware to transmit information. However, the part of the Internet which is invisible to the casual observer, the software, is actually more important than the hardware. It is the collection of rules, formats and procedures that specify how the hardware is actually used to transmit information that defines the Internet.

Formats and protocols

Different kinds of information are carried on the Internet through various physical media using different digital formats that have been optimized for both the media and the nature of the information. Thus, there must be procedures for translating between formats. The various media and associated hardware have characteristic

[1]Some of the devices that are connected to the Internet might not seem to be "computers" because they don't have keyboards and monitors. Such devices are certainly not general purpose computers, e.g. mainframe, server, desktop, or laptop. However, most computers in our world are not general purpose computers. These "imbedded" devices control our cars, elevators, rice cookers, and TV sets. Any device that connects to the Internet has the information processing power that entitles it to be called a computer.

error rates, but the rate is never zero. Thus, there must be procedures for error correction, either by just determining that an error has occurred and retransmitting the data, or by including sufficient redundancy in the original message that a true version can be constructed from the corrupted copy. The resources of the Internet are shared, and there are often cases when demand exceeds capacity. Thus there must be rules and procedures for sharing these resources.

Computers on the Internet are not all on a single wire, nor do the connections form a tree like structure. Rather, the topology of the Internet is a net, where information can be transmitted by many possible paths between any two computers. The actual path taken and the timing of data transmission on this net are controlled by thousands of routers at the nodes where the links connect. These routers have considerable freedom with data packets that are transmitted to them. If a router is given more data than it can handle, it just discards the excess. The sender (the client) learns about this when receipt of the transmission is not confirmed, and just sends the same data packet again. Proper clients decrease the rate of data transmission until the router is no longer overwhelmed, but rogue clients maintain a high transmission rate, assuming that the sacrifice of others will be sufficient. Intentional or unintentional rouge behavior can temporarily immobilize a large segment of the Internet, but up to now, human intervention has corrected the problem fairly quickly. A common method of attacking a host computer is generating and sending a huge number of fake messages to the host.

The purpose of life is to survive; survival requires replication

Since life wasn't created by humans it doesn't have a defined function in the same way as the Internet does. A few living species have been modified by humans to satisfy their needs, e.g. dogs, but most have not been directly modified by humans[2]. Our own function is a deeper philosophical or religious question than I am prepared to explore in this book. To remain on a pragmatic level, I propose that the function of life is to exist, and thus it must survive. But what are essential characteristics, properties and abilities of living organisms which enable them to survive.

The growth of organisms requires replication of information

Living organisms larger than a millimeter are typically made up of a number of cells. Each cell is enclosed by a membrane, and has a certain autonomy. The organism starts as a single cell, and thus growth into the full sized organism requires the replication of cells as well as growth of each cell. Each cell typically contains all the information needed to produce the entire organism, and thus the production of new cells requires the replication of this information.

The repair of organisms requires replication

Like any machine, living organisms are damaged, either by gradual ware and tare, or a sudden traumatic event. The damage may be visible to the eye, or only

[2] Of course humans do influence almost all species by killing them, removing habitat, and adding gases to the atmosphere that increase the surface temperature of the earth.

detectable at the molecular level. A great deal is known about repair at the macroscopic level, the healing of broken bones, torn skin, ripped muscle and tendons, ruptured blood vessels. In most cases repair consists of death and removal of damaged cells followed by replacement with replicas of the surviving cells.

Individual organisms die but are survived by progeny

All organisms eventually die. However, during their life they create replicas of themselves, progeny. Production of progeny requires replication of the genetic information that defines the organism. As long as there are living progeny of a group of similar organisms, the group, or species has survived.

The survival of a species requires change

Over a period of many generations the environment for a species will change. This change might be the weather, availability of water, type of food, the spectrum of predators, or any number of other factors. In order for a species to survive it must be able to change to adapt to the new environment. Change requires two processes; generation of diversity among progeny and selection of maximally fit individuals among those progeny.

Replication generates diversity

Replication of genetic information is not completely accurate. Mistakes at a single position of the genetic material generate mutations. Other aberrant replication events produce deletions or tandem copies of segments.

Recombination occurs between two similar molecules containing genetic information, and produces hybrid molecules in which the genetic information has been shuffled. In species containing a male and female sex, individual progeny are created containing a random selection of copies from homologous molecules containing the genetic information of each sex. The segregation of these copies produces informational hybrids of the male and female.

The production of diversity in the species is eventually as important as maintaining the size of the population. Diversity allows the selection for more fit individuals in the species by the adverse influences of the environment, with fit defined as the ability to produce more progeny. The process of selection allows the average characteristics of a species to slowly change with changes in the environment. Diversity also allows the formation of subgroups of individuals that separate and become new species and thus produce even more diverse organisms.

Death is necessary

Given the need for a species to be able to change over time, the failure of repair and survival mechanisms to enable individuals to live indefinitely should not be seen as a weakness but rather a strength. Without death of parents there would not be space for children, and without children there can be no evolution. The philosophical ideal of

immortality is presumably based on the individual's instinct to survive, and not a reasoned evaluation of the best strategy for survival of the species. If consciousness was distributed and shared among individuals of a species, the death of one individual might be not be seen as a final and tragic event. Of course, if consciousness were shared, there must still be a mechanism to shed old memory, for it to die, in order to keep the total from increasing without limit.

Life had to create itself

In addition to purpose, the creation of the Internet and life were very different; humans created the Internet, but life had to create itself. Life must have evolved from relatively simple collections of molecules generated by spontaneous chemical reactions. From the beginning the evolution of life required replication.

Replication is information transfer

Living organisms survive in vastly different environments and have very different sizes, shapes and behaviors. Many of the activities that fill the life of an organism, such as movement and eating, may not be directly related to replication, but they are required for it to occur. However, even focusing on replication we find that animals, plants and microbes have developed a huge variety of structures and behaviors to accomplish for this purpose. While the details of replication may be both interesting and important, finding a common theme requires first making the abstraction that the process basically represents transfer of information from parent to child. Examination of replication at the molecular level reveals a common method of information storage and replication.

The information that is transferred from parent to child is the information necessary to produce the organism. The growth of a living organism is the result of many specific chemical reactions that convert food into the molecules that form the organism. These chemical reactions are produced by large protein molecules called enzymes. The information necessary to make the enzymes is encoded in a digital format by giant molecules of desoxynucleic acid (DNA). DNA typically represents only a minor percent of the mass of an organism, but it encodes all the information necessary for making a new copy.

Information transfer occurs in and requires a specific environment

A DNA molecule contains all the information necessary to make an organism, but it is unable to do anything in isolation. For example, in order for it to replicate, which is a chemical reaction, the action of many enzymes is required. While the information required to make these replication enzymes is encoded in the DNA, reading the information and using it to make the replication enzymes requires many other enzymes. While the information to make these enzyme making enzymes is encoded in the DNA, without enzyme making enzymes, you can not make these enzymes or any others. The enzymatic reactions require a specific pH and temperature, and enzymes plus material made by enzymes achieve these conditions. Inside a complete organism, with all these components present, all these reactions are possible, and the system is capable of autocatylitic growth and replication.

The situation is really no different on the Internet. A message is typically stored as a pattern of charges in the memory of the sender's computer. In order to be converted into a message, the charge pattern must be read and processed by a series of logic circuits to convert it into a series of electric pulses. These pulses must be translated into a format suitable for transmission, typically TCP/IP, and then sent to the proper port on the computer and transmitted on the Internet. Getting to the destination computer requires many steps, as we will see in the following chapters. Then the message must be translated back into a form suitable for the destination computer, and stored in the memory of the target computer. Finally, it must be translated into an image on a computer screen that a human can understand. Thus the information on the sending computer memory is of little use in isolation. Only in a very specific, complex environment can it be transmitted to another computer and read by the intended human.

The theme of the book

Now the plan for the book is hopefully both more clear and more plausible. The theme is information storage and transfer, which both the Internet and living organisms need to do in different ways. Usefulness of the comparison is not the degree of similarity or divergence we find but rather the increase in the depth of understanding that results from the examination of the process. Since the main theme of this book is information, the next chapter will describe what is meant by information.