ON THE *SOUL*[*]

Ed Fredkin

The following definition of the soul is taken from the Encyclopædia Britannica, MDCCLXXI (1771 AD), Volume III, pages 618-619.

Soul, a spiritual substance, which animates the bodies of living creatures: it is the principle of life and activity within them.

Various have been the opinions of philosophers concerning the substance of the human soul. The Cartesians make thinking the essence of the soul. Others again hold, that man is endowed with three kinds of souls, *viz.* The rational, which is purely spiritual, and infused by the immediate inspiration of God; the irrational, or sensitive, which is common to man and brutes; and lastly, the vegetative soul, or principle of growth and nutrition.

That the soul is an immaterial substance appears from hence, that its primary operations of willing and thinking have not only no connection with the known properties of body, but seem plainly inconsistent with some of its most essential qualities. For the mind discovers no relation between thinking and the motion and arrangements of parts.

As to the immortality of the human soul, the arguments to prove it may be reduced to the following heads: 1. The nature of the soul itself, its desires, sense of moral good and evil, gradual increase in knowledge and perfection, etc. 2. The moral attributes of God.

Under the former of these heads it is urged, that the soul, being an immaterial intelligent substance, does not depend on the body for its existence; and therefore may, nay, and must, exist after the dissolution of the body, unless annihilated by the same power which gave it a being at first. This argument, especially if the infinite capacity of the soul, its strong desire after immortality, its rational activity and advancement towards perfection, be likewise considered, will appear perfectly conclusive to men of a philosophical turn; because nature, or rather the God of nature, does nothing in vain.

But arguments drawn from the latter head, *viz.* the moral attributes of the Deity, are not only better adapted to convince men unacquainted with abstract reasoning, but equally certain and conclusive with the former: for as the justice of God can never suffer the wicked to escape unpunished, nor the good to remain always unrewarded; therefore, arguments drawn from the manifest and constant prosperity of the wicked, and the frequent unhappiness of good men in this life, must convince every thinking person, that there is a future state wherein all will be set right, and God's attributes of wisdom, justice, and goodness, fully vindicated. We shall only add, that had the virtuous and conscientious part of mankind no hopes of a future state, they would be of all men most miserable: but as this is absolutely inconsistent with the moral character of the Deity, the certainty of such a state is clear to a demonstration.

---

[*] In this paper, we use "soul" to mean the common, current definition, and "*soul*" when we mean the new definition as given in this paper.

THE SOUL -- A SCIENTIFIC OVERVIEW

For many centuries the soul has been the subject of much philosophical, metaphysical and religious speculation. This speculation has left many ancient questions that remain largely ignored by modern science. As examples, consider the following questions:

Does the soul exist?

Is reincarnation possible?

Can the soul survive bodily death?

Can the soul exist separate from the body?

If there is an immortal soul, what kind of thing is it?

What is the relationship between the soul and the body?

Does a soul have existence before birth or before conception?

After death, might compensation for the inequities of life be visited on the soul?

Is man endowed with three kinds of souls: The rational, the irrational common to man and other animals and the vegetative, related to growth, inherited traits and evolution?

The scientific method compels us to not make hypotheses about alleged phenomena or other concepts that lack experimental evidence. However, applying a scientific approach towards the understanding of both the meanings and possibilities implicit in concepts referred to by "soul" can prove rewarding.

New scientific developments have advanced our understanding of how computational processes take place. This makes it possible to gain new insights into questions that have arisen, since antiquity, when thinking about the concept of the soul. We can now make scientifically correct quantitative theories and mechanistic models that address such questions. This understanding comes from a different direction than what has been explored in the past. Our results fit into the existing framework of philosophical, metaphysical and religious beliefs about the human soul in a new way. A modern definition of “soul” and a scientific approach will enable us to see what a soul might be and how it might work. We will attempt to answer the nine questions above in a new context.

In this paper we start with a new definition of what the meaning of “*soul*” might be. Then, based on that definition we will explain in detail what the *soul* is made of, how it works, how it might survive us in death, and much more. All of these conclusions will be obviously true to any person who makes the effort to understand what is written here. There is just one proviso, and that is the necessity of understanding the concepts about information and information processes that have been discovered through familiarity with computers and with what it is that computers can do. Geometry starts with simple definitions, axioms and postulates and proceeds to prove ever more complex theorems whose truth might not have been obvious a priori. For all such theorems, the proofs must rely on the underlying assumptions. If we assume the validity of the axioms, postulates and theorems used in the proof of a new theorem, and assuming that the proof is done correctly, we can assume that the new theorem is true.

With regard to the *soul*, we do not wish to proceed in a rigorous, Euclidean fashion, (even if we knew how) however our approach is similar in one respect. We make certain assumptions and then give an objective and simple definition of the human *soul*. We will then try to convince the reader that our definition corresponds to the common notion of what the human soul is, based on the following criteria: *The definition should make sense for all aspects of the human soul for which we have objective, scientific observations.* Once we have a definition that meets the objective criteria, we will show that such a *soul* can have many of the other properties that have been attributed to the soul despite the lack of objective evidence. The approach is to define something in a straightforward manner and then to elaborate on the properties of what has been defined. What may be controversial is that we shall call the thing that we have defined, "*soul*". We believe that the version of the *soul* defined herein covers many possibilities for differing definitions of human soul in the sense that it gives plausible explanations for most of the properties and possibilities of the soul that have been discussed in the literature.

How could a new definition of the human *soul* be in good agreement with prior definitions? Every different concept of the human soul must differ in some respects from other concepts. The definition here is perhaps much more different than the average, but it seems to capture the essence of many views of the nature of the soul. Our effort is not to decide on what properties the human soul has or ought to have, but rather to see how various properties could be true of the *soul*, once we have an understanding of a plausible model of the *soul*.

At this point, we need to define our use of a word that represents the key concept; essential to the understanding of what follows. The word is "*informational*", and the definition is: Pertaining to the arrangement of things that represent discrete, digital information, independent of what the things are. *Informational* implies that there is meaning associated with the information, and that the information is in a form amenable to being used, communicated or modified by an information process (basically some form of generalized computer). Such digital information can be static (so called Read Only Memory, printed literature, or music on a CD) or dynamic (engaged in information processing, sensing, control or communication, such as the memory contents of a working computer or music being played from a CD). Dynamic digital information is engaged in an information process (computation) where as a consequence some of the digital information may be changed. A computer, built out of silicon chips, is not an informational construct. The reason is that although the physical parts of a computer do information processing, the information represented by the parts themselves are fixed and cannot themselves be processed and changed. The digital information being processed by the computer is an informational construct. It is possible to have a computer that is purely informational; if we write a simulator for a computer that runs on another ordinary computer, then that simulated computer is an informational construct, but the ordinary computer is not.

Before we go into the detail about the *soul*, as we shall define it, we will mention some conclusions revealed by our study. In general, the conclusions are not about what is true today, but about what could be true in the future regarding the nine questions posed earlier:

- What is herein defined as the *soul* certainly does exist.
- Reincarnation is possible, but not in ways previously imagined.
- The *soul* may, in partial or nearly total form, survive the death of the body.
- The partial or nearly total *soul* may be able to exist separate from the body.
- The *soul* is potentially an immortal informational construct; thus a part of physics.
- Normally, the body is the abode of much of the *soul*, but not necessarily so.
- Certain parts of the *soul* do exist before birth or conception.
- Justice after death implies that a soul can continue in existence, more or less intact, even after death. Today, only fragments of a soul can survive the death of a person. This may change sometime in the future.
- It does seem that the *soul* does come in the very three forms mentioned in the EB. We will explain this in more detail.

We must caution the reader that the nine points above will likely be misunderstood until reading further. For example, we are certain that the laws of physics do not prohibit the survival of a person's *soul*, largely intact, after the death of that person's body. However, we do not believe that such an event has yet occurred.

Our definition of a particular human's *soul* is going to be much like the definition of a particular human's body. While most humans start out with a complete and normal body, some do not. Additionally, a person may be maimed and still survive as a human. An amputee may not possess a complete and normal human body, but he or she is still a human being. Our bodies are normally made up of arms, legs, kidneys, brain cells, a head full of hair, tonsils, etc. A bald man, someone with no kidneys, a quadruple amputee or a paraplegic is still a human. In order to understand our definition of "*souls*" we need to understand that a *soul* is made up of many parts, that not all *souls* in human bodies have all of the parts that other *souls* have and that parts of a humans *soul* can exist outside of that human's body just as one persons kidney can be removed and can remain alive, disconnected from the rest of the body, long enough to be transplanted into another person. Further, just as a tissue culture taken from a person, such as Helen Lane, may continue to live in laboratories for decades, parts of a human *soul* may continue to exist and even function outside of the human body. Finally, our bodies do not disappear with our death. Similarly our DNA can survive our death and remain intact and accessible. Some of our DNA along with many other fragments of our soul may survive in our children, we will show various other ways in which parts of a human *soul* can also survive after the death of its body.

We are now ready to give our definition of "*soul*" followed by a paraphrased (in light of our new definition) third paragraph from the quote, at the beginning of this paper, taken from the Encyclopaedia Britannica, MDCCLXXI (1771 AD), Volume III, pages 618-619.

The *soul* in every living thing is the informational parts of that thing that are purposefully engaged in the informational aspects of its ability to be conceived or germinate, grow with cells differentiating, grow further in size, move, make use of sensory information, react reflexively, learn, behave instinctually, think intelligently, communicate with other beings, teach, reproduce, evolve and in general carry out informational interactions starting with the combining of parental DNA, informational interactions with itself, with things external to itself through senses, actions, constructions, creations and communications, and with its progeny through contributed DNA. A *soul* can learn from experience, from reflection or by being taught by other *souls*. In turn, a soul can teach other *souls*.

Thus the *soul* is an immaterial substance and its primary operations of willing and thinking have not only no connection with the known properties of the elements of body, but seem plainly inconsistent with some of a body's most essential qualities. The *soul* is an informational entity, which is constructed out of the states and the arrangements of material things. Yet it is immaterial, in the sense that the particular choice of material, whose states and arrangements represent the bits of digital information, is not of great consequence. If the entire informational structure of a *soul* was transferred intact into a suitable new host, it could continue in existence largely as the same soul.

First we will define a dynamic *soul*. Second we will define a static *soul*.

A dynamic *soul* is just what we normally think of as the *soul* of a living, awake and active being. It is something that seems to inhabit a living body that is the essence of the ongoing intellectual and emotional activity of that person. It is the *soul*, with the help of the brain and the body, that contemplates, reflects, feels, thinks, learns, communicates, decides and commands the body to do its bidding. It is the home of processes that implement our ethical, moral and religious convictions. In our view, the dynamic *soul* makes use of the brain in doing its thinking and also makes use of the body to manipulate things in the real world, to receive information through the senses and to communicate with other *souls*.

A static *soul* is simply digital information that is similar to a combination of computer programs and computer data. A static *soul* cannot be understood without understanding the exact processes that interpret and process the information that corresponds to computer programs, and understanding the nature and intent of the part that corresponds to computer data, which is certainly complexly encoded in some kind of database like set of structures.

When your computer comes to life with Microsoft Windows and various programs doing various things, what is going on from the informational point of view, is similar to the concept of what is going on in a dynamic *soul*. When you buy a CD-ROM that contains various programs and data that you can use with your computer, from an informational point of view that CD is similar to the concept of part of a *static soul*.

The dynamic *soul* is the consequence of a set of computer like programs that are running in the brain-body; the dynamic *soul* is in control of the body in the same general way that a computer might be in control of a robot. If one could extract a copy of a *soul* from its brain-body, it would make no sense without also knowing the detailed structure and logic of brain-bodies in general and of its particular brain-body. Similarly, if one is given a binary copy of a machine code computer program, it cannot be understood without also understanding the exact nature of the machine and operating systems that are capable of running that code. It is quite possible (but perhaps extraordinarily difficult) that the most likely design and structure of the machine or computer capable of running the code could be deciphered from a study and analysis of the code.

When a program is running in a computer, it makes use of the hardware and operating systems of the computer in order to run properly. The running of a program is the evolution of the program and its data in time. While it is running, we may think of the program as being in a dynamic state. If we pause the computer, so that the program comes to an instantaneous stop, ready to continue at the press of a key, we would still say that the program is in a dynamic state. In actual fact, a dynamic program running a billion instructions per second could be said to be at a dead stop a billion times each second.

A program is in the static state when it has been rendered inoperable by being translated and copied onto some medium, such as onto a floppy disk that is stored for future use. At some time in the future, the program can be reinserted into the computer and caused to continue from the exact point where it had stopped. It also can be brought to life in a different computer so long as the different computer is sufficiently compatible with the original one. Thus the difference between static and dynamic is only a matter of degree. It is no more than, in this case, the difference between being able to take the next step in a few pico-seconds, or in a few hours or a few millennia.

A single picture frame from a motion picture film is a static representation of a movie at one instant in time. It’s as though the world in the movie was brought to a stop. To understand the meaning of a snapshot in relation to a movie, we need to know the position of the snapshot within the sequences of snapshots in the movie. Such a snapshot is a lot like the state of a program and its data stored on a floppy. However, the program on the floppy is complete while the snapshot is only a partial 2 dimensional representation of a 3 dimensional real world. Of course the program on the floppy can only come back to life if it is loaded into a compatible computer.

If we could capture some of what there is to a human *soul*, at one moment in time, we would have a snapshot of part of a *soul*. Of course, a real snapshot leaves a lot out. It is two-dimensional and deals only with what is visible to the camera. Looking at a single snapshot, we can’t necessarily tell how things were changing. On the other hand, our definition of a static *soul* is that we capture, at one point in time, everything about that *soul*, including all of the information needed to make it dynamic.

The most common example of a natural static soul is a seed. It has all of the informational aspects of the plant that it can grow into. While static, a seed can remain viable for decades. Then, when put into an appropriate nurturing environment, it can start to grow into whatever it is destined to be. A seed is similar to a floppy disk that has a software virus on it. It is static until it is placed in a nurturing environment, i.e. moist, fertile earth for the seed, a computer for the computer virus.

Assume that we want to solve a very hard mathematical problem. Factoring a very large integer (hundreds of digits) is such a problem that can be broken down into parts. One hypothetical way to do this would be to create a computer virus which contains the number to be factored, and some factoring software. The algorithm could involve a probabilistic search. Once in the computer, the program could make copies of itself and spread the virus. The descendents of the original virus are very unlikely to search the same space as any of their ancestors as there are simple processes that allow a program to come up with a random number. The likelihood that any descendant will come up with a random number that is the same as some other descendant's can be made arbitrarily small. Eventually, one of the descendants of the original virus might solve the problem. In this example, the every copy of virus has a *Soul*. It is the informational aspects that allow it to operate, reproduce and fulfill its destiny. Similarly a carrot seed has a soul. It is the informational aspects that allow it to germinate, grow and differentiate into a mature carrot, and reproduce its kind by producing carrot seeds. Obviously, in the 2 examples above, the computer virus reproduces asexually while the carrot reproduces sexually. I'm sure that, in the near future, we will see much improved computer viruses that reproduce sexually and that can evolve as do living things.

To be specific, we believe that a static *soul* can be represented by information. We will use a simple mathematical notation; a superscript, i, to indicate an individual person and a subscript, t, to indicate a point in time. The static *soul* is just that information that defines S_t^i the static *soul* of person, i , at time t. S_t^i can be conveniently divided into three parts:

$_t^i$, The information that describes the accessible, external world.

P_t^i, The information that defines how that brain-body processes *Q*.

Q_t^i, The information that is retained in and processed by that brain-body.

S_{t+1}^i=F($_t^i$, P_t^i, Q_t^i)

The boundaries between *W*, *P* and *Q* are not precise.

It may be useful to compare these parts of a *soul* with corresponding parts of a computer system. *W* is like the environment that a computer is in. This includes connections to external networks such as the Internet. The computer might have video or audio input devices which enlarge *W*. *P* is like the computer hardware along with the operating system and various applications programs. *Q* is like my files in the computer. Things I have typed in or modified. *Q* includes pictures I may have downloaded from the Web, my correspondences, the state of games I am playing, etc. The function *F* is the computational engine, the hardware that actually produces the new state, time=*t+1*, from the old state, time=*t*.

When a program sets a parameter in the operating system, is that parameter part of the operating system or part of the program? One can think of it either way. With some exceptions, it is easy to see what should be in each of the three categories, *W*, *P* and *Q*. Occasionally, the preferred category may be decided somewhat arbitrarily. The breakdown of the informational constructs of a *soul* into the three classes, *W*, *P* and *Q* is a useful exercise, aiding one in developing a sensible model of the informational aspects of a *soul*.

The evolution of P (the acquisition of skills and useful know how) proceeds relatively faster in a child than in an adult. On the other hand, the evolution of Q (the application of skills and know how) can proceed relatively faster in adults.

Given two people, i and j, P^i and P^j have a lot in common. Just as everyone is normally born with 2 eyes, 2 ears, general bilateral symmetry with a few exceptions, much of the abode of the *soul* is similar from one person to another. However, one person may be an Olympic Gold Medallist at the 100 Meter dash, while another person may be a couch potato. From one person to another, P^i and P^j are as similar and as different as is the body of one person compared to another. If i and j are of the same age and sex and general physical condition, P^i and P^j would have more in common. If i and j are a pair of identical twins, P^i and P^j would have even more in common though there would still be differences.

What we call “brain-body” most others would simply call the brain. However, we consider it important to realize that thinking involves more than just the brain. The brain is only part of the nervous system. Our eyes are intimately connected to the visual cortex. We often have thoughts that cause a visceral reaction that is in turn communicated back to the brain. Whenever we think, we are in an environment; our body – our world. We almost never imagine the possibility of being in some really foreign kind of world. The possibilities as to what could be different are vast: a world could have a different spatial geometry and connectedness; it might not have the hierarchical structure (biology, chemistry, molecules, atoms, particles) of our physics.

Thus, the normal abode of the *soul* is not just the brain, not just what we are calling the “brain-body”, but also the combination of the brain-body and the environment. *P* is not the brain-body, but rather it is a precise informational model of how the brain-body-world processes information. *P* has two major components: generic and specific. The generic component involves that which is in common with our world (e.g. the laws of physics) and with essentially all other normal humans. In the computer analogy, the generic component is similar to the idea of a particular computer and operating system, such as in the Mac, or in a PC running Windows; it would include software packages such as MS Office. The specific information has to do with all the things that differentiate this person from other persons. This includes genetic differences along with the memories of a lifetime, the characteristics of personality, etc. In the computer analogy, the specific component is similar to what parameters and data are in the computer and how various files such as autoexec.bat, config.sys, win.ini, the Registry etc., are setup. Of course it also includes all files peculiar to that machine such as documents written by the user, records of email exchanges, digitized photos from the user’s camera, etc.

What this means is that P is the definition of a particular, complex *interpreter* (in the computer sense). In a computer, we may write a program in Java, or one of thousands of other computer languages or application development systems. An *engine* is needed to interpret what we write in order for a computer to carry out the actions that the programmer had in mind. The *engine* is either a kind of program called an *interpreter*, or it is the computer hardware. "*Execution*" is a good name for the action of the computer hardware directly *executing* the steps of a program. Sometimes, *interpretation* is preceded by *compilation*. This is a step where the program, written in a language optimized for the human process of writing the program, is translated into a language more suitable for *execution* on a particular computer. Additionally, all of the other software that will be needed during program *execution* must also be available in any host computer that hopes to run the program correctly. A *compiler* needs to know two things: first, the actions that are supposed to result as a consequence of what the programmer wrote, and second, how to get a particular computer to carry out those actions.

A start towards understanding *information processes* had to await the computer age. But just because the computer age is upon us does not mean we already have a full understanding as to what information is. What is true is that we now merely have a better understanding of what an information process is. Information processes are unusual from the point of view of physics. They don't have the need of conventional kinds of physical dimensions. For example, a person can be described as having a certain amount of mass, normally measured as weight. A person has a nominal length, thought of as his or her height. A person can be described as a system that converts a certain amount of chemical energy into mechanical work and heat, normally thought of as having a rate of metabolism and often measured in calories/day.

From the point of view of physics, what we call our weight really is what we think of as our mass; it has dimensions M and is measured in Kilograms or perhaps in pounds. Height has dimension L and is measured in Meters or feet and inches. Time has dimension T and is measured in seconds. Energy has dimensions ML^2/T^2 and is measured in Joules, kilowatt-hours or Calories. Information processes, on the other hand, are not described by any of the current units or physical dimensions so far as we know. This means that 2 bits of information that interact in a logic element of a computer cannot now be understood in terms of amounts of Mass, Length or Time or any particular combination of such units. This is likely just a reflection of our current ignorance. Someday we may discover a unit of information that is closely related to some combination of units such as ML^2/T, which happens to be angular momentum. In physics, the unit of angular momentum is called ∇ (Planck's Constant Reduced). The 2 most fundamental constants of physics are the speed of light, c, and ∇. What seems likely today, is that those 2 constants are involved in the most basic kinds of computer operations.

Of course such physical things as printed-paper, magnetic disks, brains or DNA can represent information. Information must ultimately have a physical means of its representation, however the information itself is independent of the means of its representation.

The ages old concept of a clear distinction between the physical parts of the body and the evanescent part, the human *soul*, turns out to be correct. The fact that the *soul* can be resident in a physical brain-body does not mean that the *soul* is physical in the physics sense of Mass, Length and Time. The *soul* is physical only in the sense that information is part of the world of physics. But physics has not yet finished putting "information" into the proper pigeonhole. Information is certainly subject to the laws of physics, but it is not physical in the sense that a machine is physical. A machine can be explained in its entirety as interacting components, each of which is covered by the units of Mass, Length and Time. For a machine, the choice of materials is important; you cannot make an anvil out of warm butter. Information has to do with the arrangement of things and the processes that result in the rearrangement of things. It is totally independent of what the things are made of. This point must not be obscured by the fact that, today, books are best printed with ink on paper and integrated circuits are best made out of silicon plus carefully chosen impurities. In principle, what is used to represent information is irrelevant to the information. Those who long ago thought to describe the *soul* as immaterial made a wise and prescient choice.

The parts of the *soul* that exist before conception

While the *soul* is immaterial information, it is represented by various material systems. First there is DNA. DNA is both coded information about the design of a creature along with coded information about the process that transforms a fertilized egg through all of the stages from embryo to adult. In addition, DNA is part of the actual processes that allows for cell replication, differentiation and even the physical actions of building proteins and other components of cells and whole living things.

Just prior to conception, just prior to the fertilization of an egg by a sperm, there is a great deal of information carried to this process by the egg and the sperm. Much of this information has to do with the design of the process that constructs a human body from the sea of nutrients and hormones that also may influence the result. Some of the information has to do with things that are accessible to one's adult intellect. For example, identical twins, separated from birth till adulthood, will often make identical decisions regarding certain arbitrary preferences. We know that there are many ways that the DNA formed at conception has a profound influence on the development of the intellectual part of the *soul*. A seemingly random process chooses the exact combination of parental DNA that fuses to make a new individual. However, what is not random is the fact that all of the DNA comes from one or the other of the parents. It is easy to see many characteristics of a parent in looking at a child. It is reasonable to expect that part of the *soul* of a child is inherited from its parents through genetic information. This information exists in the parents, the grandparents and all of this information comes from ancestors of the child (with the exception of mutations). While the combination that the child gets is not determined, what is clear is that in general, we may say that the child does not normally get any genetic information other than what was present in his or her ancestors.

Because information is not the same as physical objects, one is not constrained to replicate the medium when replicating information. We can take information from a floppy disk and transfer it into computer memory (RAM) and later transfer it onto a CD-ROM. Bits are represented by the orientation of magnetic domains on a floppy disk, by physical or optical deformation in a CD and by electrostatic charges in RAM. One can think about the process of development from a fertilized egg to an embryo to a newborn to an adult from the informational viewpoint. We will try to expand on that concept.

DNA is highly organized information that allows for at least nine different functions:

1. Information storage
2. Information merging, so that offsprings inherit from both parents
3. Information replication when cells divide
4. Information controlled growth with cell differentiation
5. Enabling conversion of an informational design into a body of an adult living thing
6. Biological chemistry, e.g. protein synthesis.
7. Making available accessible information to the mind of the living thing.
8. Passing on to heirs information that was evolutionarily useful even though it was not expressed in the current creature, e.g. 2 brown eyed parents can conceive a blue eyed offspring
9. Enabling the process of inheritable mutation that allows for the evolution of new species

Let us imagine a fairly complete list of Dewey Decimal system numbers. Further, imagine two libraries where each library must have exactly one appropriate book for each different Dewey Decimal number. The two libraries need not have the same book for a given Dewey Decimal number although they might. Now a child can obtain a new library by choosing and copying just one book for each Dewey Decimal number, either from the father's library or from the mother's library. The child's library will have a complete set of books, one for each Dewey Decimal number, but about half of the books would be from the father's library and the rest from the mother's library.

Aside from information that was in either the father's library or in the mother's library, the child's library may also have new information. While the child's library will have nothing new in any of the books, the juxtaposition of books from the two libraries may have some amount of information that was not in either library. For example, under the category "tools" the father's library may have had a book that discusses how to use woodworking tools and the mother's library may have a book on how to use gardening tools. Under the category of furniture, the mother's library might have a book on designs of various kinds of wooden furniture and cabinets while the father's library might have a similar book on the designs of various metal and plastic furniture. Neither library has sufficient information to allow using woodworking tool to make a cabinet, but the two together (the father's woodworking tools book and the mother's wood furniture design book) have the all the necessary information. Conversely, if one inherits the same information from each of two parents or incompatible information, then the combination of some of the books may have less information than the sum of the information in each of the books. Thus combining information involves mechanisms that may preserve, add and subtract information. There are many different analogies that can usefully add to our comprehension of how information is involved in living creatures and that can add to our understanding of the *soul* from an informational viewpoint.

In what follows, we will show how the *soul* can survive the death of the body. We will explain characteristics of a disembodied *soul*. We will explain exactly what is necessary for a *soul* to re-inhabit a different body. We will explain exactly what is required for a static *soul* to be able to exist intact though in a suspended state without a body. Further, we will be able to understand exactly how a *soul* could think and communicate with humans, despite the fact that the *soul* no longer had an ordinary corporeal body. We will be able to think clearly about what will be possible in the future as opposed to what might not be possible today. Some things may appear to be forever impossible and we will be able to know what facts underlie our belief in their impossibility. Finally, we will answer in a definitive way, essentially all questions about reincarnation.

Before we can gain a clearer understanding about the *soul*, we must first gain familiarity with ideas common to both the *soul* and to computer systems. A computer can be a relatively simple kind of machine that uses certain principles in order to work with information. In a similar though more complex way, the brain-body also works with information. When we commit something to long term memory, it is somewhat similar to a computer system adding information to a database. Of course, we have ways of retrieving things from our memories that use processes quite different from what is currently done in computer systems. The point is only this: by understanding how the computer works with information we will be in a good position to understand further explanations as to what the *soul* is and how it works. We have relegated our short tutorial "How the Computer Works with Information" to Appendix A.

The process of capturing a human soul.

There are three basic ways that parts of a human *soul* can be captured.

- By examining and analyzing the external informational products of a *soul*;
    - Remembrances and recollections of friends, relatives and colleagues
    - Remembrances and autobiographical writings
    - Works of literature, music and art or of engineering, construction, design, inventions or discoveries etc.
    - Recorded discoveries in science, philosophy, mathematics etc.
    - Tales of deeds, accomplishments, acts of heroism or leadership
    - Books, recordings, videos, and other recorded media that is relevant.
- The decoding of the DNA of that soul.
- A future informational process that accomplishes for a living human mind something akin to getting a backup tape from an operating computer system along with the design details of the computer.

By carefully examining an integrated circuit chip with high-powered microscopes, it is possible to deduce the detailed logic design of the chip. From that design we can make a replica chip. It can be done but it is not easy. In the old Soviet Union, in the city of Zelenograd, a secret industrial enterprise existed to decode the logic design of foreign computer chips and to re-implement the design so as to produce locally made versions. One can often deduce the intent of the designers and the purpose of various subparts of the chip, however being able to do so is not necessary in order to make a faithful replica that performs exactly as the original. It is even possible to conduct a microscopic examination of a processor chip while it is operating. This includes the possibility of being able to detect and measure the electrical signals as they propagate through the circuitry. In a similar fashion, we can imagine that brain science could advance to the point where we understand the functioning of the most primitive logic and communication elements. If we could also create a technology to produce a complete wiring diagram of all such elements in the entire body nervous system of a person whose brain still retained its full complement of information, then it would be possible to create P, the information that defines how the brain-body processes information. This would allow the creation of an interpreter (program) that could process information in the identical way. Despite having such an interpreter, we might still be unable to understand how the brain did it. Understanding how the brain did it could be useful and interesting, it’s just that it is not necessary.

The methodology of getting the detailed *wiring diagram* of the brain might be similar to current MRI technology that generates 3-D pictures of the brain. In this case, the level of detail would be enormously greater, but the principle of obtaining information internal to the brain in a non-invasive fashion would be the similar. Further developments of similar technologies would allow for the extraction of contents of the brain's memories, the evanescent information of memories and learnt skills that decay with advancing age.

So, the process that creates a copy of P, might also create a copy of Q. Q is the information that represents the state of mind of an individual (a major part of the *soul*) at a point in time. Q is the information that is processed by the brain, along with sensory inputs, in order to determine the state of mind at a future time. Part of the evolution of Q with time is called “consciousness”. The rest of the evolution involves the workings of the unconscious mind. Capturing Q may be much more difficult than capturing P. Further, the information in the brain that represents Q may not last long after death. It may be like the information stored in integrated circuit memory chips. Most such chips loose the information stored very quickly if the power supply is interrupted. Of course, there are other kinds of memory chips that continue to retain information without being supplied with external power. It is most likely that the preservation of the information in Q will someday be possible despite the very recent or immanent body death of the person. Thus we have a situation just the reverse of what we are used to in normal organ transplants. Normally the doctors wait for a potential donor to become brain dead before they begin to harvest his or her organs for transplanting. It will most likely require heroic action prior to brain death in order to save the entirety of a person’s *soul*, when and if it becomes possible.

Let us assume for the moment that future technologies for obtaining high fidelity copies of both P and Q are developed in the following ways:

- Similar in application to MRI (Magnetic Resonance Imaging or NMR)
- Non-invasive and non-injurious
- Capable of determining the wiring and logic of the brain
- Capable of determining the state of all information (other than short term memory) stored in the brain
- The whole process should only take a few hours.

Having the capabilities listed above would mean that the state of almost the entire soul of a living person could be extracted. Having done so, it could be incarnated into an advanced computer system and presented to the owner of that soul so as to allow for some verification of the fidelity of the process. Being confronted with what is essentially one's self may be a strange experience, but like many things something that society and its individuals might be able to get used to. This could be done a number of times during a person's lifetime, sort of like making a backup copy of the state of a computer, so that if the person dies suddenly, recent additions and changes to that person's soul would not be lost. Of course, a person's views change over time, but given that that is recognized, it shouldn't pose a problem. It might be comforting for someone on his or her deathbed to have with them the dynamic copy of their soul, watching and listening to what is happening.

Computer Hardware Needed to Support One *Soul* in Real Time.

It is natural to ask "How much computational resources (processing and memory) are needed to handle the computational load of one person in real time?" The answer with regard to processing seems to be about equal to the power of a few high end processor chips. Today, a single CPU chip (the heart of today's PCs) can do about one billion instructions per second. However, the organization of that chip is optimized for the kinds of numeric and symbolic computation done by today's computers. One such chip would be useful but the other chips, of equal size and amount of digital logic, would need different architectures in order to work with images, sounds, and to perform certain cerebral tasks. We don't yet know how to do these tasks, but it is possible to judge the amount of silicon that will be sufficient despite our lack of detailed knowledge of what methods are to be implemented. Instead of just being a general purpose computational engine, most of the silicon would be better dedicated to specialized functions. With regard to memory, about 1 GB of RAM and 64 GB of slower (disk) memory are clearly more than sufficient. These numbers require that we understand clearly what needs to be programmed and how to program it. This is not the case today. What is often true in such cases is that the first working systems will use much more in the way of computer cycles and amounts of memory (due to our primitive understanding of how to program what's necessary) and later systems will more closely approach the amount of resources that is truly necessary.

Appendix B gives a detailed accounting in support of all of the above numbers. Today (2000), $3,000 can purchase a computer system with such capabilities. What is inescapable is that the cost for a given amount of computational resources decreases by about a factor of two every two years. Therefore, within 20 years, $3,000 will buy about 1,000 times the necessary computational power.

Once Phi has P and Q, then given the same sensory inputs, Q will evolve in Phi along the same path as Q would have evolved in the living human *soul*.

How can we know whether we have captured a human's *soul*?

This fact is important, for it lets us understand all of the possibilities of the human *soul*.

First of all, we will think about the *soul* in quantitative terms.

Assume for the sake of argument that the Pythagorean theorem was discovered and communicated to others by Pythagoras. If you have just learned the theorem, then a bit of what was the *soul* of Pythagoras becomes part of your *soul*. This transaction can be viewed quantitatively; first from the perspective of the recipient and second from the perspective of the *soul* of the source.

We need a word to represent the informational aspects of a fragment of a human *soul*. I suggest "soulelcule" to represent a single *molecule* of a human *soul*. The number of possible distinct soulecules would be similar to the number of possible distinct molecules. Examples of soulecules and their sources are:

- "To be or not to be, that is the question" – William Shakespeare
- "Given a right triangle with sides a and b and hypotenuse c, $c^2=a^2+b^2$" -- Pythagoras
- "$E=MC^2$" -- Albert Einstein
- "Blood, Sweat and Tears" – Winston Churchill
- The image in the painting *The Scream*. – Edvard Munch
- "do la fa mi so fa do do re re mi mi fa" Wolfgang Amadeus Mozart at age 7.

A measure of the informational content of a fragment of a *soul* appears necessarily complex. Consider the Pythagorean Theorem. Understanding it and being able to prove it cannot stand by themselves but must be a part of a complex of other informational constructs. First of all there are the informational definitions and procedures associated with common sense notions about space and distances, lines and points. Then there is the idea of geometry. While the axioms and postulates are at the base, the whole process of *doing* geometry can itself be viewed as an informational construct. Then, there is a handy library of already proven theorems we can draw on, so as to not need to always start from the beginning. All the Geometry you learnt in school shares in making use of certain kinds of informational constructs that are common to the various theorems you might know. These are often associated with common sense concepts, including ideas such as that of the point, straight line, arc, etc. along with axioms, postulates and procedures such as those common to constructions. There are procedures related to vision in general that assist in understanding a visual geometric diagram. There are general procedures that motivate one to start and continue a process such as learning a method, solving a problem, etc.

Much of what we know is very likely to be efficiently coded by making references to other things that we know. To measure how much information is used to learn something new, we have to not only count up the direct representational information, but we must also credit the measure with some share of the common informational resources that are being shared with other things already learnt. By that method, equally complex concepts would have the same measure whether learned early or late. On the other hand, it also means that sometimes when we learn something new, the informational measure of each related thing already known might decrease, e.g. the new information might be the key to allowing more efficient representation of previously learned knowledge. For example, one might remember a great deal of anecdotal information about how to reset the clocks for daylight saving time. Once one knows the phrase "spring forward, fall back", it becomes possible to better understand the process while using less information.

If we chose to only take into account the correct incremental information added by learning one more thing, then the sum of all the incremental information would always give the correct total but nearly identically complex concepts would have different measures depending on when they were learned.

To measure a *soul* (quantitatively), we start with the total genetic information. Just knowing the base pairs is far from sufficient. A full understanding of protein synthesis is necessary but again not sufficient. We must be able to understand the informational processes that make use of the DNA information and which translate it into all the things that are fully expressed and exist in an adult. To that we must add the proper measures of the entire set of informational content added to the brain since conception and then subtract what has been lost through forgetting. A most dramatic loss of *soul* is the process known as infantile amnesia. Its as though we must lose our intellectual baby teeth in order to make room for adult thoughts since many of the memories of early childhood may have little useful evolutionary benefit. The implication is that the brain only has so much capacity for the storage of memories. Another informational loss occurs continually as most of what is in our conscious mind and short-term memories is discarded as inessential. Only a small fraction is saved into long-term memory. Some informational gain is incredibly slow. One can serve, by carefully swinging a tennis racquet 1,000 times a year for 10 years and actually learn a little bit each time (on average). Thus the parts of one's *soul* that develops skills at serving and volleying can continue to increase in informational measure over a very long time despite the likelihood that the totality of the informational measure distilled out of all that practice is fairly small. What's important is that over time it becomes more of the right information.

Finally as we age, the process of loosing information gains headway over the process of learning new information and the measure of the totality of ones internal *soul* starts to decrease. Obviously this occurrence is sometimes pronounced and sometimes too small or too subtle to notice.

To simplify our measures of the number of bytes needed to represent a *soul* we will look at the soul of a 21 year old person blind from birth. Here we will not separately consider the amount forgotten but rather only count the net amount added to permanent long-term memory. One must remember that calculating the net amount involves estimating the rough efficiency of the low level and high level encoding schemes (data compression) that are undoubtedly at work.

1. The number of days is approximately 7500; waking seconds, 7500 days x 18 hours/day x 3600 seconds/hour=approximately 0.5 billion seconds.
2. Imagine an average learning rate, committed to long term memory, of 40 bytes per second for half the waking day. (Not the raw data rate, but the nearly optimally encoded data rate. This is a sustained rate that is probably beyond anything that can be demonstrated in an experiment.) This gives 10 GB of input.
3. We can imagine that other learning from other experiences along with mental reflection to come to another 10 GB. That gives a subtotal for non-visual inputs to long term memory of about 20 GB.
4. If our example person recovered his sight at age 20, one would expect that within a year, he could be said to have an essentially normal soul, only with respect to the amount of required memory. It is very hard to do reasonable estimates of what goes into long term memory on account of having one's eyes open, but it seems doubtful that even in this extreme case that it would result in a peak average rate greater than 10 KB per second with a sustained 18 hour per day average that is dramatically lower; probably below 400 bytes per second. These numbers could be more accurately determined by means of a few carefully designed psychophysical experiments. At 400 bytes per second, we have 365x18x3600x400=about 10GB.
5. For a normal person, because of high level encoding schemes, we would expect the total long term memory from visual inputs of a normal person to vary from the estimate given above by no more than a factor of 10 either way (which shows the amount of estimated precision in the estimates!).
6. For the DNA, the information in $3x10^9$ base pairs (2 bits each) equals about 6 gigabits or about 750 megabytes. This includes the basic design of the brain.
7. Thus we have a grand total of about 40GB as our estimate. Using the uncertainty factor mentioned in item 5 above, the best guesstimate of the bounds for about a 50% probability of having the right number might be a range from 100GB to 16GB.
8. On the issue of processing power, again it is useful to exclude visual tasks in a first estimate.

What is fascinating is to consider are measures of the contribution to the collective *soul* of mankind. Consider a Mozart soulecule as expressed in a melody remembered by millions of people. One measure that makes sense is to compute the total of all contributions over all time to all *souls* of a given Mozart soulecule. Another measure is the current value, which is the same sum but restricted to currently active (living) *souls*.

Of course, not all solecules are equal. The value of a soulecule transferred from one person to another depends on the state and future experiences of the recipient, as well as the inherent value of the solecule. Some soulecules are worth little, nothing or even less than nothing. Others can act like a key or a catalyst, bringing great value by connecting different soulecules into new informational structures. There are good examples from physics where abstract mathematical discoveries turned out to be the key to understanding processes in science. A good example is the concept of imaginary and complex numbers which started out as strange mathematical curiosities and ended up as essential to electrical engineering and quantum mechanics.

Three categories of worth are:

- Soulecules that extend the length of the recipient's life.
- Soulecules that increase qualitative aspects of the recipient's life.
- Soulecules that result in the recipient having enriched his own *soul*.

A measure of a teacher.

Perhaps one could consider the solecules that are added to the *souls* of the students.

Two functions of a teacher; a facilitator of two processes:

- The transfer of soulecules to students.
- The generation in the future of new soulecules by the students.

The informational function of art.

- Artists engage in the creation of external solecules that are the works of art.
- Art can enrich the *souls* of others through the direct absorption of soulecules from the work of art and followed by the possibility that the recipient creates new soulecules as a consequence.

Of course, an artist can aim to maximize the near term effects or create works for posterity.

Fame versus Notoriety

NOTES ON THE SOUL

From: "A History of Western Philosophy" by Bertrand Russell

Page 4, ... "The Egyptians were preoccupied with death, and believed that the *souls* of the dead descend into the underworld, where they are judged by Osiris according to the manner of their life on earth. They thought that the soul would ultimately return to the body; this led to mummification and to the construction of splendid tombs."

The Answers

1. Does the *soul* exist?

   The *soul*, as defined herein, certainly does exist. However the *soul* is not made up of matter; it is an informational construct wherein the states and arrangements of matter and energy are used to represent the information. The informational definition of the *soul* allows for the existence of both the dynamic living *soul* and the disembodied static *soul*.

2. Is reincarnation possible?

   Reincarnation is possible in almost every way that has ever been imagined. However, just as a program needs a compatible computer as its host, a *soul* needs a compatible brain-body or a compatible computer system in order to function in a complete fashion. It is possible that some of the parts of the *soul* can function in host of less or just different capabilities. Thus it is possible that a part of the human *soul* could function in a host, such as a cow, that though incapable of human thought could still be capable of other intellectual processes that would reflect parts of a human *soul*. It seems that the only high quality hosts (the brain-body) for a reincarnated *soul* would have to be either a living clone of the original brain-body, or a similarly accurate computer system that simulates the mind (brain) and body and also provides the external environment.

3. Can the *soul* survive our bodily death?

   Today, the answer is that certain aspects of the *soul* can survive bodily death. This survival is normally in the static form; writings, paintings, music, sculpture, photos, movies, videos... However, whenever a young orphan remembers his father and says to himself “What would dad have done?” then he is being the host to a fragment of his father’s dynamic *soul*. When a pianist can feel what Beethoven must have had in mind when he wrote the Appassionata, then a fragment of Beethoven’s *soul* becomes a dynamic entity while that pianist is playing Beethoven’s music. Eventually, it should be possible to capture in a computer system a sufficient fraction of a dead person’s *soul* so as to leave no doubt in the minds of those friends of that person who converse with that computer; no doubt that what they are talking to has the *soul* of their departed friend. There is no law of physics that stands in the way of recreating a clone with the original adult *soul* of the person cloned or of embodying major aspects of a *soul* in the brain of another person. The brain of another creature could be the host to certain parts of a person’s *soul*. A dog could host parts of a human *soul* that might be recognizable as reflecting parts of that person’s personality.

   To say these things are possible is not to say we know how to accomplish them. What we are saying is that the laws of physics do not stand in the way -- therefore it's a matter of engineering and technology to accomplish what is clearly possible. We discuss possible steps that might allow the capture of the soul of a living person with no ill effects to that person in XXXX.

4. Can the *soul* exist separate from the body?

   The answer is "Yes." This, again, is a matter of technology. The technology to allow some parts of a *soul* to exist separate from a body has existed for thousands of years, and that technology has grown continuously up to the present. Today, records and movies tell us more about the *soul* of a contemporary but deceased singer than what we could experience in the past when all we had was narrative descriptions. Tomorrow, someday, technology will have advanced to the stage where instead of passively looking at a movie, we will be able to visit with holographic images or with robots that clearly prove that a computer system has captured the nearly complete essence of a particular human *soul*. We believe that a contemporary high end personal computer can have the computational resources to host the dynamic soul of one person in real time. This might require the addition of a few special purpose cards, similar in technology to very high performance display cards.

5. If there is a human *soul*, what kind of thing is it?

   The human *soul* is an informational construct. It can exist in two forms, static and dynamic. We break down the human *soul* into three main components: P, Q and W. We can understand the kind of thing that a *soul* is by understanding the kind of thing that a computer program is. This is similar to understanding metabolism by understanding a steam engine. It is similar to understanding certain functions of organs in the body by understanding the functions of a chemical plant. A human *soul* has various parts. Some are dynamic, functional parts, and some are fragments that are static. When one writes, paints or records a song, static fragments of that person's *soul* are revealed in the record that is preserved. The life's works of a prolific writer or musician are insufficient to reconstruct his or her entire *soul* but they do serve to narrow the possibilities. When one person understands and absorbs and performs the work of another person, such as when a musician studies and then plays a Mozart piano sonata, then a fragment of Mozart's *soul* becomes dynamic again.

6. What is the relationship between the *soul* and the body?

   The *soul* is an informational construct that is made dynamic by the living body that serves as its host. The body, which includes the brain, is a necessary part of the definition of the *soul*. A particular body, differentiated from the generic body, is not just the host, but is reflected in the intellectual workings of the *soul*. A *soul* cannot be completely understood without understanding its host body and the environment that encompasses the body. While the *soul* of a deceased person, made dynamic through the workings of a computer host, cannot function in the exact manner that it did while the person was alive, it can be close, and surely recognizable. As the technologies required for this are developed, souls will be hosted with better and higher fidelity.

7. Does a human's *soul* have existence before birth or conception?

   In a word, yes. A dynamic *soul* grows and learns and flowers into maturity. Our definition of the *soul* encompasses the entirety of the informational construct. This implies that the part that exists before birth or conception is but a fraction of the mature *soul*. The part that exists before conception has many avenues into the nascent *soul*: through DNA, environment, nurture, education and socialization. Of course, in addition to the DNA, one's parents normally have a great impact on the development of a *soul*. Similarly with one's teachers. When you first hear a memorable new melody, a small part of the composer's *soul* becomes a part of your *soul*. A tremendous amount of information from the past can be incorporated into the *soul*. From the informational viewpoint it is correct to say that many fragments of an adult *soul* existed before conception, and were incorporated into that adult *soul* starting at conception and continuing over its lifetime. A measure of a man's or woman's life and of the progress of society may be the answer to the following question: "What soulecules from this person's *soul* survived his or her death and how many other *souls* have incorporated some of those soulecules?"

8. After death, might compensation for the inequities of life be visited on the *soul*?

   While according to EB, the "...constant prosperity of the wicked, and the frequent unhappiness of good men in this life, must convince every thinking person, that there is a future state wherein all will be set right, and God's attributes of wisdom, justice, and goodness, fully vindicated...", there is no evidence that this has been happening. What is quite remarkable is that it may yet happen; and happen in a way even superior to what was imagined.

A Disembodied Soul Resurrected Inside a Computer

If you are about to die sometime in the future when technology has advanced, it might be possible for almost all of your *soul* to be captured and preserved. It could then be made dynamic again by being put into a suitable computer; *a computerized soul*. We will explore what it might feel like to be a computerized *soul*. In the more distant future, in addition to restoring your *soul*, you might be given a robotic body that would appear very much like your own body (perhaps when you were younger!). Prior to the time of full robotic hosts, you might just have a sort of bedridden body whose only mechanical functions were limited to facial movements (eye, mouth and expressions) much like a severe paraplegic. In yet earlier times, (or, at less cost) you might just be an image of yourself on a computer screen, talking through a loudspeaker, seeing with a TV camera and hearing with a microphone.

There are two major possibilities to consider with respect to becoming a computerized *soul*. First, that the purpose is to avoid the loss of your *soul* to both yourself and the rest of mankind and to keep it functioning and active. This would mean that your *soul* would be able to conduct discourse with normal living beings and with other computerized *souls*. The second possibility would be, in essence, to be able to go to heaven. These two possibilities are not mutually exclusive. A *soul* in heaven could be allowed some form of communication with *souls* not in heaven.

Let us assume that sometime in the future you are near biological death, and your *soul* is to be computerized. It would be necessary to do quite a bit to get ready. Your DNA would be analyzed in order to help with the task of determining P, the nature of some aspects of your brain-body system that interprets Q. At a point when your brain and mind are still intact but your body is near biological death, (or perhaps at some earlier time) the process of *soul* extraction would begin. You might be put to sleep as though in preparation for major surgery. The process of computerization would might involve surgery and microsurgery on the brain, but then again it might be possible to develop non-invasive technologies such as MRI. Exactly how this could be done is not known today, but it would be similar to the problem of looking into a working RAM chip to read out the state of the bits stored in it and looking into a particular processor chip to figure out exactly how it processes data. It may also be possible to get the brain to recall many memories that could be picked up as they were transferred from one part of the brain to another. Such a method would be far too slow for the capture of data, but could give some clues as to the informational organization of the data. Most likely, some combination of these and other methods would be employed.

The information that defines P, obtained from both DNA and somehow from the brain would need to be tested and debugged. The information that defines Q, obtained from DNA, the brain and more evanescent state information would be recorded and saved. It would be entered into a computer capable of interpreting P, which would then be used to interpret Q. A great deal of testing and debugging might be required to obtain a satisfactory result. Most of this process could be automatic as the totality of the reborn *soul* comes into operation. Many questions and answers would be given to the newly reborn computerized *soul* to ensure the correctness of the process. Once all had been worked out, the computerized *soul* would be restored to its state at the time of biological death, and then brought directly into its more complete and correct state in order to avoid having the intensive process of getting it into good working order adversely affect it.

In order to not overly complicate the description of what would happen when your *soul* awoke in a new host, we will describe a single possible experience, instead of trying to explain the variety of possible experiences. Once your computerized *soul* is transferred into a proper host computer, and set into operation, it would be ready to wake up. If the process is done well, it would wake up feeling much like itself. As it come back to consciousness, it might hear sounds or conversation. When it *opened its eyes*, it would feel its eyelids opening and it would be conscious of various other bodily feelings. All these feeling would probably be coming from programs designed to provide them. A doctor might ask “Good morning, how do you feel?” He would judge the degree of success of the process from the reactions he observed. There would most likely be a continuing need, for some period of time, to make various adjustments in P and Q in order to obtain the best results. Depending upon the state of the art in robotics, the computerized *soul* would get various kinds of differing reactions from attempts to move parts of its body. Until the state of the art of *soul* computerization reached a very high level, it would no doubt notice many differences, but they would be things that would either be quickly adjusted to or complained about (for the doctor to fix). For example, if the *soul* had poor vision in its prior incarnation and had needed glasses, its vision could be fixed. It would not make sense to change too much at once, so vision improvement would probably be done in a number of small steps as it become ready and used to all the other changes. If it was blind from birth, it would be important to introduce vision in a careful and controlled manner because of all the other simultaneous changes. It would probably find that some sounds, such as people's voices, were somehow different at first. It might not recognize someone's voice the first time it spoke to them on the phone. It would also take time to adjust to the various limitations imposed by the state of robotic technologies.

Not every difference would be negative (or a loss). For example, when a phone rang, if the *soul* so desired, it could simply become aware of who was calling and think whether to answer or not. If it chose to answer, it could use real sound or it could, in a thought, opt for ESP conversation, where its unvocalized mental version of what it wanted to say would be transmitted over the phone in its voice. The *soul* could hear the other party without recourse to the need for sound. Of course, the private thoughts that it was having during the conversation would not be transmitted over the phone. There would be many possible enhancements that would require intelligent evaluation. The question is, are various changes and so-called improvements actually a good thing? It is possible that *soul* might be, in a sense, corrupted merely to partake of things that never could have been a part of it while it was alive.

As to the question as to whether or not a computerized *soul* would be conscious, what can be said is that if you asked a computerized *soul* whether or not he was conscious, he or she would reply "Definitely!" If you asked a jury to determine whether or not they thought a computerized *soul* was conscious by engaging it in conversations, they would also reply "Definitely!"

Heaven, a Place That Is Heavenly to Everyone In Heaven

Let us consider a computerized *soul* that would rather go to heaven than to continue here on Earth. We shall define "heaven" as that place that is heavenly to everyone in heaven. It should be that place where all inequities are corrected, all sins forgiven, and where one can exist fulfilled and triumphant. This will be very possible in the following way. Imagine a very large computer charged with the task of computing what might be called "General Heaven." This is the large set of things that are mostly in common to most inhabitants of heaven. Then, imagine that the computational resources given to a *soul* in heaven is quite a bit greater than needed for just that *soul*; perhaps 1,000 times greater. The added computational power would be used to recast the actions and personalities of other heavenly *souls*, and to modify other heavenly aspects, such as weather, nature of surroundings, etc. The concept would allow two heavenly *souls*, A and B, who must each beat the other at tennis at least 3/4 of the time to both be in a heaven where it is true that each of them does win 3/4 of the time, and all aspects of that inconsistency are smoothed out by the added computational power of the process that computes Q and W. Thus, in A's heaven, B is very much the same as B is in B's heaven, except for the fact that A wins 3/4 of the time. In B's heaven, A is very much the same as A is in A's heaven, except that it is B that wins 3/4 of the time. All inconsistencies that arise as a consequence are smoothed out for both immortal *souls*. And of course, it is arranged that neither can know what happens in the other's heaven.

Immortality

We often speak of "...our immortal *soul*..". This gives rise to the question of what "immortality" means. We cannot know how long the Universe will continue in existence, so a definition of "immortality" that makes reference to infinite time does not seem to make sense. A better definition of immortality would refer to the idea that the thing which is immortal could exist with no degradation in its essential characteristics so long as its environmental parameters remained reasonable. If the Sun becomes a supernova, then things on Earth that had been immortal up to that point then might cease to exist. Those things transferred to another region of space, with a more stable star system, could certainly survive the end of the Earth. If the Big Crunch occurs and the Universe collapses back to an atom like object of very high density, then it seems likely that all things thought to be immortal up to that point, would then lose their existence. Within those constraints, we can explore the concept of immortality.

What we know is that high energy, in the forms of high temperatures, high energy particle collisions, massive (meteoric) collisions and strong force fields such as electrical or gravitational tend to act in ways that limit the mechanical stability of things made of atoms. As a result, things made of atoms can only be expected to be immortal under benign environmental conditions. However, some parts of the Earth have been reasonably benign for billions of years. Thus a carefully designed, subterranean networked computer complex, that was geographically distributed, perhaps partly in lunar caverns, could expect to continue to function for billions of years. Massive volcanic activity such as that which produced the Deccan Traps in India, could wipe out all computer facilities in that area, but when something bad is happening somewhere, there have always been other parts of the Earth safe from disaster. Thus, it is reasonable to assume that some future generation should be able to find ways to implement computer systems able to operate autonomously with very long expected lifetimes, despite damage do to various local disasters and other sources of degradations. This might involve having computer systems that can make and geographically distribute additional computer systems. The Moon might be the ideal local for such futuristic systems.

Atoms come in two kinds, intrinsically stable and unstable. By stable, we mean that the expected lifetime is very long compared to the age of the universe. A good example would be an He^3 (Helium 3) atom. On the other hand, an H^3 (Hydrogen 3 or Tritium) atom is not stable, it decays with a half life of about 15 years. Atoms are little machines that are always working but never wear out. No matter how old an atom is, it is indistinguishable from every other atom with the same atomic number, weight and state. Stable atoms are near perfect examples of non-trivial machines that are immortal in the sense of never growing old and never wearing out. If one is desirous of creating an immortal thing out of ordinary matter, then it should be made up of stable atoms. When we combine atoms into molecules, then the energy needed to break apart such structures is very much less than that needed to break apart an atom. Thus, the environmental conditions required by immortal molecular structures is very much more constrained then that required by atoms. Subatomic particles are even more resistant to environmental conditions than are atoms. Neutrons can survive under conditions of great pressure where atoms cannot. However, a neutron cannot survive for long when it is isolated. The half life of an isolated neutron is about 15 minutes!

An example of an apparently immortal living thing is a one celled bacterium that divides as a means of reproduction. One could claim that a species is immortal since, despite the fact that every individual member of the species dies at the end of its lifetime, the species has no intrinsic limit on its lifetime. While it is easy to contemplate ordinary matter as components of an immortal being, there is a problem concerned with the motion of the parts of such a being. While a rock may be very much the same as it was 4 billion years ago, it is doubtful that a diesel engine will be the same after 4 billion years of operation. The reason is that local components, internal to the engine, have from time to time energies sufficient to break molecular bonds and thus to alter the system in ways that must be thought of as decay.

On the other hand, a computer may have no moving parts other than electrons or photons, and all operations may take place at energies so low as to allow for billions of years of operations with little chance of internally caused decay. Such a machine could have ways to minimize, compensate for and correct errors introduced by such external events as cosmic rays. It is possible and practical to build such a computer, powered by starlight, or by the decay of some Uranium U^{238} which has a half life of 4 billion years. Such a computer, launched into interstellar space, could have an expected lifetime of billions of years. It could be the host of a *soul* or of millions of *souls*, all of whom could then be said to be immortal.

There is one kind of mechanism that never has any necessary wear, decay or energy dissipation and that is an informational process. Just as the continual workings and motions of an atom are not dependent upon a continual supply of energy which must be dissipated, the workings of an informational process similarly are not necessarily dependant on a supply of energy which must be dissipated.

A properly designed computer, once put into operation, could continue to compute for billions of years, just as the solar system, once put into motion, could continue in motion for billions of years. In the case of a solar system, we know that there are various mechanisms slowly robbing it of energy, such as tidal effects or gravitational radiation. These effects can be so small as to allow a solar system to exist and function for many times the current age of the universe.

There is a methodology for designing computers, whose operation is very similar to modern computers, but where the necessary energy dissipation is reduced arbitrarily close to zero. This is much like the fact that we can build wires out of superconductors where the necessary energy dissipation due to the resistance of wires is reduced to zero. Such computers can be designed using Conservative Logic. There are two charming myths about reversible computers: first, that such computers, after finishing a computation, must be put into reverse to run the entire computation backwards to the beginning in order to not dissipate any energy and second, that such computers must run very slowly. These myths originate from a seminal paper by Charles Bennett where he showed that one could make a mechanical Turing Machine dissipate zero energy by such processes. He pointed out that his approach was sufficient but he never suggested that his approach was necessary! Andrew Ressler, gave an example of the detailed logic design of a dissipationless computer that was in every essential way very similar to ordinary computers, and it certainly did not have to run itself backwards from the end of the computation back to the beginning in order to not dissipate energy. On the contrary, one must be sure that the computer could, in principle, run backwards from the end of the computation back to the beginning as a form of proof that it is properly designed. There is, however, no necessity to run the whole computation backwards in order to have the system completely dissipationless.

If such a computer is effectively insulated from external energetic processes, it is not clear that there is any impediment to its continual operation based on the laws of physics; so long as its environment is reasonably benign. In addition, a complex computer system can have many mechanisms at work that allow for the correction of minor errors. Each such correction might best involve the dissipation of a very small amount of energy. In a well designed reversible computer, a few pounds of radioactive matter could provide all the power needed for billions of years of correcting whatever errors might occur. This means that errors introduced by, for example, an energetic cosmic ray, would not lead to a growth in erroneous computation, but rather the consequent errors would be detected and corrected without ever affecting the overall computational process other than by delaying things for a tiny amount of time.

Thus a *soul* could be essentially immortal. Since a *soul* is an informational construct, it has no parts to wear out or to get old. It is easy to see how a static *soul* can be immortal; it is written out on some permanent media, such as a future kind of CD-ROM and then stored in a safe place. A dynamic *soul* requires a host, but it can also exist as a static *soul* sequestered in some safe place. Just as one does a backup of a hard drive, so that after a catastrophic loss of a hard-drive one can buy and install a new hard drive and recover most of the data from the old hard-drive

For a *soul* to be immortal, it must either be resident in an immortal host, or it must be able to survive the mortality of each member of a sequence of hosts. Since the *soul* is an informational construct…. ….Of course, it all depends on the temperature….

To be continued and finished sometime soon… Ed Fredkin

Appendix A

Imagine a contemporary (2000) multimedia (audio-visual) personal computer system. We will call the system AV-1. Such a system consists of a computer containing a processor, various memories, hardware support for audio inputs and outputs, and video inputs and outputs including MPEG-2. AV-1 has Input-Output devices such as a keyboard, mouse, printer, scanner, modem, speakers, microphones, video cameras and various internal sensors. This is very much like some kind of creature that happens to be immobile and without a *soul*. Given the proper software it can see (capture and analyze images). A computer can read music and it can play that music approximately as though performed by any instrument or ensemble of instruments by synthesizing the equivalent sounds. A computer can be made to *imagine* creatures, common or strange, running around and show us what it is imagining on its display monitor, as was done in the movies *Star Wars, The Phantom Menace* and *Jurassic Park*. It can hear by capturing and recording sounds from its microphone and by performing speech recognition and music recognition. It can save and recall information from a database. Of course, in the year 2000 all of these computer capabilities are still very primitive. However, what a computer can do becomes more and more extensive with every passing year. Many of the things done routinely today, such as speech recognition or character recognition, were once thought far too difficult for an inexpensive computer.

It is possible to give an informational description of the hardware that is complete and exact, given that we only care about the informational point of view of the hardware. For example, an informational description of the processor is basically an interpreter program, SAV-1 written for some standard computer that can faithfully run any program written for AV-1. For the sake of being definite, we will call the standard computer that runs all simulations “Phi”. While absolutely any computer (that had enough memory) would do, there is a great advantage in using a concrete example so we arbitrarily pick an ordinary PC.

Imagine an AV-1 program speaking (converting text to speech), using loudspeakers and listening to its own speech using microphones (and processing the sound with speech recognition software to recover the original text). We could say that AV-1 is aware of the fact that it is speaking. On the other hand AV-1 could simulate all the transformations that occur starting from the digital textual information representing the words it sent to the speakers to the digital textual information representing the words heard from the microphones. Let’s look at all the steps in a sophisticated end to end simulation of the process:

1. Distortion and errors in the speech synthesis program that converts ASCII text, and other information regarding the emotional content of what is being said and that modulates the speech sounds, into a digital representation of the sound waveform, then
2. Distortion in the conversion from digital numbers which represent a sound waveform into electrical signals sent to a speaker, then
3. Distortion in the conversion of the electrical signals into sound waves produced by the speakers, then
4. Distortions that occur as the sound wave travels to the microphone and those due to extraneous external sounds picked up by the microphone
5. Distortions in the conversion by the microphone into electrical signals

6. Distortions in the conversion back into digital numbers
7. Finally, the processing of the digital numbers by a speech recognition program back into, hopefully, the same ASCII text along with the recovery of the emotional information that modulated the speech sounds.

Nevertheless, while not exact, such a simulation can still be faithful to high level interpretations. When doing this, AV-1 can be said to be consciously thinking about the audio aspects of what it is about to say. A similar process could model the reactions of a person who hears what AV-1 says, who thinks about it and then generates one of a number of plausible replies. If AV-1 is doing all these things, including analyzing high level consequences of other facts that come into its sensory apparatus while it is speaking to a nearby person, it can be said that AV-1 is conscious of what it is doing. Consciousness need not be a mystery.

The technology of simulating one computer with another is such that one can be confident that the results of the simulated computer are, aside from timing, substantively identical to the results of the original computer for any non-pathological program. A pathological program is one that is unnecessarily sensitive to the timing vagaries of various devices such as disk-drive seek times.

Let us imagine that, somehow, we have captured a static *soul* on a magnetic tape. This means that we have the informational content of Q. We also need P or at least a good approximation to P. Even if we had the Q part of a human *soul* recorded on a digital tape, we do not yet know how to program an interpreter for P. It is simply the case that we do not yet have sufficient knowledge and understanding as to what is required; there is no intrinsic reason that will prevent us from being able to do so in the future. This would mean that a computer such as Phi, could be engaged in a conversation wherein those conversing with it would feel that the essence of the captured *soul* had come to life.

To be very precise, consider P to be an interpreter program that runs in Phi. When Phi runs P, that interpreter processes the current state of Q in order to produce the future state of Q. In other words, the operation of P in Phi causes the temporal evolution of Q. If P and Q represent the informational aspects of a *soul*, then the operation of P in Phi which causes Q to evolve is the process by which a static *soul* becomes dynamic (which means that the *soul* can be said to be thinking and acting). Strangely enough, while we need to know more than we do to enable a disembodied *soul* to think, we do not have to understand how it is that that *soul* can think. We do not have to understand thinking or how to create thinking de nova. We can clone thousands of copies of a single tree from cells of one tree while it is not necessary for us to be able to create a tree from basic chemical synthesis. I see no need to imagine that we will be forever unable to know how to create a tree from scratch or how to create something that thinks from scratch. It's that much can be done with partial knowledge.

It is not necessary for the disembodied *soul* to be cut off from the real world and the people in the real world. With video input, with audio input and output, a disembodied *soul* operating in Phi will be able to see and recognize people and carry on normal conversations with them. In such a conversation, if you were to ask the disembodied *soul* whether or not he or she felt that he or she existed, the answer might well be that Descartes' Cogito applies. If the *soul* can think, the *soul* is.

The disembodied *soul* can be preserved for eternity in the state corresponding to its last living moment, and/or additionally it can be allowed to evolve as time goes by.

The PC Analogy

Say we observe an actively used PC (Personal Computer). Aside from use in the normal fashion, it might also be working simultaneously as a network server and as a bulletin board system. If we wish to characterize what we see and wish to understand it in general terms, we can divide it into the following entities:

1. The physical computer -- the microprocessor, the random access memories (RAM, ROM, and cache), the hard disks, floppy drive, modems, network interfaces, backup tape drive, CD, etc.
2. The software and data that are available to the PC -- this would include the operating system (e.g. Window NT), utilities, applications, games, network, BBS, etc.
3. The nature of what the PC system is connected to -- power, the network, the Internet, phone lines, ISDN, keyboard, mouse, monitor, audio, video, and indirectly other computers and various people. Here, we are more interested in the nature of the information exchanged rather than with hardware details.

Item 1 has a precise definition in terms of the static information that defines an interpreter that can exactly mimic the operation of the PC system. In a formal sense, the interpreter could be a very simple Turing machine. It would make more sense to define it in terms of a very simple canonical computer that is organized much like a simple RISC architecture computer.

Appendix B

The John Cocke Theory of Dreams

There are such things as optimal methods of encoding. In a computer, a fragment of textual information is usually represented by one 8 bit byte for each ASCII character. More modern computer systems use 16 bits (2 bytes) for each Unicode character (which might be a character from any one of a great deal of human languages). Thus, a short story made up of a few thousand words might take 6 (characters per word) times 8 bits per character times 3,000 words equals 144,000 bits. There are various methods of encoding that reduce the number of bits needed to represent such a story. Encoding schemes fall into 2 broad classes: lossless and lossy. A message encoded by a lossless encoding scheme can always be reconstructed exactly. A message encoded by a lossy encoding scheme can only be reconstructed approximately. Lossy encoding schemes are commonly used to encode pictures (JPEG or MPEG) or speech or music (MP3). An example of lossy encoding for textual messages could involve the loss of all information as regards to font, formatting and capitalization while some words are replaced by synonyms and some word order might be swapped. Further, a really intelligent lossy encoding scheme might encode a story into the same story as told by another person quoting the story from memory (but not memorized verbatim), as opposed to a verbatim copy.

The idea is to reduce the number of bits required as much as possible. Let us imagine that somehow, we have at our disposal an optimal lossy encoding scheme. The encoded text might require only a few percent of the bits used in a Unicode representation. Normally, the encoded text would resemble a random string of bits. In fact it should be able to pass a test for randomness. The encoding method would undoubtedly involve various global steps such as defining the context, the time frame, the participants, and the nature of the text. Fragments of the text might already be in memory. For example, if we are encoding a fairy tale, the beginning of the encoded version might correspond to a more compressed version of the following: "*fairy tale*, *s1s…*" Given that it's a fairy tale, s1s might stand for "standard 1st sentence fragment" meaning in this case "Once upon a time…". As in a play, the characters and some things about them could be listed at the beginning along with the shorthand nicknames that would refer to them in the encoded version.

The reason that an optimally encoded story would look like a random sequence of bits is that any lack of randomness would usually mean a less than optimal encoding. When reconstructing the original from a lossy, optimally encoded story, the meaning of the bits in the middle of the story is totally dependent on information from earlier in the story. What is fascinating about a truly optimal story encoding scheme is that the decoding of every possible random sequence of input bits must yield a reasonable story! If not, then it wasn't optimal encoding. Of course, "reasonable" allows for improbable. In any case, the decoded story would have continuity. In the middle of decoding a story, the next events would almost always have continuity with the prior events as a natural consequence of an optimal encoding for a person's memories of various time intervals. The obvious thing is that the inputs to the optimal decoder are multiple. They include sensory inputs, and various sources of memories. This just makes sense in the design of a mental process that would help ensure the survival of any creature with a brain, such as a person or a dog.

The reason that the decoder needs to be awoken first whenever any sensory input occurs is obvious: The creature needs to take the sensory input to memory and retrieve information that might be vital to survival. If the sounds are those of a predator sneaking up, the information retrieved has to signal the creature to wake up and flee or fight. Thus there is good reason for the decoder to both control the level of consciousness and feed information to the higher levels of semi-consciousness or consciousness.

Now, let us assume that a person or a dog is asleep and in a state when a dream is likely and that the machinery for optimal decoding is operational. But, there is nothing to drive the input other than random noise combined with the following: some sensory input such as the distant sound of a dog barking and the sensation of being too warm. A dream starts up in a human. He is at a barbecue, it's a hot sunny day, children are playing with a dog that is barking. The scene and events are familiar but evolve down random choices of reasonable paths; the human is having a dream. For the dog, her dream might be of being in an overheated house and, hearing the bark of the neighbor's dog, wanting to run outside. The dog tries to run to see if the back door is open, and while dreaming, the dog's legs actually go through abbreviated running motions and little muffled barking sounds are actually made as the dog tries, in her dream, to attract someone who might open the door.

The John Cocke Theory of Dreams was told to me, on the phone, late one night back in the early 1960's. John's complete description was contained in a very short conversation approximately as follows:

"Hey Ed. You know about optimal encoding, right?"

"Yup."

"Say the way we remember things is using a lossy optimal encoding scheme; you'd get efficient use of memory, huh?"

"Uh huh."

"Well the decoding could take into account recent memories and sensory inputs, like sounds being heard, right?"

"Sure!"

"Well, if when you're asleep, the decoder is decoding random bits (digital noise) mixed in with a few sensory inputs and taking into account recent memories and stuff like that, the output of the decoder would be a dream; huh?"

I was stunned.