TEKA. COMMISSION OF MOTORIZATION AND ENERGETICS IN AGRICULTURE – 2012, Vol. 12, No.4, 210-217

Outline of a theory of semantic information and misinformation

Oleg Pogorelov

Volodymyr Dahl East-Ukrainian National University, Lugansk, Ukraine

Summary. The update version of semantic information theory, in which the not bona fide source of messages is considered and the absurdness of some sentences is assumed, is presented. Ternary logic is used. This logic includes along with the values of truth: "true" and "false" – also "absurd" value with respect to nonsense sentences. The quantities of semantic information and of misinformation, which are contained in messages, are defined.
Key words. Quantity of semantic information, logic of predicates, probability

INTRODUCTION

Mathematical theory of communication [1] developed in 1948 by Claude Shannon stimulated the development of the semantic theory of information. In 1952 the first work on this subject by Bar-Hillel and Carnap in a form of a report on investigations made [2] appeared. It was later published in Bar-Hillel's book [3]. It operated the Leibniz's idea of the plurality of "possible worlds" which exist because of our lack of awareness of the "real world" realities. Receiving new information, we reject some "possible worlds" which do not correspond to it and, thus, we narrow the number of possible combinations. The number of rejected "worlds" is used as a measure of the information message content.

Let us assume that we use the language system Λ_2^2, which contains two individual constants a and b, and two predicates, a one-place predicate M and a two-place predicate L. The first predicate may have the meaning "man", and the second predicate may mean "loves". We will denote the inversion $\neg M$ with the symbol W, "woman". The table 1 represents all sixteen descriptions of "possible worlds" within the specified language system.

Individual constant a may be interpreted as "Tom is a man", and the constant b like "Mary is a woman". And moreover, "Tom loves Mary" and "Mary loves Tom". Then the combination of the number 3 in the table 1 corresponds to the "real world". All other worlds are "false". It is necessary to get some semantic information contained in the sentences to understand it. Each of 16 combinations is a model of the world created by means of mathematical logic and constitutes a conjunction of four atomic statements. Bar-Hillel and Carnap called such atomic sentences (and their inversions) the *basis* sentences, similar to the system of coordinate vectors. According to the established tradition [4, 5], we denote the plurality of all "possible worlds" with a letter W.

Table 1. Descriptions of "possible worlds"

1.	$M(a) \wedge M(b) \wedge L(a,b) \wedge L(b,a)$
2.	$W(a) \wedge M(b) \wedge L(a,b) \wedge L(b,a)$
3.	$M(a) \wedge W(b) \wedge L(a,b) \wedge L(b,a)$
4.	$W(a) \wedge W(b) \wedge L(a,b) \wedge L(b,a)$
5.	$M(a) \wedge M(b) \wedge \neg L(a,b) \wedge L(b,a)$
6.	$W(a) \wedge M(b) \wedge \neg L(a,b) \wedge L(b,a)$
7.	$M(a) \wedge W(b) \wedge \neg L(a,b) \wedge L(b,a)$
8.	$W(a) \wedge W(b) \wedge \neg L(a,b) \wedge L(b,a)$
9.	$M(a) \wedge M(b) \wedge L(a,b) \wedge \neg L(b,a)$
10.	$W(a) \wedge M(b) \wedge L(a,b) \wedge \neg L(b,a)$
11.	$M(a) \wedge W(b) \wedge L(a,b) \wedge \neg L(b,a)$
12.	$W(a) \wedge W(b) \wedge L(a,b) \wedge \neg L(b,a)$
13.	$M(a) \wedge M(b) \wedge \neg L(a,b) \wedge \neg L(b,a)$
14.	$W(a) \wedge M(b) \wedge \neg L(a,b) \wedge \neg L(b,a)$
15.	$M(a) \wedge W(b) \wedge \neg L(a,b) \wedge \neg L(b,a)$
16.	$W(a) \wedge W(b) \wedge \neg L(a,b) \wedge \neg L(b,a)$

In fact, the authors of the work [2] used the idea of Shannon about information as a removed uncertainty [1], although this idea is not expressed explicitly in their article. At the initial moment, when there is no information about the realities of the "real world", we have 16 alternatives, so the uncertainty is high. When we obtain the most possible information, we have only one option, and the uncertainty is equal to zero. According to the formula of R. Hartley [6]

$$\inf(\sigma_3) = \log_2(16) - \log_2(1) = 4\ bit,$$

where: $\inf(\sigma_3)$ is the amount of semantic information contained in the message σ_3 (it is written in the third line of the table 1).

Suppose that another, less informative message, for instance: "$M(a) \vee M(b)$", came instead of the most informative message σ_3. On the basis of the information contained in this sentence, we cross off from table 1 four "possible worlds": σ_4, σ_8, σ_{12} and σ_{16}. The amount of the obtained semantic information makes then

$$\inf(M(a) \vee M(b)) = \log_2(16) - \log_2(12) = \log_2 \frac{4}{3}\ bit.$$

Thus, the message σ_3 has in about ten times more information than the message "$M(a) \vee M(b)$". In the case of an arbitrary message σ we have:

$$\inf(\sigma) = \log_2 \frac{1}{m(\sigma)} = \log_2 \frac{1}{1 - cont(\sigma)}, \quad (1)$$

where: $cont(\sigma) \le 1$ is *the measure of the content of the message* [2, 3] (the quotient obtained by dividing the number of the descriptions of the "possible worlds", inconsistent with σ, by the total number of "possible worlds"), $m(\sigma) = 1 - cont(\sigma)$. If we know $cont(\sigma)$, then we can clearly define $\inf(\sigma)$, and vice versa. For example:

$$cont(\sigma_3) = \frac{15}{16}, \qquad cont(M(a) \vee M(b)) = \frac{4}{16} = \frac{1}{4}.$$

The disadvantage of the theory of Bar-Hillel and Carnap is that the logical contradiction in their concept contains an infinite amount of information, because it does not match any of the "possible worlds". This fact is counterintuitive and is known as the "Bar-Hillel-Carnap paradox".

For this reason, other theories appeared, in which the number of semantic information is determined otherwise. A.A. Kharkevich proposed to measure the value of information through changing the possibility of achieving a certain goal, arising under the influence of the received message. [7] In Y.A. Schreider's work it is proposed to estimate semantic information as the degree of changing the system of knowledge (thesaurus) of an addressee as the result of the perception of the received message. [8] Although this concept had some success, and continues to be popular even in modern times [9], the insufficient formalization of the concept of thesaurus makes this theory too vague and does not allow to subject it to mathematical analysis.

Luciano Floridi in his work [10] presented his theory of "strongly semantic information" (this concept is developed in his later articles [11, 12, 13]). In this paper he calls the theory of Bar-Hillel and Carnap the theory of "weakly semantic information". Floridi introduced the function $\vartheta(\sigma)$, which describes the difference between the message σ and the real situation. This function changes within the interval [-1.0, +1.0]. The negative values of this function describe the degree of the "inaccuracy" of the received message, and the positive functions describe the degree of its "emptiness". He also introduced the function of the "degree of the message information value".

$$\iota(\sigma) = 1 - \vartheta^2(\sigma). \quad (2)$$

The amount of the meaningless information in the message σ can be defined using the formula

$$\vartheta^*(\sigma) = \frac{\int_0^{\vartheta} \iota(\sigma)dx}{1 sbit}, \quad (3)$$

where: $1 sbit = \int_0^1 \iota(\sigma)dx = \log_2 \frac{2}{3} = 1 bit - 1 trit$.

And the amount of semantic information in a message σ is defined by the formula

$$\iota^*(\sigma) = 1 sbit - \vartheta^*(\sigma). \quad (4)$$

Since the values of the function $\vartheta(\sigma)$ for logically controversial sentences are equal to "-1", then according to the formula (4) they do not contain semantic information:

$$\iota^* = 1 sbit - \frac{\int_{-1}^{0} \iota(\sigma)dx}{1 sbit} = 0.$$

Thus, the Bar-Hillel-Carnap paradox is solved under the new concept. However, the approach of Floridi is not notable for its clarity and simplicity typical for the classical theory. The article [14] says on this subject: "Unlike Hartley and C.E. Shannon, who tried to apply simple models of Boolean algebra for the description of information, L. Floridi uses more complex mathematical models ...".

It should be noted, that the paradox of Bar-Hillel and Carnap is not the only problem of the theory of semantic information. One of them is related to the possibility of misinformation. The fact is that the source of the message in classical works [2, 10] is interpreted as the conscientious (*bona fide* source of information). It means that only true messages are accepted, what does not always correspond to the real situation. Another problem refers to the completeness of the description of the "possible worlds".

Let us consider a rather complex language system Λ_n^π containing n number of individual constants (objects, events or locations) and π number of predicates of different arity. To take into account all the possible relationships between the constants in the description of the possible state of the world it is necessary to use the following number of atomic sentences:

$$l = \sum_{k=0}^{n} \pi_k \frac{n!}{(n-k)!}, \qquad (5)$$

where: π_k is the number of k- place predicates in the language system Λ_n^π.

Naturally, not all l of basic sentences will make sense. For example, if in the simple language system Λ_2^2 the individual constant a is interpreted not as "Tom", but as "brick" (building material), then the expression $M(a)$ "A brick is a man" confuses. Let us assume, that it is false. Then its inversion $W(a)$ "A brick is a woman" is a false sentence as well. It leads to breaking the law of the excluded middle.

Of course, every adult knows that a woman and a man are people and a brick is not a man. Then using the method of deduction it is easy to conclude the absurdity of the sentences "A brick is a man" and "A brick is a woman". However, it should be remembered, that in the real life the most of the knowledge (especially at the initial stage) is acquired through the induction. Since often we can not know in advance (before obtaining the relevant information), which sentences make sense, and which do not, we can not exclude a priori the meaningless sentences from the descriptions of the possible states of the world.

Thus, the question arises: what truth value should be ascribed to a meaningless sentence?

THE AIM OF THE WORK

The paper presents a new version of the theory of semantic information, which consideres not necessarily a conscientious message source and in which senselessness of some atomic and molecular sentences is conceded. The theory uses the ternary logic, including, along with the traditional truth values "true" and "false", an additional value "absurd" applied to the meaningless sentences. The definition of the amount of semantic information and misinformation contained in messages is given in the work.

TERNARY LOGIC

Why do children love absurd statements so much and Lewis Carroll and other writers indulge them in this? Because such statements help to understand better, what is allowed in this world, and what is forbidden. Moreover, absurd sentences bear fascination, i.e. they are emotionally appealing [15, 16].

Having assumed the existence of meaningless sentences, let us assign them a new truth value, "*absurd*". In this case, all of the classic definitions and equivalences of binary logic (with the exception of the law of the excluded middle) remain in force. Let φ and ψ be the different sentences of ternary logic. Let us formulate five axioms that seem to be intuitively obvious.

Axioms

1. For each φ only one of three expressions: "$\varphi \equiv true$", "$\varphi \equiv false$", "$\varphi \equiv absurd$" is true, the other two are false.
2. $(\varphi \equiv absurd) \to (\neg\varphi \equiv absurd)$;
3. $((\varphi \equiv absurd) \wedge (\psi \equiv absurd)) \to (\varphi \wedge \psi \equiv absurd)$;
4. $((\varphi \equiv absurd) \wedge (\psi \equiv true)) \to (\varphi \wedge \psi \equiv absurd)$;
5. $((\varphi \equiv absurd) \wedge (\psi \equiv true)) \to (\varphi \vee \psi \equiv true)$.

Based on these axioms, we can prove the following theorems.

Theorem 1.

$((\varphi \equiv absurd) \wedge (\psi \equiv false)) \to (\varphi \wedge \psi \equiv false)$.

Proof: If $(\varphi \equiv absurd) \wedge (\psi \equiv false)$, then, based on the axioms 2 and 5 (and using the law of De Morgan), we can write: $\neg(\varphi \wedge \psi) \equiv \neg\varphi \vee \neg\psi \equiv true$. Therefore: $\varphi \wedge \psi \equiv false$.

Theorem 2.

$((\varphi \equiv absurd) \wedge (\psi \equiv false)) \to (\varphi \vee \psi \equiv absurd)$.

Proof: If $(\varphi \equiv absurd) \wedge (\psi \equiv false)$, then, based on the axioms 2 and 4 (and using the law of De

Morgan), we can write: $\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi \equiv absurd$. Therefore: $\varphi \vee \psi \equiv absurd$.

Theorem 3.
$((\varphi \equiv absurd) \wedge (\psi \equiv absurd)) \rightarrow (\varphi \vee \psi \equiv absurd)$.

Proof: If $(\varphi \equiv absurd) \wedge (\psi \equiv absurd)$, then, based on the axioms 2 and 3 (and using the law of De Morgan), we can write: $\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi \equiv absurd$. Therefore: $\varphi \vee \psi \equiv absurd$.

Theorem 4.
$((\varphi \equiv absurd) \wedge (\psi \equiv true)) \rightarrow (\varphi \rightarrow \psi \equiv true)$.

Proof: If $(\varphi \equiv absurd) \wedge (\psi \equiv true)$, then, based on the axioms 2 and 5 (and using the known equivalence), we can write: $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi \equiv true$.

Theorem 5.
$((\varphi \equiv true) \wedge (\psi \equiv absurd)) \rightarrow (\varphi \rightarrow \psi \equiv absurd)$.

Proof: If $(\varphi \equiv true) \wedge (\psi \equiv absurd)$, then, based on the axiom 2 and the theorem 2 (and using the known equivalence), we can write: $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi \equiv absurd$.

Theorem 6.
$((\varphi \equiv absurd) \wedge (\psi \equiv false)) \rightarrow (\varphi \rightarrow \psi \equiv absurd)$.

Proof: If $(\varphi \equiv true) \wedge (\psi \equiv false)$, then, based on the axiom 2 and the theorem 2 (and using the known equivalence), we can write: $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi \equiv absurd$.

Theorem 7.
$((\varphi \equiv false) \wedge (\psi \equiv absurd)) \rightarrow (\varphi \rightarrow \psi \equiv true)$.

Proof: If $(\varphi \equiv false) \wedge (\psi \equiv absurd)$, then, based on the axioms 2 and 5 (and using the known equivalence), we can write: $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi \equiv true$.

Theorem 8.
$((\varphi \equiv absurd) \wedge (\psi \equiv absurd)) \rightarrow (\varphi \rightarrow \psi \equiv absurd)$.

Proof: If $(\varphi \equiv absurd) \wedge (\psi \equiv absurd)$, then, based on the axiom 2 and the theorem 3, we can write: $\varphi \rightarrow \psi \equiv \neg\varphi \vee \psi \equiv absurd$.

Definition 1. We will define an equivalence in ternary logic as follows:

$$(\varphi \equiv \psi) \equiv ((\varphi \equiv true) \wedge (\psi \equiv true)) \vee$$
$$\vee ((\varphi \equiv false) \wedge (\psi \equiv false)) \vee$$
$$\vee ((\varphi \equiv absurd) \wedge (\psi \equiv absurd)).$$

From this definition follows, that expression $\varphi \equiv \psi$ is true only in those cases, when the values of truth of sentences φ and ψ coincide. It is false in all other cases. Sentense $\varphi \equiv \psi$ can not be absurd.

On the basis of the accepted axioms and the proved theorems it is possible to make the truth table of ternary logic (table 2), in which the sign "1" corresponds to the value of truth "$true$", the sign "0" corresponds to the value "$false$" and the sign "-1" corresponds to the value "$absurd$".

Table 2. The truth table of ternary logic

φ	ψ	$\neg\varphi$	$\varphi \wedge \psi$	$\varphi \vee \psi$	$\varphi \rightarrow \psi$	$\psi \rightarrow \varphi$	$\varphi \equiv \psi$
0	0	1	0	0	1	1	1
0	1	1	0	1	1	0	0
0	-1	1	0	-1	1	-1	0
1	0	0	0	1	0	1	0
1	1	0	1	1	1	1	1
1	-1	0	-1	1	-1	1	0
-1	0	-1	0	-1	-1	1	0
-1	1	-1	-1	1	1	-1	0
-1	-1	-1	-1	-1	-1	-1	1

THE SET OF THE POSSIBLE WORLDS

The cardinal number of the "possible worlds" set is equal to $|W| = s^l$, where s is the number of truth-values in the used logic [10]. The description of a separate world (its logical model) is the conjunction, containing l of basic sentences. The basic sentence will be denoted by the Greek letter β_{ij}, the index i denotes the number of the "possible worlds", and the index j denotes the sequence number of the sentence in a conjunction. State-descriptions are presented in the table 3.

Table 3. State-descriptions

$$\beta_{11} \wedge \beta_{12} \wedge \cdots \wedge \beta_{1l}$$
$$\beta_{21} \wedge \beta_{22} \wedge \cdots \wedge \beta_{2l}$$
$$\beta_{31} \wedge \beta_{32} \wedge \cdots \wedge \beta_{3l}$$
$$\beta_{41} \wedge \beta_{42} \wedge \cdots \wedge \beta_{4l}$$
$$\beta_{51} \wedge \beta_{52} \wedge \cdots \wedge \beta_{5l}$$
$$\beta_{61} \wedge \beta_{62} \wedge \cdots \wedge \beta_{6l}$$
$$\beta_{71} \wedge \beta_{72} \wedge \cdots \wedge \beta_{7l}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\beta_{s^l 1} \wedge \beta_{s^l 2} \wedge \cdots \wedge \beta_{s^l l}$$

Basic sentences with the same index j consist of the totality of periodically recurring s sentences. In ternary logic the first of them asserts the truth of a certain atomic sentence (which is denoted as α_j), the second sentence asserts its falsity and the third one asserts its absurdity. Using numbers 1, 0, -1 to indicate the truth, the falsity and the absurdity of the sentence α_j respectively, we can make an image of the descriptions of the world's states as a set of pseudo-ternary codes (for convenience presented in a specular reflection). Now β_{11} indicates the sentence "$\alpha_1 \equiv true$", β_{21} indicates the sentence "$\alpha_1 \equiv false$", and β_{31} indicates the sentence "$\alpha_1 \equiv absurd$".Then the truth values are repeated: β_{41} indicates the sentence "$\alpha_1 \equiv true$", β_{51} indicates "$\alpha_1 \equiv false$", β_{61} indicates "$\alpha_1 \equiv absurd$" and so on. In the second column of the model structure the first three sentences β_{12}, β_{22}, β_{32} state the truth of the other atomic sentence α_2, the next three sentences state its falsity and the following three sentences state its absurdity, etc. We shall call *the core of basic sentences* β_{ij} $(1 \le i \le s^l)$ the atomic sentence α_j for the fixed j. Though the core can take three values of the truth: "*true*", "*false*" and "*absurd*", all basic sentences are either true or false.

THE PROCESS OF COGNITION

Let us assign to the i-th "possible world" $(1 \le i \le s^l)$ the assessment of the probability $p_i(t)$, that for $t \to \infty$ it will be the "real world", made at the moment of time t. The time t is measured with the number of the messages received (some of them may be repeated). The messages may be both "external" and "internal". The "external" come from the outside world, the "internal" are the result of the deductive conclusion made on the basis of the earlier obtained information. Introducing probabilities into the logical system, we come to the probabilistic logic, in which multiplying of the probabilities corresponds to conjunction, and *comultiplication* corresponds to disjunction [17-22]. So we can summarize

$$p_i(t) = \prod_{j=1}^{l} p_{ij}(t) \qquad (6)$$

where: $p_{ij}(t)$ is the made in the moment of time t assessment of the probability, that for $t \to \infty$, the basic sentence β_{ij} is true.

Definition 2. The set W with the given on it probabilities $p_{ij}(t)$ will be called *thesaurus.*

At the initial moment of time, when the uncertainty is maximal, for atomic sentences presented in the conjunctions of the set W we have:

$$p_{ij}(0) = \frac{1}{s}\ , \qquad (7)$$

and for a separate "possible world" we have

$$p_i(0) = \frac{1}{s^l}\ . \qquad (8)$$

Upon the receipt of the message σ (atomic or molecular), some estimated probabilities from the general totality $p_{ij}(t)$ will change. Changes can occur in the following algorithm.

Algorithm 1. Let us denote the estimated probability, that for $t \to \infty$ the sentence σ will be true, as $p_t(\sigma,t)$; the estimated probability, that for $t \to \infty$ it will be false, as $p_f(\sigma,t)$; the estimated probability, that for $t \to \infty$ it will be absurd, as $p_a(\sigma,t)$. Let $m_t(\sigma,t)$, $m_f(\sigma,t)$ and $m_a(\sigma,t)$ be the number of the messages indicating the truth, the falsity and the absurdity of the sentence σ respectively. Then for $t > 0$ we have:

$$p_t(\sigma,t) = \frac{m_t(\sigma,t)}{m_t(\sigma,t)+m_f(\sigma,t)+m_a(\sigma,t)+m_{tf}(\sigma,t)}, \qquad (9)$$

$$p_f(\sigma,t) = \frac{m_f(\sigma,t)}{m_t(\sigma,t)+m_f(\sigma,t)+m_a(\sigma,t)+m_{tf}(\sigma,t)}, \qquad (10)$$

$$p_a(\sigma,t) = \frac{m_a(\sigma,t)+m_{tf}(\sigma,t)}{m_t(\sigma,t)+m_f(\sigma,t)+m_a(\sigma,t)+m_{tf}(\sigma,t)}, \qquad (11)$$

where: $m_{tf}(\sigma,t) = \begin{cases} m_t(\sigma,t),\ if\ m_f(\sigma,t) \ge m_t(\sigma,t), \\ m_f(\sigma,t),\ if\ m_t(\sigma,t) \ge m_f(\sigma,t). \end{cases}$

The variable $m_{tf}(\sigma,t)$ is used, because the sentence «$\sigma \wedge \neg\sigma \equiv true$» is not true. If the sentence σ is basic, then using formulas (9), (10) and (11) we can directly calculate estimated probabilities $p_{ij}(t)$.

Algorithm 2. Algorithm 1 can be complicated. The process of cognition has the subjective nature. Let us suppose that the subject A at the moment t_1 has received not too many messages and does not have much trust in his assessments of probabilities. He prefers to use the estimates of the subject B, who has received much more messages and therefore has more experience. More precisely, the subject A leads a "double bookkeeping", that is, he continues to count messages and calculate probabilities according to the formulas (9), (10), (11), but "for the time being" he prefers somebody else's data. Over time, his trust in his estimates grows and he starts using his own experience more often.

The presented in this paper theory of semantic information is based on the following hypothesis.

Hypothesis. The random process of changing the estimated probabilities $p_{ij}(t)$, done according to the algorithm 1 or the algorithm 2, for $t \to \infty$ asymptotically leads to the fact that for some natural number i the following equality will be satisfied: $\lim_{t\to\infty} p_{ij}(t) = 1$ for all j over the range $(1 \le j \le l)$.

The hypothesis states that the cognition is possible and is not restricted. It is clear that certain conditions are necessary to accomplish it, the received messages should be diverse and possess a certain degree of trustworthiness, and the subject of cognition should have logic abilities. We will assume that these requirements are fulfilled and the algorithm tallies.

The peculiarity of our case is that the law of the excluded middle is replaced by the law of the excluded fourth:

$$p_t(\sigma,t) + p_f(\sigma,t) + p_a(\sigma,t) = 1. \quad (12)$$

Next expressions are thus just

$$p_t(\neg\sigma) = p_f(\sigma), \quad (13)$$

$$p_f(\neg\sigma) = p_t(\sigma), \quad (14)$$

$$p_a(\neg\sigma) = p_a(\sigma). \quad (15)$$

Let us suppose that σ is a molecular sentence obtained from the two sentences φ and ψ (atomic or molecular) connected by logical connectives. Formulas (17-27) on which it is possible to expect probabilities of truth ensue from formulas (12-16), to falsity and absurdity of the compound sentences by the known probabilities of truth, falsity and absurdity of the sentences φ and ψ (the sign of the argument t is omitted everywhere). If $p_a(\bullet) = 0$, they are converted into regular expressions of the probabilistic logic.

$$p_t(\varphi \wedge \psi) = p_t(\varphi) \cdot p_t(\psi \mid \varphi), \quad (16)$$

$$p_f(\varphi \wedge \psi) = \\ = p_f(\varphi) + p_f(\psi \mid \varphi) - p_f(\varphi) \cdot p_f(\psi \mid \varphi), \quad (17)$$

$$p_a(\varphi \wedge \psi) = \\ = p_a(\varphi) + p_a(\psi \mid \varphi) - p_a(\varphi) \cdot p_a(\psi \mid \varphi) - \\ - p_f(\varphi) \cdot p_a(\psi \mid \varphi) - p_a(\varphi) \cdot p_f(\psi \mid \varphi), \quad (18)$$

$$p_t(\varphi \vee \psi) = \\ = p_t(\varphi) + p_t(\psi \mid \neg\varphi) - p_t(\varphi) \cdot p_t(\psi \mid \neg\varphi), \quad (19)$$

$$p_f(\varphi \vee \psi) = p_f(\varphi) \cdot p_f(\psi \mid \neg\varphi), \quad (20)$$

$$p_a(\varphi \vee \psi) = p_f(\varphi) \cdot p_a(\psi \mid \neg\varphi) + \\ + p_a(\varphi) \cdot p_f(\psi \mid \neg\varphi) + p_a(\varphi) \cdot p_a(\psi \mid \neg\varphi), \quad (21)$$

$$p_t(\varphi \to \psi) = 1 - [p_t(\varphi) + p_a(\varphi)] \cdot [1 - p_t(\psi \mid \varphi)], \quad (22)$$

$$p_f(\varphi \to \psi) = p_t(\varphi) \cdot p_f(\psi \mid \varphi), \quad (23)$$

$$p_a(\varphi \to \psi) = p_t(\varphi) \cdot p_a(\psi \mid \varphi) + \\ + p_a(\varphi) \cdot p_f(\psi \mid \varphi) + p_a(\varphi) \cdot p_a(\psi \mid \varphi), \quad (24)$$

$$p_t(\varphi \equiv \psi) = p_t(\varphi) \cdot p_t(\psi \mid \varphi) + \\ + p_f(\varphi) \cdot p_f(\psi \mid \neg\varphi) + p_a(\varphi) \cdot p_a(\psi \mid \varphi), \quad (25)$$

$$p_f(\varphi \equiv \psi) = 1 - p_t(\varphi \equiv \psi), \quad (26)$$

$$p_a(\varphi \equiv \psi) = 0. \quad (27)$$

Let us define the uncertainty of the message recipient at the moment of time t as an entropy:

$$H(t) = \sum_{i=1}^{s^l} p_i(t) \log_s \frac{1}{p_i(t)}, \quad (28)$$

$$H(0) = \sum_{i=1}^{s^l} \frac{1}{s^l} \log_s s^l = l. \quad (29)$$

Definition 3. The volume of thesaurus $V_t(t)$ at the moment of time t is equal to:

$$V_t(t) \overset{Def}{=} H(0) - H(t). \quad (30)$$

Definition 4. The change of the thesaurus volume $\Delta V_t(t)$ at the moment of time t is equal to:

$$\Delta V_t(t) \overset{Def}{=} H(t-1) - H(t). \quad (31)$$

The value of the change of the thesaurus volume $\Delta V_t(t)$ corresponds to the definition of the amount of semantic information by Schreider [8]. The disadvantage of this definition is that in the conditions of possible misinformation (which leads to the necessity of the correction of the thesaurus, when the falsehood is exposed), the quantity $\Delta V_t(t)$ may take both positive and negative values. Also, $\Delta V_t(t)$ depends not only on the content of the message, but on the thesaurus state at moment of time $t-1$. Bar-Hillel and Carnap (together with Floridi) try to give an objective meaning to the definition of the amount of semantic information. The classical definition assumes that the recipient of information before obtaining the tested message σ is extremely naive, and the volume of his thesaurus is equal to zero (or the message content does not depend on the content of the previously received messages). We also will follow this approach.

Let us assume that the sentence σ is true. Let m be the number of the possible worlds, which do not contradict this sentence. Then the value of the uncertainty removed by this sentence is equal to $\log_s s^l/m$. However, the assumption of the truth of the sentence σ requires confirmation. If the assumption is not confirmed, there will be no removed uncertainty.

On the other hand, the ratio m/s^l can be treated as an a priori estimated probability of the truth of the sentence σ. In the case of the basic sentence this interpretation is equivalent to the formula (7), in the case of the conjunction of basic sentences it is equivalent to the formula (8).

THE AMOUNT OF SEMANTIC INFORMATION AND MISINFORMATION

In definition of the amount of semantic information and misinformation in the message we will be guided by the following principles:

- a true message has a certain amount of semantic information;
- a false message has a certain amount of misinformation;
- an absurd sentence has neither semantic information nor misinformation;
- the amount of semantic information in a true message is proportional to the value of the uncertainty, removed in the result of the receiving of this message.

Definition 5. The amount of semantic information $\inf(\sigma)$ contained in the message σ equals to:

$$\inf(\sigma) \overset{Def}{=} \lim_{t\to\infty} p_t(\sigma,t)\cdot\log_s\frac{1}{p_t(\sigma,0)}, \qquad (32)$$

where: $p_t(\sigma,t)$ is the estimated probability of the truth of the sentence σ at the moment of time t.

Definition 6. The amount of misinformation $\text{mis}(\sigma)$ contained in the message σ equals to:

$$\text{mis}(\sigma) \overset{Def}{=} \lim_{t\to\infty} p_f(\sigma,t)\cdot\log_s\frac{1}{p_t(\sigma,0)}, \qquad (33)$$

where: $p_f(\sigma,t)$ is the estimated probability of the falsity of the sentence σ at the moment of time t.

SOME RESULTS OF THE THEORY

Let us consider several simple examples. It is easy to ensure that a true atomic sentence has one unit of semantic information and no misinformation. A false atomic sentence has no semantic information and one unit of misinformation. Using the data of the table 2 (or table 4), we can calculate that the true sentence "$\alpha_1\wedge\alpha_2$" ("α_1" and "α_2" are atomic sentences) has two units of semantic information. The true sentences "$\alpha_1\vee\alpha_2$" and "$\alpha_1\to\alpha_2$" both have $\log_3 9/5\approx 0{,}535$ units of semantic information, and the true sentence "$\alpha_1\equiv\alpha_2$" has one unit of semantic information.

Now we will consider tautology: $\varphi\vee\neg\varphi$. Suppose in the beginning, that sentence φ either truly or falsely (but not absurdly). Then according to a formula (19) for any t $(0\le t<\infty)$:

$$\begin{aligned}&p_t(\varphi\vee\neg\varphi,t)=\\&=p_t(\varphi,t)+p_t(\neg\varphi\mid\neg\varphi,t)+p_t(\varphi,t)\cdot p_t(\neg\varphi\mid\neg\varphi,t)=\\&=p_t(\varphi,t)+1-p_t(\varphi,t)=1,\end{aligned}$$

$$p_f(\varphi\vee\neg\varphi,t)=p_a(\varphi\vee\neg\varphi,t)=0,$$

$$\begin{aligned}&\inf(\varphi\vee\neg\varphi)=\\&=\lim_{t\to\infty}p_t(\varphi\vee\neg\varphi,t)\cdot\log_s\frac{1}{p_t(\varphi\vee\neg\varphi,0)}=\log_s\frac{1}{1}=0,\end{aligned}$$

$$\begin{aligned}&\text{mis}(\varphi\vee\neg\varphi)=\\&=\lim_{t\to\infty}p_f(\varphi\vee\neg\varphi,t)\cdot\log_s\frac{1}{p_t(\varphi\vee\neg\varphi,0)}=0\cdot\log_s\frac{1}{1}=0.\end{aligned}$$

If sentence φ is absurd, then

$$p_t(\varphi\vee\neg\varphi,t)=p_f(\varphi\vee\neg\varphi,t)=0,\ p_a(\varphi\vee\neg\varphi,t)=1.$$

In this case we get the same result:

$$\inf(\varphi\vee\neg\varphi)=\text{mis}(\varphi\vee\neg\varphi)=0.$$

Thus, tautology carries neither information nor misinformation, and absurd sentence reminds this property.

THE SOLUTION OF THE BAR-HILLEL-CARNAP PARADOX

Now we will consider contradiction $\varphi\vee\neg\varphi$. Suppose in the beginning, that sentence φ either truly or falsely (but not absurdly). Then according to a formulas (16) and (17) for any t $(0\le t<\infty)$:

$$\begin{aligned}&p_t(\varphi\wedge\neg\varphi,t)=\\&=p_t(\perp,t)=p_t(\varphi,t)\cdot p_t(\neg\varphi\mid\varphi,t)=p_t(\varphi,t)\cdot 0=0,\end{aligned}$$

$$\begin{aligned}&p_f(\perp,t)=\\&=p_f(\varphi)+p_f(\neg\varphi\mid\varphi)-p_f(\varphi)\cdot p_f(\neg\varphi\mid\varphi)=\\&=p_f(\varphi)+p_t(\varphi\mid\varphi)-p_f(\varphi)\cdot p_t(\varphi\mid\varphi)=1,\end{aligned}$$

$$\begin{aligned}&\inf(\perp)=\\&=\lim_{t\to\infty}p_t(\perp,t)\cdot\log_s\frac{1}{p_t(\perp,0)}=p_t(\perp,0)\cdot\log_s\frac{1}{p_t(\perp,0)}=\\&=\lim_{x\to 0}\left[x\cdot\log_s\frac{1}{x}\right]=0,\end{aligned}$$

$$\begin{aligned}&\text{mis}(\perp)=\\&=\lim_{t\to\infty}p_f(\perp,t)\cdot\log_s\frac{1}{p_t(\perp,0)}=\log_s\frac{1}{p_t(\perp,0)}=\infty.\end{aligned}$$

If sentence φ is absurd, then

$$p_t(\varphi\wedge\neg\varphi,t)=p_f(\varphi\wedge\neg\varphi,t)=0,\ p_a(\varphi\wedge\neg\varphi,t)=1.$$

In this case we get the same result:

$$\inf(\varphi\vee\neg\varphi)=\text{mis}(\varphi\vee\neg\varphi)=0.$$

Thus, contradiction does not carry semantic information, however, if φ is not absurd sentence, contradiction carries an infinite amount of misinformation. Although this conclusion is some unexpected, he does not conflict with intuition. A role of contradiction in logic and in science is such, that, taking its truth for a true, we destroy all bases of thought.

CONCLUSIONS

In this paper we tried to strengthen the Bar-Hillel and Carnap's theory of "weakly semantic information", which, despite the well-known paradox, is attractive for its simplicity and in the framework of which a lot of interesting results have been already received. The use of such new truth-value as "absurd" actually leads to the division of the set of all possible sentences into three classes that have fundamentally different properties. True sentences bear semantic information, false sentences bear misinformation and absurd sentences bear neither information nor misinformation. In the update version of theory the contradiction does not carry semantic information. The Bar-Hillel-Carnap paradox is explained by an illogical attempt to ascribe to false sentence a truth value.

REFERENCES

1. **Shannon C.E., 1948.:** A Mathematical Theory of Communication. Bell Syst. Tech. J., 27: 379-423, 623-656.
2. **Bar-Hillel Y., Carnap R., 1952.:** «An Outline of a Theory of Semantic Information», Technical Report No. 247, October 27, Research Laboratory of Electronics. – 49.
3. **Bar-Hillel Y., Carnap R., 1964.:** «An Outline of a Theory of Semantic Information», rep. in Bar-Hillel Y., pp. 221-274.
4. **Kripke S. A., 1963.:** «Semantical Considerations on Modal Logic», Acta Philosophica Fennica 16:83-94.
5. **Kripke S. A., 1963.:** «Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi», Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 9:67-79.
6. **Hartley R.V., 1928.:** «Transmission of Information», Bell Syst. Tech. J., 7:535.
7. **Kharkevich A.A., 1960.:** At the values of information, "Problems of cybernetics", v. 4: 586-601.
8. **Schreider Y.A., 1965.:** About one model of semantic information theory, "Problems of cybernetics", v. 13: 60-68.
9. **Chursin N.N., 2010.:** Concept of thesaurus in the informative picture of the world: monograph. - Luhansk: Publ. "Knowledge". - 305 p.
10. **Floridi L., 2004.:** «Outline of a Theory of Strongly Semantic Information», Minds and Machines, 14(2), 197-222.
11. **Floridi L., 2005.:** «Is Semantic Information Meaningful Data», Philosophy and Phenomenological Research, 70(2), 351-370.
12. **Floridi L., 2006.:** «The Logic of Being Informed», Logique et Analyse, 49(196), 433-460.
13. **Floridi L., 2007.:** «In Defence of the Veridical Nature of Semantic Information», The European Journal of Analytic Philosophy, 3(1), 1-18.
14. **Tsvetkov V. Ya., 2012.:** Semantic Information Units as L. Floridi's Ideas Development, European Researher, Vol.(25), N 7: 1036-1041.
15. **Knorosov Y.V., 1973.:** To the question about the classic signaling// the Basic problems of africanystic. - M.: Science: 324-334.
16. **Muxelyshvily N.L., Schreider Y.A., 1997.:** Information and fascination in a line and indirect communication// of STI, 2, Inf. Processes and systems, № 8: 1-7.
17. **Nilsson, N. J., 1986.:** «Probabilistic logic», Artificial Intelligence 28(1): 71-87.
18. **Jøsang, A. and McAnally, D., 2004.:** «Multiplication and Comultiplication of Beliefs», International Journal of Approximate Reasoning, 38(1), pp.19-51.
19. **Jøsang, A., 2008.:** «Conditional Reasoning with Subjective Logic», Journal of Multiple-Valued Logic and Soft Computing, 15(1): 5-38.
20. **Gerla, G., 1994.:** «Inferences in Probability Logic», Artificial Intelligence 70(1–2):33–52.
21. **Norbert Chamier-Gliszczynski, 2011.:** "Modelling of Traffic Flow in an Urban Transportation System", TEKA, Commission of Motorization and Power Industry in Agriculture. – OL PAN, 2011, 11v.: 12-18.
22. **Norbert Chamier-Gliszczynski, 2011.:** "Selected Aspects of Modelling of Sustainable Urban Transportation System", TEKA, Commission of Motorization and Power Industry in Agriculture. – OL PAN, 2011, 11v.: 19-24.

НАБРОСОК ТЕОРИИ СЕМАНТИЧЕСКОЙ ИНФОРМАЦИИ И ДЕЗИНФОРМАЦИИ

Олег Погорелов

Аннотация. Представлена новая версия теории семантической информации, в которой рассматривается не добросовестный источник сообщений и допускается бессмысленность некоторых предложений. Используется трехзначная логика, включающая наряду со значениями истинности «true» и «false» дополнительного значения «absurd», применяемого по отношению к бессмысленным предложениям. Дается определение количеств семантической информации и дезинформации, содержащихся в сообщениях.
Ключевые слова. Количество семантической информации, логика предикатов, вероятность