数理解析研究所講究録
第 1964 巻 2015 年 88-97

Structural and Quantitative Characteristics of Information and Complexity

Marcin J. Schroeder
Akita International University, Akita, Japan
mjs@aiu.ac.jp

1. Introduction

Idols of Francis Bacon did not include an idol of the number. This idol entered scientific inquiry too recently to be of any concern for the Age of Reason. The idol is hidden in the common conviction about the superiority of quantitative methods of inquiry over those qualitative. The problem is not in the answer to the question whether quantitative methods are superior or qualitative, but in the very fact of asking it, as if the question was meaningful. One of the fundamental questions, which should stimulate philosophical, methodological discussion, is: What actually do we know about some concept, when we describe it in terms of some magnitude with numerical values? [1]

It is obvious that the description of an object of study in terms of magnitudes with numerical values plays a fundamental role in science, but only under the condition that these magnitudes have foundation in a structural, and therefore qualitative, analysis of the object. Not always, or even not frequently this happens. Very often more or less ad hoc magnitude appears first, and then there are some attempts to associate with this magnitude a concept for which it apparently is a quantitative description.

The struggle with the concept of information gives an excellent illustration of the problem. Information became one of the most frequently used words in the contexts of all types of inquiry, but at the same time it continues to be one of the most elusive concepts. It entered the stage almost unnoticed, so today there is a common belief that theory of information was initiated in 1948 by the famous article of Claude E. Shannon "A Mathematical Theory of Communication," which one year later was re-published in the book format with the title "The Mathematical Theory of Communication,". [2,3]

Historically more accurate birth certificate for information should give as a date and place the ten year earlier article "Transmission of Information" by Ralph Hartley. [4] Shannon quoted this article in his work, but because Hartley's article introduced a measure of information which is a special case of Shannon's entropy, a measure "of information, choice and uncertainty" [3], the article is frequently considered premature and non-deserving attention. However, actually it was Hartley's article that pushed the study of information in its direction of development.

Hartley derived his formula $\mathbf{H = n \log_m(s)}$ (where s is the number of symbols available in all selections and n is a number of selections, m arbitrarily chosen according to the preferable choice of the unit of information) for the quantitative measure of information in a symbolic form from the assumption of invariance. He considered the invariance with respect to the grouping of the "primary symbols" (here associated with physically distinct states of the physical system) into "secondary symbols" representing psychologically determined and therefore subjective symbols carrying meaning. The choice of the formula makes values of H independent from grouping but was dependent on the assumption that all primary symbols are equally likely to be used in the formation of secondary symbols.

Shannon did not refer in his derivation of entropy to any non-mathematical assumptions and only one of the assumptions has a direct interpretation. He referred to the fact that "With equally likely events there is more choice, or uncertainty, when there are more possible events." [3] The argument is convincing for uncertainty, but its relationship to information is counterintuitive. Why should the choice of encoding with equally likely symbols be distinguished? Why should the way we encode information influence the amount of information? After all, Hartley was using as a point of departure the independence from the way of encoding.

2. Quantitative Description

Whatever guided Shannon, his entropy became an orthodox measure of what started to be commonly considered information:

$$\mathbf{H(p_1,p_2,\dots,p_n) = -\sum_{i=1}^{n} p_i \log_2 p_i} \text{ where } \mathbf{\sum_{i=1}^{n} p_i = 1 \;\&\;} \forall i: p_i \geq 0$$

The indices i = 1, 2,..., n represent the elements of the information carrier S, or what Shannon would rather call elements of the alphabet. We can see that every transformation T of the set S (bijective function $T:S\to S$) preserves the value of H, as long as $\mathbf{p^S{}_i = p^{T(S)}{}_{T(i)}}$. Since we have no restriction on the choice of T, we get suspiciously high level of symmetry. Entropy is an invariant of all transformations of S, and therefore it cannot reflect any structural characteristics of the carrier of information.

Entropy is a functional which assigns a nonnegative real number to each class of equivalent probability distributions. Thus, there is nothing in entropy that cannot be directly derived from the probability distribution. Since entropy cannot be calculated without some pre-determined probability distribution, it is simply a partial characterization of whatever the distribution characterizes or describes. It is only a partial characterization, because two different distributions may give rise to the same entropy. If we believe that entropy is a measure of information, then we have to accept that actual description of information, more complete than entropy, is given by one of probability distributions producing given value of this magnitude.

This brings us to unacceptable consequences. Every geometric structure definitely carries out some information. For instance we can think about the geometric structures of the alphabet letters or numerals. Is it possible to describe geometric structures in terms of probability theory? The answer is obviously not, although of course we can define probability measures within geometric structures. Thus, we cannot expect that probability distributions are sufficient to describe or characterize information. If information cannot be characterized exclusively in terms of probability distributions, entropy is not sufficient characteristic of information.

This can explain multiple attempts to generalize the concept of entropy in order to have a mathematical model of information better suited to serve as characterization of information.

First step was the transition to the continuous probability distributions for random variables. It is already a convenient, but suspicious fact that when we define entropy in terms of random variables, the actual values of entropy depend exclusively on mass functions not on the values of the random variable. Transformations of random variables may change entropy by changing mass functions, and for instance entropy is not an invariant of linear transformations of random variables. But the values of a random variable are irrelevant for entropy.

Transition from finite, discrete random variables to continuous ones is equally disappointing. In the limit transition to the infinite case entropy is divergent. Thus entropy's continuous counterpart is an *ad hoc* functional called "differential entropy", not derivable from any limit process.

$$\mathbf{H(X) = -\int_a^b f(x) \log_2(f(x))\, dx}$$

The difficulties are of theoretical character and they did not inhibit the development of the theory of communication based on this form of entropy. However, it is difficult to say whether the use of entropy was successful because of its relationship to information or simply because of the involvement in the study of powerful methods of probability theory.

Analogy to the original form of entropy led to von Neumann's entropy in the context of quantum mechanics: $\mathbf{S = - Tr\ (\rho \log \rho)}$, where ρ is the density operator describing a mixture of quantum states.

There were many attempts to involve generalized forms of entropy guided by the recognition that Shannon's entropy is a convex function. A function f(x) is convex over an interval (a,b) if for all x_1, x_2 in the interval (a, b) and for every λ such that $0< \lambda <1$ we have $f(\lambda x_1 + (1- \lambda)x_2) < \lambda f(x_1) + (1- \lambda)f(x_2)$. The association of entropy with convexity stimulated attempts to search for generalized forms of entropy which have a similar property of convexity, but which involves one or more parameters, and which for some values of the parameters or in the limit when parameters approach some value the generalized function collapses to Shannon's entropy.

One of the earliest such generalizations was proposed by Alfred Renyi. [5] This generalization is twofold. First, but marginal is that the probability distribution does not have to be complete, which means that the probabilities of elementary events do not have to add up to one. More important was the involvement of a positive parameter α, such that we have a class of generalized Renyi entropies:

$$\mathbf{H^{\alpha}{}_{n}(p) = (1/(1-\alpha)) \log_2 (\sum_{i=1}^{n} p_i^{\alpha})/(\sum_{i=1}^{n} p_i) \quad (\alpha > 0, \alpha \neq 1, \sum p_i \leq 1)}$$

When the parameter α approaches 1, and $\sum p_i = 1$, Renyi's parametrized entropies have Shannon's entropy as a limit. In the following years a series of generalizations were proposed involving an increasing number of parameters. [6,7] Finally, P. N. Rathie [8] introduced the generalized entropy with n+1 parameters for arbitrary n:

$$\mathbf{H^{\alpha,\beta_1,\beta_2 \ldots \beta_n}{}_{n}(p) = (1/(1-\alpha)) \log_2 (\sum_{i=1}^{n} p_i^{\alpha+\beta_k-1})/(\sum_{i=1}^{n} p_i)} \text{ where } \mathbf{(\alpha>0, \alpha\neq 1, \beta_k>1, \sum p_i \leq 1)}$$

Constantino Tsallis went in different direction abandoning the requirement of convexity, but with physical applications in consideration. His parametrized by real numbers q entropies

$\mathbf{S_q(p) = (k/(q-1)) [1- (\sum_{i=1}^{n} p_i^{q})], (\sum p_i = 1)}$ involved Boltzmann constant k, as the concept was developed in the physical context. When we eliminate the physical constant k, parametrized entropies of Tsallis become Shannon's entropy in the limit when q approaches 1.

All these attempts seem like drawing concentric circles around the arrow shot by Shannon, or actually Hartley, in order to justify claim that the arrow hit the bull's eye.

That the arrow hit another, actually opposite part of bull's body (when the bull is a symbol of information) was quite clear from the very beginning. One of the sources of confusion was in the famous small book "What is Life? The Physical Aspect of the Living Cell" published by Erwin Schrödinger in 1944. [10] Schrödinger did not refer directly to the concept of information in his "meditation on life", but his reflection on the mechanisms of heredity, metabolism, and life in general made it clear that when we consider them in terms of information, we should associate information with the negation of entropy (he called it "negative entropy"), not entropy itself.

Leon Brillouin in his book "Science and Information Thory" [11] abbreviated "negative entropy" into "negentropy" and merged Schrödinger's physical explanations and models explicitly with Shannon's theory of information. What was a marginal oversimplification in Schrödinger's book (that with the sun light life on earth gets negative entropy, instead of that the low entropy light arriving with sunshine is entering the balance of entropy in biosphere allowing the increase of organization without contradiction with the Second Law of Thermodynamics excluding decrease of entropy in isolated systems) became a bizarre concept of the magnitude, which is a negation of the positive value of entropy, but itself is positive. This too would not be harmful, if it was not commonly interpreted as an expression of the conformity between physics and Shannon's measure of information.

The puzzle can be easily solved when we observe that Shannon was interested exclusively in measuring of the transfer of information in the process of communication. Thus, he was using as his reference frame the relationship between information source and destination. His entropy can be easily interpreted as a measure of the increase of information within destination, from the original one based on the probability distribution of possible messages, to the final, which is given by the 0-1 distribution probability, which characterizes full information. Of course, if the destination contains full information (one outcome has probability 1, all other 0), there is no increase of information, which corresponds to the fact that entropy is 0. This means an alternative measure of information is more appropriate, and the need for negentropy is eliminated: [12]

$$\mathbf{Inf(n,p)} = \sum_{i=1}^{n} \mathbf{p_i \log_2 (np_i)}, \quad \sum_{i=1}^{n} \mathbf{p_i = 1}\ \forall i\colon \mathbf{p_i} \geq 0$$

Obviously $\mathrm{Inf}(n, p) = H_{max} - H(n, p)$, where $p = (p_1, p_2, \ldots, p_n)$.

Then we have a characteristic of the degree of determination of information or relative measure of information:

$$\mathbf{Inf^*(n,p)} = \sum_{i=1}^{n} \mathbf{p_i \log_n (np_i)}, \quad \sum_{i=1}^{n} \mathbf{p_i = 1}\ \forall i\colon \mathbf{p_i} \geq 0$$

Or simpler $\mathrm{Inf}^*(n,p) = \mathrm{Inf}(n,p)/\mathrm{Inf}_{max}$, giving the range $\mathbf{0 \leq Inf^*(n,p) \leq 1}$.

The alternative measure has many useful features of entropy, but also is an invariant of linear transformations and is giving a smooth transition from the finite, discrete probability distributions to the case of infinite, continues probability distributions. [12] For our purpose the two most important features are as follows. [12]

Let S be a disjoint union of the family of probability spaces $\{A_i\colon i = 1,\ldots,m;\ A_i \cap A_k = \varnothing, \text{ if } i \neq k\}$, each with probability distribution $p(i)$. Let n indicates the number of elements in S, and n_i of elements in A_i. We can define a probability distribution $p(x)$ on S the following way.

For every x in S, $p(x) = a_i\, p(i)(x)$, where i is selected by the fact that x belongs to A_i and $a_1 + \ldots + a_m = 1$. Of course, $a_i = p(A_i)$ and we can write $p(x) = p(A_i)\, p(i)(x)$.

Then,

$$\mathbf{Inf(n, p)} = \sum_{i=1}^{m} \mathbf{p(A_i)\, Inf(n_i, p(i))} + \sum_{i=1}^{m} \mathbf{p(A_i) \log_2[(n/n_i)\, p(A_i)]}.$$

If all sets A_i have the same size k, then the formula for Inf (n,p) becomes much simpler:

$$\mathbf{Inf(n, p)} = \sum_{i=1}^{m} \mathbf{p(A_i)\, Inf(k, p(i))} + \sum_{i=1}^{m} \mathbf{p(A_i) \log_2[m\, p(A_i)]}.$$

We can interpret this as an assertion that the total information amount Inf(n,p) can be separated into information identifying the element of the partition A_i, plus the average information identifying an element within subsets of the partition.

Let $\mathbf{S = S_1 \times S_2}$ with the probability distribution given by $\mathbf{(p \times q)_{ik} = p_i p_k}$. Let S_1 consists of n elements, S_2 consists of m elements. Then $\mathbf{Inf(nm, p \times q) = Inf(n, p) + Inf(m, q)}$.

It has to be emphasized that this approach is a purely probabilistic study of information as in the work of Hartley or Shannon. Thus, the alternative measure can be interpreted as a measure of information provided we can define the concept of information and develop its structural theory justifying the formula.

3. Structural (Qualitative) Description

The attempts related to the quantitative methods initiated by Shannon in his study of communication ignored the structural (qualitative) aspects of information. No wonder that they had to give up the study of semantics for information, if in such perspective information is an amorphous aggregate with the description exclusively in terms of the probability of meaningless components. However, the attempts focusing on the qualitative aspects of information were even less successful, since they did not go much beyond the relationship with the philosophical reflection on the concept of form.

In order to combine both aspects of information and to place this concept in the context of non-trivial philosophical conceptual framework, the present author introduced his definition of information in terms of the one - many categorical opposition with a very long and rich philosophical tradition. [13] Thus, information is defined as a resolution of the one-many opposition, or in other words as that, which makes one out of many. There are two ways in which many can be made one, either by the selection of one out of many, or by binding the many into a whole by some structure. The former is a selective manifestation of information and the latter is a structural manifestation. They are different manifestations of the same concept of information, not different types, as one is always accompanied by the other, although the multiplicity (many) can be different in each case.

This dualism between coexisting manifestations was explained by the author in his earlier presentations of the definition using a simple example of the collection of the keys to rooms in a hotel. It is easy to agree that the use of keys is based on their informational content, but information is involved in this use in two different ways, through the selection of the right key, or through the geometric description of its shape. We can have numbers of the rooms attached to keys which allow a selection of the appropriate key out of many other placed on the shelf. However, we can also consider the shape of key's feather made of mechanically distinguishable elements or even of molecules. In the latter case, geometric structure of the key is carrying information. The two manifestations of information make one out of very different multiplicities, but they are closely interrelated.

The definition of information presented above, which generalizes many earlier attempts and which due to its very high level of abstraction can be applied to practically all instances of the use of the term information, can be used to develop a mathematical formalism for information. It is not a surprise, that the formalism is using very general framework of algebra. [14]

The concept of information requires a variety (many), which can be understood as an arbitrary set S (called a carrier of information). Information system is this set S equipped with the family of subsets $\Im$ satisfying conditions: entire S is in $\Im$, and together with every subfamily of $\Im$, its intersection belongs to $\Im$, i.e. $\Im$ is a Moore family. Of course, this means that we have a closure operator defined on S (i.e. a function f on the power set 2^S of a set S such that:

(1) For every subset A of S, $A \subseteq f(A)$;

(2) For all subsets A, B of S, $A \subseteq B \Rightarrow f(A) \subseteq f(B)$;

(3) For every subset A of S, $f(f(A)) = f(A)$).

The Moore family $\Im$ of subsets is simply the family f-Cl of all closed subsets, i.e. subsets A of S such that $A = f(A)$. The family of closed subsets $\Im$ = f-Cl is equipped with the structure of a complete lattice L_f by the set theoretical inclusion. L_f can play a role of the generalization of logic for not necessarily linguistic information systems, although it does not have to be a Boolean algebra. In many cases it maintains all fundamental characteristics of a logical system. [15]

Information itself is a distinction of a subset $\Im_0$ of $\Im$, such that it is closed with respect to (pair-wise) intersection and is dually-hereditary, i.e. with each subset belonging to $\Im_0$, all subsets of S including it belong to $\Im_0$ (i.e. $\Im_0$ is a filter in L_f).

The Moore family $\Im$ can represent a variety of structures of a particular type (e.g. geometric, topological, algebraic, logical, etc.) defined on the subsets of S. This corresponds to the structural manifestation of information. Filter $\Im_0$ in turn, in many mathematical theories associated with localization, can be used as a tool for identification, i.e. selection of an element within the family $\Im$, and under some conditions in the set S. For instance, in the context of Shannon's selective information based on a probability distribution of the choice of an element in S, $\Im_0$ consists of elements in S which have probability measure 1, while $\Im$ is simply the set of all subsets of S.

This approach combines both manifestations of information, but the relationship between articulations of these manifestations within the formalism thus far was based on the intuitive interpretation. It was not clear in what sense we can talk about dualism. What exactly do we mean by dualism? How can we use the formalism of information to describe two information systems in the dual relationship?

The example of information carried by a key shows that the dualism of manifestations of information can be associated with the transition from the level of elements of a set S representing a variety to the level of sets belonging to the power set representing another, higher level of variety. [16]

The basic idea of the relationship between dual information systems is that within a variety (set S) an element x can be selected by a specification of its properties, which can be interpreted in the terms of the set theory as a distinction of the family of subsets whose intersection has x as its element. On the other hand, the structure built on the variety can be characterized by the family of substructures defined within S. In both cases information system is associated with a closure operator f on S.

Closed subsets for this closure operator (i.e. sets satisfying $f(A) = A$ which can be associated with actual properties of objects) form a complete lattice L_f which can be considered a generalization of the concept of logic for information (see above). It is reducibility or irreducibility of this lattice that shows the level of integration of information.

The point of departure in the formalization of the duality of information manifestation can be found in the way how we are associating information understood in the linguistic way with the relation between sets and their elements formally expressed by "$x \in A$". The informational aspect of the set theory can be identified in the separation axiom schema, which allows interpretation of $x \in A$, as a statement of some formula $\varphi(x)$ formulated in the predicate logic which is true whenever $x \in A$. The set A consists then of all elements which possess the property expressed by $\varphi(x)$.

If we are interested in a more general concept of information, not necessarily based on any language, we can consider more general relationship than $x \in A$ described by a binary relation R between the set S and its power set 2^S: xRA *if* $x \in f(A)$.

If this closure operator is trivial (for every subset A its closure $f(A) = A$) we get the usual set-theoretical relation of belonging to a set xRA *if* $x \in A$. In more general case, only closed subsets correspond to properties.

Let S, T be sets, and $R \subseteq S \times T$ be a binary relation between sets S and T. R^* is the converse relation of R, i.e. the relation $R^* \subseteq T \times S$ such that $\forall x \in S\ \forall y \in T$: xRy *iff* yR^*x. Then we define $R^a(A) = \{y \in T: \forall x \in A: xRy\}$, $R^e(A) = \{y \in T: \exists y \in A: xRy\}$.

If R is a binary relation between sets S and T, the pair of functions $\varphi: 2^S \to 2^T$ and $\psi: 2^T \to 2^S$ between the power sets of S and T defined on subsets A of S by $\varphi: A \to R^a(A)$ and on subsets B of T by $\psi: B \to R^{*a}(B)$ forms a Galois connection, and therefore their both compositions, defined on subsets A of S by $f(A) = \varphi\psi(A) = R^a R^{*a}(A)$ and on subsets B of T by $g(B) = \varphi\psi(B) = R^{*a} R^a(B)$ are transitive closure operators. Also, the functions φ, ψ are dual isomorphisms between the lattices L_f and L_g of closed subsets for the closure operators f and g. [17]

Now, let $<S,f>$ be closure space which represents an information system. Define for $x \in S$ and $A \in 2^S$ a binary relation $R_f \subseteq S \times 2^S$ such that $xR_f A$ *iff* $x \in f(A)$. If no confusion is likely, we will write simply R instead of R_f.

One way of the Galois connection defined by this relation and described above will return us back to the original closure f, as for every subset A of S: $R_f^{*a} R_f^a(A) = f(A)$. For us, more interesting is the other way which generates the closure operator g defined on 2^S by: $\forall \beta \subseteq 2^S$: $g(\beta) = R^a R^{*a}(\beta) = \{A \subseteq S: \cap\{f(B): B \in \beta\} \subseteq f(A)\}$.

We know from the properties of Galois connections that the complete lattice lattices L_f of f-closed subsets of S is dually isomorphic to the lattice lattices L_g of g-closed subsets of the power set 2^S. Thus, we have that every information system on a set S is associated with an information system on the set of all subsets of S, in such way that their logics are dually isomorphic.

This correspondence is expressing in mathematical language the duality of information manifestations in hierarchically related information systems, which links the consecutive levels in such a way that information structure is preserved in a (lattice-theoretic) dual way. Since we can repeat the reasoning for the "upper level" closure operator g, our construction can produce unlimited number of levels and we have the formal description of multi-level hierarchic information systems.

4. Quantitative Description of Structural Characteristics

The most important characteristic of the structural manifestation of information is its level of information integration describing the specific type of the structure imposed on the variety. This structure may have different levels of integration related to its decomposability into component structures. Decomposability of the structure can be described in terms of irreducibility of the logic L_f of information system into a direct product of component lattices. Quantum mechanics provides examples of completely integrated information systems, but there are many other examples, for instance geometric information systems. [18]

In an earlier article the author presented an analysis of complexity linking this concept to irreducibility of information. [19] Traditional attempts to define and analyze complexity put the emphasis on the large number of components. Frequently this was combined with the qualification "large number of closely interrelated components". However, this "interrelation" was never explicitly defined. In this approach, the emphasis is on interrelation defined as the level of information integration.

Analysis of the level of information integration can utilize extensive mathematical knowledge of irreducibility (or reducibility) of partially ordered sets. We will need only fundamentals whose details can be found in the classic monograph on lattice theory written by Garret Birkhoff.[17]

Our goal is to find a quantitative description of the structural characteristics of information.

In the following slides the references will be made to the numbering of theorems in Chapter III, Section 8 of Birkhoff's monograph:

The main tool for reducibility/irreducibility of the posets is the concept of a center:

Def. *The center of a poset P with 0 and 1 is the set C of elements (called "central elements") which have one component 0 and the other 1 under some direct factorization of P.*

Thm. 10. *The center C of a poset P with 0 and 1 is a Boolean lattice in which joins and meets represent joins and meets in P.*

Def. *An element a of a lattice L with 0 and 1 is neutral iff (a,x,y)D for all x,y in L, i.e. the triple , x, y generates a distributive sublattice of L.*

Thm. 12. *The center of a lattice with 0 and 1 consists of its complemented, neutral elements.*

Fact. *0 and 1 are central elements in every poset with 0 and 1.*

It follows from Thm. 12 above that the lattices M_5 and N_5 are irreducible and that every Boolean lattice is identical with its center.

We can observe that although the Exchange Property of Steinitz
(wE) $\forall A\subseteq S \forall x,y\in S$: $x\notin f(A)$ & $x\in f(A\cup\{y\}) \Rightarrow y\in f(A \cup \{x\})$ itself does not imply complete irreducibility of the corresponding logic L_f, but it does if every two element set has closure with at least three elements.

From now on, it will be assumed that the logic of information L_f (i.e. the complete lattice of f-closed subsets) is finite.

Lemma: If lattices L_1 and L_2 have their centers C_1 and C_2 respectively, then the direct product $L_1 \times L_2$ has $C_1 \times C_2$ as its center.

It is a simple corollary of Thm. 11.

We will write $|L|$ for the number of elements in set L.

Now we can show using Thm. 10 and Thm 11 that the number of irreducible components of the logic L_f is $\log_2(|C|)$. This number is giving us some indication regarding reducibility of the logic (complete irreducibility is for value 1, the increase indicates that the number of irreducible components is increasing). But as long as we do not know the size of the logic, the value of such measure is limited. It is better to consider first a measure of complexity m(L).

Def. Measure of complexity of logic L

$\mathbf{m(L) = \log_2(|L|/|C|) = \log_2(|L|) - \log_2(|C|)}$

Then, **if L is a Boolean lattice (completely reducible), then m(L) = 0,** but when **L is completely irreducible, then $m(L) = \log_2(|L|) - 1$.**

Since the center is preserved by all lattice automorphisms, so is m(L).

Also, m is **"semi-additive"** in the sense that: $\mathbf{m(L_1 \times L_2) = m(L_1) + m(L_2)}$.

In particular for a logic $\mathbf{L = L_1 \times L_2 \times L_3 \times \ldots \times L_k}$ where all L_i are irreducible, in agreement with the definition we have $\mathbf{m(L) = \log_2(|L|) - k = \log_2(|L|) - \log_2(|C|)}$.

On the previous slide m(L) was called a measure of complexity, not of irreducibility, as it is increasing to infinity with the size of the logic L. We have simple irreducible two-element logic with m(L) = 0, as it is a Boolean lattice. Also, for the completely irreducible logics with lattices M_5 and N_5 we have:

$\mathbf{m(M_5) = m(N_5) = \log_2(5/2) \approx 1.32}$ while there are many reducible (although not completely) logics with higher values of m(L). So, in order to have a measure of irreducibility we can introduce a relative measure m* which is an invariant of transformations preserving information structure:

Def. For lattices with at least two elements: $\mathbf{m^*(L) = m(L)/(m_{max}+1) = m(L)/\log_2(|L|) =}$

$\mathbf{\log_2(|L|/|C|)/\log_2(|L|)}$, **where m_{max} is the maximum complexity m(L) for a logic of size |L|.**

From the definition $\mathbf{m^*(L) = \log_2(|L|/|C|)/\log_2(|L|) = 1 - [\log_2(|C|)/\log_2(|L|)]}$.

Then it is easy to see that

If L is a Boolean lattice, m*(L) = 0, but **when L is completely irreducible, then:**
$\mathbf{m^*(L) = 1 - [1/\log_2(|L|)]}$.

So $\mathbf{0 \le m^*(L) < 1}$ and m* is an increasing function of the size of L with limit 1 at infinity.

$m^*(M_5) = m^*(N_5) = 1 - 1/\log_2(5) \approx 0.57$; $m^*(D_{10}) \approx$ ca. 0.70

It is not a surprise that m* is not semi-additive (in the sense in which m is), because m* measures irreducibility. When we have a product of logics, we cannot expect increase of irreducibility. Instead we have a logarithmic weighted mean:

$\mathbf{m^*(L_1 \times L_2) = \alpha\, m^*(L_1) + \beta\, m^*(L_2)}$, where $\mathbf{\alpha + \beta = 1}$ and

$\mathbf{\alpha = \log_1(|L_1|)/\ [\log_2(|L_1|) + \log_2(|L_2|)]}$

$\mathbf{\beta = \log_1(|L_2|)/\ [\log_2(|L_1|) + \log_2(|L_2|)]}$.

Both measures, measure of complexity **m(L)** and measure of information integration **m*(L)** are invariants of all transformations which preserve the logic of information (isomorphisms of L_f).

There are many cases, when there is a high level of information integration (information logic is not Boolean), but still we can define a generalized form of probabilistic measure. For those cases, we can consider both types of the measure of information. An example can be found in quantum mechanical information systems.

Further question is, how to measure selective manifestation for information system whose logic does not admit orthocomplementation, and therefore the concept of probabilistic measure does not make sense.

5. Conclusion

The measures of complexity and of information integration give an example of quantitative characteristics of structural properties of information. We can find some analogy of the relationship between of the concepts describing in a quantitative way selective manifestation of information in the form of author's alternative measure of information Inf (n, p) and the degree of determination Inf* (n, p), and for the structural manifestation, between the measure of complexity m (L) and measure of information integration m*(L).

References

[1] M. J. Schroeder, Crisis in Science: In Search for New Theoretical Foundations. *Prog. Biophys. Mol. Biol.*, 113(1), 2013, 25-32.

[2] E. C. Shannon, E. C. (1948). A mathematical theory of communication. *Bell Sys. Tech. J.*, 27, 1948, 379-423, 623-656.

[3] E.C. Shannon, W. Weaver. *The Mathematical Theory of Communication.* University of Illinois Press, Urbana, IL, 1949.

[4] R. V. L. Hartley. Transmission of Information. *Bell Sys. Tech. J.*, 7(3), 1928, 535-563.

[5] A. Renyi, On measures of entropy and information. *Proc. Fourth Berkeley Symposium on Mathematics, Statistics and Probability 1960*, Vol. 1, Univ. of Calif. Press, Berkeley, 1961, pp. 547-561.

[6] R.S. Varma. Generalizations of Renyi's entropy of order α. *J. Math. Sci.*, 1, 1966, 34-48.

[7] J. N. Kapur. Information of order α and type β. *Proc. Indian Acad. Sci.* A68, 1968, 66-75.

[8] P. N. Rathie. On generalized measures of inaccuracy, information and errors in information. *Statistica*, 30, 1970, 340-349.

[9] C. Tsallis. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.*, 52, 1988, 479-487.

[10] E. Schrödinger. *What is Life?* Cambridge University Press, Cambridge, 1944.

[11] L. Brillouin. *Science and Information Theory*. New York: Academic Press, 1956.

[12] M. J. Schroeder. An Alternative to Entropy in the Measurement of Information. *Entropy, **2004**,* 6, 388-412.

[13] M.J. Schroeder. Philosophical Foundations for the Concept of Information: Selective and Structural Information. In *Proceedings of the Third International Conference on the Foundations of Information Science, Paris 2005*, http://www.mdpi.org/fis2005/proceedings.html/.

[14] M.J. Schroeder. From Philosophy to Theory of Information. *International Journal Information Theories and Applications*, 18 (1), 2011, 56-68.

[15] M. J. Schroeder. Search for Syllogistic Structure of Semantic Information. *J. Appl.Non-Classical Logic*, 22, 2012, 83-103.

[16] M. J. Schroeder. Algebraic Model for the Dualism of Selective and Structural Manifestations of Information. In M. Kondo (Ed.), *RIMS Kokyuroku*, No. 1915. Kyoto: Research Institute for Mathematical Sciences, Kyoto University, 2014, pp. 44-52.

[17] G. Birkhoff, *Lattice Theory, 3rd. ed.* American Mathematical Society Colloquium Publications, Vol XXV, Providence, R. I., 1967.

[18] M. J. Schroeder, Quantum Coherence without Quantum Mechanics in Modeling the Unity of Consciousness, in: Bruza, P. et al. (Eds.) *QI 2009, LNAI 5494*, Springer, Berlin, 2009, pp. 97-112.

[19] M.J. Schroeder, The Complexity of Complexity: Structural vs. Quantitative Approach. In: *Proceedings of the International Conference on Complexity, Cybernetics, and Informing Science CCISE 2013 in Porto, Portugal*, 2013.
Available at: http://www.iiis-summer13.org/ccise/VirtualSession/viewpaper.asp?C2=CC195GT&vc=58/