FROM PHILOSOPHY TO THEORY OF INFORMATION

Marcin J. Schroeder

Abstract: *This is an attempt to develop a systematic formal theory of information based on philosophical foundations adequate for the broad context of pre-systematic concept of information. The existing formalisms, in particular that commonly called information theory, consider only some aspects of information, such as its measure. In spite of spectacular successes of Shannon's entropy and its generalizations, the quantitative description did not help in the development of the formal description of the concept of information itself. In this paper, the brief review of the contexts in which the term information is being used is followed by similarly brief presentation of philosophical foundations incorporating such aspects of information as its selective and structural manifestations, information integration and semantics of information presented in more extensive form in other publications of the author. Finally, based on these foundations, a mathematical formalism is proposed with an explanation of its relationship to the philosophical concepts associated with information. The formalism utilizing mathematical concepts from the theory of closure spaces and associated with them complete lattices of closed subsets playing the role of generalized logic of information is taking into consideration the selective and structural manifestations of information. Since the original source of inspiration in the development of the formalism was in quantum logics, an outline of concepts in this domain is included in the appendix.*

Keywords: *Theory of information; Philosophy of information; Selective and structural information; Information integration; Semantics of information;*

ACM Classification Keywords: *H.1.1 Systems and Information Theory – Information Theory*

Introduction

Development of every systematic theory of a concept such as information requires some philosophical foundations and practical experience in dealing with this concept at a pre-systematic level. The latter includes in the case of information its quantitative characteristics in the process of information transmission provided by Claude E. Shannon's entropy and its subsequent generalizations, and the analysis of manipulation of information in the process of computing initiated by Alan Turing [Shannon, 1949; Turing, 1936].

Philosophical foundations for the theory should take into consideration other concepts which have been used to build the contexts in which reference has been made to the concept in question in order to reflect all already accumulated knowledge in the conceptual framework for the theory. For instance, the context for Shannon's entropy has been built by the probability theory, in particular probability distribution involved in the formula for entropy. Turing's analysis of computation is dependent on the idea of a state (of the processing unit, e.g. human computer or the machine, but indirectly also of the tape) and of its change. Neither of the original sources of the study of information have provided description of information or introduced its structural analysis. Shannon's entropy has been simply declared as a measure of information. Similarly, computation has been commonly called information processing, but both references to information have only character of interpretation going beyond the actual formal consideration. If we want to utilize the experience of these two domains, before a systematic theory of information is formulated, we have to find sufficiently broad conceptual framework which allows for consideration of the role of the concepts of a probability distribution and of a state.

It does not mean that the systematic theory has to repeat all conclusions of the studies made at the pre-systematic level, but that its philosophical foundations should allow for the reflection on advantages and disadvantages of the earlier results. The importance of the systematic approach is just in providing tools for the critical analysis of the concept as it was used in multiple contexts. In the case of information, there are many points where the earlier studies show weakness, such as apparent irrelevance of the meaning of information, and its theory should be able to overcome at least some of the earlier problems.

In the present paper, there is an attempt to select suitable philosophical foundations grounded in the broad range of contexts of the pre-systematic concept of information, and to propose a formalism which can be used for the purpose of the development of genuine theory of information.

Contexts of Information

The main two contexts of the pre-systematic concept of information have been already mentioned above together with their deficiencies for playing the role of theory of information, in particular their dissociation from meaning. Shannon's information theory has been criticized by the authors of attempts to develop semantic theories of information [Bar-Hillel & Carnap, 1952/1964], but these attempts have been no more successful in developing semantics of information, nor in formulation of adequate theory of information, than the orthodox approach.

It is not only that the measures of information, entropy or its generalizations do not refer to the meaning. More serious problem is that measuring information in terms of entropy seems inconsistent with the intuitive meaning of meaning. Even when we talk about information in a message, we construct the measure of information by the sum of entropy values for characters, while characters are units of the message below the level where we can talk about meaning. One character message is carrying measurable information, while in many languages it cannot have any meaning. Of course, we can introduce some meaning to one character by a new convention, but it is completely independent from the value of entropy.

Suppose we should avoid association of meaning with information. In the orthodox studies of information explicitly declaring irrelevance of the meaning for its goals there are frequent references to information as a reduction of uncertainty. How can meaningless information reduce uncertainty? If for instance it is understood that one character message, when arrives, is reducing uncertainty regarding which character would arrive, we are saying that it does not matter which character arrives, as the entropy is the same. But if it does not matter which character arrives, where is this uncertainty? Thus, it is not an issue *which* character arrives, but *that* a character is arriving. Thus, in such understanding each character of the alphabet is carrying the same information.

Sometimes, in the context of transmission of the one character message the measure of information is applied not to a message, but separately to each of the characters of the alphabet (as the logarithm with base two of the inverse of its probability) and entropy is interpreted as the mean value of measures for all characters. Here of course, we have clear reference to the meaning of the message, but meaning understood as the choice of a letter. Each letter automatically has different meaning. However, this is not what we understand by meaning of the language. If you do not know given language, you can access all letters of the message without knowing its meaning at all.

It is not much better in algorithmic information theory which has been developed by Andrey Nikolaevich Kolmogorov and independently by Gregory Chaitin within the framework of Turing machines. Here we have a curious result which tells us that the greatest measure of information is in random sequences of symbols. Even worse, we have the same algorithmic measure of information (the length of the shortest sequence of characters between its first left and first right non-blanks which produces the sequence under consideration in the process of computation) for sequences of different length and therefore of different entropy. The close relationship between the expected value of complexity measure for some special probability distributions and entropy which seems to connect the two measures from the perspective of algorithmic complexity theory [Li & Vitányi, 2008] does not mend the conceptual distinction. When in Shannon's approach the measure of information in a message is associated with the number of characters and the probability distribution of the selection of particular type of characters, algorithmic complexity is being derived from the structural characteristics of a sequence of characters based on the interaction between the states of the processing unit and that of the sequence.

At first, we could suspect that the algorithmic complexity may better reflect the meaning of information encoded in a sequence of characters, as it depends on the structural characteristics of the sequence. But actually, it tells us only about equivalence between sequences which does not even have to preserve the meaning sometimes associated with such sequences. For instance, if we use as a tape for Turing machine a message encoded in Morse alphabet with zero and one characters and run the machine, the meaning of the message most likely would be lost. Of course, there could be meaning of different type which can be preserved in the process of calculation, but there is no obvious way to find it.

Obviously, someone could say that the meaning and measure of information are two completely independent characteristics of the same entity. But then, what is the value of such measure for our understanding information? What actually is being measured? Moreover, we can find example showing that the same mathematical instrument, probability distribution of characters in the language, which determines the value of entropy (and all its generalizations) is a fundamental cryptographic tool (frequency analysis) for decryption of ciphers, i.e. for finding their meaning. It is true, that the knowledge of the value of entropy is not sufficient for decryption, we need all probability distribution, but we cannot expect complete independence of the meaning and measure of information.

While the algorithmic theory of complexity is clearly interested in the structural characteristics of information understood as sequences of characters, in Shannon's approach this interest is much less obvious. It can be seen in his interest in the description of messages in terms of probability distributions of subsequences (doubles, triples, etc.). What is common for both approaches, it is the assumption that the complexity or measure of information is related to the principle of linearity which requires that simpler units are put in a sequence to form units of higher rank. Certainly, it is a consequence of the original source of information studies in natural or artificial languages. The difference is in the fact, that implementation of the computing information systems related to algorithmic complexity theory involves propositional logic, while logical considerations in Shannon's approach are present only in the use of probability theory. However, it is clear that the study of information requires some involvement of logical or grammatical considerations, but in a generalized form appropriate for the level of generality transcending that of natural or artificial languages.

Thus far only the two main approaches to the study of the information related matters which have assumed paradigmatic role in science have been taken into account in reviewing the context for the concept of information.

However, the context is much wider. For the purpose of this article only three themes out of many in the disciplines which influenced the common sense view of information in the highest degree will be mentioned here.

The first is physics. Shannon hesitated to associate his measure of information with physics in spite of the decision to borrow the name of the measure of information from physical magnitude described by the same formula (if we disregard Boltzmann’s constant). However, in time information was becoming increasingly “physical” [Szilard, 1929/1983; Schrödinger, 1945; Brillouin, 1956; Landauer, 1991, 1998] to end up in the view of many physicists as a concept of equally or even more fundamental character than matter and energy [Wheeler, 1990]. In addition to bringing the physical theories into the context of information, its study has been enriched by new ideas of quantum theoretical form of information and computation. The intuitive association of information with the state of its carrier has been formalized due to the fact that the state in quantum mechanics is described by probability distribution, which in turn is the key concept in the construction of the measures of information. Even more important was the fact that the probability measures in quantum theory are defined on a more general structure than the Boolean algebra of sets.

The second theme belonging to biology had its source in the question about genetic inheritance, but in some sense has been already present in the discussion of the meaning and origin of life for long time. The study of biological inheritance has been influenced by the same small book “What is Life?” written by Erwin Schrödinger [Schrödinger, 1945], which made information a subject of interest for physics. Its reading prompted Francis Crick to move his scientific interests from physics to molecular biology, which ultimately led to cooperation with James Watson, the discovery of the structure of DNA and the mechanism of information transfer in living material. Genetics is not the only chapter of biology in which information plays crucial role. If we combine two concepts, of information and structural organization, we could say that all biology and evolution theory are about explanation of life in their terms, although not all biologists link their discipline with information as close as François Jacob [Jacob, 1973]. Biology could add to the physical point of view the need for explanation of the high level of organization and integration of information observed in living organisms, which in the past was interpreted in the form of the vitalism as a categorically different status of living matter and more recently as emergentism [Schultz, 1998; Emmesche, 2001].

The third theme is the study of cognitive processes in terms of information processing in the brain. It started from the work of W. S. McCulloh and W. H. Pitts [McCulloh & Pitts, 1943] and has been accelerated by the contributions of the authors of great authority [Wiener, 1948; von Neumann, 1951] leading through the times of great hope for the ultimate resolution of the mysteries of the brain in terms of artificial neural networks to the attempts of utilizing quantum mechanical description. Although the hope for explanation of consciousness in the distributive form of neural networks and integrated form of quantum mechanical mechanism has not been justified by the actual outcomes, information has become the key concept in the studies of cognition and consciousness. The latter has been even identified with the integrated information in one of the most active research programs [Tononi & Edelman, 1998a,b; Velmans & Schneider, 2007].

Thus, the context for the development of philosophical foundations which can serve as the point of departure for the systematic theoretical approach to information includes such concepts as a system (at different levels of organization or integration), its state, symbol and meaning, measure of information and its relationship with probability distribution, integration of information, selective and structural information, ontological status of information and its relationship to matter and energy, its relationship to consciousness and cognition.

In this paper, due to limitation of its scope, not all elements of the context are directly involved in the presentation of the philosophical foundations and the theoretical framework. However, they have been analyzed in more extensive way in the other papers of the author listed in references.

Philosophical Foundations

The main objective of this paper is to present a formal theoretical model of information based on the particular choice of philosophical foundations for the concept of information proposed by the author in the earlier articles. To make the article self-contained an outline of these foundations is included below together with brief references to some of the concepts from the context of information described above. More extensive explanation can be found elsewhere [Schroeder, 2005, 2009].

The approach is built upon the very old philosophical tradition of the reflection on the relationship between unity and variety (one-many opposition). Information is understood here in this conceptual philosophical framework as an identification of the variety i.e. that which makes one out of the many (or creates unity out of variety). It presupposes some variety (many) which can be identified as a carrier of information, and some form of unity (one) which is predicated of this variety. Since the relationship (opposition) of one to many is relative, so is the concept of information understood this way.

There are two most basic ways the many can be made one, by a selection of one out of many (selective manifestation of information) or by a structure introduced in the many which unites it into a whole (structural manifestation of information). These are two complementary manifestations of information, not separate types of information, as either of them requires the presence of the other, although possibly with respect to a different information carrier, i.e. different variety. If the elements of the variety are devoid any structure, it is difficult to expect any information involved in the selection of one of them. The selection of one out of many is purely random. On the other hand, every particular structure imposed on the elements of the variety can be considered an outcome of the selection of one of a variety of possible structures. In the first case, the original variety of the elements is different from the variety formed by the structural subcomponents of each of the elements. In the latter case, the original variety of elements bound into a structure is different from the variety of potential structures. So the transition between different manifestations of information requires a change of the information carrier.

As a consequence of this understanding of information, there are two its main characteristics. One is quantitative, referring to the selective manifestation. If the selection of one out of many can be described by probability distribution, a measure of information reflecting the size of the variety and the level of determination of the selection can be the familiar entropy of Shannon which in the finite case is given by formula (1):

$$H(n,p) = -\sum p_i \log_2(p_i), \ \sum p_i = 1, \ \text{for } i=1,2,\ldots,n. \qquad (1)$$

or, to be consistent with the definition considered above, rather the alternative, but closely related measure given by formula (2) advocated by the author [Schroeder, 2004]:

$$Inf(n,p) = H(n,max) - H(n,p) = \sum p_i \log_2(np_i), \ \sum p_i = 1, \quad \text{for } i=1,2,\ldots,n. \qquad (2)$$

It should be observed that the use of probability theory or the mathematical formula for the measure does not constitute any choice of the formal concept of information. It does not give us any knowledge of the properties (qualities) of information. Also, we have to remember that the measure can be introduced under the condition that a probability distribution describing the selection has been already defined, which not for all instances of information is possible.

The other is qualitative characteristic (but possibly admitting a quantitative form) referring to the structural manifestation, of a level of information integration which reflects the mutual interdependence of the elements of a variety [Schroeder, 2009]. This characteristic has been derived by the author from the formalism of quantum theory and within the more general philosophical considerations can be understood as indecomposability of the structure of the variety into independent components.

Before we will proceed to the formalization of the concept of information, a few further philosophically significant concepts will be presented. First of them is a concept of an information system which divides information carrier into portions of completely integrated information which can be identified with the identities of objects, and non-integrated information which can be identified with the states of the objects and their relations. Risking possible confusion, this division can be compared to the distinction between the essential and accidental properties of objects in Aristotelian philosophy. However, here objects do not possess properties or participate in them, but they are constituted by information in its integrated form. It is important to observe that this view is not necessarily leading to transcendental philosophical position. We can assume that information has epistemological status, consider identity of objects as different from “things as they are” and continue further reflection about the relationship (e.g. causal) between an object as it is and the integrated information constituting its identity, or we can stop at this point and assume that there is nothing beyond information integrated into identities of objects.

The second philosophical concept important for the context of the presented here theory of information is that of symbol and its meaning. It seems quite clear that symbolic relationship between a sign and its meaning appeared in the process of humanization of the ancestors of modern humans. In the view of the author it was a way to overcome the limitations of information processing in the brain which even now do not allow for handling more than about seven items at a time, which gave the title to the famous paper of George Miller “The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information.” [Miller, 1956/1994]. Most likely the magic is in the number three, as the number of seven objects gives us a variety of the size eight, when we remember that our brain has to handle the case of the absence of any items, and these eight choices require three bits of information. In any case, one of possible solutions to the limitation was a process of information nesting, which requires that the system carrying big volume of information is replaced in the processing by a system with much smaller volume. Thus, symbol is not pointing from a word (information unit) to its meaning (object of different ontological status,) but from an information system of small volume to one of a big volume. Therefore, in such understanding of symbolic representation it is a purely informational relationship. Someone committed to the position of the epistemological status of information can associate the meaning of information as described above (i.e. the information system represented by the symbol) with an interpretant in the tripartite approach of Peirce, but in the opinion of the author, it does not help at all in the understanding of the function of a symbol in which “object as it is” does not participate at all.

Summarizing, little bit more formally we could think about the relationship between a symbol and its denotation as a relationship between two information systems such that we have a correspondence between the identities of objects and a parallel relationship between their states.

Formalism for the Theory of Information

Now we can proceed to the formalization of the concept of information in reference to the philosophical definition of information as that which gives the unity to a variety. Thus the starting point could be to associate a set S with the variety and to consider simply the relationship resulting from set membership: $x \in S$. Although the membership in set S is determined by some (informally understood) information characterizing these and only these elements which belong to S, but it may have many different forms, for instance by a property predicated about x, or simply by listing of elements. Also, it would describe only the selective manifestation of information.

The set S may have, and usually actually does have some structure, as the varieties which we encounter in our experience are structured. Thus, formalization of information may be started from some very general structure defined on the set S, or better on the family of all its subsets. If we start from the power set of S, we can consider both manifestations of information. We can consider membership of some of subsets of S in a distinguished family of subsets which gives us association with the selective information through the membership, and at the same time we can assume that the distinguished family of subsets consists of exactly these subsets which inherit from S its possible, original structure, which gives us an association with the structural manifestation of information.

Information itself can be understood in this framework as a collection of the subsets of this distinguished family which form filter (in the algebraic sense). This means that together with every subset in the filter, all subsets of the family including it also belong to the filter, and additionally the filter is closed with respect to intersection. The reason for using filters as a description of information is the fact that both a selection of one out of many and making one of the many has its consequences which also have features of information. White horse is a horse, so information consisting in selection of a white horse is also selecting a horse. On the other hand, the structural characteristics of a white horse include structural characteristics of a horse.

In a slightly more formal way, for the purpose of defining information we distinguish a family $\Im$ of subsets of S, such that $S \in \Im$, and which is closed with respect to arbitrary intersections. This family can be easily recognized as a Moore family of subsets of S which defines a (transitive) *closure operator* on S, i.e. set S with a function *f* called a closure operator assigning to every subset A its closure *f*(A), such that

$$\text{for all } A, B \subseteq S: A \subseteq f(A),\ \ A \subseteq B \Rightarrow f(A) \subseteq f(B), \text{ and }\ f(A) = f(f(A)).$$

The set S with a closure operator *f* form a *closure space* $<S, f>$. Every closure operator on a set S is uniquely defined by the Moore family of its *closed subsets* $f\text{-Cl} = \{A \subseteq S: A = f(A)\}$, and every Moore family $\Im$ of subsets of S, i.e. family of sets which includes S and is closed with respect to arbitrary intersections, is the family of closed sets for the unique closure operator defined by $f(A) = \cap\{B \in \Im: A \subseteq B\}$. It is easy to see that for every closure operator its family of closed sets forms a complete lattice L_f with respect to the set inclusion [Birkhoff, 1967]. Information is now defined as a filter (or dual ideal) in a set theoretical sense within $\Im$, or equivalently a filter in a lattice sense in L_f.

The semantic relationship between two information systems (i.e. between a system playing the role of a symbol and another system playing the role of denotation) in this formalism is given by a continuous function (in sense of generalized continuity for closure spaces more general than topological) from the latter to the former. It can be associated with the concept of a random variable for an extreme case of a Boolean algebra of all subsets of given set, and with an observable in another very special case of quantum logics. After all, magnitudes expressed in numerical form are symbolic representations of information inherent in physical systems.

This complete lattice of closed subsets L_f with respect to the set inclusion is of special interest, as it can serve as a tool to characterize and classify the structures (e.g. algebraic, geometric, topological, etc.) introduced in the set S, when the closure is defined by the Moore family of its substructures. We can call this lattice the *logic of an information system* described by the closure space <S, *f*>.

Even more important is the role of the lattice of closed subsets in the characterization of the level of integration of information in the system. There are two extreme possibilities. In one this lattice is completely irreducible, this means it is not isomorphic to the direct product of any pair or collection of lattices. This corresponds to completely integrated information. An example of such lattice can be found in the class of so called quantum logics, i.e. lattices of closed subspaces of a Hilbert space (here too we have a closure space) in the formalism of quantum mechanics [Jauch, 1968]. Irreducibility of quantum logic is equivalent to unlimited applicability of the Superposition Principle, which is the core characteristic of quantum mechanical systems.

The other extreme case is of a trivial closure space in which all subsets are closed, i.e. we do not have any structure on the set S. Then the lattice of closed subsets is distributive and is a Boolean algebra of all subsets. In physics this case is associated with a purely classical mechanical system. Boolean algebras are completely reducible to the direct product of trivial two-element structures. This corresponds to the case of completely disintegrated information.

There are possible intermediate cases when the lattice of closed subsets is reducible to a direct product of non-trivial irreducible lattices. In this case we can distinguish so called center, a subset of elements which form a distributive sublattice and has properties similar to the completely reducible case (disintegrated information,) but the factor lattices into whose direct product the lattice is reduced are completely irreducible. Each minimal element (atom) of the center corresponds to one factor. Thus, we have separation of the system into purely integrated portions of information which can be interpreted as identities of objects, and the part of information which is purely disintegrated and can be interpreted as that which describes the state of object or objects.

At this point it is important to notice the terminological discrepancy between the use of the term “state” here and in quantum theory, where there is no distinction between the state and identity of the quantum-mechanical object.

Closure spaces which have as their lattices of closed subsets purely quantum logics are very special examples of irreducibility. Quantum logics are defined with an additional structure of orthocomplementation (see Appendix) imposed over the lattice structure which allows for the definition of a generalized probability measure necessary to define a physical state, and from our point of view it is important that only then we can define entropy or other familiar measures of information.

While orthocomplementation is an independent structure from that of the lattice (in the sense that we can define two different orthocomplementations on the same lattice producing non-isomorphic ortholattices), this additional structure can be introduced without going beyond the language of closure spaces. However, the condition for closure spaces to admit orthocomplementation is quite restrictive. This means that the quantitative description of information is possible for a restricted class of information systems.

There is a legitimate question whether we need such a high level of generalization to include the cases when the familiar quantitative description is impossible. The answer is positive, as there are many important information systems described in terms of closure spaces which do not admit it, such as for instance geometric or topological information.

Appendix

In the following, the basic concepts of quantum logics are reviewed for the convenience of the reader, as they provide a very important special instance of the more general structures of the logic of information (as described above) in which it is possible to introduce probability measure, and therefore familiar forms of information quantification. More details can be found in a variety of books on the subject [Jauch, 1968].

The quantum logic formalism of quantum mechanics (in its most conservative form) can be understood as reformulation of the Hilbert space (“standard”) formalism in the more abstract terms of complete orthocomplemented lattices and probability measures defined on them, generalizing the “classical” probability theory on Boolean algebras. There are some further generalizations of the concept of quantum logic which are of little interest for us, as they require separation from the conceptual framework of closure spaces. Thus, the conceptual basis of quantum logic presented here consists of the partial order relations with the increasing level of completeness (lattices, complete lattices) and their dual automorphisms.

The fundamental structure of a purely quantum logic is defined as an orthocomplemented, orthomodular, complete atomic and atomistic lattice with the atomic covering property and exchange property.

Thus, it is a complete lattice $\langle L, \wedge, \vee, 0, 1\rangle$ with the meet operation “$\wedge$”, and join operation “$\vee$”, the least element “0” and greatest element “1.” The lattice structure is associated with the partial order defined by:

$$a \leq b \text{ iff } a \wedge b = a \text{ iff } a \vee b = b.$$

In a lattice L with the least element 0 we can distinguish a subset of elements called atoms

$$At(L) = \{p \in L: \forall x \in L: 0 \leq x \leq p \Rightarrow x = 0 \text{ or } x = p\},$$

whose elements are all minimal non-0 elements.

The lattice defining quantum logic is atomic, i.e. every element of the lattice is greater than or equal to at least one atom. It is atomistic, i.e. every element is a join of atoms smaller than it.

Also, it has the atomic covering property which means that $\forall a,b \in L \forall p \in At(L)$: $a \leq b \leq a \vee p \Rightarrow a = b$ or $a = a \vee p$.

Finally, it has the exchange property which means that $\forall a \in L \forall p,q \in At(L)$: $a \wedge p = 0$ and $p \leq a \vee q \Rightarrow q \leq a \vee p$, or equivalently $a \vee p = a \vee q$.

In an atomistic lattice the last two properties are equivalent, and such a lattice is called simply an AC lattice.

Now, we can introduce a concept which goes beyond the formalism of lattices.

An orthocomplementation on L, is an involutive anti-automorphism, i.e. a bijective mapping of L on itself ($a \rightarrow a^*$) such that: (1) $a^{**} = a$, (2) $a \leq b \Rightarrow b^* \leq a^*$, and $a \wedge a^* = 0$ and $a \vee a^* = 1$. Frequently, the fact that $a \leq b^*$ is written $a \perp b$ and is read “a is orthogonal to b.” A lattice with an orthocomplementation is called an ortholattice. The separate term “ortholattice”is justified by the fact that the same lattice can admit two non-isomorphic orthocomplementations.

In an ortholattice the properties of being atomic and atomistic are equivalent.

Boolean algebra is an example of an ortholattice, with the complementation playing the role of orthocomplementation ($a \rightarrow a^*$ can be simply defined by $a \wedge a^* = 0$ and $a \vee a^* = 1$, as for every element a there is exactly one a^* which satisfies this condition, while in general there may be many such elements).

Atomic Boolean algebras play a very special, and in some sense trivial role of the purely “classical logic” characterized by the distributive property. Lattice L is distributive if $\forall a,b,c \in L$: $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$.

Distributive equality is not satisfied in quantum logics of quantum mechanics, and actually, the violation of this rule is critical for the distinction between the classical and quantum mechanics. The “global” classical character of the logic for classical mechanics can be localized to a pair of the elements in any ortholattice. We call two elements compatible aCb if $a = (a \wedge b) \vee (a \wedge b^*)$. Of course, in a Boolean algebra every two elements are compatible, while in the purely quantum case we will have that the only elements compatible with all other are 0 or 1.

In every ortholattice L, $\forall a,b \in L$: $a \le b \Rightarrow$ aCb, but in general not necessarily bCa. Quantum logics are required to have the relation of compatibility symmetric which is equivalent to the requirement of orthomodularity.

An ortholattice L is orthomodular if $a \le b$, then $b = a \vee (b \wedge a^*)$.

This property is equivalent to the condition that for any pair of elements $c,d \in L$ such that $c \le d$, the interval [c,d] of L, i.e. the set of elements $[c,d] = \{x \in L: c \le x \le d\}$ which always forms a sublattice, is a sub-ortholattice, with respect to the orthocomplementation defined by $a^{\#} = (c \vee a^*) \wedge d$.

While the relation of compatibility describes the classical aspect of the structure, the atomic bisection property or in other words irreducibility property refers to the quantum character. The condition of irreducibility or coherence is defined in terms of atoms. If p and q are different atoms of L, then there exists a different third atom r, such that $r \le p \vee q$, or equivalently that $p \vee q = p \vee r = r \vee q$. It is called the Superposition Principle, as it is a counterpart of this principle in the conventional formalism.

It is easy to see that this condition is never satisfied in Boolean algebras, but is an axiom for quantum logics (exchange property.) Moreover, if it is satisfied by an omolattice L (short for orthomodular lattice,) then the only elements compatible with all other elements of L are 0 and 1. Thus, it is a characteristic of a purely quantum or purely coherent system.

The name irreducibility condition comes from the fact that it is equivalent to the condition that the quantum logic is not isomorphic to a direct product of other quantum logics. On the other hand every atomic complete Boolean algebra is isomorphic to the product of two element Boolean algebras, each of them trivially satisfying the condition of irreducibility. The intermediate case (quantum, but not purely quantum logic) will be the case when given quantum logic is isomorphic to a product of its nontrivial, coherent, irreducible components. There is a bijective correspondence between the coherent components of L and the atoms of the center of L, i.e. of the set of the non-0 elements of L which are compatible with all elements of L.

The following representation theorem is irrelevant for the objectives of the article, but it is included here to explain the relationship of the quantum logic formalism to the standard formulation of quantum mechanics.

Every purely (i.e. irreducible) quantum logic, i.e. an orthocomplemented, orthomodular, complete atomic AC lattice (the redundant conditions have been eliminated) can be represented by a lattice of closed subspaces of a general Hilbert space. There is some additional condition necessary to have representation in a Hilbert space over a field such as of complex numbers in the standard formalism of quantum mechanics, but it is of no interest in this context.

It has to be emphasized that in the usual case of quantum mechanical physical systems which admit superselection rules their quantum logics belong to the partially reducible intermediate type.

On a quantum logic (not necessarily irreducible,) we can build full physical formalism of quantum mechanics by defining states of the physical system by generalized probability measures. The axioms are basically the same as in classical probability on Boolean algebras, so it is the underlying ortholattice structure which makes quantum probability different.

The state on the quantum logic L is a probabilistic measure μ: L$\rightarrow\Re$, such that

(1) $\mu(0_L) = 0$, $\mu(1_L) = 1$,

(2) For every $a \in L$: $\mu(a) \geq 0$,

(3) For every countable family of mutually orthogonal elements of L,

$$\vee \{a_i \in L, i \in N: \forall i,j \in N: i \neq j \Rightarrow a_i \perp a_j\} = a \Rightarrow \mu(a) = \sum \{\mu(a_i), i \in N\}.$$

To complete the physical formalism, we can define observables (physical magnitudes) as ortho-homomorphisms from the Boolean algebra of Borel subsets of the set of real numbers to the quantum logic. It is easy to see that observables are equivalent to random variables, but defined in the more general context of quantum logics. In classical probability, the random variables are usually defined as functions from the set of outcomes to real numbers, but they can be defined in an equivalent way as above.

When a purely quantum logic is represented by an ortholattice of closed subspaces of a Hilbert space, or as an ortholattice of projections in a Hilbert space, the observables defined for the quantum logic become selfadjoint linear operators defined by their projection valued spectral measures.

What is a natural description of quantum coherence in terms of irreducibility cannot be reintroduced in terms of a Hilbert space, as each coherent component (or rather factor) of a reducible quantum logic requires a separate Hilbert space representation. This explains why superselection rules corresponding to the reducibility conditions for a quantum logic do not have easy interpretation in the conventional formalism.

As we could see, the concept of an orthomodular lattice is a generalization of the concept of a Boolean algebra, so quantum theory can be considered a generalization of probability theory. There is a natural question in what degree the structure of orthocomplementation imposed on the lattice of quantum logic is restricting the choice of closure operators for which this lattice is the lattice of closed subsets. The answer is that even if we eliminate the condition of orthomodularity the restriction is quite strong. It turns out that such closure operator can be defined by an appropriate symmetric binary relation [Ore, 1943].

Many lattices interesting for the purpose of the study of information, for instances lattices generated by some geometric or topological structures, do not satisfy this condition. Thus, the formalism of general closure spaces which seems to be a good for the study of the structural manifestation of information is too general to be used in the quantitative analysis of selective manifestation of information, as long as we have to use some form of probability measures.

Conclusion

The formalism presented above is only a first step towards authentic theory of information. Its strength is in giving the framework for considering a very wide range of contexts in which the pre-systematic concept of information has appeared (to the best knowledge of the author virtually all contexts). Its another strength is in being based on a philosophical foundations which are very general, but also highly nontrivial in the sense that in the form of the reflection on the opposition of one and many they have been studied in connection with all fundamental problems of philosophy. On the other hand, the concept of information formulated within this conceptual framework is very promising in giving answers to philosophical problems belonging to most difficult in the intellectual traditions of many cultures.

Since the presentation of the formalism is very concise, many of its aspects require more detailed exposition which is included in forthcoming papers of the author. In better understanding of the formalism two very special instances are very useful, that of completely disintegrated information described within Boolean algebras of subsets, and that of completely integrated information described in quantum logics. However, specifics of these two instances (for instance existence of orthocomplementation in the lattice of closed subsets giving an easy association with traditional logic) can be misleading. The need for consideration of geometric or topological types of information requires much higher level of generality.

Bibliography

[Bar-Hillel & Carnap, 1952/1964] Y. Bar-Hillel and R. Carnap. An Outline of a Theory of Semantic Information. Technical Report No. 247, Research Laboratory of Electronics, MIT, 1952; reprinted in Y. Bar-Hillel, Language and Information: Selected essays on their theory and application. Reading, MA: Addison-Wesley, 1964, pp. 221-274.

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

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

[Emmeche, 1997] C. Emmeche. Does a robot have an Umwelt? Semiotica, 134, 653-693, 1997.

[Jacob, 1973] F. Jacob. The Logic of Life: A History of Heredity. Transl. B. E. Spillman. New York: Pantheon Books, 1973.

[Jauch, 1968] J. M. Jauch. Foundations of Quantum Mechanics. Reading, Mass.: Addison-Wesley, 1968.

[Landauer, 1991] R. Landauer. Information is Physical. Physics Today, May, 23-29, 1991.

[Landauer, 1998] R. Landauer. Information is Inevitably Physical. In A. J. G. Hey (ed.) Feynman and Computation: Exploring the Limits of Computers. Perseus, Reading, Mass., 1998, pp. 77-92.

[Li & Vitányi, 2008] M. Li and P. Vitányi. An Introduction to Kolmogorov Complexity and Its Applications, 3-rd ed.; New York: Springer, 2008.

[McCulloch & Pitts, 1943] W. S. McCulloch and W. H. Pitts. A logical calculus of the ideas immanent in nervousactivity. Bull. Math. Biophy, 5, 115-133, 1943.

[Miller, 1956/1994] G. Miller. The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. Psychological Review, 63, 81-97, 1956 (reprinted in 101 (2), 343-352, 1994).

[Ore, 1943] O. Ore. Some Studies On Closure Relations. Duke Math. J., 10, 761-785, 1943.

[Schrödinger, 1945] E. Schrödinger. What is Life? Cambridge: Cambridge University Press, 1945.

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

[Schroeder, 2005] 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.

[Schroeder, 2006] M. J. Schroeder. Model of structural information based on the lattice of closed subsets. In Y. Kobayashi, T. Adachi (eds.) Proceedings of The Tenth Symposium on Algebra, Languages, and Computation, Toho University, 2006, pp.32-47.

[Schroeder, 2007] M. J. Schroeder. Logico-algebraic structures for information integration in the brain. Proceedings of RIMS 2007 Symposium on Algebra, Languages,and Computation, Kyoto: Kyoto University, 2007, pp. 61-72.

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

[Schultz, 1998] S. G. Schultz. A century of (epithelial) transport physiology: from vitalism to molecular cloning. Am. J. Physiol. Cell. Physiol., 274, C13-23, 1998.

[Shannon, 1949/1989] C. E. Shannon. The Mathematical theory of communication. In: The Mathematical Theory of Communication. Ed. C. E. Shannon and W. Weaver. University of Illinois Press, Urbana, 1949, reprinted 1989.

[Szilard, 1929/1983] L. Szilard. On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings. Z. Phys.53, 840. Engl. Transl. in J. A. Wheeler and W. H. Zurek, Quantum Theory and Measurement. Princeton, N. J.: Princeton Univ. Press, 1983.

[Tononi & Edelman, 1998a] G. Tononi and G. M. Edelman. Consciousness and Complexity. Science, 282, 1846-1851, 1998.

[Tononi & Edelman, 1998b] G. Tononi and G. M. Edelman. Consciousness and the Integration of Information in the Brain. Advances in Neurology, 77 (1998), 245-280, 1998.

[Turing, 1936] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. Proc. London Math. Soc., Ser. 2, 42, 230-265, 1936, with a correction, ibid., 43, 544-546, 1936-7.

[Velmans & Schneider, 2007] M. Velmans and S. Schneider, Eds. The Blackwell Companion to Consciousness. Malden, MA: Blackwell, 2007.

[von Neumann, 1951] J. von Neumann. The general and logical theory of automata. In L. A. Jeffres, (ed.) Cerebral Mechanisms in Behavior: The Hixon Symposium. New York: Wiley, 1951, pp. 1-32.

[Wheeler, 1990] J. A. Wheeler. Information, physics, quantum: The search for links. In W. H. Zurek (ed.) Complexty, Entropy, and the Physics of Information. Redwood City, CA: Addison-Wesley, 1990.

[Wiener, 1948] N. Wiener. Cybernetics: or Control and Communication in the Animal and the Machine. Cambridge, MA: MIT Press, 1948.

Authors' Information

Marcin J. Schroeder *– Professor and Dean of Academic Affairs, Akita International University, Akita, Japan; 193-2 Okutsubakidai, Yuwa, Akita-shi, 010-1211 Akita, Japan; e-mail: mjs@aiu.ac.jp.*

Major Fields of Scientific Research Related to Information: Philosophical and formal aspects of information; Theoretical models of consciousness; Information integration;