The K.E. Shannon and L. Floridi's amount of information

Victor Yakovlevich Tsvetkov

Moscow State Technical University of Radio Engineering, Electronics and Automation MGTU MIREA, Vernadsky Prospekt, 78, Moscow, 119454, Russia

Abstract. Paper compares the estimates of the information amount using the K.E. Shannon L. Floridi's approaches. The difference in these approaches to the content of the term information is shown. The concept of "information design" as an extension of the term "communication" is introduced. The difference in the assessment of the Shannon L. Florida's informativeness is shown. The difference between the concepts of "not knowing" and "ignorance" is shown. It is shown that clear characteristics' values are specified by point numeric values. It is shown that nonclear characteristics' values are designated by interval numerical values. The significance of information units for the interpretation of information structures is revealed. The Shannon massages' range equivalence condition of one meaningful massage by L. Floridi is given. It is shown that the theory of Shannon defines the area that should be called "statistical information theory". It is shown that the L. Floridi's approach defines foundations of the semantic information theory.
[Tsvetkov V.Y. **The K.E. Shannon and L. Floridi's amount of information.** *Life Sci J* 2014;11(11):667-671] (ISSN:1097-8135). http://www.lifesciencesite.com. 123

Keywords: information, information construction, the amount of information, information theory, statistical information theory, semantic theory of information, informativeness, informative measurability, information units

1. Introduction

In the information theory, there are different directions. One approach is based on the work of K. Shannon [1]. In this theory, information is understood as "something" that reduces uninformedness and uncertainty. Such information reduces uncertainty or ignorance. Wiener calls the K.E. Shannon's theory as "statistical theory of the amount of information" [2]. Initially, K.E. Shannon did not speak about his work as about a theory of information. He noted that "*the semantic aspects of the information are not relevant to the technical problems of communication*" [1]. However, in 1964 there was the collaborative work of Shannon and Weaver [3], in which the point of view of "information theory based on the mathematical theory of communication", was developed as the main one.

In the work [3] was considered "three levels" of communication problems: a technical problem, a semantic problem, a problem of an effective messages' transfer. It should be emphasized that it is spoken about a message transfer, and not about the information transfer. The problem of semantics reduces to the question of "How precisely do the transmitted symbols convey the desired meaning?". But the answer to this question was not full and touched the accuracy of the received symbols in comparison with the original ones.

This view of information theory was ambiguously accepted by the scientific community and other work began to appear to reflect different points of view. Another approach has been developed by N. Winner [2], Carnap R. [4], Floridi, L. [5], A.N. Kolmogorov [6], Y. Schrader [7], V.Y. Tsvetkov [8], etc. In this approach, the information is understood as the content of the information structure regardless of its volume, and often consider it in terms of semantics. This theory is often called semantic theory of information. In this theory the information is understood as "something, that contains knowledge and as a result increases the knowledge of the recipient".

Shannon's information reduces "our not knowing", Winner and Florida's information increases "our knowledge". The knowledge increasing and the ignorance reducing are not equivalent; that will be shown below. The term "not knowing" differs from "ignorance". Ignorance means one entity or set element. "Not knowing" denotes the set of entities or many. Dichotomous pair [9] "not knowing - knowledge" is an opposition variables [10] and sets the scale of analysis for the study of the information messages meaningfulness or other information structures. In his theory, Shannon used the concept of "message", that appling to the sciences of information, means "information message".

2. Methods.

2.1. Technologies

The Shannon's theory is connected to the information transmission. It considers one information technology – information transfer.The basic information object of these technologies is the message. In relation to information science, this term

means "an informational message". In reality, the information technologies are more. These are the following technologies: collection, processing, storage and presentation of information. In these technologies the term "information construction" should be used instead of the term "massage". This term is a generalization of the term "message" and suggests the presence of the structure and information models' components in the information construction. Complex information structures are formed from simple elements, which are often called information units. Information structures can be clear and unclear.

2.2. Clearness and unclearness

In "clear logic" and "clear" physics or mathematics one qualitative phenomenon is characterized by a quantitative value. For example, they say that the train passed 250 km; at the current time it is at a speed of 80 km/h; it slows down with the acceleration of 0.1 m/s^2. This means a paradigm: "one factor – a single numeric value". This means that the person is confident and with probability of 1 defines the value of any factor or characteristic**.**

With the absence of confidence or presence of uncertainty a man sets the interval of possible values of such characteristics. For example, the train passed not less then 200 km and not more than 300 km; at the current time it is at a speed from 45 to 85 km/h; it slows down with the acceleration of 0.1 ± 0.01 m/s^2. This corresponds to the paradigm of "one factor – the number of numeric values" or "one factor – the many of numerical values." Consequently, with the presence of uncertainty, to describe the characteristics a range of values is used .

2.3. Florida's Approach

L. Floridi considers the information not only as a tool to remove the uncertainty, like Shannon L. does. Floridi shares and develops the N.Vinera's view, that the most important thing in information is a content-richness and knowledge. He defines it as term of semantics. L. Floridi determines the main criterion for the presence of semantic in information, that is the truth of information from the point of view of the epistemology positions [11]. L. Floridi does not introduce distinguishing criteria between plausible and true knowledge. This weakens his theory. The main weakness of the L.Floridi's theory is the consideration of information as a kind of a homogeneous total.

Mainly the Wiener's approach, and not the Shannon's one, is developed by Luciano Floridi [12], when he made an attempt to introduce a "the General Definition of Information – GDI". He makes the following provisions for this:

σ is an instance of information, understood as a semantic content, if and only if:

- (GDI.1) σ consists of one or more *data*;
- (GDI.2) data in σ are *well formed*;
- (GDI.3) well-formed data in σ are *significant* [12].

A Floridi notes a polysemy and ambiguity of information. He says that the meaning of the information depends on an environment and context.

2. Experiment

To explain the differences between the approaches of the Shannon and Florida let's consider three situations.

3.1. Situation 1

Suppose the point *M* is in the plane and is in the first quadrant. According to the Shannon's theory all massages, which reduce uncertainty, contain information i.e. it is information measurable. Measurable information we call such an information structure, which in addition to qualitative assessment may have one or a number of quantitative assessments, which allow comparing it with other information structures.

These messages (information structures) are as follows: "The point M is not in the second quadrant"; "The point M is not in the third quadrant"; "The point M is not in the fourth quadrant." Each of these messages inform the message recipient and reduces its information uncertainty. However, having received three messages, the recipient within a process of elimination (since there are four quadrants total) determines that the point is in the first quadrant.

According to L. Floridi's theory, true and meaningful message is one "point M is in the first quadrant". All other messages such as "point may be in the first and or second", "point is not found…" are meaningless.

In this example, three messages according to Shannon, reducing uncertainty, are equivalent to one meaningful massage according to L. Floridi. This means that the interval (interval value) of messages according to Shannon are equivalent to one (point value) massage according to L. Floridi.

3.2. Situation 2

We may complicate the experiment. We may divide the plane into 360 sectors. Let the point is conventionally in sector 360. In this case, 359 massages according to Shannon "the point is not in the sector i", i=1 ... 359, will be equivalent to one massage according to L. Floridi "a point is situated in sector 360". By this the difference between "not knowing" and knowledge is emphasized. That is, for n states, the interval of *n*-1 massages, reducing information uncertainty, is equivalent to one

massage, which contains knowledge. The interval of the *n*-1 massages, we call a complete message, since it removes uncertainty information.

Comparing the situation 1 and 2 the following can be stated. According to L. Floridi there will be one information meaningful message in both cases.

Acoording to Shannon a complete number of informational messages will be different for the first and second situation. For the first situation, the number of messages is equal to 3, for the second situation is equal to 359. But these series are equivalent in terms of information.

You can input the condition of messages' number equivalence according to Shannon to a single massage according to L. Florida. Shannon's set of messages will be equivalent to a meaningful massage according to L. Floridi, if it forms a complete series of messages' interval. The term "communication" can be replaced by the term "information construction".

3.3. Situation 3

It is possible to pass from the discrete model to the continuous model. Let a point on the plane has the following coordinates X=3, Y=4. According to L. Floridi one message, let's call it conditionally (A), is true, that is informative "X=3, Y=4", all others are uninformative.

According to K.E. Shannon, infinite number of statements will contain the amount of information, which reduces uncertainty, that is to be informative Shannon. There are three groups of informative massages may be pointed out according to Shannon:

- The first group of X=3, Y= any, except 4;
- The second group X= any, except 3, Y = 4;
- The third group of X= any, except 3, Y = any, except 4.

But in these groups won't be any statement which is equivalent to the statement (A)

3. Discussion

A common thing for the approaches of Shannon and Wiener is an implicit use of information units [13], although in different aspects of consideration. The simplest data message may include a single word. This word can also be regarded as an information unit. Word is the symbol. The essence of the word or its meaning is revealed through the interpretation of the word, through the totality of its interpreting sentences.

Word (as an information unit) according to Shannon contains as much information as it is defined by information symbol capacity characters included in this word. If you rearrange letters in a word, the meaning of it disappears, but according to Shannon the amount information does not change. Consequently, a Shannon's measure defines the amount of messages' information or information construction.

According to L. Floridi, information that contains the word, is determined by the number of its interpreting sentences. The amount of interpreting sentences is determined by the number of semantic units, and not by meaningless symbols. If you rearrange symbols in a words, the meaning of it disappears and this implies the disappearance of interpreting sentences, i.e. the amount of information becomes zero. Consequently, the Wiener and Florida's measure defines the semantics of the message.

Currently, in majority of cases, the amount of information is defined on a basis of the statistical approach, which arose long before the K.E. Shannon's work [14] and developed later [15]. The main reason for this is the easiness of the quantitative assessment obtaining and unwillingness to pass to the complex information models. This is partly due to the fact that the term "information theory" is a general, and the content of this theory is not the general one.

In our opinion it is necessary to distinguish between the concepts: information theory, statistical information theory [1, 3], semantic information theory [4], the general theory of information [12]. The K.E. Shannon's approach should be attributed to the statistical theory of information. The disadvantage of Shannon's theory is the exclusion of the possibility to build the information construction, using the information units. Wiener and Florida's approaches allow doing it.

5. Results

The K.E. Shannon's information theory on explicitly excludes the notion of information units and semantic information units. Information uncertainty is described by the messages' interval and defines "not knowing." Information definition is described by one meaningful message or by full series of messages according to Shannon. These series remove uncertainty and define knowledge.

The Shannon approach solves two problems. He explores the information uncertainty and allows us to estimate the information capacity of information construction. Shannon's approach sets the interval number of information structures for a complete description of a certain event. Shannon's approach sets certain information construction, defining the "not knowing". Shannon's approach can be used to obtain true knowledge, but it is necessary to use the full number of massages' interval, which reduces uncertainty. Information theory approach, given by K.E. Shannon is statistical and not a semantic one.

The K.E. Shannon's information theory

allows you to define the information content of only those information structures that contain symbols, which are measured in bits. This excludes the informative content graphical symbols analysis, topological structures and pictograme. Researchers introduce alternative concepts, such as graphic information content to distinguish it from the "bit" informativeness. The Wiener and Florida's approach allows you to explore different types of information content.

Shannon's information certainty is associated with the ability to measure the informational message or to determine the status of the recipient of this message. According to Shannon, the amount of information is connected with the number of states, in which the recipient of the information is. This makes assessment of the amount of information subjective.

The K.E. Shannon's information theory uses only one data object. This object is a message. This makes it difficult to use such a theory in other information technologies.

According to L. Floridi information content is related to the area of interpretability of information unit. It is determined qualitatively by the application area of information units. According to Floridi [16] information content is determined by the number of interpretable sentences.

Florida's information certainty to is not connected with the "recipient of the message" subject. Information certainty is related to the truth of information construction. This makes the given assessment of the amount of information more objective.

The information unit concept introduction enriches the information theory considerably and extends the possibilities of information structures' analysis.

6. Conclusion

The current theory of information is not a complete theory. It has different directions. In our opinion, the development of information theory with the introduction of such concepts as "information construction" and "information unit" is promising. Studies, carried out in this paper are related only to technical computer science area and information technologies**.** The range of application of information structures and information units is much wider. This is the field of linguistics and computational linguistics, cognitive semantics, psychology, search systems and artificial intelligence. The information theory extension, with the use of information structures and information units allow combining listed research directions with the information theory. This is the subject of further researches.

Amount of information in the K.E. Shannon's information theory is the messages' information capacity or other information construction without regard to a content. It is measured in bytes and bits, regardless of the fact that these bytes contain something or they do not contain anything. Accordingly, according to K.E.Shennon, information content is a measurement in bits of an informational capacity or information construction.

Amount of information in Florida's information theory is a content-richness of an information construction. Winner-Florida's approach explores the task of a meaning transfer. Winner-Florida's approach sets the information structures, defining "knowledge" and in fact is the basis of semantic information theory.

Corresponding Author:
Dr.Tsvetkov Victor Yakovlevich
Moscow State Technical University of Radio Engineering, Electronics and Automation MGTU MIREA
Vernadsky Prospekt, 78, Moscow, 119454, Russia

References

1. Shannon, C.E., 1948. A Mathematical Theory of Communication, Bell System Technical Journal, 27: 379-423 & 623-656.
2. Winner, N., 1948. Cybernetics or Control and Communication in the Animal and the Mashine. The Technology Press and John Wiley & Soris Inc. New York – Herman et Cie, Paris, pp: 194.
3. Shannon, C.E. and W. Weaver, 1964. The Mathematical Theory of Communication. The University of Illinois Press – URBANA, pp: 132.
4. Carnap, R. et al., 1952. An outline of a theory of semantic information. Research Laboratory of Electronics, Technical Report #247, MIT: 221-274.
5. Floridi, L., 2011. Semantic conceptions of information. Date Views 24.05.2014 http://plato.stanford.edu/entries/information-semantic.
6. Kolmogorov, A.N., 1965. Three approaches to the information quantity defining. Problems of information transfer. Book 1, issue 1: 3-11.
7. Shreider, Y.A., 1967. About semantic aspects of information theory. Information and cybernetics. M.: Soviet radio: 15-47.
8. Tsvetkov, V.Y., 2007. The introduction to the information theory. M.: MaksPress, pp: 115.
9. Tsvetkov, V.Y., 2014. Dichotomous Systemic Analysis. Life Science Journal, 11(6): 586-590.

10. Tsvetkov, V.Y., 2014. Opposition Variables as a Tool of Qualitative Analysis. World Applied Sciences Journal, 30(11): 1703-1706.
11. Floridi, L., 2007. In defence of the veridical nature of semantic information. European Journal of Analytic Philosophy, 1(3): 31- 41.
12. Floridi, L., 2005. Semantic Conceptions of Information. First published Wed Oct 5, 2005; substantive revision Fri Jan 28, 2011. Date Views 25.02.2014 http://plato.stanford.edu/entries/ information-semantic.
13. Tsvetkov, V.Ya., 2014. Information Units as the Elements of Complex Models. Nanotechnology Research and Practice, 1(1): 57-64.
14. Bhattacharyya, A., 1946. On some analogues of the amount of information and their use in statistical estimation. Sankhyā: The Indian Journal of Statistics, pp: 1-14.
15. Gel'fand, I.M. and A.M., Yaglom, 1957. Computation of the amount of information about a stochastic function contained in another such function. Uspekhi Matematicheskikh Nauk, 12(1): 3-52.
16. Whitehurst, G.J. and S., Sonnenschein, 1981. The development of informative messages in referential communication: Knowing when versus knowing how. Children's oral communication skills: 127-142.

7/13/2014