Information as mapping

József Holovács[a], Gábor Kusper[b]

[a]Uzhgorod National University

[b]Eszterhazy Karoly University of Applied Sciences

Abstract

As information is a basic amount, it doesn't have an exact definition. Because of this, information is explained in a number of ways. One can speak of five hierarchical levels of information. The Shannon information-theory only deals with the lowest level, the codes. In this article we introduce the concept of information mapping, which gives us the ability to understand the essence of information. We analyze the characteristics of information mapping as well as basic information operations, disinformation, and its causes. Informational operations, just as the creation of information mapping, can only be the result of a particular intellectual activity. Matter is needed as the carrier of information. Let's investigate the connections between information and the second law of thermodynamics.

The genetic information, as a dominant part of biological information, corresponds to the structure of information mapping, and describes the compliance between amino acids and DNS. Biological information is the result of intelligent design, and cannot come into existence by itself. The paradigm of evolution, which doesn't take the essence of information into account, is widespread in our society. Instead of it the informational paradigm has to be accepted and applied. A paradigm-change is needed.

1. Introduction

The information, in the same way, as matter and energy, can be considered a basic amount. Because of this, information cannot be exactly defined, and this is the reason of why people try to explain it in several ways.

Finding out the essence of information is of scientific and philosophical importance. For example, a clear answer has to be found for an important question: how did biological information come into existence and among other things the genetic information? (As we all know, there is no life without information. The living world can be primarily distinguished from the lifeless world, taking into account, that all biological processes are guided by information.)

The role and importance of information was understood only from the middle of the first century. N. Wiener, whose book called Cybernetics appeared in 1948 [1], plays an important role in this. It became more comprehensible, that the foundation of the operation of all the technical and biological systems – is the information, and only through this can a control which meets a certain goal be accomplished.

A comprehensive analysis regarding the problems of information was performed by Gitt. Among others he introduced the five hierarchical levels of information [2]:

1. *statistic* (describes the used code-system)

2. *syntax* (a potential structure of the language which describes information)

3. *semantics* (the essence, central aspect of information)

4. *pragmatics* (purpose which the transmitter wants to achieve at the receiver)

5. *apobetics* (purpose of the sender).

All the five levels of informational tasks are of intellectual nature, material components are just of secondary importance. To be able to understand the essence of information, all five hierarchical levels are needed. Therefore, concepts which try to explain information just on a statistical, code level, are not complete.

In this respect Shannons information theory [3] can be mentioned. Shannon practically just deals with the lowest level, the code. Shannon started to apply the *bit* unit as the first person to "measure" information. But information cannot be "measured". Only lines of code can be measured, which are used to save and transmit information. The statistical analysis of code lines can be important for the practical implementation of information transfer, but Shannon doesn't deal with semantics of information, as a central aspect of information. This was pointed out by many. *Karl Steinbuch*: "The classical information theory can be compared to a person, who counts a kilo of gold and a kilo of sand as equal." [4]. *Warren Weaver*: "Two news can be regarded exactly equal in Shannon's point of view, even if one is rich in meaning and the other is total non-sense." [2]. The lack of the statistical level of information can be expressed with "the meaning of text cannot be found in the chemistry of paper and ink" [5].

2. Information mapping

Based on the analysis of the essence of information a conclusion can be made, that the concept of information can be approached in a way, where information is being considered as a special mapping. We introduce the term *information mapping.*

Information mapping (IL) – an $f\colon X \to Y$ (or $Y = f(X)$) matching, of which the set of objects in the *domain* of arbitrary nature (including non-material as well) X, and the Y *value range* is represented by material objects or objects which are representable in the material world.

For example, the simplest information mapping can be created in the following way. We introduce the set $X = \{$"There is no danger, coming in is possible!", "Danger, coming in is not possible!"$\}$, and the set $Y = \{$"There is a flower pot in the window", "There is no flower pot in the window"$\}$.

Then we can create the following information mapping:

a) $\{$"There is no danger, coming in is possible!"$\} \rightarrow \{$"There is a flower pot in the window"$\}$

b) $\{$"Danger, coming in is not possible!"$\} \rightarrow \{$"There is no flower pot in the window"$\}$.

Of course, the "there is a flower pot in the window" state (or the opposite one) is not information by itself. It is just a possible value of a code. The information mapping has to be created by an *intellectual person* – any matching can only come to existence by an intellectual process. This means that there is no material process which would be able to create information by itself.

Regarding informational operations, and inversed information mapping *(IIL)* is needed as well, $g\colon Y \rightarrow X$ (g is the inverse of f).

Information mapping, as any other mapping in general, can be:

- *clear mapping* – every element of X can be matched to only one element of set Y

- *multivalued mapping* – one element of X can be matched to more elements belonging to set Y

- *injective mapping* – different elements of X correspond to different elements of Y ($|X| \leq |Y|$)

- *surjective mapping* – every element of Y corresponds to at least one element of X ($|X| \geq |Y|$). In this case several X elements can be mapped to a Y element, and the inverse mapping cannot be unequivocal.

A *mapping* is *bijective*, if it is injective and surjective simultaneously. For the coding of information only *bijective information mapping* is used. In this case there is an unequivocal inverse mapping. Otherwise the decoding (inversed information mapping) is not unequivocal and it is complicated to restore the coded message.

If two IL value ranges are identical: $f_1 : X \rightarrow Y_1$ and $f_2\colon X \rightarrow Y_2$, then this *information mappings* are *equivalent*. We can put it this way: the choice of the coding system (Y) is *invariant* in the viewpoint of information semantics (X). The information is determined by the semantics, which is described by the X range, and does not depend on the material carrier.

It can be spoken of three kinds of information mapping:

$(1 \rightarrow 1)$ – correlation, *bijection* (an inversed mapping exists).

$(1 \rightarrow \mathrm{M})$ – *multivalent mapping* (an unequivocal inversed mapping exists).

(M → 1) – *surjective mapping* (an unequivocal mapping doesn't exist). This, depending on the context, *IL*, which inversed mapping can be cleared, an "external", *higher level IL* can be considered.

This creates a problem of uncertainty, of which a solution can be described by semantics.

Assume that $g: Y \to X$ os a context-dependent inverse IL, in which the value of the x prototype not only depends on the y element of the prototype-set, but on a higher level $F: U \to V$ information mapping as well (assume G mapping is the inverse of F). As this is a surjective mapping, assume x_1, x_2 is the g result of the inversed information mapping, and besides that the $(u_1, x_1), (u_2, x_2)$ pairs are semantically compatible, where u_1, u_2 are within the meaning of the F mapping in this way: $u_1 \to v_1$, $u_2 \to v_3$. When the receiver gets the message, which includes the $(\ldots, v_1, y, \ldots)$ values, then based on G the value of the prototype set will be u_1 , and as (u_1, x_1), the value of x_1 can be clearly determined. In the other case we get the value of x_2.

In a lot of cases a dynamical information mapping (DIL) is needed, as a sequence of information mappings

$$f: X(k) \to Y(k), \quad k = 0, 1, 2, \ldots,$$

which is determined in k points in time. Then the inversed dynamical information mapping (IDIL) will be $g: Y(k) \to X(k)$, $k = 0, 1, 2, \ldots$.

Algorithms are the most frequently used dynamical information mappings.

Regarding the control problems, informational processes are needed: (x_k, y_k), $k = 0, 1, 2, \ldots$ pairs, where $x_k \in X$, $y_k \in Y$.

The control of objects of different kinds of nature happens through informational processes.

In order to be able to execute the informational basic functions a sender is needed, who communicates the information and one or more recipients, who get the information. The sender, as well as the recipient can be an intelligent being or an object created by intelligence. In this regard the connection between the sender and the recipient is possible in four shapes [2]:

intelligent being → intelligent being,
intelligent being → object created by intelligence,
object created by intelligence → intelligent being,
object created by intelligence → object created by intelligence.

Information mappings are used for two basic functions: *direct* and *inversed informational operations. Direct informational operation* (executed at the *sender* side): the evaluation of the information mapping based on a particular X element, of which the result is – an X element.

It is understandable, that informational operations, as well as the creation of an information mapping can only be the result of an intellectual activity. The matter is just the *carrier*, *holder* of the information.

Disinformation. In practice, potential disinformation has to be considered as well. Let y be the result of the evaluation of the $f \colon X \to Y$ "accurate" information mapping. Then any difference from the y value (accidentally or deliberately) is called *disinformation*. The reasons of disinformation can be:

- error in the selection of the X element in the value range ($f \colon X^* \to Y$)

- error in the application of the f information mapping ($f^* \colon X \to Y$)

- error in the fixation, "calculation" of the value range element ($f \colon X \to Y^*$).

The disinformation can lead to contortion of the informational semantics. Because of this the mutations (as disinformation) of DNS, can only lead to the loss of genetic information. The characteristics of information refute the idea, that "new and more complex species can come into existence by mutations and natural selection", as there are "informational gaps" between the species.

Denton M. writes about this in his book *Evolution: A Theory in Crisis* (1986) [7]: "If complex computer programs cannot be changed by random mechanisms, then surely the same must apply to the genetic programs of living organisms. ... *We are very close to the formal refutation of the whole Darwin paradigm.* By which abilities are living organisms defy the laws of probability, which every complex system obeys?"

The characteristics of information do not depend on the carrier of the information, therefore those characteristics are true for all technical, economic, biological, etc. objects. As we know, there is no material process, which would be able to create creative information. This is true for the genetic information stored in the DNS. Therefore, if a theory only tries to explain the origin of life by material processes, then this is wrong based on the characteristics of information.

3. Genetic information as mapping

Genetic information has to be regarded as a special information mapping as well, which defines the *amino acids* $(X) \to$ *DNS codes* (Y) compliance. In total 21 codes are needed: 20 amino acids (and a STOP code) has to be marked. The *functional proteins* consist of the given 20 amino acids.

In the DNS the coding is based on the four-letter alphabet $\{0, 1, 2, 3\}$, of which the material (chemical) representation happens through the molecules *adenine* (A), *cytosine* (C), *guanine* (G), *thymine* (T). A-T, C-G – build complementary pairs, which make the DNS able to replicate. One amino acid is coded by the three-letter "words", *triplets* found in the DNS. For instance, "220" codes the *glycine* amino acid.

This means that in total 64 triplets exist, which is significantly more than the needed 21 codes, therefore the *information mapping* is *polyvalent*: 20 amino acids (+ «Stop») $\to$ 64 triplets. The inversed mapping (DNS triplets $\to$ amino acids) are unequivocal. Because of this there is a way to build a kind of protection into the

DNS against potential mutations. For instance, the *alanine* amino acid is coded in the information mapping in four ways:

Alanine → 210 (GCA);
Alanine → 211 (GCC);
Alanine → 212 (GCG);
Alanine → 213 (GCT).

As the letter of the triplets coding alanine can be any letter, on this place one letter replacement, as a mutation is not able to deform the content, semantics of the genetic "text". Similarly the other amino acids correspond to more than one triplet. The selected coding system is optimal in terms of matter demand. The next table contains the other possibilities of coding [2]

	N = 2	N = 3	**N = 4**	N = 5	N = 6
D = 2	4	9	16	*25*	*36*
D = 3	8	*27*	*64*	125	216
D = 4	16	*81*	256		
D = 5	*32*	243			
D = 6	*64*	729			

(D – length of words, N – number of letters in the alphabet)

The number of letters can only be even in the chosen alphabet, otherwise the complementary copy could not be performed.

The complexity of the information stored in living creatures and the complexity of the led processes for a brilliant design necessarily leads to a creator of the living world.

Biological information – excellent engineering and informational performance.

4. The connection between the second law of thermodynamics and information

The "movement" of a guided object in a phase space in an ordinary form (in a discrete case) can be described in the following way:

$$x(t + 1) = \varphi(x(t), u(t)),$$

where $x(t)$ phase vector, $u(t)$ controlling vector in t point in time.

As the second law of thermodynamics deals with objects which are not under control, only after switching of the $u(t)$ control the object becomes uncontrollable, which leads to the object entering the impact of the second law of thermodynamics.

The "correction" of such an object's state is only possible through a controlling vector, which is calculable only based on appropriate information. The external energy, of any amount, is not able to restore the objects state. The execution of

the control needs energy, but this is only a necessary condition for the reduction of entropy (in terms of thermodynamics).

The application of information describing the object's control excludes the process provided by the second law of thermodynamics, and vice versa, the lack of it starts this process. That is, the growth of controlling information and entropy (in terms of the second law of thermodynamics) mutually exclude each other. Taking into account the information factor contributes to the correct application of the second law of thermodynamics. Because of this the criticism of the claims of the evolution theory which don't take the connection between the second law of thermodynamics and information into account, are well founded.

Morris Henry pointed it out this way: "Because of the complexity of the living world the second law of thermodynamics clearly doesn't make the evolution possible to hold. The tendency of becoming a disorganized system is present in every system. This tendency is only stoppable through external energy sources which are guided by an *appropriate informational program*. Based on this informational mechanism ("*input → storage → transformation*") the energy transforms to a specific work which is essential for the building of a system's complicated structure." [10].

5. Informational paradigm

Under the evolutional paradigm everything evolves in a natural way, it evolves, gets better, and all this happens without the intervention of a mind, intelligence. This means that this paradigm doesn't take the information's essence into account. The paradigm of evolution became dominant in science and society.

For instance, in biology and researches dealing with the creation of life it "suits" to only give evolutionary explanations to the facts, which lead to faulty conclusions. For instance, *Richard Dawkins* defines biology like this: "... study of complicated things that give the appearance of having been designed for a purpose " [7]. *R. Dawkins* and his followers don't want (or are not able to) admit, that life is indeed the "result of a purposeful design". Recently many start to understand that in terms of researches the *informational paradigm* has to be applied, which puts the information into the center of everything: "The creation and appropriate operation of any system (complex object) is only possible through *creative information*. This information is the result of a proper *intellectual process*."

It is obvious that the change of paradigms became timely. This can be useful, for instance, to understand the origin of life. Currently a lot of scientists perform researches in the spirit of the informational paradigm. A few quotes follow in this connection.

R. Junker, S. Scherer, German microbiologists: "The full, required information was already present in living creatures from the beginning. Any creature belongs to a certain basic type, which had to appear already finished." [8]

Denton M., Evolution: from his book *A Theory in Crisis* (1985):

"If complex computer programs cannot be changed by random mechanisms, then surely the same must apply to the genetic programs of living organisms. ... We

are very close to the formal refutation of the whole Darwin paradigm. By which abilities do living organisms defy the laws of probability, which every complex system obeys?"

"Is it believable that a reality could be created randomly in which the smallest element – a *functional protein* or *gene* has a complexity which goes beyond our creative abilities, a reality, which is the full antithesis of randomness, and goes beyond anything which human intelligence has created?" [6]

Meyer Stephen, American philosopher: "We only know one source which is able to create information – the mind. The fact that DNS contains information leads us to the conclusion that it is the result of design. . . The understanding, that the DNS molecule is an evidence based on which we can accept the existence of a Higher Intelligence, intensifies. . . The DNS contains characteristics, informational content which is similar to computer programs and text written in natural languages, which are the products of minds" [9].

Followers of the evolutional paradigm using it, come to wrong conclusions. Many try to prove that information can be created randomly. For example, *R. Dawkins* "proves" it like this: "(R. Dawkins) imagines that monkeys are given a target sentence, e.g. Shakespeare's sentence: "Methinks it is like a weasel". This sentence's length is 28 letters (counting space as a letter as well). The probability that the first letter of this sentence will be right by random typing, is $1 : 27$, and that the first two letters will be right, $1 : 27^2$ etc. So the probability that we get the right sentence by typing randomly, is $1 : 27^{28} \approx 1 : 10^{40}$, stunningly small. Now Dawkins has a solution to the problem. Every time a monkey types a letter, a computer (or the main monkey, as *David Berlinski* mathematician humorously notes!) compares it with the target sentence, and keeps every correct letter which the monkeys type. This way they get to the target sentence quickly – in Dawkins' simulation in 43 steps" [5].

Dawkins contradicts himself: in his book [7] he states, that evolution is *blind and without a destination*, information comes to existence spontaneously, but in the computer simulation he uses two "forbidden" mechanisms – he *compares* the result with the *target sentence*, and *stores* the correct *result*.

That's the typical faulty logic, and the result of a reasoning disfigured by the evolutional paradigm. Concepts, like *target, comparison, storing, correct result* are meaningless without the mind. With this attitude people like Dawkins deceive themselves, as well as the reader. The evolutional paradigm is untrue, and that's why a change of paradigms is needed.

References

[1] Wiener N.: Cybernetics. Control and Communication in the Animal and Machine. Cambridge, MIT Press, 1948.

[2] Gitt W.: Kezdetben volt az információ, Evangéliumi Kiadó, 1998.

[3] C.E. Shannon, W.Weaver: A kommunikáació matematikai elmélete, OMIKK, Budapest, 1986.

[4] Steinbuch K.: Falsch programmiert, Deutsche Verlags-Anstalt, Stuttgart, 1969.

[5] Gooding D.W., Lennox J.C.: Küzdelem az élet értelméért, Evangéliumi Kiadó, 2001.

[6] Denton M.: Evolution: A Theoty in Crisis, Bethesda, Marylend.: Adler and Adler, 1986.

[7] Dawkins R.: A vak óramester, Arfdémia Kiadó – Mezőgazdasági Kiadó, 1994.

[8] Junker R., Scherer S.: Entstehung Gesiche Der Lebewesen, Weyel, 1986.

[9] Meyer St.: The Return of the God Hypothesis, Seattle, Discovery Institute Center for the Renewal of Science and Culture, 1998.

[10] Morris H.: Kreacionizmus: a teremtéselmélet. Budapest. KI Alapítvány. 2000.