3
The physics of computing

Luca Gammaitoni

ICT-Energy Summer school 2016, Aalborg

A computer process information

A computer is a machine = physical system

Its functioning has to obey the laws of physics

What is the relation between **energy**, **entropy** and **information** ?

Information

From latin/italian: INFORMARE

informo, informas, informavi, informatum, informāre

FORMA = SHAPE

Meaning: “to **give shape** to something”

extended meaning“to instruct somebody (give shape to the mind)”

Relation between information and communication

Claude Elwood Shannon
(Gaylord, Michigan 1916 -
Medford, Massachussets 2001),

Reprinted with corrections from *The Bell System Technical Journal,*
Vol. 27, pp. 379–423, 623–656, July, October, 1948.

A Mathematical Theory of Communication

By C. E. SHANNON

INTRODUCTION

THE recent development of various methods of modulation such as PCM and PPM which exchange bandwidth for signal-to-noise ratio has intensified the interest in a general theory of communication. A basis for such a theory is contained in the important papers of Nyquist[1] and Hartley[2] on this subject. In the present paper we will extend the theory to include a number of new factors, in particular the effect of noise in the channel, and the savings possible due to the statistical structure of the original message and due to the nature of the final destination of the information.

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning*; that is they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. As was pointed out by Hartley the most natural choice is the logarithmic function. Although this definition must be generalized considerably when we consider the influence of the statistics of the message and when we have a continuous range of messages, we will in all cases use an essentially logarithmic measure.

The logarithmic measure is more convenient for various reasons:

1. It is practically more useful. Parameters of engineering importance such as time, bandwidth, number of relays, etc., tend to vary linearly with the logarithm of the number of possibilities. For example, adding one relay to a group doubles the number of possible states of the relays. It adds 1 to the base 2 logarithm of this number. Doubling the time roughly squares the number of possible messages, or doubles the logarithm, etc.

2. It is nearer to our intuitive feeling as to the proper measure. This is closely related to (1) since we intuitively measures entities by linear comparison with common standards. One feels, for example, that two punched cards should have twice the capacity of one for information storage, and two identical channels twice the capacity of one for transmitting information.

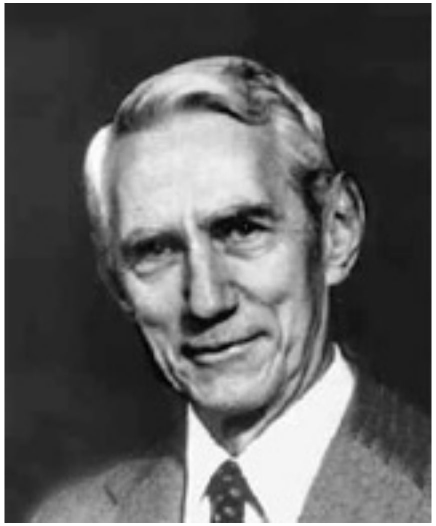
3. It is mathematically more suitable. Many of the limiting operations are simple in terms of the logarithm but would require clumsy restatement in terms of the number of possibilities.

The choice of a logarithmic base corresponds to the choice of a unit for measuring information. If the base 2 is used the resulting units may be called binary digits, or more briefly *bits*, a word suggested by J. W. Tukey. A device with two stable positions, such as a relay or a flip-flop circuit, can store one bit of information. N such devices can store N bits, since the total number of possible states is 2^N and $\log_2 2^N = N$. If the base 10 is used the units may be called decimal digits. Since

$$\begin{aligned}\log_2 M &= \log_{10} M / \log_{10} 2 \\ &= 3.32 \log_{10} M,\end{aligned}$$

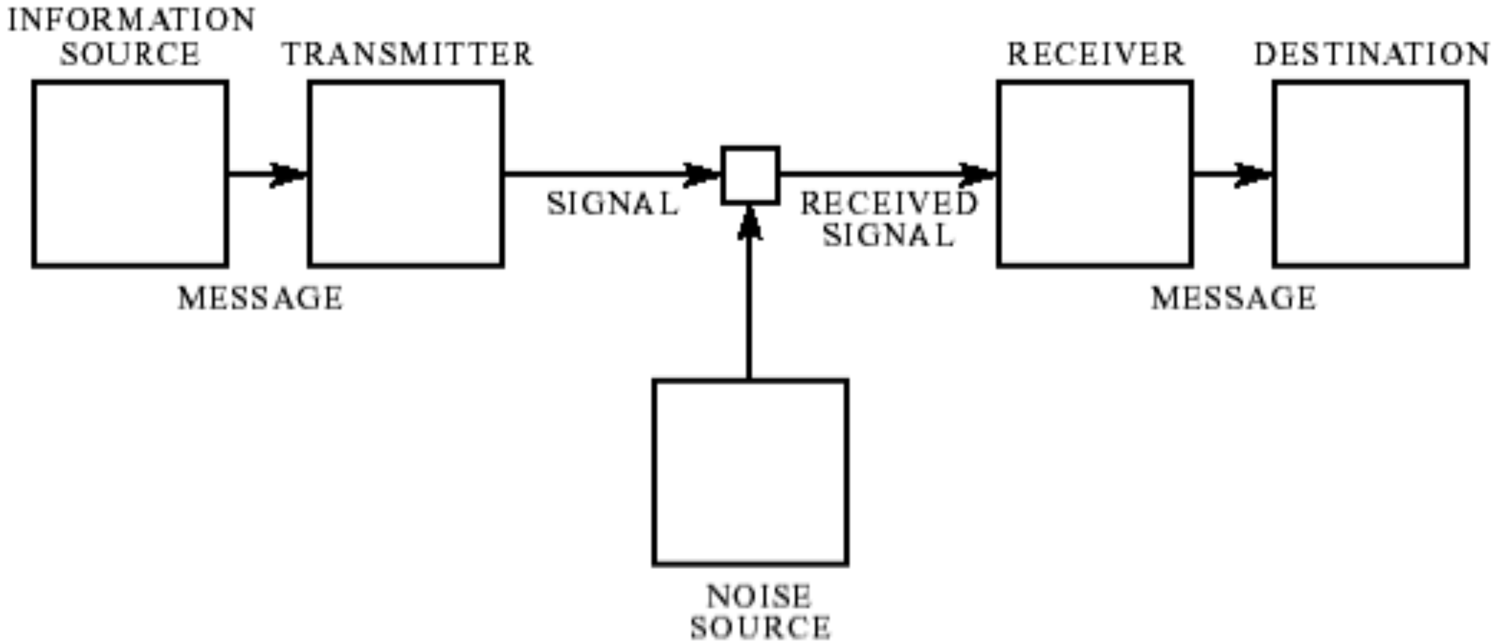
[1]Nyquist, H., "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal,* April 1924, p. 324; "Certain Topics in Telegraph Transmission Theory," *A.I.E.E. Trans.,* v. 47, April 1928, p. 617.
[2]Hartley, R. V. L., "Transmission of Information," *Bell System Technical Journal,* July 1928, p. 535.

1



C. Shannon, 1948 - *A Mathematical Theory of Communication*

Information and communication



C. Shannon, 1948 ***A Mathematical Theory of Communication***

Available at: http://www.fisica.unipg.it/~gammaitoni/info1fis/documenti/shannon1948.pdf

Information: what is it?

It is a property of a message.
A message made for communicating something.

We say that the information content of a message is greater the greater is its *casualty*.

In practice the less probable is the content od the message the more is the information content of that message.

Let’s see examples…

Information: what is it?

Let’s suppose we are waiting for an answer to a question.
The answer is the message.

Case 1:	answer yes	(probability 50%)
	answer no	(probability 50%)

The two messages have the same information content.

Information: what is it?

Let's suppose we are waiting for an answer to a question.
The answer is the message.

Case 2:	answer yes	(probability 75%)
	answer no	(probability 25%)

The two messages have the different information content.

Information: how do we measure it?

Let's suppose we want to transmit a text message:

My dear friend….

We have a number of symbols to transmit… 25 lower case letters + 25 upper case letters + punctuation + …

Too large a number of different symbols… it is unpractical.

The advantage is that we have a small number of different symbols:

0,1

But the message becomes longer…
Example: My dear--------------> 0100010111010111010101110101

Information: how do we measure it?

We send the message: 0100010111010111010101101 01

How much information are we sending?

We assume that information is an additive quantity, thus the information of the message is the sum of the information of the single components of the message, i.e. the symbols.

Now: if I send the symbol “0” how much information is in it?

Answer: it depends on the probability of that symbol, meaning the probability that the specific symbol “0” happens to be in my message.

Information: how do we measure it?

If we call p_0 the probability of having “0”and generically
p_x the probability of having the symbol “x” (a given number) we have:

$$I = - K \log p_x$$

Amount of information associated with symbol “x”.
This is technically known also as “Self-information” or “Surprisal”.

Information: how do we measure it?

We send the message: 010011001110101010101011100

If we have a message with n_0 symbol "0"; n_1 symbol "1":

$$I = -K\,(n_0 \log p_0 + n_1 \log p_1)$$

Information is an additive quantity

Entropy

The entropy of a discrete message space M is a measure of the amount of uncertainty one has about which message will be chosen. It is defined as the average self-information of a message x from that message space:

$$H = - K\, p_x \log p_x$$

Amount of information associated with symbol “x”.
This is technically known also as “Entropy”.

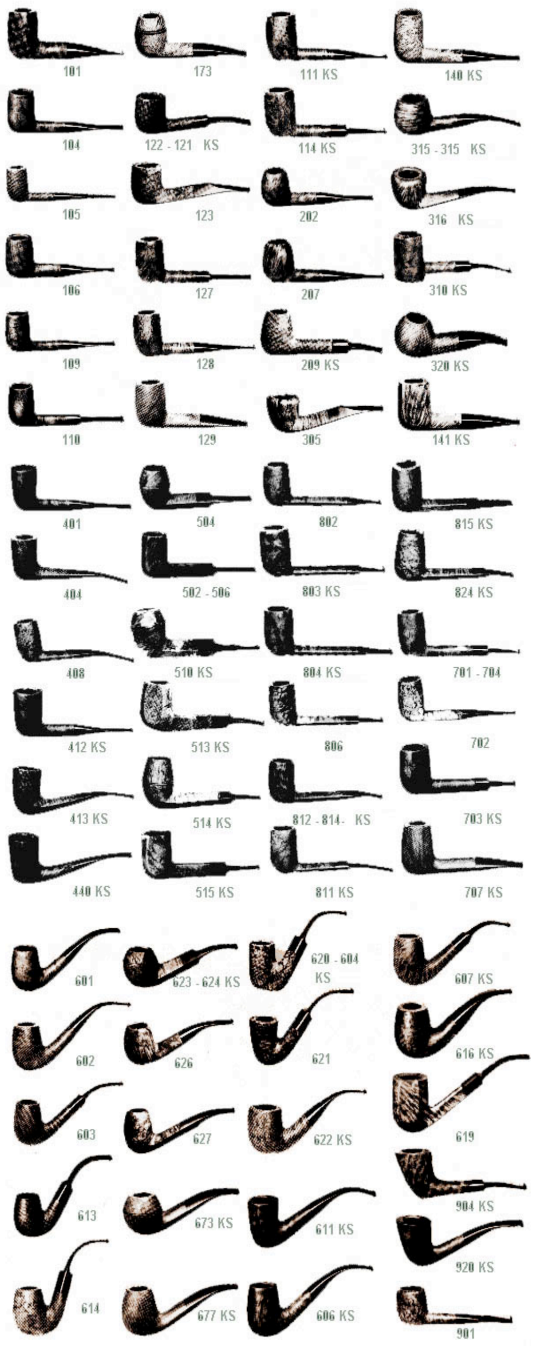
Information: binary is better

If it is long m characters (with m large), the probability $p_0 = p_1 = 1/2$

$$H = - K\, m\, 1/n \log (1/n)$$
$$= - K\, m/n\, (- \log n) = K\, m/n \log n$$

$$H = K\, m/n \log n = 2\, m/2 \log_2 2 = m$$

Thus $\quad H = m =$ number of bits



information

In-forma = in shape

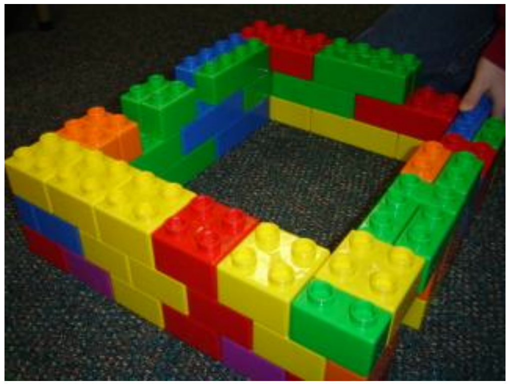
Information =
to put something in shape

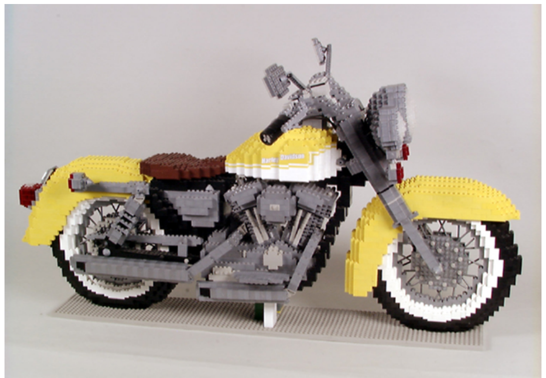
Forma = shape

The shape of an object is
a visual maifestation of the amount of
Information encoded in that object…

Example with LEGO bricks

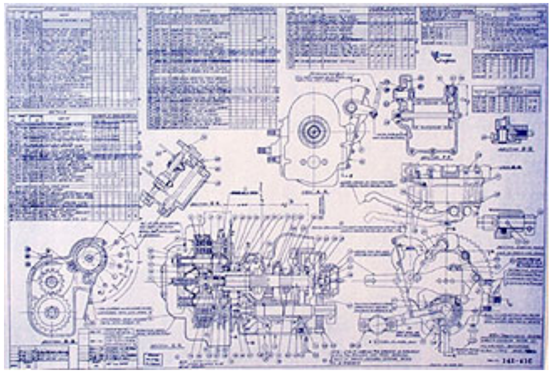
Shape = Pattern = Configuration… FORMA



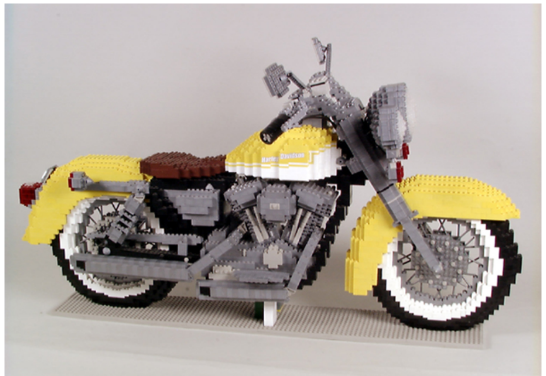


Object = bricks elements + information

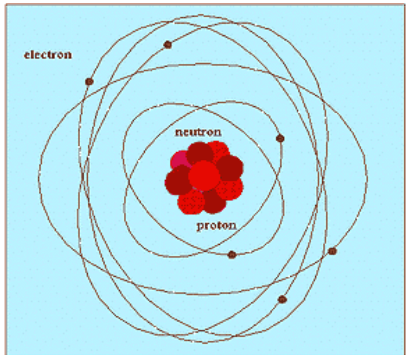
+



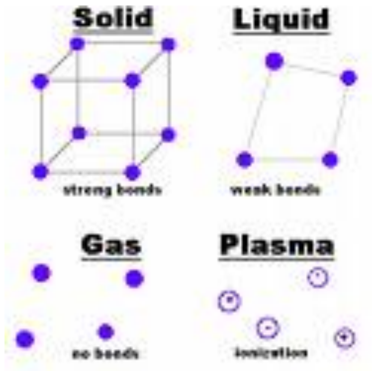
=



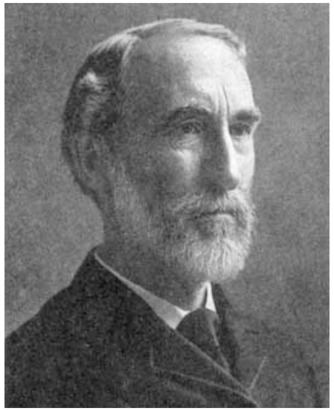
Atoms + information = matter



+



=



Shape or configurational Entropy

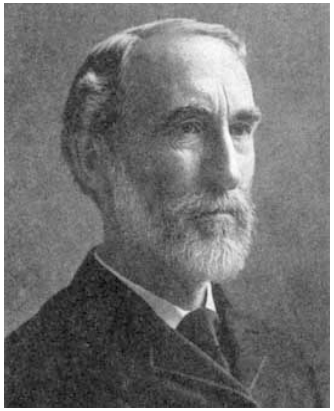
We define *shape entropy* the quantity

$$S_i = K \ln N_i$$

where K is an arbitrary constant. This quantity coincides with the microscopic form given by Boltzmann and Gibbs of the thermodynamic entropy initially introduced by Clausius, if we interpret the number of configurations N_i for a given shape as the number of accessible microstates for a given state of the thermodynamical system. Specifically, Gibbs entropy is given by

$$S_G = -K \sum_l p_l \ln p_l$$

p_l is the probability of the microstate of index l and the sum is taken over all the microstates.



Shape and information

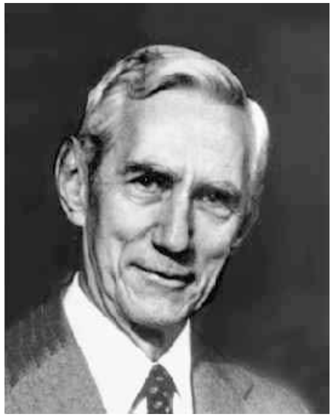
If the probability of the microstates are all the same, then the Gibbs entropy reduces to the Boltzmann entropy.

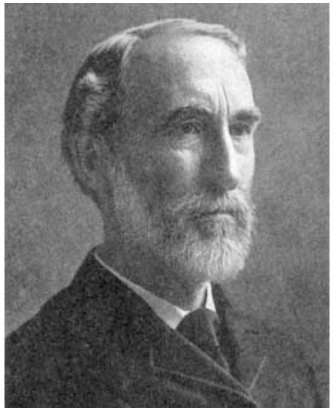
Thus if we identify the microstate of a physical object with a configuration that realizes one shape we have that **the shape entropy** IS **the Boltzmann entropy** of our object.

Shape and information

Thus we have seen that the configurational (shape entropy) of Boltzmann – Gibbs and the Information Entropy introduced by Shannon have similar formulations.

$$S_G = -K \sum_l p_l \ln p_l$$





Probability of a given symbol within a given message

Probability of a given configuration within a given shape

What about computers ?

By the moment that information processing/computing can be associated with the change of bits, in order to perform this activity we need two very important ingredients:

a) a physical system capable of assuming two different physical states: S_0 and S_1
b) a set of forces that induce state changes in this physical system: F_{01} produces the change $S_0 \rightarrow S_1$ and F_{10} produces the change $S_1 \rightarrow S_0$.

A simple system to do computation

the physical system, made by a pebble* and two bowls.

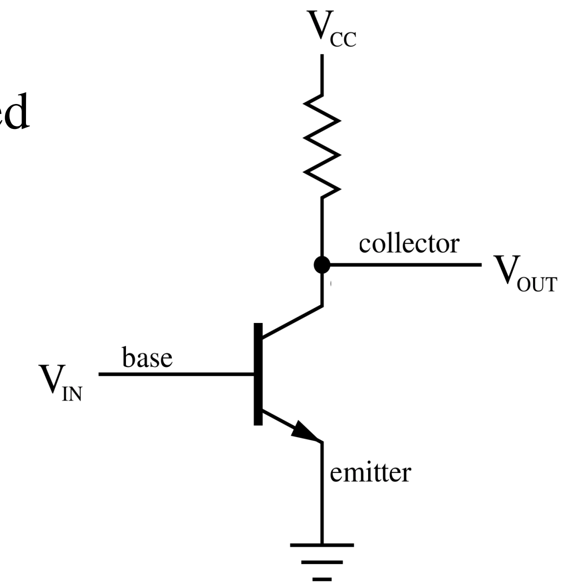
a) The two states are represented by the measurable quantity “position of the pebble”: state “0” = pebble in left bowl; state “1” = pebble in right bowl;
b) the way to induce state changes represented by a force that brings around the pebble.

* “Calculus” is the Latin word for pebble

a) a physical system capable of assuming two different physical states: S_0 and S_1

b) a set of forces that induce state changes in this physical system: F_{01} produces the change $S_0 \rightarrow S_1$ and F_{10} produces the change $S_1 \rightarrow S_0$.

Devices that obeys the rules a) and b) are called *binary switches*.



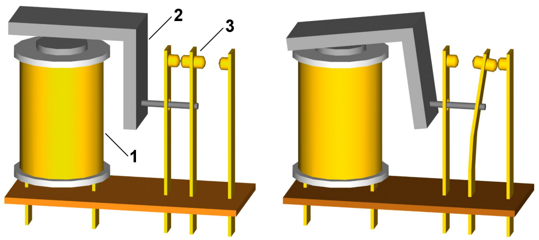
In modern computers binary switches are made with transistors. These are electronic devices that satisfy the two conditions required:

a) The two states are represented by the measurable quantity “electric voltage” at point V_{OUT}. As an example state “0” = $V_{OUT} < V_T$; state “1” = $V_{OUT} > V_T$; with V_T a given reference voltage.

b) The way to induce state changes represented by an electromotive force applied at point V_{IN}.

Binary switches

There exist at least two classes of devices that can satisfy the rules a) and b). We call them ***combinational*** and ***sequential*** devices.

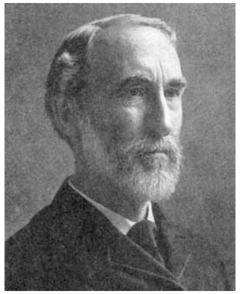
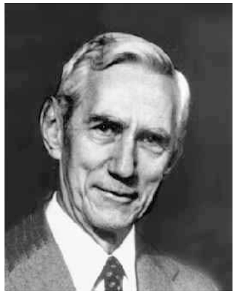
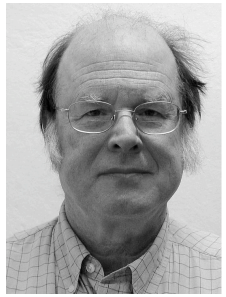


Conbinational:
in the absence of any external force, under equilibrium conditions, they are in the state S_0. When an external force F_{01} is applied, they switch to the state S_1 and remain in that state as long as the force is present. Once the force is removed they go back to the state S_0.

Sequential:
They can be changed from S_0 to S_1 by applying an external force F_{01}. Once they are in the state S_1 they remain in this state even when the force is removed. They go from S_1 to S_0 by applying a new force F_{10}. Once they are in S_0 they remain in this state even when the force is removed.

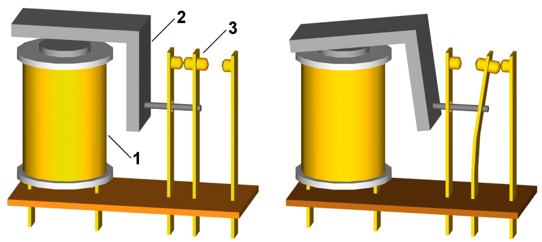
Questions

- What is the minimum energy we have to spend if we want to produce a switch event ?
- Does this energy depends on the technology of the switch ?
- Does this energy depends on the instruction that we give to the switch ?
- ….

Some answers are still controversial…

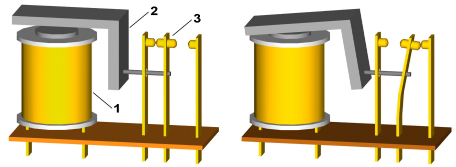
1) What is the minimum amount of energy required to operate a combinational switch?

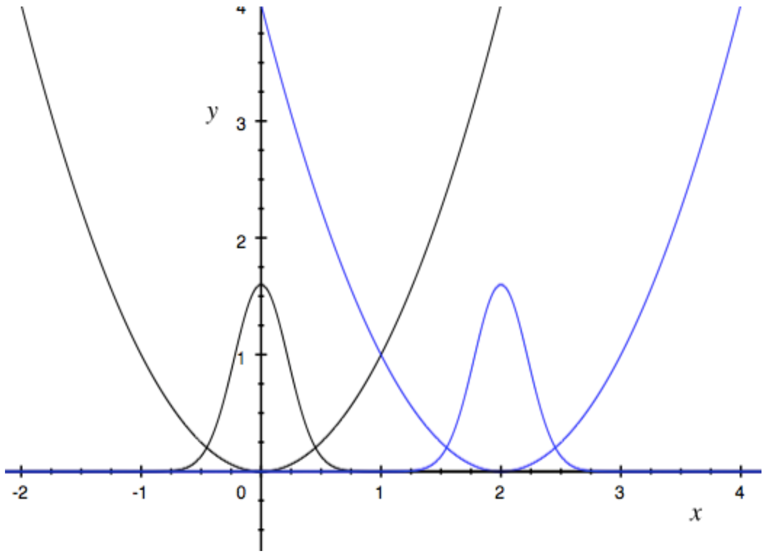


2) What is the minimum amount of energy required to operate a sequential switch?

In order to answer these questions we need a physical model

The combinational switch

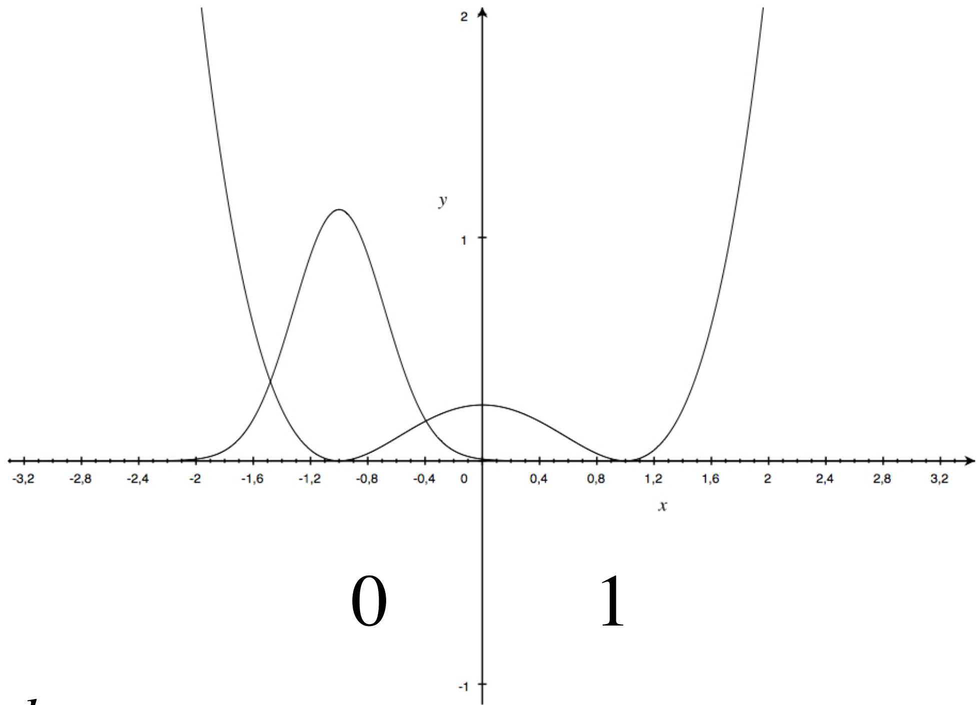




0 1

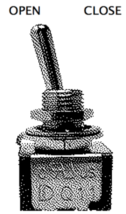
$$m\ddot{x} = -\frac{d}{dx}U(x) - m\gamma\dot{x} + \xi(t) + F \qquad U(x) = \frac{1}{2}x^2$$

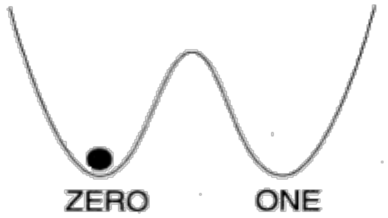
The sequential switch

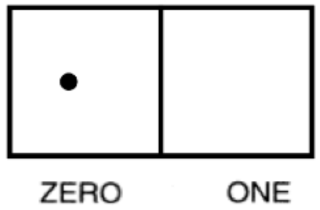


$$m\ddot{x} = -\frac{d}{dx}U(x) - m\gamma\dot{x} + \xi(t) + F$$

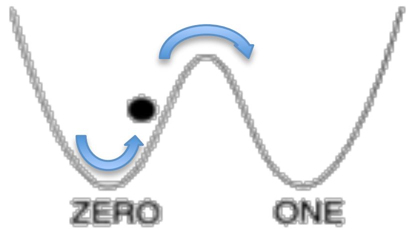
$$U(x) = -\frac{1}{2}x^2 + \frac{1}{4}x^4$$







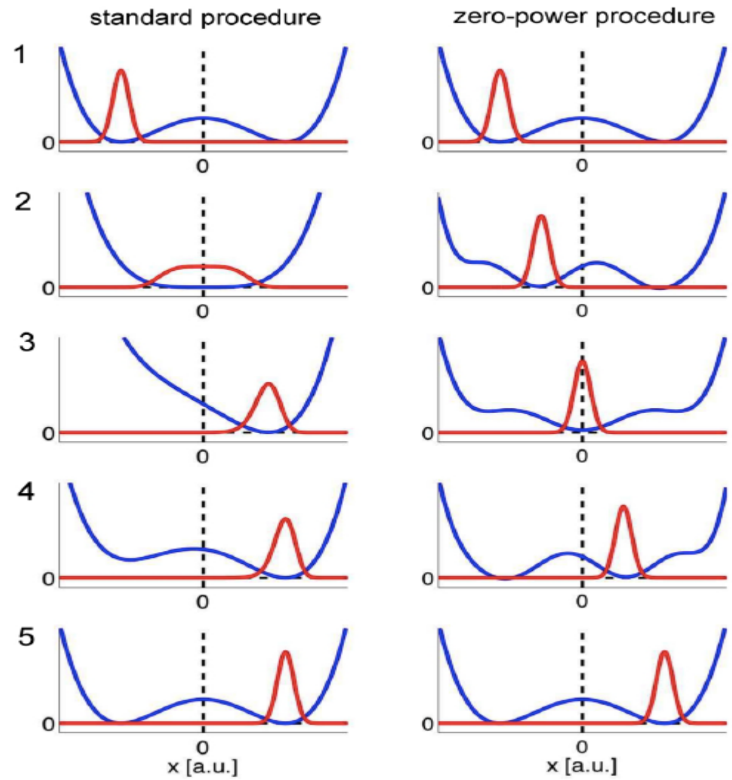
The switch operation (i.e. the change of state)



A simple protocol for sequential switches

Entropy changes

Entropy stays constant



Summary

1) Information is formally connected with entropy
2) Computers obey the laws of physics
3) Computing is altering information and thus may take energy

To learn more:

The book "ICT - Energy - Concepts Towards Zero - Power Information and Communication Technology” InTech, February 2, 2014