© 2006 Imprint Academic Mind & Matter Vol. 4(2), pp. 141–167

Pragmatic Information: Historical Exposition and General Overview

Dieter Gernert
Technische Universität München
München, Germany

Abstract

Pragmatic information, understood as the impact of a message upon a receiving system, represents a matured and comprehensive concept of which earlier proposals are special cases. The different kinds of recipients and reactions to incoming message are characterized. In a historical exposition the principal approaches to the definition and operationalization of information are critically reviewed. From a modern point of view, the measurement of pragmatic information is possible but requires novel and specific procedures. As a perspective notion, pragmatic information will be analyzed in its relationships with other perspective notions, particularly "meaning" and "interpretation". Main fields of application include physics, general systems theory and cognitive science. Together with some reflections on information and meaning, these areas of application point forward to connections with the mind-matter problem.

1. Pragmatic Information: an Informal Approach

The great variety of applications and cross-references – from modern physics to epistemology – makes an advanced and up-to-date concept of *information* indispensable. Information is generated as soon as a channel has been passed and the process of signal transmission is accomplished. Information exists as soon as the structure or the behavior of a recipient have been altered. The concept of information epitomized in this way has become known as *pragmatic information.* It is the purpose of this paper to expound and to analyze this concept under multiple aspects, including:

- its historical development and the motivation behind the concept,
- its relationship with other (earlier) concepts of information,
- an approach towards a mathematical formalization,
- an overview of possible applications (with a final outlook on the mind-matter problem).

The current concept of pragmatic information unifies earlier and alternative proposals (details will be given in Section 3). Pragmatic information is not in contrast or contradiction to other information-based approaches, which have their own merits. Unfortunately, one early approach (Shannon and Weaver 1949) has been dubbed and popularized as "information theory" such that it may be erroneously understood as *the* unique and authoritative concept, defining the standard for an evaluation of competing proposals.

As will be shown in detail, the time-honored Shannon theory – still an indispensable tool in its traditional domains of application – is included in the comprehensive concept of pragmatic information as a special case. (A first attempt to show this can be found in Kornwachs and Lucadou (1982).) Therefore, no prior knowledge about the "classical" theory is required in order to understand the modern concept (nor would such prior knowledge make the access easier[1]). The following exposition starts immediately with the notion of "impacts upon a recipient" and will later proceed to a specific mathematical treatment.

2. Information as an Impact upon a Recipient

2.1 Receiving Systems and Their Possible Reactions

According to Carl Friedrich von Weizsäcker (1985, p. 351), "information is only that which is understood", and such an understanding "is possible only in sufficiently large composite systems" (p. 355f). A *receiving system* or *recipient* may be an individual, a social system (team, group, organization), an animal, a part of a living organism, a technical information-processing system, or even a component of a physical process. Understanding – or, more generally, the impact upon the receiver – can manifest itself in an immediate reaction or in the enhancement of a repertory if, for instance, something learned today can be utilized later. If the receiving system is a programmable computer then some new software can carry pragmatic information if it brings about a superior performance of the system.

But the capability of an incoming message to trigger a change within the receiving system goes deeper: A message can modify the receiver's structure or behavior with the consequence that the impact of another message arriving later (or the same message coming in again) will be different. This is frequently expressed by saying that "pragmatic information can alter the basis of its own valuation" (Weizsäcker (1974, p. 99), Kornwachs and Lucadou (1982).

An illustration is given by a message consisting of two parts, where the second part is an exact duplicate of the first one. Having done with the first part, a human reader will be bored and ignore the rest – except

---

[1] A modern textbook (Roederer 2005) comprehends both the classical theory and the concept of pragmatic information; the latter is marked as "the primary, fundamental concept" (Roederer 2005, p. 122).

for the special situation that the second part is needed to check whether some dubious entries are due to noise in the channel or not. Thus, the meaning of a message depends on the receiver's individual requirements and disposition, a central point to be studied in the sequel.

2.2 Novelty and Confirmation

Two notions which are fundamental for the understanding of pragmatic information were proposed by Ernst and Christine von Weizsäcker (1972): *novelty* and *confirmation.* (Note that occasionally the term *primordiality* is used instead of *novelty.*)

If a message only repeats material already known to the recipient, then it only conveys confirmation, but no novelty (as in the case of exact duplication). On the other hand, a "breaking news" in an unknown foreign language may contain a great deal of novelty, but because there is nothing in common with the receiver's prior knowledge, nothing will be understood. This is expressed by the statement that "there is no confirmation". The latter term is used here in a specific technical sense: It addresses the requirement that a message must be understandable to the receiver, refer to the receiver's previous knowledge, and contribute at least partially to an existing information demand.

A measure of the information contents – to be developed later in Section 4 – must be zero if either novelty or confirmation is zero; in all other cases it must take on positive values. With some reservations (Section 4.3), both novelty and confirmation can be determined quantitatively, and the same holds for pragmatic information as a function of these variables.

It should be mentioned already here that the notion of pragmatic information is a *perspective notion*: The information contents of a message cannot be determined "absolutely" – regarding only the message itself – but requires an analysis comprehending the receiver's previous understanding, interests, and subjective information demand as well (see Section 5).

2.3 Special Effects in the Reception of a Message

In order to avoid an inadequate "mechanistic drive", it should be kept in mind that some strange effects can occur when a message is received by an individual or a social system.

- A message may be inconsistent if parts of it are incompatible with other parts, or the impacts produced by some parts are weakened or neutralized by other parts.
- Possibly the effect of an incoming message can be correctly recognized only in retrospect (e.g. somebody regards a message as a trick and stores it together with that reservation).

- The way in which a message is understood by the recipient can be arbitrarily different from the sender's intentions.
- In social systems some proportion of a message may reach the wrong recipient.

3. Historic and Alternative Approaches

3.1 Early Attempts and Early Criticism

Historically, the attribute "pragmatic" has been strongly utilized by the American philosopher Charles William Morris (1901–1979). Starting mainly from specific parts in the extensive work of Charles Sanders Peirce (1839–1924), Morris was the first to formulate semiotics, the theory of signs, as an explicit and elaborated concept. The well-known triple of syntactics, semantics, and pragmatics is essentially due to Morris (1938), who gives the following definitions:

- *syntactics* is the theory of the relations between signs;
- *semantics* is the theory of the relations between signs and the objects designated by them (designata);
- *pragmatics* is the theory of the relations between signs and their users.

Already early in the era of Shannon's theory it was realized that the concept of information has a Janus face. "It is surprising but true that, from the present viewpoint, two messages, one heavily loaded with meaning and the other pure nonsense, can be equivalent as regards information" (Weaver 1949, p. 12, similarly in Shannon and Weaver 1949, p. 99). Here the term "information" refers to *syntactic information*, whereas the word "meaning" and the sceptical phrase "from the present viewpoint" may be regarded as a premonition of the concept of *pragmatic information.*

As a simple example, consider an original "normal" text and its "scrambled version", a randomized permutation of the letters. On the basis of simple information measures,[2] the syntactic information (as, e.g., quantified by transmission fees), is the same, whereas (much of) the pragmatic information is destroyed.

Originally, the theory mainly created by Shannon was intended as a theory of *information transmission.* As the linguist Bar-Hillel (1969) pointed out, the original term "theory of information transmission" was abridged to "theory of information" and then converted into "information theory". This happened although the basic structure and the scope of the theory remained unchanged.

---

[2] Special information measures which are sensitive to permutations of letters are not relevant here.

In the years following the original publications by Shannon and Weaver (1949), various endeavors were made to build and modify the theory, and to try applications to many scientific problems. In an astounding parallelism with cybernetics, which grew up and was fashionable at about the same time, the Shannon-Weaver approach was, e.g., mistaken for a "universal key" to a diversity of problems in science and society, or even as a "philosophy of scientific knowledge"(Kampé de Fériet 1974, p. 2). Evidently, we cannot understand a society unless we also understand its way of dealing with information, but it is unclear how that theory could advance such an understanding. Similar to the history of cybernetics, exaggerations and false expectations, together with foreseeable disappointment, could not be avoided.

Shannon himself felt urged to argue against the public overestimation of a discipline essentially created by himself, and published a leading article entitled "The Bandwagon" in a distinguished journal. The spirit at that time can be best summed up by a quotation from this publication (Shannon 1956, p. 3):

> Information theory has, in the few last years, become something of a scientific bandwagon. Starting as a technical tool for the communication engineer, is has received an extraordinary amount if publicity ... In part, this has been due to connections with such fashionable fields as computing machinery, cybernetics, and automation ... As a consequence, it has been ballooned to an importance beyond its actual accomplishments. Our fellow scientists in many different fields, attracted by the fanfare ..., are using these ideas in their own problems. Applications are being made to biology, psychology, linguistics, fundamental physics, ... , and many others ... While we feel that information theory is indeed a valuable tool ..., it is certainly no panacea for the communication engineer or, a fortiori, for anyone else.

It must be emphasized that Shannon and Weaver (1949) had clearly marked the limitations of their own theory by stating that the proposed measure of information does not take the meaning and usage of a message into account. Nevertheless, misunderstandings and subsequent disappointments could not be precluded.

In 1968, the criticism of misuses of the Shannon-Weaver approach was taken to extremes at a conference in Starnberg (cf. Ditfurth 1969, or Weizsäcker and Weizsäcker 1972, p. 540). A number of renowned experts, representing psychology, biology, linguistics, and computer science, frankly declared the Shannon-Weaver theory as "irrelevant" and "useless" with respect to their individual fields. Their arguments were based on the insight that no scientific discipline can ultimately restrict itself to the syntactic level of information alone.

3.2 The Limited Scope of the Concept of Entropy

In the classical Shannon-Weaver approach a function H, termed "entropy", was defined to measure the degree of uncertainty of a system before its state has been measured. It is a fundamental principle of this theory that a gain in information is equivalent with a reduction of uncertainty. Later on, it turned out that this choice of terminology became a source of errors and misunderstandings, and sometimes even a mindblocker; so some corresponding remarks may be adequate.

Indeed, talking about entropy obscures the fact that there are really two concepts of entropy: *thermodynamic entropy* and *informational entropy.* This ambiguity (or terminological sloppiness) became a source of shaky analogies and conclusions, which could be avoided by more precise terms (such as in Ebeling *et al.* 1998). The association with a familiar field, thermodynamics, and a couple of handsome formulas provoked the illusion that all open questions about information were settled with the Shannon-Weaver theory. Occasionally, this theory is even considered a standard by which any competing proposals would have to be assessed.

As emphasized by Shannon and Weaver, their theory is restricted to syntactic aspects of information. Thermodynamic entropy explicitly refers to a class of systems (or processes) which is traditionally studied by thermodynamics. Analogies are possible in other branches of science, but they become dubious already in biology (Ebeling *et al.* 1998, p. 32), and are even less useful in the study of social systems. The classical example is a human person who receives a very short message, but reacts in a spectacular manner due to the exceptional meaning transmitted.

Furthermore, the classical theory is based on the model of a telegraph station, which accepts incoming messages – series of signs from a fixed "alphabet" – and impartially prints them on paper, without any permanent change of its state. It is no longer tenable that entropy is *in general* a measure of uncertainty and disorder. The example of crystal growth shows that "order out of disorder" is possible, and disorder can be reduced (Weizsäcker 1985, pp. 175–178).[3] To sum up, we should leave the term "entropy" to those areas from which it originated.

3.3 Intermediate Proposals

In face of the limited scope of the classical theory of information, it was a natural development that various efforts were made to modify and extend that theory, mainly with the intention to quantify the meaning of a message.

---

[3] In such cases, an excess of entropy is produced somewhere in the surroundings of the system considered, such that the second law of thermodynamics remains valid; but here only disorder and entropy within the original system (in the case of crystal growth: the crystal itself) are relevant.

The triple of syntactics, semantics, and pragmatics was rather popular and frequently quoted in those attempts. So it is not surprising that proposals came up for a consideration of the semantic aspects, partially under labels like "theory of semantic information" (see Zoglauer 1996). If one wants to go beyond purely syntactic information, one can speak about "information" only if a set of signs carries a meaning. But regarding the meaning of some signs or signals, it turns out that meaning can never be identified or ascribed without reference to a specific – present or future – "user" or "receiver", i.e. a receiving system that interprets and makes use of the signs. For instance, our letter "C" will be understood as the letter "S" by anyone using Cyrillic scripture. This dependence of the meaning on the receiver's prior disposition can be disregarded only if there is a "tacit agreement", fixing a certain set of possible receivers. Hence, any concept of semantic information will be included in the concept of pragmatic information as the special case in which a class of possible receiving systems has been fixed, such that their uniform way of interpreting the signals is constant and predictable.

Marko (1966) modified the original concept by including bidirectional communication (signal transmission in both directions). Jeffreys (1946) and Kullback and Leibler (1951) proposed a special version of Shannon's theory by introducing a *distance* between statistical populations and based their measure of information contents on this distance. In a similar manner, Jumarie (1990) started from a statistical description of the receiver's previous information and proposed a concept of "relative information". Other approaches try to replace the usual measures of probability by measures of possibility or plausibility, or fuzzy measures (cf. Kampé de Fériet 1974, Klir and Folger 1988, Klir 1991, Kornwachs 1992).

Several proposals were based upon an estimate of "useful information", e.g. trying to assess the "utilities" of possible outcomes (together with their probabilities) or the enhancement of the probability to reach a given goal (Sharma *et al.* 1978, Straubel 1980, Copper 1992). By way of contrast, Weizsäcker (1985, pp. 189–199) scrutinizes the possible connection between information and utility, and comes to the result that an identification of both notions depends upon previous assumptions (e.g. the theoretical concept of the *homo oeconomicus*). But a satisfying definition of "utility" is also inaccessible and trapped in circularity in economic theory (the entrepreneur maximizes utility, and utility is what can be taken from the entrepreneur's maximizing behavior).[4] Another reservation is given by the necessity to estimate the probabilities of future events. Further remarks on the proposals compiled here are to follow in Section 4.1.

---

[4] This does not exclude utility functions which are defined for a restricted area of applications. For example, it may be possible to assess the utility of an additional cue in a searching task, in proving a mathematical theorem, or in interpreting a text (see van Rooij 2006).

3.4 Other Facets of the Concept of Information

The concept of information presented here includes the case that an impact upon a receiver manifests itself only in the future. Therefore, the following cases, which are mentioned as special types of information in the literature, are covered by the definition of pragmatic information:

- *Information on the way*: A sequence of signals just being transmitted in a channel has a chance to reach its addressee, and hence contains information. If it is valuated under the aspects of telecommunication engineering, exactly the classical Shannon-Weaver theory will be regained. This shows that the latter theory is indeed included as a special case. The same holds for any kind of entries in books, data bases, etc.
- *Structural information* (also termed *latent information* or *potential information*): In crystals or fossils the represented structure may be interpreted by experts after millions of years.
- *Active information*: This is a concept proposed by Bohm and Hiley (1993) mainly for purposes of interpreting quantum theory (see also Hiley 2002). It will be addressed in detail in Section 6.1.

4. How to Find Measures of Pragmatic Information

4.1 Necessary Preconditions for Measuring Procedures

Most of the earlier proposals outlined above, together with some of their combinations and ramifications, have a significant shortcoming: they do not open a pathway towards an analysis of meaning. The central points of criticism are:

- All occurring variables and terms are restricted to the syntactical level.
- All meanings of signs, valuations, probabilities, etc. are presupposed to be time-invariant. It is excluded to deal with time-dependence or with valuations altered by incoming messages.
- Those proposals assume that probability functions, valuations, etc. are known or at least can be derived from the situation. However, in fact they depend on the current interest and prior knowledge of the recipient.

In view of all the requirements discussed so far it is clear that a procedure for the measurement of pragmatic information cannot have the same structure as a measurement in classical physics. Consequently, some introductory reflections are adequate. In a modern understanding, *measurement* means that each object of a given class is mapped onto an element of a formal system $\mathcal{F}$.

- If the elements of $\mathcal{F}$ that are assigned to the given objects are real numbers – with the restriction that the usual arithmetical operations make sense – then we have the case of *cardinal measurement.*
- If the given objects are mapped onto an order or a semi-order, this is called an *ordinal measurement.* The character of an ordinal measurement is often camouflaged by the widespread habit to express the results of such a measurement by numbers – but these numbers only reflect a greater/less relation, and arithmetic (e.g. quotients) makes no sense.
- In a *nominal measurement*, the formal system $\mathcal{F}$ consists of a finite list of attributes, and one of these attributes is assigned to each of the given objects.

As demonstrated by successful applications – for instance in physics (see Section 6.1) – a cardinal measurement of pragmatic information is possible in special cases. In face of the broad spectrum of manners, however, in which individuals can react to incoming messages, a more comprehensive definition of "measurement" should be accepted that includes both ordinal and cardinal measurements.

As mentioned before, pragmatic information does never depend only on a message itself, but also on the *current state* of the receiving system. If a message happens to arrive a little bit later than expected, then the system state can have been altered meanwhile, and the pragmatic information will be different. This context-dependence of pragmatic information has significant consequences:

- If the reception of a message is regarded as an operator (which mirrors the fact that the system state is altered), then two such operators can be non-commutative. This possibility indicates that (non-trivial) acts of information processing are generally non-commutative (Section 6.3).
- Any formula for the information contents, any equation describing the transition of a system to another state due to new information, must take into account that the occurring variables are time-dependent. Rigorously speaking, such formulas are "only valid for short time intervals" (*snapshot principle*).

4.2 The Product Formula

After an early attempt by Gäng (1967), a significant step towards a quantification of pragmatic information was made when Ernst and Christine von Weizsäcker (1972) proposed the fundamental notions *novelty* and *confirmation.* In their original work it was presupposed that a cardinal measurement of both novelty N and confirmation C is possible; this "standard" case will be discussed in Section 4.3 (for the "non-standard" case see Section 4.4.). Under this assumption, the product formula[5] for pragmatic information P can be written as

$$P = N \cdot C \tag{1}$$

The usual arguments to motivate this formula are as follows: If either N or C equals zero, this means that a message conveys no pragmatic information. There are qualitative reasons that P will take on a maximal value if both N and C (each of which is limited by a natural bound) lie in an intermediate range.

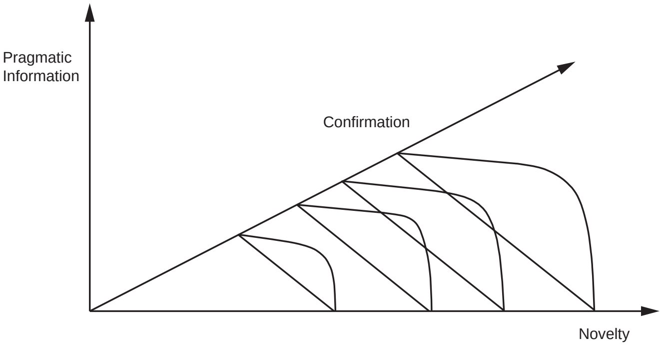

Figure 1: Pragmatic information as a function of novelty and confirmation (qualitatively), according to Weizsäcker (1984, p. 176).

Figure 1 shows, qualitatively, P as a function of the two independent variables N and C (with $N \geqslant 0$, $C \geqslant 0$). The characteristic "bulge" stands for the *concavity* of P according to Eq. (1). Each plane which stands orthogonally (vertically) on the (x,y)-plane and meets both the x- and y-axis under an angle of 45^o intersects the two-dimensional surface in a concave curve. This mathematical property is plausible since it would be absurd to assume a dent within that curve.

Yet a more profound analysis is necessary. First of all, Eq. (1) is not a unique formula with the requested properties. Rather, there are infinitely many functions defined for $x \geqslant 0$, $y \geqslant 0$ with

---

[5] As Weizsäcker (1972, pp. 549f) declared, they had not even a hypothesis "about the kind of a multiplication of novelty and confirmation". This formulation indicates that there need not be a unique shape of this formula. It seems that the first hint to the product formula stems from Kornwachs and Lucadou (1982).

$$\begin{aligned} f(x,y) &= 0 \text{ if } x = 0 \,, \\ f(x,y) &= 0 \text{ if } y = 0 \,, \\ f(x,y) &> 0 \text{ if } x > 0 \text{ and } y > 0 \,. \end{aligned}$$

An example for a formula with all these properties, but different from Eq. (1), is obtained by scaling the right-hand side of Eq. (1) with a positive constant $c \neq 1$. There are also pairs of admissible functions which do not simply differ by a scale factor.

Another source of confusion may be due to the dimension in which pragmatic information can be specified. The linguistic coincidence that the German word "Wirkung" means both the *impact* upon a recipient and *action* in the sense of physics (e.g., the product of energy and time) can, at most, be regarded as a heuristic cue (or a historical reminiscence), but not as a scientific argument. For instance, an observation in physics requires an activity on the observer's part. As considered by Atmanspacher (1989), the transfer of amount ΔI of syntactic information from an observed system S to the observer is achieved by means of a corresponding energy transfer ΔE from the observer to S. The rate $\Delta I/\Delta t$ of information transfer is limited as the rate of energy transfer is; ΔI can be transferred only in a finite time interval Δt. Under these conditions the transferred amount ΔI of information is given by $\Delta E \cdot \Delta t$, which has the dimension of an action. However, this action must be distinguished from the impact that the meaning delivered together with ΔI leaves upon the observer/recipient.

At this point another *caveat* seems to be adequate as well. In the literature, pragmatic information is often illustrated by a diagram showing the above-mentioned concave curve defined by the intersection of the bulge and a vertical plane. The vertical axis is labelled as pragmatic information, whereas the abscissa represents novelty and confirmation in such a way that one of them takes on values from 0 to 100%, while the other drops from 100% to 0. This is explained to mean that N and C add up to 100% and, to make things still worse, this feature is dubbed "complementarity". Apart from terminological sloppiness, such a diagram is neither a heuristic cue nor a didactic reduction, but simply misleading. It is not true that, under suitable conditions, N and C can vary maintaining a constant sum; nor can a text be partitioned into two subsets such that the words or phrases in one subset would convey novelty whereas the other subset would stand for confirmation.

In quantum theory, *complementarity* is closely related to the noncommutativity of variables, which means that the temporal order of measurement makes a difference. In the literature, arguments for the noncommutativity of N and C were given rather early (Weizsäcker 1974, Kornwachs and Lucadou 1982); occasionally the dimension of P is used as a cue. However, it is hard to imagine a thought experiment in which

– under identical conditions – N and C are measured twice, the second time in reverse order. A possible scenario would be the following:

- First ask a human subject to evaluate the degree of novelty of a particular text.
- Then, check the extent of text understanding and the interest taken by the same subject.

Alternatively, these two tests would have to be performed in reverse order. Evidently such a design seems rather artificial.

A different argument to support non-commutativity may start from the statement that cognitive processes can be described by sequences of operators which are, except for trivial cases,[6] non-commutative (Gernert 2000). The same property holds for human valuations underlying measurements of N and C: any valuation can alter the basis for a subsequent one – in perfect analogy with the fact that an incoming message can modify the basis for messages arriving later. After all, the product formula Eq. (1), the complementarity of N and C, and the dimension of P can be considered as plausible. This is in agreement with the arguments given by Weizsäcker (1974).

4.3 Measurement of Similarity

For a reasonable measurement of similarity or dissimilarity, respectively, two assumptions are usually made:

1. A cardinal measurement of N and C is possible.
2. Both a message M and a receiver's expectation E can be formally described, such that – after identifying a suitable dissimilarity function d – the dissimilarity between M and E can be expressed.

An example in which both requirements are fulfilled will follow soon. Of course, a universal procedure cannot be formulated.[7] First, the two problems of how to quantify novelty on the one hand and confirmation on the other hand must be treated. Between these two problems there is a structural analogy:

- In order to measure the novelty of a message M we must quantify the similarity between M and the recipient's prior knowledge.
- In order to measure the confirmation contained in a message M we must quantify the similarity between M and the receiver's information requirement and expectation.

---

[6] An example for trivial cognitive activities is given by tasks of memorizing unstructured material, such that only rote learning is possible. Changing the order within the material has no influence on the effort and efficiency of learning.

[7] In complex self-organizing systems (see Section 6.2), system components can act as emitters and receivers simultaneously, thus supplying pragmatic information to each other. But for the sake of simplicity, terms like "message" and "recipient" are maintained here. If necessary, they should be understood in a metaphorical sense.

The two problems have in common that the *similarity* between a message and a state of the receiving system must be measured. Problems of this type can be equivalently formulated in terms of *distance* or *dissimilarity*. If there is a high degree of similarity between two objects, then their distance (dissimilarity) is low and *vice versa*. As soon as either a similarity or a distance measure is known, the transformation between them is trivial.

For technical reasons it is often easier to work with *distance*. In the sequel, the procedure for the quantification of distance will be demonstrated for the example of *confirmation* (quantification of novelty is done correspondingly). Thus we have to quantify $d(M,E)$, the distance between a message M and the expectation E of a receiving system. As an example, take a receiver's expectation materialized in a questionnaire and the answers given. Then one can run a series of plausibility checks (completeness, correct data type, permitted numerical range, internal inconsistencies, etc.). The result supplies a measure for the deviation from a correctly completed questionnaire.

The distance function or dissimilarity function d has the usual properties of a metric, as specified by the well-known system of axioms:

$$\begin{aligned}
&\text{non} - \text{negativity}: \quad d(x,y) \geqslant 0\\
&d(x,y) = 0 \;\text{ if and only if }\; x = y\\
&\text{symmetry}: \quad d(x,y) = d(y,x)\\
&\text{triangle inequality}: \quad d(x,y) \leqslant d(x,z) + d(z,x)
\end{aligned}$$

It is possible that $d(M,E) = 0$: If somebody wants to know a particular telephone number and if the message exactly consists of that number, than there is a perfect matching between the message M and the expectation E, such that $d(M,E) = 0$. Generally there will be no such correspondence, and hence $d(M,E) > 0$. If a message has nothing in common with the expectation, then $d(M,E)$ will be large.

The distance function formulated here has nothing in common with the ordinary (geometric) distance between two points in space. Rather, we have to quantify the distance between two complex objects, both equipped with an internal structure. Although in the example of the questionnaire a distance function is easy to find, there is a great variety of styles for adapting such a function to the individual problem considered.

A special, but not exclusive technique which may be suited also for complex systems (particularly with a differentiated internal structure) represents each object by a finite connected graph, possibly with vertex labels. These vertex labels may have the form of graphs themselves. By recursion we arrive at the concept of "hierarchical graphs", with which objects with a hierarchical internal structure can be treated.

In the easiest case, a finite set $\mathbf{G}=\{G_1, G_2, \ldots, G_n\}$ of finite connected graphs is presupposed. The ideal tool for defining a function $d(G_i, G_k)$ with the required properties is supplied by *graph grammars.* A graph grammar is given by a start-graph and a finite number of production rules. Each production rule permits the generation of a new graph from one of the already existing graphs. A production rule is a triple $\{S_l, S_r, R\}$, where S_l is the subgraph to be replaced (left-hand side), S_r is the subgraph to be substituted for it (right-hand side), and R denotes the embedding rule which governs the way in which S_r has to be inserted.[8]

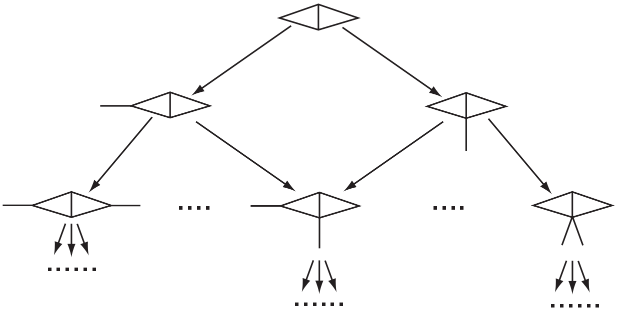

Figure 2: Example of a graph grammar, with some of the graphs generated in the first steps. See also Gernert (1997).

Figure 2 shows an example of a start-graph and a subset of those graphs which can be generated in at most two transformation steps. Here only one production rule exists: the addition of a new edge such that it has exactly one vertex in common with the previous graph. Generally, the "continuation" is not unique because, at a certain spot, one out of several admitted production rules must be chosen, or one production rule can operate with respect to different subgraphs. Hence there will be "branching", as indicated by the arrows in Figure 2.

For the present application a graph grammar Γ is required which generates *at least* all graphs in the given set $\mathbf{G}$. If Γ has been fixed, then the requested distance function can be defined by

$$d(G_i, G_k) = \min\ L(G_i, G_k),$$

where $L(G_i, G_k)$ is the length of a "path" that leads from G_i to G_k by applying production rules from Γ and the inverse transformations. Each such step contributes 1 to the length L (L corresponds to the number of steps "upward" and "downward" in the tree-like diagram representing Γ).

---

[8] For technical details of graph grammars see Gernert (1997), with diagrams and references.

For any given finite set **G** of graphs it is always possible to set up a graph grammar Γ such that at least all graphs in **G** are generated by Γ. But, apart from trivial cases, a graph grammar specified in this way cannot be unique. Rather, there is a multitude of graph grammars which are all suited to represent **G**. The reason is that "similarity" is a *perspective notion* – a fundamental concept to be discussed later in Section 5. For the moment, we can state that no definition of similarity or distance can be formulated without a reference to the purpose (or goal etc.) pursued with the individual measurement. Similarity between structures is never a property of the structures alone; rather, it arises due to an interpretation by an observer. Different positions taken by two observers will manifest themselves in two disparate graph grammars, but as soon as a graph grammar is specified, the individual understanding of "similarity" will be clear.

As mentioned before, the quantification of N can be done by the same method as explained for C. If a cardinal measurement for both N and C has been accomplished, then, with some reservations (Section 4.2), a quantification of P through the product formula Eq. (1) is possible.

4.4 Measurement in Non-Standard Cases

Even if a separate measurement of N and C is impossible, there may be a chance to "directly" measure P. In the evaluation of physical experiments and astrophysical observations (Atmanspacher and Scheingraber 1990, Kurths *et al.* 1992; see also Section 6.1) a procedure was applied which focuses on the relative increase of process efficiency through the acceptance of pragmatic information. If η_1 and η_2 (with $\eta_1 > 0$, $\eta_2 \geqslant \eta_1$) denote the efficiency before and after some distinguished change in behavior, e.g. a transition or instability, then one has[9]

$$P = (\eta_2 - \eta_1)/\eta_1 \ . \tag{2}$$

In a similar way, an increase of efficiency may be assessed by a comparison of prior and subsequent behavior with respect to a particular event, and P can take on a concrete dimension.

- A new engineering concept can help to improve a telecommunication system, and so P is expressed in GBytes or GBytes/sec.
- In a socio-economic system advice may lead to an increase of productivity (if it is adequately put into practice), and then P can be expressed, e.g., in "barrels per day".

Incoming messages can have the effect that the internal organization of an information-processing system is modified such that, due to this

---

[9] It must be remarked that this refers to a special case and must not be taken as a general equation for pragmatic information.

change, the future data storing or processing will run more efficiently. In this case, the reduction of the system complexity – defined just in this specific manner – can supply a measure of pragmatic information.

Also "negative information" is occasionally mentioned in literature. Apart from the case of counter-productive advice, the following situations can occur in social systems:

- A single human recipient or a team can no longer cope with abundant information.
- A message opens (or even stimulates) a way to circumvent an overdue organizational change.

Finally, a nominal or ordinal measurement (see Section 4.1) can be imagined, too. An individual reaction to a message can be described in qualitative terms (e.g., expected/unexpected, relaxed/excited), and the impressive strengths of words can be arranged ordinally (e.g., strange – weird – terrible – horrifying).

5. Pragmatic Information, Meaning, and Other Perspective Notions

Perspective notions are terms which – beyond the well-known context-dependence of word meanings in general – require an *explicit* statement of the context. A simple example is the term "classification": A lot of books can be arranged or classified according to their contents, size, color, weight, price, etc. Any task to classify a single object or to subdivide a given set can be accomplished only after the purpose of the classification or the relevant criteria have been exposed. Evidently, such indispensable contexts are often self-understood among the persons involved. On the other hand, just this fact can turn out to be a source of misunderstanding (to be discussed below).

Four eminent perspective notions (or pairs of such notions) are pragmatic information, meaning and interpretation, similarity and dissimilarity, and complexity. They can be represented by the four vertices of a tetrahedron in three-dimensional space or by a plane drawing as in Figure 3. The six connecting lines illustrate that each entry is interrelated with each other.

Clearly, *pragmatic information* is such a perspective notion. The *meaning* of a word, phrase, sentence, text, symbol, diagram, etc. as well as the process of *interpretation* and its outcome, depend on the involved persons, the situation, the historical and social context, and the purpose of the analysis. *Similarity* and *dissimilarity* were discussed in Section 4.3: Within a given set, two objects can be more or less similar depending on their size, shape, appearance, structure, function, etc. Finally, the

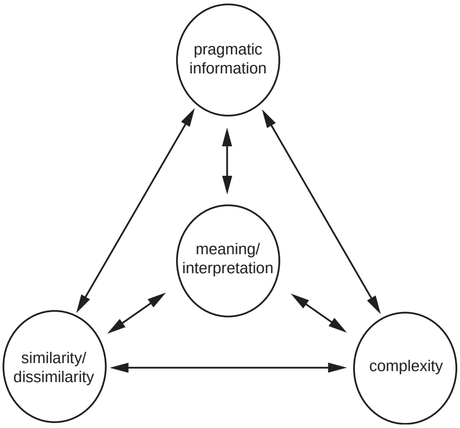

Figure 3: Four eminent perspective notions and their mutual relatedness.

complexity of a system depends (among other things) on the individual education and training. What seems to be very complex to a newcomer may be evidently simple for an expert. The specialist literature offers a multitude of definitions for complexity, all tailored to specific purposes. Even within the same branch of research we occasionally find more than one definition.[10]

Fortunately, perspective notions do not necessarily exclude a formal mathematical treatment. But such a treatment needs to disclose the rationale underlying the procedure demonstrated above (Section 4.3) by the example of similarity measurements.

Any observation – and, consequently, any measurement – is indeed a *two-stage process*. What is really happening can be briefly characterized by the following two phases due to Roederer (2005, p. 166):

1. The initiator decides *what* is to be observed and *how* this shall be done, also defining the context, e.g., the purpose of the observation, the relevant features, etc. (definition phase).
2. If necessary, technical preparations are cared for, and then the proper observation is performed.

It must be accentuated here that the role of the decision phase (1) is frequently ignored or neglected. The situation in which that context is

---

[10] For an overview of definitions of complexity see Grassberger (1991), Atmanspacher *et al.* (1992), Wackerbauer *et al.* (1994), Kurths *et al.* (1996), Atmanspacher and Wiedenmann (1999).

(tacitly) self-evident can be illustrated as follows. If somebody in a laboratory is told to "measure the temperature", then the location and the time are clear, as well as other circumstances like the required precision, the instruments at hand, their error rates, and the customs of documentation. The frequency of situations like this can easily foster the illusion of a straightforward procedure, assuming that only a variable to be measured must be mentioned, and then the measurement can start. It seems plausible that this tacit assumption is one of the reasons why the concept of pragmatic information is only being accepted with reluctance.

Every measurement requires preparatory steps according to phase (1), which must be formulated explicitly when dealing with perspective notions. In the example of similarity measurements the preceding step can be implemented by exposing a graph grammar with the requested properties. With respect to other perspective notions, specific mathematical tools must be designed.[11]

6. Applications

6.1 Pragmatic Information in Physics

The concept of pragmatic information can provide insight even in contexts where one would not expect this in the beginning. It turned out that the application of pragmatic information in down-to-earth experimental physics and in observational astrophysics supplies an interesting key to an understanding and even to a quantitative analysis of empirical observations.

As Atmanspacher and Scheingraber (1990) demonstrated, the behavior of a specific class of lasers (multimode continuous-wave dye lasers) around instabilities can be described and evaluated on the basis of pragmatic information. The change in efficiency at a critical value of a control parameter, where an instability occurs, can be directly used to measure the pragmatic information gained during the instability. For an initial efficiency η_1 and a final efficiency η_2 one can calculate P by Eq. (2). Quite a similar analysis was performed by Kurths *et al.* (1992) for time series representing solar activity.

A first hint suggesting that the presuppositions underlying the conventional Shannon-Weaver theory are no more valid in the quantum case was indicated by Connes *et al.* (1987). In a different context, Brukner and Zeilinger (2001) argued more explicitly that the Shannon theory is not appropriate in defining the information contents of quantum systems. The reason is that quantum measurement is completely different from classical

---

[11] A tool for the mathematical handling of stage (1) is the concept of operators called "preselectors" (Gernert 2000) which assume the form of preselector matrices in simple cases.

measurement, particularly as far as the complementarity of observations is concerned.

Another proposal, which differs explicitly from Shannon-Weaver theory, has been termed *active information* (Hiley 2002, Hiley and Pylkkänen 2005, and references therein). Hiley explains this by the example of a quantum system, where information is provided by the experimental conditions, e.g. the environment of a particle. Information is not "information for us" (for the experimenter or any observer), but "objective information for the particle"; the activity takes place in the system itself.[12] The idea that a particle accepts "active information" from its environment, and acts correspondingly, is only terminologically different from the position of Atmanspacher and Scheingraber (1990), according to which a component of a physical system, or process, is influenced by received pragmatic information.

All these contributions, formulated under different aspects, illustrate a relationship between pragmatic information and *meaning.* The fact that a physical system can perform a transition from its present state to a specified subsequent state, as well as the details of such a transition, demonstrate that the experimental conditions are meaningful for the system. This meaning can be measured on the basis of pragmatic information. In other words, meaning or, at least, terms like "having meaning" can be operationalized. If a message has a meaning for a receiver, then this must be observable by interpreting its behavior after receiving the message (Kornwachs 1998, p. 184).

The same applies to the notion of *interpretation.* A particle *interprets* the surrounding field by following a specific path (cf. Roederer 2005, pp. 161f). It is not important that a description of such processes is also possible without the use of perspective notions (e.g., on the basis of field equations or wave functions only). The crucial fact is that a description and analysis based upon perspective notions *is possible*, and that this is likely to open a pathway to further understanding.[13]

6.2 Pragmatic Information and General Systems Theory

Rather early in its history, the concept of pragmatic information was connected with general systems theory. One of the first questions was how a modified concept of information could be expressed on this basis.

---

[12] According to Hiley, this also holds for biological systems, where the information provided by the environment is defined by soil conditions, humidity, etc. Processes in classical physics can be described by active information as well (Bohm and Hiley 1993, p. 62). Hence this concept does not depend on a specific interpretation of quantum theory.

[13] Evidently, the notions of meaning and interpretation sound somewhat anthropomorphic when they are applied to physical systems. For further remarks on the admissibility of such terms in science see Section 7.1.

Pragmatic information can be generated by extremely simple devices. Consider a cash register which stores the data of all sales over the day. After closing time the shopkeeper is not interested in the list of numbers, but in the total. By performing the addition the cash register produces pragmatic information, since the total sum comes closer to the user requirements than the list of raw data would do.

By way of contrast, systems which *react* on pragmatic information cannot be adequately described by deterministic automata.[14] If pragmatic information is received and the effect of this reception can be observed then we speak of a "non-classical system". In classical systems, all essential properties remain unchanged during observation. Their behavior is deterministically predictable, and it is always possible to decide whether or not an element or a subsystem has a certain property (in quantum systems this is in general not possible). If at least one of these features is missing then we have a non-classical system. Further characteristics of non-classical systems in this sense were outlined in Kornwachs and Lucadou (1982) and Kornwachs (1998).

A particularly important kind of non-classical systems is represented by evolutionary biological systems, which are typically self-organizing. Signals from the environment and the conditions of this environment are pragmatic information if they have a long-term effect in the development of the species (Kornwachs and Lucadou 1982).[15] As Atmanspacher and Scheingraber (1990, p. 731) stated:

> In self-organizing systems each constituent acts as a transmitter and receiver of information simultaneously. Since pragmatic information changes the state (of knowledge) of any receiver, it does equivalently change its state as a potential transmitter. The corresponding self-reference or circular causality is a key feature of self-organization.

6.3 Connections with Cognitive Processes

An immortal and recurring argument concerning mathematical (and logical) insight is that mathematics can never supply something new, but its results are "tautological". This is intended to mean that any result has already been existing beforehand in the form of the original data.[16] But a second thought reveals that this argument is questionable, as can be demonstrated for the case of the cash register mentioned above. If it

---

[14] Here the special case must be excluded that a computer is equipped with new software and thus gets new capabilities.

[15] For further details on information and self-organization see Eigen (1971, 1976) and, with explicit reference to pragmatic information, Weinberger (2002).

[16] Arguments against this improper use of the term "tautological" are given by Morgenstern (1965).

is done well, the result of a mathematical derivation is closer to the user requirements, and hence pragmatic information has been generated.

Every cognitive process can be described by a combination of operators which belong to specific classes of "elementary operators". Each operator characterizes an elementary cognitive act, like learning, drawing conclusions, valuating, etc.[17] There exist special situations in which the combination (consecutive application) of two operators is commutative. For instance, in a mathematical proof it can be possible that two acts of deriving a conclusion can be done in arbitrary order. But in the general case the next step utilizes the state created by the former one, and the combination of two operators is *non-commutative.*

This non-commutativity of operators is related to the very nature of information. The reception and spreading of information are, to a very large extent, irreversible processes (leaving aside the risk of later forgetting). The spoken word cannot be made unspoken. Not only cognitive processes, but all nontrivial acts by which information is received, transformed or transmitted are necessarily non-commutative and irreversible.

7. Concluding Remarks and Outlook

7.1 Anthropomorphic Terms and Parallel Theories

As expounded before, in specific situations a state transition within a physical system can be described as a utilization of *meaning* or, equivalently, as an act of *interpretation.* The use of these two notions has, with some regularity, provoked contradiction and even bewildering debates.

The propensity to exorcize terms with an "anthropomorphic" flavor in natural science can be understood from the history and sociology of science. In the early era of modern science the rejection of too speculative terms was clearly justified. Later on, this led to a more and more clear-cut split between "Naturwissenschaften" (natural sciences) and "Geisteswissenschaften" (letters, arts, humanities, cultural sciences).[18] This tradition is still present in our days. Terms like "interpretation" are regarded as vague, and as a monopoly of cultural sciences, hence foreign to natural sciences. Another argument claims that the range of meaning of such terms would be overstretched in the natural sciences, implying all kinds of troubles concerning precise communication.

However, there is no more reason today to maintain that dichotomy of scientific disciplines (any such subdivision of disciplines mainly serves administrative and related purposes outside scientific logic). Interpretation

---

[17] A complete list of elementary operations, as well as stochastic elements, concurrency of operations, and the algebra formed by the operators are beyond the scope of this article. See Gernert (2000) and Atmanspacher and Filk (2006) for more details.

[18] The motives and influences, clearly outside scientific logic, which fostered the overaccentuation of this split, are characterized by Topitsch (1965).

does not imply arbitrariness. In the case of classical mechanics, where identical settings lead to identical outcomes, interpretations for identical settings will simply be the same.

From time to time scientific progress requires the insight that certain apparently disparate phenomena have essential features in common and, therefore, should be expressed in one common terminology. It was an achievement to recognize that the morning star and the evening star are identical, or that the spectrum of a glass prism and a rainbow have a common origin. Of course, new terms alone can never be an answer to a scientific problem, but they can make future work easier.

In the theory of self-organizing systems terms like "emergence of meaning" and "self-creation of meaning" are often used (e.g., Haken 1988, pp. 23–29). Evidently, the notion of meaning is vital to this field. Wheeler (1989) even speaks about "meaning physics" and distinguishes it from branches of physics not requiring the notion of meaning.

The argument that we could easily do without new concepts of this kind, that the empirical findings can be handled in the traditional style, requires a comment. We propose the notion of a *parallel theory* to denote a new scientific framework that describes and predicts just the same empirically verified facts as already existing theories. The notion of a "parallel theory" (rather than "alternative theory" or something like that) is deliberately chosen in order to avoid biases in favor of one among several parallel theories. One might suppose that a new approach is unnecessary, but nevertheless it may be suited to incorporate novel phenomena, which will be detected in the future and which are outside the scope of older theories.

A candidate for a new parallel theory should satisfy the following qualifying features:

- *General requirement:* The usual properties of a scientific theory must be guaranteed, e.g. internal consistency.
- *Continuity of terms:* As far as possible, established terms should be used in their traditional meaning.
- *Empirical correctness:* The theory must be compatible with empirical observations.
- *Explanation:* There must be at least one class of empirical facts which can be explained only through the new theory, or which finds a better explanation yielding improved insight.

The last feature serves to prevent an abundance of useless theories according to Mach's principle of the economy of thinking.

An illustration is the co-existence of the Schrödinger representation and the Heisenberg representation in the early years of quantum theory. Both representations were created nearly simultaneously, such that none of them could be blown up to a dominating theory. Today it is known that they are unitarily equivalent, i.e. there is a one-to-one translation from one to the other. Nevertheless each of them has a specific domain where it is pragmatically superior.

If the conditions formulated above are fulfilled, then it is wise to admit a parallel theory. The material compiled in Section 6.1 shows that the concept of pragmatic information is qualified in this sense, hence admonitions like "this could also be done by a wave function" miss the point. This position is furthermore accentuated by possible applications to the mind-matter problem, e.g. along the lines suggested in a recent paper by Hiley and Pylkkänen (2005).

7.2 Conclusions

The concept of pragmatic information provides a unifying, comprehensive framework which includes earlier proposals – the traditional Shannon-Weaver theory with its extensions and ramifications – as special cases. Such special cases refer to recipients as telecommunication engineers who assess an incoming message by the criteria of just that discipline (speed, reliability of the transmission, etc.).

Due to an unfortunate historical development, the general acceptance of pragmatic information is still marred by remnants of obsolete positions, by a paradoxical terminology dubbing the special case as "information theory", whereas the comprehensive concept needs the diacritical epitheton "pragmatic" (or "active"). In view of this history, we should beware of rash exaggerations and misinterpretations. Omnès (1990, p. 509) warns us in this sense when he objects to "the view according to which quantum mechanics is a part of information theory". Any "information theory" – whatever this term will mean in the future – must necessarily be a *formal structural science* and, hence, belong to the same category of scientific disciplines as mathematics, logic, general systems theory, etc.[19]

As soon as some branch of an *empirical science* is addressed, a suitable substructure must be selected, and the objects of the empirical domain must be assigned to the terms of the formal system. The fundamental difference between both types of scientific procedures must never be ignored (nor can it vanish through future developments).

Pragmatic information is an example of a *perspective notion.* Therefore, concrete models for pragmatic information may teach us something about working with perspective notions. In particular, perspective notions do not necessarily exclude a mathematical treatment, although this may require novel and specific procedures.

---

[19] For further remarks on the notion of structural sciences see Weizsäcker (1995, pp. 22f).

Two other, closely linked perspective notions are *meaning* and *interpretation.* Arguments were given for a broader use of these notions, even in situations where this is not uncontroversial. The fact that some physical and some cognitive processes can be described in a similar manner may lead to futile speculations, but it also may turn out as a heuristic cue.

In the context of “a bridge across the Cartesian cut”, Atmanspacher (1994) asked for conditions under which the concept of pragmatic information will provide a surplus in explanatory power beyond purely physical models. He concluded that such an explanative surplus can be expected as soon as “an explicit account of *res cogitans* becomes unavoidable”. For bridging the Cartesian cut, as well as for the mind-matter problem in general, the concept of pragmatic information will not supply an immediate answer. But it can scarcely be imagined how such an answer could be obtained without an advanced concept of information as an essential tool.

Acknowledgments

The author would like to thank Peter beim Graben and the referees for their valuable comments and suggestions.

References

Atmanspacher H. (1989): The aspect of information production in the process of observation. *Foundations of Physics* **19**, 553–557.

Atmanspacher H. (1994): Complexity and meaning as a bridge across the Cartesian cut. *Journal of Consciousness Studies* **1**(2), pp. 168–181.

Atmanspacher H. and Filk T. (2006): Complexity and non-commutativity of learning operations on graphs. *BioSystems* **85**, 84–93.

Atmanspacher H., Kurths J., Scheingraber H., Wackerbauer R., and Witt A. (1992): Complexity and meaning in nonlinear dynamical systems. *Open Systems & Information Dynamics* **1**, 269–289.

Atmanspacher H. and Scheingraber H. (1990): Pragmatic information and dynamical instabilities in a multimode continuous-wave dye laser. *Canadian Journal of Physics* **68**, 728–737.

Atmanspacher H. and Wiedenmann G. (1999): Some basic problems with complex systems. In *Large Scale Systems: Theory and Applications*, ed. by N.T. Koussoulas and P. Groumpos, Elsevier, Amsterdam, pp. 1059–1066.

Bar-Hillel Y. (1969): Wesen und Bedeutung der Informationstheorie. In *Informationen über Information*, ed. by H. von Ditfurth, Hoffmann und Campe, Hamburg, pp. 13–42.

Bohm D. and Hiley B.J. (1993): *The Undivided Universe. An Ontological Interpretation of Quantum Theory*, Routledge, London.

Brukner C. and Zeilinger A. (2001): Conceptual inadequacy of Shannon information in quantum measurements. *Physical Review A* **63**, 022113.

Connes A., Narnhofer H., and Thirring W. (1987): Dynamical entropy of C*-algebras and von Neumann algebras. *Communications in Mathematical Physics* **112**, 691–719.

Copper M.C. (1992): The utility of information. *Technique et Science Informatiques* **11**(2), 85–101.

Ditfurth H. von, ed. (1969): *Informationen über Information*, Hoffmann und Campe, Hamburg.

Ebeling W., Freund J., and Schweitzer F. (1998): *Komplexe Strukturen: Entropie und Strukturen*, Teubner, Stuttgart.

Eigen M. (1971): Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften* **18**, 465–523.

Eigen M. (1976): Wie entsteht Information? Prinzipien der Selbstorganisation in der Biologie. *Berichte der Bunsen-Gesellschaft* **80**, 1059–1081.

Gäng P. (1967): Pragmatische Informationen. *Grundlagenstudien aus Kybernetik und Geisteswissenschaft* **8**, 77–90.

Gernert D. (1997): Graph grammars as an analytical tool in physics and biology. BioSystems **43**, 179–187.

Gernert D. (2000): Towards a closed description of observation processes. *BioSystems* **54**, 165–180.

Grassberger P. (1991): Information and complexity measures in dynamical systems. In *Information Dynamics*, ed. by H. Atmanspacher and H. Scheingraber, Plenum Press, New York, pp. 15–33.

Haken H. (1988): *Information and Self-Organization*, Springer, Berlin.

Hiley B.J. (2002): From the Heisenberg picture to Bohm: A new perspective on active information and its relation to Shannon information. In *Quantum Theory: Reconsideration of Foundations*, ed. by A. Khrennikov, Växjö University Press, Växjö, pp. 141–162.

Hiley B.J. and Pylkkänen P. (2005): Can mind affect matter via active information? *Mind and Matter* **3**(2), 7–27.

Jeffreys H. (1946): An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society London (Series A)* **186**, 453–461.

Jumarie G. (1990): *Relative Information*, Springer, Berlin.

Kampé de Fériet J. (1974): La théorie généralisée de l'information et la mesure subjective de l'information. *Lecture Notes in Mathematics* **398**, 1–35.

Klir G.J. and Folger T.A. (1988): *Fuzzy Sets, Uncertainty and Information,* Prentice-Hall, Englewood Cliffs NJ.

Klir G.J. (1991): Measures and principles of uncertainty and information: recent developments. In *Information Dynamics*, ed. by H. Atmanspacher and H. Scheingraber, Plenum Press, New York, pp. 1–14.

Kornwachs K. and Lucadou W. von (1982): Pragmatic information and nonclassical systems. In *Cybernetics and Systems Research*, ed. by R. Trappl, North-Holland, Amsterdam, pp. 191–197.

Kornwachs K. (1992): Information und der Begriff der Wirkung. *Informatik-Fachberichte* **306**, 46–56.

Kornwachs K. (1998): Pragmatic information and the emergence of meaning. In *Evolutionary Systems. Biological and Epistemological Perspectives on Selection and Self-Organization,* ed. by G. van de Vijver, S.N. Salthe and M. Delpos, Kluwer, Dordrecht, pp. 181–196.

Kullback S. and Leibler R. (1951), On information and sufficiency. *Annals of Mathematical Statistics* **22**, 79–86.

Kurths J., Feudel U., and Jansen W. (1992): Pragmatic information in nonlinear dynamo theory for solar activity. In *Evolution of Information Processing Systems*, ed. by K. Haefner, Springer, Berlin, pp. 185–199.

Kurths J., Voss A., Saparin P., and Witt A. (1996): Complexity measures and their applications. In *Information: New Questions to a Multidisciplinary Concept,* ed. by K. Kornwachs and K. Jacoby, Akademie, Berlin, pp. 129–136.

Marko H. (1966): Die Theorie der bidirektionalen Kommunikation und ihre Anwendung auf die Nachrichtenübermittlung zwischen Menschen. *Kybernetik* **3**, 128–136.

Morgenstern O. (1965): Logistik und Sozialwissenschaften. In *Logik der Sozialwissenschaften*, ed. by E. Topitsch, Kiepenheuer & Witsch, Köln, pp. 315–336.

Morris C.W. (1938): Foundations of the theory of signs. In *International Encyclopedia of Unified Science I/2*, ed. by O. Neurath *et al.*, University of Chicago Press, Chicago, pp. 77–137.

Omnès R. (1990): Some progress in measurement theory: the logical interpretation of quantum mechanics. In *Complexity, Entropy, and the Physics of Information*, ed. by W.H. Zurek, Addison-Wesley, Redwood City CA, pp. 495–512.

Roederer J.G. (2005): *Information and its Role in Nature*, Springer, Berlin.

Rooij R. van (2006): Pragmatic value and copmplex sentences. *Mind and Matter* **4**, 195–218.

Shannon C.E. and Weaver W. (1949): *The Mathematical Theory of Communication*, University of Illinois Press, Urbana IL.

Shannon C.E. (1956): The bandwagon. *IRE Transactions on Information Theory* **IT-2**(1), March, p. 3.

Sharma B.D., Mitter J., and Mohan M. (1978): On measures of useful information. *Information and Control* **39**, 323–336.

Straubel R. (1980): Bewertungsmaß pragmatischer Information. *Messen-Steuern-Regeln* **23**, 583–584.

Topitsch E., ed. (1965): *Logik der Sozialwissenschaften*, Kiepenheuer & Witsch, Köln.

Topitsch E. (1965): Das Verhältnis zwischen Sozial- und Naturwissenschaften. In *Logik der Sozialwissenschaften*, ed. by E. Topitsch, Kiepenheuer & Witsch, Köln, pp. 56–71.

Wackerbauer R., Witt A., Atmanspacher H., Kurths J., and Scheingraber H. (1994): A comparative classification of complexity measures, *Chaos, Solitons & Fractals* **4**, 133–173.

Weaver W. (1949): The mathematics of communication. *Scientific American* **181**(1), July, pp. 11–15.

Weinberger E.D. (2002): A theory of pragmatic information and its application to the quasispecies model of biological evolution. *BioSystems* **66**, 105–119.

Weizsäcker C.F. von (1985): *Aufbau der Physik*, Deutscher Taschenbuch Verlag, München.

Weizsäcker C.F. von (1995): *Die Einheit der Natur*, Deutscher Taschenbuch Verlag, München.

Weizsäcker E. von and Weizsäcker C. von (1972): Wiederaufnahme der begrifflichen Frage: Was ist Information? In *Nova Acta Leopoldina* **206**, Band 37/1, J.A. Barth, Leipzig, pp. 535–555.

Weizsäcker E.U. von (1974): Erstmaligkeit und Bestätigung als Komponenten der pragmatischen Information. In *Offene Systeme I*, ed. by E.U. von Weizsäcker, Klett, Stuttgart, pp. 82–113.

Weizsäcker E.U. von and Weizsäcker C. von (1984): Fehlerfreundlichkeit. In *Offenheit – Zeitlichkeit – Komplexität. Zur Theorie der offenen Systeme,* ed. by K. Kornwachs, Campus, Frankfurt, pp. 167–201.

Wheeler J.A. (1989): Bits, quanta, meaning. In *Festschrift in Honour of Eduardo R. Caianiello*, ed. by A. Giovannini *et al.*, World Scientific, Singapore, pp. 133–154.

Zoglauer T. (1996): Can information be naturalized? In *Information: New Questions to a Multidisciplinary Concept*, ed. by K. Kornwachs and K. Jacoby, Akademie, Berlin, pp. 187–207.

Received: 25 January 2006
Revised: 10 July 2006
Accepted: 12 July 2006

Reviewed by Klaus Kornwachs and another, anonymous, referee.