Structural Analysis of Information: Search for Methodology

Marcin J. Schroeder
Global Learning Center, Tohoku University, Sendai, 980-8576, Japan; mjs@gl.aiu.ac.jp

The apparent terminological simplicity is the most deceiving disguise of complex concepts and questions whose real depth is obscured by our habits of thinking. There are many words which we use every day believing that we know well their meaning until someone asks us for their explanation. This applies to the question about the meaning of the concept of information. The term "information" belongs to the most often used in a myriad of contexts but the concept which it represents escapes all efforts to provide a commonly acceptable definition. There is nothing unusual about information being elusive. There have been many fundamental concepts generating never ending discussions. Maybe, as it was suggested already by E. C. Shannon, the most celebrated pioneer of the study of information, we need more than one concept of information. Another possibility is that the study of information understood in multiple and not always precisely defined ways and in very diverse contexts should continue and the future will show which definition provides the most adequate and inclusive description of information or how to integrate the multiple definitions into one acceptable for everyone.

However, if we want to maintain the identity and integrity of the study of information carried out in its further development in the absence of the uniform definition, we have to establish methodological tools not necessarily identical for all forms of inquiry of informational phenomena, but at least being consistent and preferably allowing comparisons of the results of inquiries. Thus, the methodological unity of the study of information even if it may not be complete should serve as a guiding ideal for its inquiries.

This work is not intended as a study of universal methodological tools for all possible forms of inquiry in diverse disciplines. Its main objective is to search for methodological tools for the study of information with the sufficient level of universality to relate studies of information within different disciplines. However, even with this much more restricted objective it is necessary to clarify some misunderstandings present in methodological analyses of practically all scientific disciplines and in all contexts.

The title of this contribution refers to the structural analysis of information as a distinctive methodological tool for two reasons. The first is that this form of inquiry is clearly underrepresented and inadequate in the study of information. The second, closely related reason is that there are many misconceptions about the distinction between different forms of inquiry with surprisingly little attention paid to the role of structural analysis not only in the study of information, but in the majority of scientific and intellectual domains.

The latter, more general issue that is present not only in the study of information can be identified in the representative example of the relationship between quantitative, qualitative, and structural methodologies. The popular conviction of the apparent complementary, dichotomic opposition of the first two methodologies is based on the misconceptions of the role of mathematics in general and of the numbers in particular which are perpetuated in virtually all scientific inquiries. This mistaken view of the two methodologies, their exclusive and universal role in all inquiries obscures the fact that they both are just instances of structural analysis, in which mathematics can offer methodological tools going well beyond the present toolkit.

The fallacy of the opposition and complementarity of the quantitative and qualitative methodologies has its source in the hidden assumptions that are very rarely recognized in the scientific practice. Another source is in the overextension of the mathematical concepts which have very specific and restricted meaning in mathematics to the applications in science where the conditions of their definitions are not satisfied or not considered.

An outstanding example of this type of confusion is in the use of the concept of measure, which frequently in scientific applications is understood as an arbitrary assignment of real numbers to some set of not always clearly defined objects. This use of the term measure is very far from the understanding of the concept of a measure in mathematics. It would have been just a terminological inconsistency with mathematics, not an error, if this non-mathematical concept of measure was not mixed up with the mathematical concept in making conclusions regarding the results of inquiry. Very often the meaning of the term measure is simply not clarified. Sometimes the intention of the use of the term is consistent with the measure theory, but there is nothing about the related concepts of the theory whose absence makes the central concept meaningless. The reference to a measure without any clarification of how it is defined has as its consequence the hidden import of the structure on which it has to be defined to retain its mathematical meaning (a sigma ortho-algebra of measurable subsets of the measure space). Thus, there is usually a hidden structure associated with the subject of each study which serves as a tool for inquiry, but which is excluded from the overt considerations.

If we decide to disregard the conditions in the mathematical concept of a measure and consider it simply as a real-valued function on some set S, then we define on S just an equivalence relation defined by the partition of S into subsets of elements with the equal values of the function. However, in this case we have a pure case of the qualitative methodology based on partitions of a set into equivalence classes which can be identified with qualities or properties of the elements of S, but which equally well can be identified with numerical values. This shows that the distinction between the quantitative and qualitative methodologies is fuzzy. In both methodologies we assume overtly, or most frequently covertly an essentially the same structure of an equivalence relation imposed on the universe of our study. More importantly, in both cases by imposing hidden mathematical structures on the subjects of our study we actually carry out structural analysis involving equivalence relations. As long as the concept of a measure is not the one from measure theory and a measure is simply an assignment of a numerical value the distinction between the two methodologies is rather conventional and is based on the way how equivalence relations are presented. The engagement of the mathematical concept of a measure adds to the consideration an additional, structure of a non-trivial ortho-lattice of measurable subsets.

This work goes beyond the critical review of the hidden but omnipresent elements of structural methodology in the study of information. There is a legitimate question about the positive, creative aspect of the recognition of the role of structural analysis. The source of the conceptual tools necessary for further development of the structural methodology of information can be identified in the invariance with respect to transformations, the main methodological strategy of physics and several other natural sciences. Surprisingly, this idea was completely missing in the work of Shannon, but was already present in the 1928 paper by R.V.L. Hartley cited by Shannon in the footnote to the first page. Hartley did not refer directly to structural analysis, but used invariance as a tool to derive and to interpret his simpler than Shannon’s formula for the amount of information.