Minimum energy required to copy one bit of information

Marcin Ostrowski*

May 17, 2010

Abstract

In this paper, we calculate energy required to copy one bit of useful information in the presence of thermal noise. For this purpose, we consider a quantum system capable of storing one bit of classical information, which is initially in a mixed state corresponding to temperature $T$. We calculate how many of these systems must be used to store useful information and control bits protecting the content against transmission errors. Finally, we analyze how adding these extra bits changes the total energy consumed during the copying.

1 Introduction - Landauer principle

What is the relationship between information and heat (energy)? Does a machine that performs calculations (such as a computer) have to produce heat? These questions are purely fundamental but they are also important from a technical point of view. Firstly, as computer users, we want to pay less for the electricity we consume. Secondly, the problem of cooling modern electronic systems with high packing density of components is a great challenge for engineers.

The relationship between heat $Q$ (and the thermodynamic entropy $S$) and information had occupied the attention of researchers for a long time before the invention of the computer. Let us recall here the problem of Maxwell’s demon separating faster particles from the slower which resulted in the flow of heat from a colder to a hotter reservoir. These considerations led to the suggestion that the measure (gain of information) always requires the increase of thermodynamic entropy [1], [2].

According to the present state of knowledge, information processing (calculations) may take place without the production of heat. In this case, we are talking about reversible computations. The main feature of reversible computations is the ability to recover input data based on output data. The machine fully capable of implementing this type of task would be a quantum computer [3], [4].

However, the modern digital machine is not reversible. Moreover, apart from the calculations, it must perform other functions (associated with the calculations) as well as reading of data, copying of data, etc.. In 1961, Rolf Landauer showed in his work [5] that any logically irreversible manipulation of information, such as erasing the information, must result in an increase in entropy of the apparatus or the environment. Erasing one bit of information should result in the production of heat at the order of $kT \ln 2$. This is often called the "Landauer principle".

The present paper relates to those considerations, but it does not recalculate the Landauer’s result using a different method. In our work, we consider another (though related) issue. We wonder how much energy should be used to make a copy of one
bit useful information when we do not have a pure medium on which we could save a copy, that is, if there is a necessity, to overwrite the original thermal noise appearing on the medium. Therefore, we assume that information is protected against errors due to thermal noise by the presence of additional control symbols (redundancy). We investigate whether it is better to use a small number of control symbols with a greater energy expenditure per bit (higher signal to noise ratio) or the other way round (i.e. use a large number of control symbols at low signal to noise ratio). For this purpose, we introduce the concept of quantum copier that makes copies of information between subsystems. The main result of the work (presented in the last section) is limitation for the minimum energy required to copy one bit similar to Landauer’s one.

2 Quantum copier

Consider a physical (quantum) system that stores one bit of classical information. Let logic state "0" correspond to the quantum state $|0\rangle$ and logic state "1" correspond to the quantum state $|1\rangle$. States $|0\rangle$ and $|1\rangle$ are orthogonal to each other (which ensures their distinguishability in the measurement), normalized and form the basis of a quantum system.

Let us consider two such systems which will be called $A$ and $B$, respectively. Next, we construct a quantum operation that copies the states "0" and "1" from the system $A$ to system $B$ according to the following scheme:

$$
|0\rangle_A |\text{pm}\rangle_B \rightarrow |0\rangle_A |0\rangle_B \\
|1\rangle_A |\text{pm}\rangle_B \rightarrow |1\rangle_A |1\rangle_B
$$

where the state $|\text{pm}\rangle_B$ means a given state of the system $B$. As we shall see, copy could be done properly only if the system $B$ is initially in a particular state, which here is denoted by the abbreviation "pm", which stands for "pure medium".

The choice of states $|0\rangle$ and $|1\rangle$ as basis vectors (and not as some combination of the basis states) is only a simplification of the notation does not restrict the generality of our considerations.

In order to fully determine the unitary operation acting on subsystems $A$ and $B$ we have to introduce vector $|\text{um}\rangle_B$ (from the words "unprepared medium") orthogonal to $|\text{pm}\rangle_B$ and determine the action of the operation on it. If we did it as follows:

$$
|0\rangle_B |\text{um}\rangle \rightarrow |0\rangle_B |0\rangle \\
|1\rangle_B |\text{um}\rangle \rightarrow |1\rangle_B |1\rangle
$$

then our operation would be able to copy the initial state of subsystem $A$ to subsystem $B$, regardless of the initial state of medium $B$. Unfortunately, this type of operation is not unitary. This would be an irreversible process erasing information about the initial state of $B$. Thus, we have to design our system as follows:

$$
|0\rangle_B |\text{um}\rangle \rightarrow |0\rangle_B |1\rangle \\
|1\rangle_B |\text{um}\rangle \rightarrow |1\rangle_B |0\rangle
$$

This means that for copying operation (which will be called the quantum copier) to work correctly, the second entry must always be in the state $|\text{pm}\rangle$. Otherwise errors will occur.

Equations (1)-(4) fully define the operation of the copier as a unitary operation acting over $H_A \otimes H_B$. Let us emphasize that the copier can copy only the base states ($|0\rangle_A$ and $|1\rangle_A$), not their superpositions. Copying of any state is impossible by the no-cloning theorem.

1 In the further part of the work, we will use shorter notation of states (i.e. without lower indices $A$, and $B$). States $|0\rangle_A|1\rangle_B$ will be denoted by $|01\rangle$ or even $|01\rangle$. 

2
3 Copy to the unprepared medium (in the presence of noise)

Let us examine what the result of copying will be when the medium $B$ is initially unprepared. For this purpose, we choose the initial state of $B$ as a mixed state in the form:

$$\rho_B = (1 - b)|pm\rangle\langle pm| + b|um\rangle\langle um|$$  \hspace{1cm} (5)$$

where $b$ can be identified with a probability of incorrect preparation of the medium (and therefore erroneous copying).

In the case of $A$, we assume that with probability $p_0$ the system stores a "0", and with probability of $p_1$ it stores a "1". Therefore, we assign the system $A$ a mixed state in the following form:

$$\rho_A = p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1|$$  \hspace{1cm} (6)$$

Subsystems $A$ and $B$ are together in a state which can be written as:

$$\rho_{AB} = \rho_A \otimes \rho_B = p_0(1 - b)|0\rangle|pm\rangle\langle 0|\langle pm| + p_1|1\rangle|pm\rangle\langle 1|\langle pm| + p_1 b|0\rangle|um\rangle\langle um|\langle 0| + p_1|1\rangle|um\rangle\langle um|\langle 1|$$  \hspace{1cm} (7)$$

After copying information (after applying the unitary operation $U_c$ described by the equations (1)-(4)) we obtain:

$$\rho'_{AB} = U_c \rho_{AB} U_c^+ = p_0|0\rangle|0\rangle_A[(1 - b)|0\rangle|0\rangle_B + b|1\rangle|1\rangle_B] + p_1|1\rangle|1\rangle_A[(1 - b)|1\rangle|1\rangle_B + b|0\rangle|0\rangle_B]$$  \hspace{1cm} (8)$$

and after counting the partial trace:

$$\rho'_A = p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1|$$  \hspace{1cm} (9)$$

$$\rho'_B = [p_0(1 - b) + p_1 b]|0\rangle\langle 0| + [p_0b + p_1(1 - b)]|1\rangle\langle 1|$$  \hspace{1cm} (10)$$

Let us look at specific cases:

$$\rho'_B = p_0|0\rangle\langle 0| + p_1|1\rangle\langle 1| \quad \text{for} \quad b = 0$$  \hspace{1cm} (11)$$

$$\rho'_B = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \quad \text{for} \quad b = \frac{1}{2}$$  \hspace{1cm} (12)$$

$$\rho'_B = p_1|0\rangle\langle 0| + p_0|1\rangle\langle 1| \quad \text{for} \quad b = 1$$  \hspace{1cm} (13)$$

Thus, we have obtained what we expected. For $b = 0$ we obtain a faithful copy, for $b = \frac{1}{2}$ we obtain a complete lack of information copied from $A$ to $B$, for $b = 1$ we get a faithful copy with negation (the states $|pm\rangle_B$ and $|um\rangle_B$ turn its roles).

4 Relationship with Shannon’s binary information channel

We will treat states of $A$ ("0" and "1") as the realization of a stochastic process $X$ and states of $B$ (after copying) as the realization of stochastic process $Y$. The relationship between these two processes can be depicted by the graph in Figure 1.

It is a complete analogy with the Shannon’s, binary, noisy information channel. The process $X$ correspond to the process at the entrance to the channel, the process $Y$ in the output.

3
In the next step, we calculate transinformation \( I(X, Y) \) for these processes (i.e. the amount of knowledge about \( X \) which is contained in \( Y \)):

\[
I(X, Y) = \sum_{i,j} P(X_i, Y_j) \log_2 \left( \frac{P(X_i, Y_j)}{P(X_i)P(Y_j)} \right) 
\]

where the probabilities for the process \( X \) are given by:

\[
P(X_0) = p_0 \quad P(X_1) = p_1
\]

while for the process \( Y \) are in the form:

\[
P(Y_0) = p_0(1-b) + p_1b \quad P(Y_1) = p_0b + p_1(1-b)
\]

while the total probabilities:

\[
P(X_0, Y_0) = p_0(1-b) \quad P(X_0, Y_1) = p_0b
\]

\[
P(X_1, Y_0) = p_1b \quad P(X_1, Y_1) = p_1(1-b)
\]

Let us calculate transinformation (14) for the case \( p_0 = p_1 = \frac{1}{2} \), since then in \( A \), we have exactly the 1bit of information. Then:

\[
I(X, Y) = 1 + (1-b) \log_2(1-b) + b \log_2 b.
\]

5 Medium as a system with two states of energy

For simplicity, we assume that \( |p_m\rangle_B = |0\rangle_B \) and \( |u_m\rangle_B = |1\rangle_B \). In addition, we assume that they are stationary states of the free hamiltonian of \( B \). Otherwise, these states would evolve in time and reading of information (stored in them) at any moment of time would be difficult. Thus, let \( E_0 \) and \( E_1 \) denote energy of state \( |0\rangle_B \) and energy of the \( |1\rangle_B \), respectively.

Our main assumption is that before copying the system \( B \) was in contact with the thermostat at a temperature of \( T \) long enough to reach thermodynamic equilibrium. In practice, this is a natural assumption, since any real device works at a temperature different from the absolute zero.\(^2\)

Probability of filling the levels \( |0\rangle \) and \( |1\rangle \) at temperature \( T \) are respectively given by:

\[
P_0 = \frac{1}{1 + \exp(-\beta \Delta)}, \quad P_1 = \frac{\exp(-\beta \Delta)}{1 + \exp(-\beta \Delta)} = b
\]

\(^2\)In practice, some systems may be in a quasiequilibrium state, while a state of full equilibrium is possible to achieve only after a very long time. However, this situation will not be considered here.
where $\Delta = E_1 - E_0$.

It is worth noting that $b = P_1$ only for $b \in (0; \frac{1}{2})$. For $\beta \Delta \gg 1$ error rate $b$ is very low, while for $\beta \Delta \rightarrow 0$ error rate $b$ tends to $\frac{1}{2}$. The most important conclusion is that medium $B$ in thermal equilibrium is useful for copying if the energies of the levels differ (i.e., $\Delta \neq 0$). Otherwise, $P_0 = P_1 = \frac{1}{2}$ and (under considerations presented in Section 3) the medium is totally useless.

6 Redundancy

In order to copy useful information, when $b \neq 0$, we have to protect transmission against possible errors. To this end, we must encode useful information by adding a certain amount of control symbols. Let us denote the number of all symbols by $n$ (one symbol carrying one bit of useful information and $n - 1$ symbols are control symbols). In practice, we should have not only one system $A$ but the whole series of systems in an amount of $n$. We also need $n$ systems $B$ in the states corresponding to a temperature of $T$, where we finally put ”dirty” copies from systems $A$.

According to the Shannon theorem [6], we can conclude that for transfer of one bit of useful information, $n = I(X, Y)^{-1}$ bits are necessary, where $I(X, Y)$ is given by formula (15). We do not examine here any specific coding procedure (the values that control bits must take are not defined).

7 Calculations, graphs, conclusions

Let us calculate the energy that must be used to copy $n$ symbols from systems $A$ to systems $B$. It is calculated as follows:

$$W = n(\tilde{E} - \bar{E})$$

where $\tilde{E} = \Delta P_1 = \Delta b$ is the average energy of $B$ before copying (i.e., at a temperature of $T$) calculated for $E_0 = 0$ and $b \in (0; \frac{1}{2})$. Energy $\bar{E}' = \frac{1}{2} \Delta$ is the average energy of $B$ after copying. We assume here that the probabilities of occurrence of zeros and ones in the $B$ are equal and amount to $1/2$. This gives the energy (18) equal to:

$$W(\beta, \Delta) = \frac{\Delta(\frac{1}{2} - b)}{1 + (1 - b) \log_2(1 - b) + b \log_2 b},$$

where $b$ is given by Eq. (17).

Energy (19) is an increasing function of $\Delta$. This means that it is less expensive to encode the information at low signal to noise ratio with a large amount of control symbols $n$. In the borderline case, we obtain:

$$W_{\text{min}} = \lim_{\Delta \rightarrow 0} W(\beta, \Delta) = \frac{\log 4}{\beta}$$

Energy (19) does not correspond to the scattered heat (as in Landauer). Our copy machine works reversibly. The copy obtained in the $B$ still contains information about the state of medium $B$ before copying. This information can be recovered by comparing the ”dirty” copy with original one saved on $A$. Thus, there is no erasure of information. Erasure may occur in the next phase - on machine decoding and removing copying errors.

\[3\] This follows from the following facts: $p_0 = p_1 = \frac{1}{2}$. In addition, we assume that among the control symbols added before copying is the same number of ”0” and ”1”. This type of code are usually the most optimal. During copying from $A$ to $B$ probability of distortion from ”0” to ”1” and from ”1” to ”0” are the same (Fig. 1). This results in conservation of symmetry between the zeros and ones on the $B$.
Figure 2: Energy used for copying one bit of useful information as function of $\Delta$ for $\beta = 1$. The straight line (added for comparison) corresponds to the value $\Delta/2$.

References

[1] L. Szilard, Z f Physik, 53, 840, (1929),
[2] C. H. Bennett, Int. J. Theor. Phys., 21, 905, (1982),
[3] Feynman R. P. *Simulating physics with computers*, Int. J. Theor. Phys. 21:6/7, pp. 467-488 (1982),
[4] Feynman *Feynman lectures on computation*, (1996),
[5] R. Landauer, IBM J. res. Develop., 5, 183, (1961),
[6] Shannon C. E. *The Mathematical Theory of Communication*, The Bell System Technical Journal, 27, 1948