HAL open science

Analysis on Limitation Origins of Information Theory

Yong Wang, Huadeng Wang, Qiong Cao

▶ To cite this version:

Yong Wang, Huadeng Wang, Qiong Cao. Analysis on Limitation Origins of Information Theory. 7th International Conference on Intelligent Information Processing (IIP), Oct 2012, Guilin, China. pp.50-57, 10.1007/978-3-642-32891-6_9 . hal-01524976

HAL Id: hal-01524976

https://inria.hal.science/hal-01524976v1

Submitted on 19 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Analysis on Limitation Origins of Information Theory

WANG Yong[1] WANG Huadeng[1] CAO Qiong[2]

[1] School of Computer Science and Engineering, Guilin University of Electronic Technology, Guilin 541004, Guangxi Province, China

[2] Computer Science and Engineering，Chongqing University of Technology, Chongqing 400050, China

hellowy@126.com

Abstract. The limitations of Shannon information theory are pointed out from new perspectives. The limitations mainly exist in the neglects of the information reliability and completeness. The significances of the information reliability to the information measurements are further illustrated through example analysis. It is pointed out that such limitations originate from neglects of multi-level information uncertainties, uncertainty of the model and other objects of information system, and insufficient knowledge on uncertainties of probability values.

Keywords: information theory, communication, reliability, model, uncertainty, probability.

1 Introduction

Shannon information theory is aimed at the communication issues, and is not quite necessarily applicable to the information issues in the reality [1], and his infor-mation theory is called the special information theory by the later researchers. Aimed at the general information, some researchers have developed comprehensive information theory, generalized information theory, unified information theory etc [2-5], in which some limitations of the special information theory are identified, but no one of these information theories can eliminate all the limitations of information theories. In this paper, we attempt to analyze the limitations of the special information theory from some new perspectives.

2 Limitations of Shannon information theory in the reality

The currently recognized limitations of Shannon information theory are mainly as follows: Firstly, only the random uncertainty is considered in his information theory, while the uncertainties such as the limitations of sets in information expressions and the information fuzziness etc are not considered. Aimed at this issue, some researchers have developed the theories of fuzzy sets and rough sets. Secondly, neither the semantic nor the pragmatic aspects are considered in Shannon information theory,

which were considered as the main origin of the limitation of information theory by some researchers [2, 3]. ZHONG Yixin developed his comprehensive information theory which includes syntactic information, semantic information and pragmatic information [2].

We find out the following limitations of information theory:

Firstly, the information reliability is not considered in information theory, while most information in the reality is unreliable. The reliability of information is the precondi-tion to the information values, for example, the reliability of the intelligence informa-tion is very important. The reliability is a key indicator of information. however, it is not enough considered in information theory, which mainly emphasizes the uncertainty of information and the reliability of information transmission.

Secondly, the information completeness is not fully considered in information theory, while in the reality information is always incomplete (insufficient) and needs to be fused. When the more perfect information is unavailable, people cannot help but expediently take partial information temporarily as complete information. In this case, the information that may be deemed that the incomplete information is unreliable compared with the complete information [6].

Thirdly, in information theory, some simple channels may be combined into a channel through parallel or serial connections, for example, the channel matrix of two simple serial channels may be directly multiplied into one channel. However, the complex and multi-level information transmissions are not considered in information theory. for example, the information may be transmitted from one source to an intermediate sink, and forwarded from the intermediate sink to a final sink, the information may be fused from many sources to many sinks, the transition probability matrix may be uncertain, furthermore, during such transmissions, the information may be converted, during multi-level transmissions or complex transmission, multiple uncertainty may occur. In that case, not only the message or code is uncertain, but the probabilities that express the uncertainty are uncertain. The information in the reality often needs to be conveyed through this kind of multiplex transmissions, during which multiple uncertainties may occur. If the above-mentioned fuzzy sets etc are also considered, such multiple uncertainties will become more complicated. The complexity of the parameters such as the channel matrix transmission probability etc is not considered in information theory, while in the reality, such transmission property may not be necessarily maintained certain and unchanged, it may be a random variable or more complex variable. Therefore the uncertainty of information system model should be considered.

Fourthly, the objects studied in information theory are the communications, whose transmission signals are certain. While in the reality, there are many uncertainties. During communications, it does not matter to define the information as something which reduces uncertainty. However, when confronted with the intrinsically uncertain information, the information is distorted if its uncertainty is eliminated. And the trifles are attended but the essentials are neglected during such eliminations. The quantum information theory serves as one kind of popularizations of the classical information theory. while the quanta bytes may include 0, 1 and superposed numbers between 0 and 1. In the reference [4], the information reliability is considered, and it is pointed

out that the information reliability is more important than its certainty. In the reality, people tend to select more uncertain but more reliable information, rather than selecting more unreliable but more certain information.

Fifthly, the conditions in information theory are comparatively simple, and mostly expressed in the probability of conditions. However, int the reality the conditions of information are often very complicated. For example, the given conditions may be knowledge, laws etc. Based on a known and prior probability, another law may be acquired, but it is impossible to easily calculate any corresponding probability of conditions through this law.

Sixthly, the known information is expressed with the prior probability in information theory, however in the reality, a lot of known information may not be expressed with the prior probability, for example, an unknown variable may be included, which may be a restrictive condition, or may be a law.

Seventhly, as the semantic aspects are not considered in information theory, it is not considered that the information may be mutually incompatible and contradictive. In the reality, a large quantity of information may be inconsistent and contradictive each other.

.

3 Analyses on information examples

Considering the above limitations, and in order to make them look more obvious, we analyzed the following examples.

Example 1: If a sender sends a receiver an important intelligence saying: “Some two nations are going to start a war”. When speaking from an information theory perspective, such information itself is the information that involves a large quantity of information. But if this information is not true, it may cause great disasters. People tend to emphasize more on the information reliability, if the information is unreliable, further communications become insignificant, and the information becomes valueless. In information theory, efforts are extended only to ensure that a receiver receives the original information. while it is not studied whether the information from an information source is truly reliable. If the above-exemplified information is unreliable, then the probability of a war between such two nations may be not at 1, but a number between 0 and 1. in this case, the probability value itself is randomly uncertain. In order to avoid confusing the probability uncertainty and the information uncertainty, it may be assumed that the sent intelligence is rewritten as: “the probability of a war to be started between some two nations is at 0.7”. If this intelligence is not absolutely reliable, then this probability value is most likely not at 0.7, but close to 0.7. in this case, the probability itself is not certain, and the value at 0.7 may be just an average. This kind of reliability issues may be more complex. for example, in addition to that the war possibility is unreliable, it is possible that the subjects are unreliable. for example, maybe the war is possible between three nations instead of two nations, or between two persons etc.

Example 2: The probability of an event m is determined by some conditions, supposing that these conditions are c_1, c_2, …, c_n, and supposing that the probability of the event m may be expressed as

$P(m) = f(c_1, c_2, \ldots, c_n)$

If the only condition c_1 is unknown to us, while all of the others are known, maybe the average value of P(m) can be calculated according to probability distributions of the condition c_1. Based on specific cases, the condition c_1 is known and certain. But when this condition is unknown, we can only expediently replace the true probability with the average probability value calculated when this condition is unknown. In this case, the two probability values are close, but they are not equal. When the average value is used to replace a specific value, the information becomes obviously unreliable. Therefore, when the information in this example is incomplete, it may become an information unreliability issue. Generally speaking, the more conditions become known, the more complete the conditions are. and the probability well become more closer to the probability when the conditions are fully complete, therefore the information becomes more reliable. In this example, the case may become more complex. For example, the probability of m still may be a random variable even when all of the conditions are certain. It is like the inaccurate measurement principle in quanta dynamics, which is not caused by existences of hidden parameters or the incompleteness of quanta dynamics, but is one kind of natural uncertainties.

Example 3: The occurrence probability of an event m may be tested through experiments. When the known condition is t: "the experiment result is $P(m|t) = 0.7$", we can't affirm $P(m|t)$, which is the occurrence probability of the event m, we can only be sure that $P(m|t)$ is a random variable close to 0.7. If this random variable has to be replaced with the fixed value 0.7, this value will cause reliability issues. The joint probability distributions and condition probabilities frequently appear in the probability theory and information theory. but it is not considered that under many conditions, the probability value itself in the condition probabilities or joint probability distributions may be a random variable, or partially unknown, or even completely unknown, and then the uncertain probability value will be improperly taken as a certain probability value.

Example 4: A student always gets very high scores with an excellence probability at 0.875, but he gets sick right before an exam. As his learning is delayed due to his sickness, the probability of his excellence may reduce to 0.75. Before we know that the student gets sick influencing his learning, the uncertainty of the prior probability that we known is less than the uncertainty of the posterior probability after we know that his learning is influenced. From an information theory perspective, the latter is reversely reduced in the information contents. In the reality, people will not select an prior probability on the ground that its uncertainty is low, but will select a posterior probability determined when the conditions become more complete as it is closer to the reality. This example shows that measurements on information per uncertainties are limited, and such measurements are not applicable to the uncertain issues. in addition, it also shows that the information needs to be specified in measurements on its reliability and completeness, and that this specification is more important than the information entropy used to measure the information uncertainties.

Example 5: A receives from B an intelligence saying: “it is 99% sure that the enemy will attack us on tomorrow morning”. later, A receives the same intelligence from C. If seen from information theory perspective, the uncertainty of the question “will the enemy attack us on tomorrow morning?” is the same in such two cases. therefore the information contents equal to each other, while it seems that C does not provide any new information. However, people will still feel that they receive the information from C, and this kind of information makes A feel more certain “it is 99% sure that the enemy will attack us on tomorrow morning”. this example further shows that the information reliability shall be taken as one specification to measure the information.

Example 6: When the information that “all events occur at equivalent probabilities” is received, and if speaking from what contents are included in this sentence, or the question “how are the occurrence probabilities of all events distributed?” the uncertainties hereof are eliminated. However, if speaking from what events will occur, it is impossible that such information becomes more certain. the information contents cannot be increased, but only be decreased. This point shows that the information contents are only measured on the information uncertainties, while the uncertainties of the questions derived from this information are not certainly related to the information contents. therefore, the measurements through an information entropy are applicable only to limited scopes, but not to the daily information issues.

Example 7: Brillouin once developed one paradox: If one information paragraph is sent in a text whose last part tells the receiver otherwise that all of the previous information is not true. In this case, is any information transmitted? Brillouin recommends that attentions be made on the negative information. In this paradox, there are two information parts, the second of which negates the first. I believe that such two information parts contradict each other and their reliabilities are both relative. But in general cases, if the sender is honest and not making purposeful jokes, then according to the context and reasoning analysis, the possibility of the latter information part being true becomes much higher. In this case, when the former and the latter information parts are summed up, it may be deemed that no useful information has been sent. Of course, the possibility that the sender made a mistake in the latter part or sent the mistaken information cannot be absolutely eliminated. In information theory, the associations between various information symbols are expressed with redundancies. but in general cases, this information may be integrally looked as one kind of redundancy codes, namely the codes that are impossible to be sent (the codes with a sending probability at 0). As time delays exist during communications (if no time delays exist. in general cases, the information will be self appropriate), such delays enable sending the information that is originally incompatible or impossible to be sent through the complete information.

Through the above analyses, some limitations of information theory are revealed to provide foundations for identifying the origins of information theory limitations. Of course, there are many other limitations, which will not be described here further.

4 Limitations origins of information theory

As analyzed through the above examples, it may be concluded that in information theory, the information reliability is not considered, while the information reliability is a very important indicator. During communications, as the information is certain, a certain relationship exists between eliminations of uncertainties and increases of reliabilities. In fact, it is easy for us to eliminate the uncertainty. while in Shannon information theory, an uncertainty is eliminated on the foundation that the information reliability and completeness are ensured. for example, error correction codes are used to correct errors, and the posterior probability is used to strengthen the information completeness. If the information certainty is taken as the sole specification, and the information reliability is disregarded. then the probability of an event may be randomly determined as 1, while the probabilities of the other events may be randomly determined as 0. Furthermore, if we take the information certainty as the primary target to be considered, and the reliability as the secondary one. then we may also designate the probability of the most probable event as 1, and of the other events as 0. In this case, the certainty is firstly ensured. while the reliability is also satisfied to certain extents. If so, information theory and the information processing become quite simple. obviously in the reality, people do not treat them in this way. According to many of the above analysis, the reliability is one of the primary information specifications.

The above information reliability, information completeness, and the incompliance of classical sets with the reality may be all summed up into the neglects of the multi-level information uncertainties. For example, during example analysis, we discovered the unreliable information, whose expression itself was not fixed, and whose probability value might be a random variable. while the incomplete information is also similar. The non-classical sets in the categories of fuzzy sets and rough sets then may be deemed to have been caused by one set that includes uncertain objects. For example, in rough sets, an object a may or may not belong to the set X. then a random uncertainty exists as to whether the object a belongs to the set X. When some uncertainties are accumulated on the original uncertainties in information theory, the multi-level uncertainties will occur. The uncertainties in this case may include the other uncertainties in more forms, in addition to the random uncertainty and the fuzzy uncertainty. they also include some uncertainties caused by incomplete restrictive conditions. It can be seen that the neglects of the multi-level information uncertainties are important origins of information theory limitations. while the neglect of the information reliability is also an important reason why it is impossible for information theory to be widely applied. Whereas all of the information is rarely reliable and complete, so we may attribute the information reliability and completeness to the information relativity. In fact, it is very rare that people in the reality receive the completely reliable information, they can only expediently apply the comparatively reliable information. when the more reliable information becomes available, people will apply the more reliable information to replace the previously experienced information. As the reliability is also related to the uncertainty of the probability values, the information reliability may be measured with reference to the measurements by Shannon on the

information uncertainties. However, the calculations of the probability uncertainties will be more complex than the calculations of the information entropies, as the probabilities have to satisfy more restrictive conditions.

From Brillouin paradox, example 4 and example 6, we can know that the formula of Shannon information contents is used only to measure the minimum transmission when the special information is being transmitted. during communications, it also reflects the reliability to some extents. However, when the meanings correspondingly derived from the information are considered, neither the information contents nor the information meaning certainties are certainly related to the sizes of the information contents. while the certain information may end up increasing the uncertainties of the other information. Some information may even increase the uncertainties of some information and meanwhile decrease the uncertainties of the other information. for example, the information, “it rains today” may decrease the uncertainty of “are the roads wet today?”, but meanwhile increase the uncertainty of “will the students come late today?”. Obviously, it deserves our deliberations on whether the negative information needs to be considered.

Of course, information theory is highly similar to the information issues in the reality. the methods applied in information theory provide very valuable references for researches on the information issues in the reality (including researches on the information reliabilities). For example, the mutual contradictions in Brillouin paradox are very similar to the inconsistence caused by the mistaken error correction codes. similarly, the most probable events may be taken as the events that really occur. while when all probabilities are close to each other, the information may be “deleted” in the similar way how the information channels are deleted.

Information theory limitations originate from the nature that information theory is oriented to communications issues, and are the intrinsic limitations of information theory models. Of course, such limitations are also related to the limitations of the probability theory. Due to insufficient researches on the randomness and multi-level random uncertainties of the probability values, people are likely to fall into the thinking tendencies that “the probabilities (including distributions of the joint probabilities) are the certain values which never happen to be random variables in any case”, that “the condition probabilities may be determined whenever the conditions are given” etc. while these thinking tendencies are only applicable to some parts of the probability theory issues in the reality. As these restrictive conditions in information theory can better satisfy the communication demands, consequently information theory has been successfully applied in the communication fields. If we wish to popularize it in the general information fields, then corresponding restrictive conditions need to be removed according to its limitations.

Most of the limitations is rooted in the overlooking of the ubiquitous uncertainty of the objects about the information system, such as uncertainty of model, parameters, algorithms, functions, participator and process. We can simplify the information problem by using certain object to replace uncertain objects, this is very beneficial for research and can transform the reality problems to mathematical model so as to be processed by mathematical method, but this poses limitations and restricts the freedom of problems.

5 Conclusion

This paper analyzes some limitations of Shannon information theory. Through the researches on its limitations, on one hand, its applicable scopes can be determined to avoid abusing information theory and to apply information theory in suitable fields. On the other hand, directions may be provided for popularizing and referring to information theory. For example, the information expressions may be improved according to the limitations that the probabilities in the information expressions are fixed values, but not random variable or more complex variable, so as to accommodate the needs to research the information reliabilities, and to further popularize information theory. In addition, new directions are also proposed for the developments of the probability theory. Information theory is also similar, in many aspects, to the information issues in the reality, such as the afore-mentioned considerations of the information reliabilities and completeness. And some of these issues can be solved by referencing information theory.

.

References

1. Shannon C E. A mathematical theory of communication [J], Bell System Technical Journal, 27 (1948), 379—429, 623—656
2. ZHONG Yixin: Principles of information science [M], Fuzhou: Fujian People's Publishing House, 1988.
3. LU Chenguang: Generalized information theory [M], Publishing House of, University of Science and Technology of China, 1993
4. Klir, G.J. Uncertainty and Information：Foundations of Generalized Information Theory [M], Wiley,2006
5. Daniel Federico: Contributions towards a unified concept of information [M], Doctorial Thesis of University Bern, 1995.
6. Yong WANG, Analysis and Betterment of Shannon's Information Definition[J], Journal of Information, 2008, 27(8):57-60.