Chapter 1

Bits

Information is measured in bits, just as length is measured in meters and time is measured in seconds. Of course knowing the amount of information is not the same as knowing the information itself, what it means, or what it implies. In these notes we will not consider the content or meaning of information, just the quantity.

Different scales of length are needed in different circumstances. Sometimes we want to measure length in kilometers, sometimes in inches, and sometimes in Ångströms. Similarly, other scales for information besides bits are sometimes needed; in the context of physical systems information is often measured in Joules per Kelvin.

How is information quantified? Consider a situation that could have any of several possible outcomes. An example might be flipping a coin (2 outcomes, heads or tails) or selecting a card from a deck of playing cards (52 possible outcomes). How compactly could one person (by convention usually named Alice) tell another person (Bob) this outcome?

First consider the case of the two outcomes of flipping a coin, and let us suppose they are equally likely. If Alice wants to tell Bob the result of the coin toss, she could use several possible techniques, but they are all equivalent, in terms of the amount of information conveyed, to saying either "heads" or "tails" or, for that matter, to saying 0 or 1. We say that the information so conveyed is one bit.

If Alice flipped two coins, she could say which of the four possible outcomes actually happened, by saying 0 or 1 twice. Similarly, the result of an experiment with eight equally likely outcomes could be conveyed with three bits, and more generally 2^n outcomes with n bits. Thus the amount of information is the logarithm (to the base 2) of the number of equally likely outcomes.

Note that conveying information requires two phases. First is the "setup" phase, in which Alice and Bob agree on what they will communicate about, and exactly what each sequence of bits means. This common understanding is called the code. This is done before the outcome is known. Then, there is the "communication" phase, where actual sequences of 0 and 1 are sent. These sequences are the data. Thus to convey the suit of a card chosen from a deck, their code might be that 00 means clubs, 01 diamonds, 10 hearts, and 11 spades. Using that code, Alice draws the card, and tells Bob the suit by sending two bits of data. She could do so repeatedly for multiple experiments, using the same code.

After Bob knows that a card is drawn but before receiving Alice's message, he is uncertain about the suit. His uncertainty, or lack of information, can also be expressed in bits. Upon hearing the result, his uncertainty is reduced by the information he receives. Bob's uncertainty rises during the setup phase and then is reduced during the communication phase.

Note some important things about information, some of which are illustrated in this example.

- Information can be learned through observation, experiment, or measurement.

Author: Paul Penfield, Jr.

- Information is subjective, or "observer-dependent." What Alice knows is different from what Bob knows. (If information were not subjective, there would be no need to communicate it.)
- A person's uncertainty can be increased upon learning that there is an observation about which information may be available, and then can be reduced by receiving that information.
- Information can be lost, either through loss of the data itself, or through loss of the code.
- The physical form of information is localized in space and time. As a consequence ...
  - Information can be sent from one place to another.
  - Information can be stored and then retrieved later.

1.1 The Mathematical Bit

As we have seen, information can be communicated by sequences of 0 and 1 values. This very powerful abstraction lets us ignore many of the details associated with specific information processing and transmission systems.

Bits are simple, having only two possible values, and the mathematics used to manipulate single bits is not difficult. It is known as Boolean algebra, after the mathematician George Boole (1815 - 1864). In some ways it is similar to the algebra of integers or real numbers which is taught in high school, but in some ways it is different.

Algebras deal with variables that have certain possible values, and with functions which, when presented with one or more variables, return a result which again has certain possible values. In the case of Boolean algebra, the possible values are 0 and 1.

There are exactly four Boolean functions of a single variable. One of them, called the identity, simply returns its argument. Another, called not, or negation, or inversion, or complement, changes 0 into 1 and vice versa. The other two simply return either 0 or 1 regardless of the argument. The easiest way to describe these functions is to simply give a table with their results:

x	$f(x)$			
Argument	$IDENTITY$	NOT	$ZERO$	ONE
0	0	1	0	1
1	1	0	0	1

Table 1.1: Boolean functions of a single variable

Note that Boolean algebra is much simpler than algebra dealing with integers or real numbers, each of which has infinitely many functions of a single variable.

How many Boolean functions are there of two variables A and B? Each of the two arguments can take on either of two values, so there are four possible input combinations. There are 16 different ways of assigning the two Boolean values to four inputs. Of these 16, two simply ignore the input, four assign the output to be either A or B or their complement, and the other ten depend on both arguments. The most often used are AND, OR, XOR, $NAND$, and NOR, as shown in Table 1.2.

It is tempting to think of the Boolean values 0 and 1 as the same as the integers 0 and 1. Then AND would correspond to multiplication and OR to addition, sort of. However, familiar results from ordinary algebra simply do not hold for Boolean algebra, so such analogies are dangerous. It is absolutely necessary, though sometimes difficult, to distinguish the integers from the Boolean values.

This task is made more difficult by the standard notation used for Boolean algebra. (We will use this notation here, even though it can be confusing, because less confusing notations are awkward in practice.) The AND function is represented the same way as multiplication, by writing two boolean values next to each

x	$f(x)$				
Argument	AND	$NAND$	OR	NOR	XOR
00	0	1	0	1	0
01	0	1	1	0	1
10	0	1	1	0	1
11	1	0	1	0	0

Table 1.2: Boolean functions of a two variables

other with a dot in between: $A\ AND\ B$ is written $A \cdot B$. The OR function is written using the plus sign: $A + B$ means $A\ OR\ B$. Negation, or the NOT function, is denoted by a bar over the symbol or expression, so $NOT\ A$ is $\overline{A}$. Finally, the exclusive-or function XOR is represented by a circle with a plus sign inside, $A \oplus B$.

NOT	$\overline{A}$
AND	$A \cdot B$
$NAND$	$\overline{A \cdot B}$
OR	$A + B$
XOR	$A \oplus B$

Table 1.3: Boolean logic symbols

There are several general properties of Boolean functions that are useful. These can usually be proven by simply demonstrating them for all of the finite number of possible combinations of values.

A function is said to be **reversible** if, knowing the output, the input can be found. Two of the four functions of a single variable are reversible in this sense (and in fact are self-inverse). Clearly none of the functions of two (or more) inputs can by themselves be reversible, since there are more input variables than output variables, but some combinations of two or more such functions can be reversible; for example it is easily demonstrated that the exclusive-or function $A \oplus B$ is reversible when augmented by the function that returns the first argument.

For functions of two variables, there are many properties to consider. For example, if A, B, and C are Boolean variables, able to be either 0 or 1, then the function AND is **commutative** because $A \cdot B = B \cdot A$. Many of the other 16 functions are also commutative. Some properties of Boolean algebra are summarized in Table 1.4.

There are several notations used for Boolean algebra. The one used here is the most common. Sometimes AND, OR, and NOT are represented in the form $AND(A, B)$, $OR(A, B)$, and $NOT(A)$. Sometimes infix notation is used where $A \wedge B$ denotes $A \cdot B$, $A \vee B$ denotes $A + B$, and $\sim A$ denotes $\overline{A}$. Boolean algebra is also useful in mathematical logic, where the notation $A \wedge B$ for $A \cdot B$, $A \vee B$ for $A + B$, and $\neg A$ for $\overline{A}$ is commonly used.

1.2 The Physical Bit

If information is to be stored or transported, it must have a physical form. The device that stores the bit must have two distinct states, one of which is interpreted as 0 and the other as 1. A bit is stored by putting the device in one or another of these states, and when the bit is needed the state of the device is measured. If the device has moved from one place to another then communications has occurred. If the device has persisted over some time then it has served as a memory. If the device has had its state changed in a random way then it has forgotten its original value.

We are naturally interested in physical devices which are small. The limit to how small an object can be and still store a bit of information comes from quantum mechanics. A quantum bit, or qubit, is an object

Idempotent:	$A \cdot A = A$ $A + A = A$	Absorption:	$A \cdot (A + B) = A$ $A + (A \cdot B) = A$
Complementary:	$A \cdot \overline{A} = 0$ $A + \overline{A} = 1$ $A \oplus A = 0$ $A \oplus \overline{A} = 1$	Associative:	$A \cdot (B \cdot C) = (A \cdot B) \cdot C$ $A + (B + C) = (A + B) + C$ $A \oplus (B \oplus C) = (A \oplus B) \oplus C$
Minimum:	$A \cdot 1 = A$ $A \cdot 0 = 0$	Unnamed Theorem:	$A \cdot (\overline{A} + B) = A \cdot B$ $A + (\overline{A} \cdot B) = A + B$
Maximum:	$A + 0 = A$ $A + 1 = 1$	De Morgan:	$\overline{A} \cdot \overline{B} = \overline{A + B}$ $\overline{A} + \overline{B} = \overline{A \cdot B}$
Commutative:	$A \cdot B = B \cdot A$ $A + B = B + A$ $A \oplus B = B \oplus A$ $\overline{A \cdot B} = \overline{B \cdot A}$ $\overline{A + B} = \overline{B + A}$	Distributive:	$A \cdot (B + C) = (A \cdot B) + (A \cdot C)$ $A + (B \cdot C) = (A + B) \cdot (A + C)$

Table 1.4: Properties of Boolean Algebra

that can store a single bit but is so small that it is subject to the limitations quantum mechanics places on measurements. In particular, a measurement cannot be made without possibly altering the object being measured. On the other hand, if a bit is represented by many objects acting together, a measurement can be made and enough objects left unchanged so that the same bit can be measured again. And if not all objects measure in the same way the result might be intermediate between the two possible Boolean values.

1.3 The Classical Bit

In today's electronic systems, information is carried by many thousands of objects, all prepared in the same way (or at least that is a convenient way to look at it). Thus a semiconductor memory stores a single bit using the presence or absence of perhaps a thousand electrons. Similarly, a large number of photons are used in radio communication.

Because many objects are involved, measurements on them are not restricted to a simple yes or no, but instead can range over a continuum of values. Thus the voltage on a semiconductor logic element might be anywhere in a range from, say, $0V$ to $5V$ The voltage might be interpreted to allow a margin of error, so that voltages between $0V$ and $1V$ would represent logical 0, and voltages between $4V$ and $5v$ a logical 1. The circuitry would not guarantee to interpret voltages between $1V$ and $4V$ properly.

If the noise in a circuit is always smaller than $1V$, and the output of every circuit gate is either $0V$ or $5V$, then the voltages can always be interpreted as bits without error. Circuits of this sort display what is known as "restoring logic" since small deviations are eliminated as the information is processed. The robustness of modern computers depends on the use of restoring logic.

A classical bit is an abstraction in which the bit can be measured without perturbing it. As a result copies of a classical bit can be made (a qubit cannot be copied). This abstraction works well for circuits using restoring logic.

Because all physical systems ultimately obey quantum mechanics, the classical bit is always an idealized approximation to reality. However, even with the most modern, smallest devices available, it is an excellent one.

An interesting question is whether the classical bit abstraction will continue to be useful as semiconductor technology reduces the size of components. Ultimately, as we try to control bits with a small number of atoms or photons, the limiting role of quantum mechanics will become important. It is difficult to say exactly when this will happen, but some people believe it will be before the year 2015.

1.4 The Quantum Bit

According to quantum mechanics, it is possible for a small object to have two states which can be measured. This sounds perfect for storing bits. However, if these two states have the same energy associated with them, then it is possible to prepare the object so that it has a combination of these two states. So what is it that would be measured?

In a classical context, a measurement could determine exactly what that combination is. Furthermore, for greater precision a measurement could be repeated, and multiple results averaged. However, the quantum context is different. In a quantum measurement, the question that is asked is whether the object is or is not in some particular state, and the answer is always either "yes" or "no," never "maybe" and never, for example, "27% yes, 73% no." Furthermore, after the measurement the system ends up in the state corresponding to the answer, so further measurements would not yield additional information. The result of any particular measurement cannot be predicted, but the likelihood of the two answers can, in terms of probabilities. This peculiar nature of quantum mechanics offers both a limitation of how much information can be carried by a single qubit, and also an opportunity to design systems which take special advantage of these features.

We will illustrate quantum bits with an example. Let's take as our qubit a photon, which is the elementary particle for electromagnetic radiation, including both radio, TV, and light. A photon is a good candidate for carrying information from one place to another. It is small, and travels fast.

A photon has an electric and magnetic field oscillating simultaneously. The direction of the electric field is called the direction of polarization (we will not consider circularly polarized photons here). Thus if a photon is headed in the z-direction, its electric field can be in the x-direction, in the y-direction, or in fact in any direction in the x-y plane or the "horizontal-vertical plane."

The polarization can be used to store a bit of information. Thus Alice could prepare a photon with horizontal polarization if the bit is 0 and with vertical polarization if the bit is 1. Then when Bob gets the photon, he can measure its vertical polarization (i.e., ask whether the polarization is vertical). If the answer is "yes", then he infers the bit is 1.

It might be thought that more than a single bit of information could be transmitted by a single photon's polarization. Why couldn't Alice send two bits, using angles of polarization different from horizontal and vertical? Why not use horizontal, vertical, half-way between them tilted right, and half-way between them tilted left. The problem is that Bob has to decide what angle to measure. He cannot, by quantum mechanical limitations, ask the question "what is the angle of polarization" but only "is the polarization in the direction I choose to measure." And the result of his measurement can only be "yes" or "no", in other words, a single bit. And then after the measurement the photon ends up either in the plane he measured (if the result was "yes") or perpendicular to it (if the result was "no").

If Bob wants to measure the angle of polarization more accurately, why couldn't he repeat his measurement many times and take an average? This does not work because the very act of doing the first measurement resets the angle of polarization to the angle he measured or the angle perpendicular to it. Thus subsequent measurements will all be the same.

Or Bob might decide to make multiple copies of the photon, and then measure each of them. This approach does not work either. He can only make a copy of the photon by measuring its properties and then creating a new photon with exactly those properties. All the photons he creates will measure the same.

What does Bob measure if Alice had prepared the photon with an arbitrary angle? Or more to the point, if the photon had its angle of polarization changed because of random interactions along the way? Or if the photon had been measured by an evil eavesdropper (typically named Eve) at some other angle and therefore been reset to that angle? In these cases, Bob always gets an answer "yes" or "no", whatever direction of polarization he chooses to measure, and the closer the actual polarization is to that direction the more likely

the answer is yes. To be more specific, the probability of the answer yes is the square of the cosine of the angle between Bob's angle of measurement and Alice's angle of preparation. It is not possible to predict the result of any one of Bob's measurements. This inherent randomness is an integral aspect of quantum mechanics.

Qubits have other interesting properties, not mentioned so far, when there are two or more of them and they are prepared together in particular ways. One such property, which we will not discuss now, known as "entanglement," allows two photons to go to different places yet have a single correlated state such that measurement of one of the photons influences subsequent measurements of the other.

Note that not all quantum systems exhibit the peculiarities discussed here. The classical bit model may be sufficient for some quantum systems. For example, if the angles of polarization are constrained to be horizontal and vertical, and there are no noise perturbations, and the appropriate angle for measurement is known, and a measurement can always be made without perturbing the photon. Thus in this special case copying is possible. If there is a very small amount of noise, so small that the perturbations in angle of polarization do not appreciably affect the probabilities of measurement, then there is a sort of restoring logic in the sense that after a measurement the polarization is always reset to horizontal or vertical.

1.5 An Advantage of Qubits

There are things that can be done in a quantum context but not classically. Some are advantageous. Consider again Alice trying to send information to Bob using polarized photons. She can prepare the photon at any angle, and could tell Bob, at the start of the setup phase, what the angle will be. Now let us suppose that a saboteur Sam wants to spoil this communication by processing the photons at some point in the path between Alice and Bob. He constructs a machine that will reflect the polarization about an angle he selects. Thus if he selects 45°, every horizontal photon becomes a vertical photon and vice versa. Knowing Alice codes bits as either horizontal or vertical photons, Sam sets his angle to 45°and employs the machine on half the photons in the message, selected at random.

Since Alice sends Bob messages coded horizontally and vertically, half the bits will be inverted by Sam, and no useful communication is possible, because for every photon which Bob measures, it is equal likely to be what Alice sent, or the other Boolean value.

Alice learns about Sam's scheme and wants to reestablish reliable communication with Bob. What can she do?

She tells Bob (using a path that Sam does not overhear) to measure photons at 45° and 135°. Sam's machine reflects the angle of polarization about 45°, so it does not affect either of the two states chosen by Alice. Of course if Sam discovers what Alice is doing, he can rotate his machine back to vertical. Or there are other measures and counter-measures that could be put into action.

This scenario relies on the quantum nature of the photons, and the fact that single photons cannot be measured by Sam except along particular angles of polarization. Thus Alice's technique for thwarting Sam is not possible with classical bits.