Information Studies Without Information

Jonathan Furner

Abstract

In philosophy of language, the phenomena fundamental to human communication are routinely modeled in ways that do not require commitment to a concept of "information" separate from those of "data," "meaning," "communication," "knowledge," and "relevance" (inter alia). A taxonomy of conceptions of information may be developed that relies on commonly drawn philosophical distinctions (between linguistic, mental, and physical entities, between objects and events, and between particulars and universals); in such a taxonomy, no category requires the label "information" in order to be differentiated from others. It is suggested that a conception of information-as-relevance is currently the most productive of advances in theoretical information studies.

Unsurprisingly, the nature of information has long been a topic of central concern for scholars of information studies (IS).[1] The body of literature in which authors have attempted to provide answers to the question "What is information?" may be viewed in any (or any combination) of the following ways: (i) as contributing to *science*—if information is cast as a naturally occurring phenomenon; (ii) as contributing to *social science*—if information is considered to be a product of human artifice; or (iii) as contributing to *philosophy*—if "information" is treated primarily as a fundamental concept existing at the same level as, for example, meaning, knowledge, and truth.

Although this body of literature is sizeable when taken as a whole, the quantity of work that may be classed under the third heading is small. One approach that is often taken in studies representative of the third class is to compare theories of information with theories of knowledge. Commonly,

Jonathan Furner, Graduate School of Education and Information Studies, University of California, Los Angeles, 300 Young Drive North, Box 951520, Los Angeles, CA 90095-1520
LIBRARY TRENDS, Vol. 52, No. 3, Winter 2004, pp. 427–446

the everyday conception of knowledge (as the content of mental states) is contrasted with a philosophical conception of knowledge (typically, as justified true belief); and information is identified as knowledge (that is, knowledge in the first, everyday sense) that has been recorded or that is in some sense objective, external, or public. Somewhat oddly, given the nature of the conception of knowledge that is typically adopted in such accounts, a tendency has been for authors to go on to use the results of this kind of analysis to locate IS with respect to *epistemology*. A more productive approach might instead be to relate work in IS to that in *philosophy of language*, since it is the latter branch of philosophy that is concerned more exclusively with the content of mental states (i.e., thoughts); with the ways in which such content may be expressed, represented, or recorded; and with the ways in which such expressions may be interpreted or their meaning understood.

My suggestion is that if this alternative direction is taken, we shall find that philosophers of language have modeled the phenomena fundamental to human communication in ways that do not require us to commit to a separate concept of "information." Indeed, we can conclude that such a concept is unnecessary for IS. Once the concepts of interest have been labeled with conventional names such as "data," "meaning," "communication," "relevance," etc., nothing is left (so it may be argued) to which to apply the term "information." One corollary of such a conclusion is the equally negative judgment that the field of IS is itself misnamed, and that its subject matter should more appropriately be treated as a branch of communication studies, semiotics, or library studies.

In this paper, I will present a simple taxonomy of common conceptions of information—a taxonomy in which no single category seems to unequivocally require the label "information" to differentiate it from others. Before reaching that point, however, I wish to review some terminological distinctions that are commonplace in the literature of philosophy of language and that may correspond to certain distinctions that lie at the core of philosophy of information. And I would like to begin by attending to two preliminary questions that immediately present themselves when embarking on any attempt to develop a philosophical theory of information or to distinguish between competing theories.

The Desirable Properties of a Philosophical Theory of Information

First: What is the distinctive nature and scope of a philosophical theory of information (as distinguished from a theory of information of any other kind)?

A philosophical theory of information—or, more precisely, a metaphysical or ontological theory of information—is assumed here to be a specification of the necessary and sufficient conditions under which a phenomenon may be identified as "information." Arriving at such a specification

involves conceptual analysis—that is, analysis of the meaning of the concept(s) referred to by the word "information." Conceptual analysis emerged as the primary method used by philosophers in the particular paradigm ("analytic" philosophy) that rose to dominance in anglophone countries in the twentieth century. In our present case, the analysis also will involve conceptual classification, since it happens that several different conceptions of information have risen to different levels of prominence, and it is often helpful to construct a taxonomy highlighting the differences perceived to be most important.

Our second preliminary question is, On what criteria may a philosophical theory of information be evaluated?

Given the parallel existence of multiple (and perhaps mutually exclusive) conceptions of information, it would be helpful to choose from among them on the basis of some agreed-upon criterion (or set of criteria). The possibilities include the following:

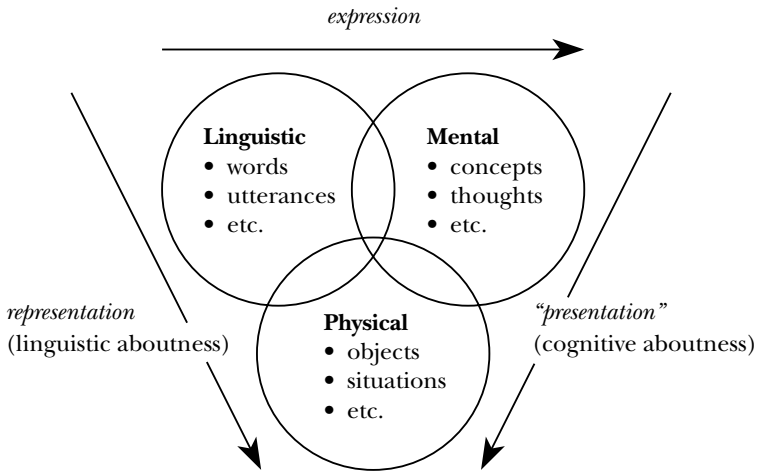
1. *Coherence.* Our understanding of information should plausibly cohere with our understanding of other related concepts such as knowledge, meaning, truth, cognition, relevance, etc.
2. *Parsimony.* According to the principle of Occam's razor, a theory should not posit the existence of unfamiliar kinds of entities, unless it proves impossible to account for certain phenomena in terms of primitive or familiar concepts.
3. *Utility.* The primary purpose of a theory, one might argue, is to enhance our understanding of the object of study. Understanding may be improved by many means; one of the most productive is the process by which we come to recognize the simultaneous validity of multiple perspectives on a single issue. If a theory is suggestive of an agenda for future work, either through further theorizing or through empirical research, then it is doing its job.

The Physical, the Mental, and the Linguistic

A simple model that one may use to show how some of the concerns of modern philosophical inquiry relate to one another is depicted in Figure 1. In this model, entities of interest are divided among three categories: (i) physical entities, such as objects and situations; (ii) mental entities, such as concepts and thoughts; and (iii) linguistic entities,[2] such as words and utterances.[3] The nature of entities in the physical world is the concern of *metaphysics;*[4] the nature of entities in the mental world is the concern of *philosophy of mind;* and the nature of entities in the linguistic world is the concern of *philosophy of language.*

Where these separate areas of inquiry overlap is in their shared interest in the nature of the relationships between entities in different categories. We might wish to say, for instance, that linguistic entities (e.g., utter-

Figure 1. Entities of interest to modern philosophical inquiry.



ances) both "express" mental entities (e.g., thoughts) and are "about" physical entities (e.g., situations). We might also wish to say that mental entities are, in some similar sense, "about" physical entities. This simple rendering of the triangular structure immediately raises questions that seem to require more complex answers—answers upon which, ideally, there would be consensus in multiple branches of philosophy.

For example, we might ask, What, more precisely, is the nature of the linguistic (or semantic) aboutness that relates words and objects? Is this linguistic aboutness different from the cognitive aboutness that relates concepts and objects? If one of these kinds of relationship is to be characterized as "meaning," how is the other to be differentiated? Each of these kinds of relationship has been the object of extended analysis; indeed, the central project of the philosophy of language in the twentieth century may be viewed as an extended exploration of the meaning of "meaning." What is it to say (i) that something has meaning (i.e., is meaningful) and (ii) that something has the particular meaning *p*? A comprehensive review of the contributions to the literature on these topics is well beyond the scope of the present paper.[5] Instead, what is provided here is a brief introduction to some of the more basic issues, with the aim simply of reaching the point where we may comprehensibly suggest definitions of the categories that, it is argued, are most relevant to a philosophy of information.

Situations and Utterances

One way of beginning a review of some of the very early steps to be made in any pursuit of a theory of meaning is to distinguish between physical situations and linguistic utterances.

A *situation* is a particular state of affairs in the physical world. At first sight, the phrase "state of affairs" may seem vague and obscure. The essence of the idea is that a situation is composed of certain things, that it has a structure of some kind. There are a number of different ways in which we might wish to describe the composition of a situation. Our choice will depend on our basic ontological views on the existence of certain fundamental categories of physical entities—categories such as objects, events, classes, properties, and relations. On one account, we might say that a situation is made up of a set of objects arranged in a certain way; which (some would argue) is tantamount to saying that any situation may be considered (in mathematical terms) as a graph made up of (i) a set of objects, together with (ii) a set of object-pairs denoting the relationships among objects. On this account, it is also common (but not necessary) to view each individual object as a set of properties or attribute/value pairs.[6] For example, we might imagine a situation in which I am standing in front of my car. In this particular situation, two of the important objects are me and my car; these two objects are arranged in a particular way (i.e., with me in front of my car rather than behind it, in it, or jumping over it, etc.); and, to take just one of the two objects in question, my car may be described as being small in size, red in color, Japanese in origin, and so on.

An *utterance* (or expression) is a particular vehicle of meaning. "Meaning," of course, is a complex concept of central interest to us and requires further analysis. In the meantime, we can be somewhat more specific and suggest that an utterance is the product of a human decision to act by using words in a way that has meaning. An example of an utterance is the particular string of words that I spoke at 6 p.m. yesterday: "I am standing in front of my car." At 6 p.m. yesterday, I decided to act in a way that would have a particular meaning, and I acted on that decision by voicing a particular sequence of sounds. (Of course, not all utterances are spoken; many are written. And although utterances may be said to occur at the moment they are first spoken or written, they can be—and often are—recorded for future consideration.)

As well as being something that has meaning, an utterance is something that happens, that takes place. In other words, an utterance is an event—something that occurs on a particular occasion (or, we might say, over a particular period). As an event, an utterance is *datable*, in the sense that we may, in principle, determine its temporal properties; we may specify, for instance, its time of commencement and/or of completion. Indeed, the category of events is itself a plausible candidate for inclusion in any list of fundamental categories of entities, and thus our categorization of utter-

ances as events—in virtue of their physicality rather than their meaning or semantic content—would commit us to identifying utterances themselves as, potentially, components of situations.

We also need to bear in mind that an utterance is the product of a human decision to act. Utterances occur because people make them. As a human artifact, an utterance is *authenticatable,* in the sense that we may, in principle, determine the identity of its author or speaker. We also may speculate as to the intentions of the speaker of an utterance when making the decision to utter. We may ask: What was the goal of the speaker? What did the speaker hope to achieve through uttering the set of words that were uttered? Is it sensible to talk about differences between the meaning that was intended and the meaning that was actually expressed?

Finally, we must be aware that, typically, at least one component of a speaker's set of intentions in making an utterance is to affect a hearer in some specific way—to bring about a change of some kind in the hearer's mental state. With respect to any given utterance, we may thus additionally ask, Is it sensible to talk about differences between the meaning that was expressed by the speaker, and the meaning that was understood by the hearer?

We now can return to what might be conceived as the basic issue: to specify ways in which utterances have (or are perceived to have) meaning. We might proceed by considering the suggestion that, in general, we talk about situations; that human discourse is about the world; that the things we say (i.e., our utterances) are *about* things (i.e., situations) that exist in the physical world. A suggestion of this kind seems to be an argument for recognition of a particular kind of relationship: one that is exemplified by an utterance and the situation that that utterance is about. We could call this relationship "aboutness." We could even call it "meaning" and propose that my utterance (at 6 p.m. yesterday) of the words "I am standing in front of my car" stands in a relationship of meaning with the situation in which (at 6 p.m. yesterday) I was standing in front of my car. Just as we could say that the utterance *u* is about the situation *s,* so we could say that *u* means *s* or that *s* is the meaning of *u.* If we were to use the latter formulation, we would need to be careful to note the distinction between the two different senses of "meaning" we would already be using: (i) our name for a kind of relationship between *u* and *s* and (ii) our name for a category in which, with respect to *u, s* falls. Whenever we say that an utterance "has" meaning (or semantic content, or indeed aboutness), we seem to be saying both that it stands in a relationship of a certain kind with another entity of a certain kind and that that other entity "is" its meaning (or content, or aboutness).

In the account just given, utterances mean situations. We might identify such an account as a *referential* theory of meaning, since there is a sense in which, by this account, utterances are deemed to be meaningful in virtue of their reference to situations. Moreover, it might seem possible to

enrich such an account by suggesting both that the components of utterances (e.g., subject terms and predicates) refer to the components of situations (e.g., objects and properties) and that it is in virtue of this reference that the components of utterances are meaningful. Referential or compositional theories of meaning are commonly associated with correspondence theories of truth, whereby an utterance is evaluated as true if it corresponds with (i.e., refers to) "the facts"—in other words, if it corresponds with a situation that "obtains," that is "actual," that is "the case."

Sentences

Linguists, lexicographers, information retrieval system designers, and others interested in the use of words often need to be careful to distinguish between word-tokens and word-types. In the previous sentence, for example, the single word-type "to" is instantiated by two word-tokens or occurrences, as is the word-type "and." Similarly, we should be careful to distinguish, in the current context, between utterance-tokens and utterance-types. If I utter the words "I am standing in front of my car" at 6 p.m. on Monday, and then again at 6 p.m. on Tuesday, a single utterance-type has been instantiated by two utterance-tokens. To avoid confusion, we might reserve the term "utterance" to stand for "utterancetoken," and use another term to stand for "utterance-type." From now on, in this paper, I shall use "sentence" instead of "utterance-type," while fully recognizing that such usage conflicts with the ordinary, everyday usage of "sentence."[7] Some examples of sentences are listed in Table 1.

Two utterances of the same sentence can have different truth-values: for example, my utterance on Monday might correspond with the facts, in which case I would be telling the truth, whereas on Tuesday I might be lying. Two utterances of the same sentence also can have different references. Suppose I own two cars. My utterance on Monday might refer to the situation in which I am standing in front of my Honda; my utterance of the same sentence on

Table 1. Some examples of sentences.

1
1101
1101 11 10101101
f
fpqhe
fpqh hpt gvehepp
bag green fury the
I am standing in front of my car
Je me trouve devant ma voiture

Tuesday might refer to the different situation in which I am standing in front of my Toyota. Alternatively, the two utterances may have different references simply because they are spoken by two different people.

We also might wish to consider that two different sentences can have the same reference. If, at 6 p.m. yesterday, I utter the sentence "I am standing in front of my car," and simultaneously my friend Hank utters the sentence "Jonathan is standing in front of his car," the two utterances are of different sentences, but they share the same reference. Similarly, if I were to utter the three sentences "I am standing in front of my car. I am positioned upright before my vehicle. Je me trouve devant ma voiture," all three would have the same reference.

But there is another general way in which two utterances of the same sentence can have different meanings and in which two different sentences can have the same meaning. Just as a single word-type can have different meanings depending on the context in which it is used or instantiated as a word-token, a single sentence can have different meanings depending on the context in which it is uttered. For example, there seems to be a sense in which the meaning of the utterance "I am standing in front of my car" has a different meaning when spoken in response to the question "Are you ready to go?" than when in response to the question "Where are you," no matter what is understood to be the physical reference of the utterance.

Many would argue that referential theories fail to give a comprehensive account of meaning. One cause of failure (it is suggested) is the inability of such theories to deal with utterances that seem to refer, if to anything at all, to abstract concepts rather than to physical objects. Another problem is the suggestion that, on many occasions, our intention in making an utterance is not to refer to anything at all, but rather to have an effect of some other kind on the world (hence, the common characterization of utterances as *speech acts*). In these cases, our intuitive notion of "meaning" might lead us to wish to commit to a sense in which utterances with nonreferential functions have meaning. If the meaning of a sentence is not its referent (i.e., the situation that it represents or for which it stands), then what is it?

Thoughts

A common proposal in response to this kind of failure is to exploit the idea of a third broad category of entities existing alongside linguistic entities and physical entities. This third category is one of mental entities—a category of thoughts or ideas. In the same way that we can conceive of situations as comprising objects and properties, we might wish to treat an individual thought as a composite of concepts of different kinds.

We would then have an alternative to the idea that, whenever we talk, we refer to situations. This alternative is to say that, whenever we talk, we express our thoughts; that we think about the world, and our talk represents our thoughts.[8] The relationship of meaning that is now proposed is

one between an utterance and a thought. The meaning—the semantic content, the message—of an utterance is the thought that it expresses.

It might then appear to be a simple step to retain both ideas—that utterances express thoughts and that utterances refer to situations—by considering the additional suggestion that thoughts themselves represent (or perhaps "present") situations in some way. In other words, the suggestion is that linguistic entities represent mental entities, and any representation that occurs of physical entities is carried out through the mediation of mental entities. A typical composite account of this nature runs as follows: that there exists a world of physical objects, among which are included people and recorded utterances; that each individual person forms a mental image of that physical world; that the content of each mental image is expressed or "reflected" in the form of a person's utterances, which are, in turn, considered by other minds in the formation of new mental images.

Propositions

In the philosophical literature, the term "thought" is sometimes replaced in this context by the term "proposition." The meaning of the utterance "I am standing in front of my car" is the proposition that I am standing in front of my car; in general, the meaning of utterance *u* is proposition *p*. Theories of meaning that assume the existence of propositions might be referred to as *propositional* theories, to distinguish them from referential theories.

We might consider, however, that it is additionally important to recognize a distinction between the particular thought entertained by an individual person—which is unique—and the class of thoughts of which that particular thought is a member by virtue of its similarity to others within that class (which others may or may not be entertained by the same person). Just as we can distinguish between utterance-token and utterance-type, we also may wish to distinguish between thought-token and thought-type, specifically by substituting "proposition" for the latter category. And just as we may characterize two utterances of the same sentence as sharing the same *form*, we might characterize two thoughts of the same proposition as sharing the same *content*.

So far, we have shied away from consideration of an obvious and crucial issue for any theory of meaning, which derives from the essential communicative function of utterances. One of the primary reasons for our going to the trouble of expressing our thoughts is our desire that others should have access to those thoughts. Typically, our primary intention when making an utterance is that our audience should interpret the utterance by assigning to it the same meaning—the same proposition—as the one that is instantiated by the thought we are expressing. Unfortunately, however, our success in achieving this goal of perfect understanding never can be guaranteed, primarily as a result of the underdetermination of the speaker meaning or hearer meaning of any utterance by its form.[9] Whatever the

proposition that is conventionally understood as the meaning of a given sentence, it is possible for the speaker to intend that an utterance of that sentence has a different meaning from the conventional one. Similarly, whatever the proposition that an utterance is intended by its speaker to express, it is possible for the hearer to interpret the utterance in a different way, i.e., to understand the utterance as having a different meaning from the intended one.

If only for analytical purposes, three separate components of the communication process can be isolated at this point. One is *expression,* i.e., the act of the speaker in producing an utterance, which involves making decisions of the following kinds (inter alia): (i) a decision to select, from the stock of thoughts making up the speaker's mental state, a particular thought to express at time *t;* (ii) a decision to select a particular set of utterances as the language from which an utterance will be chosen; and (iii) a decision to select a particular utterance from that set. Another component is the establishment of *convention,* i.e., the process by which speakers and hearers reach an intersubjective consensus on the ordinary meanings of wordtypes and sentences. The third component of the process is *interpretation,* i.e., the act of the hearer in assigning meaning to a heard utterance. This act should be recognized as one that is essentially creative, in that the hearer's knowledge of any meaning that is conventionally assigned to utterances of the given sentence is only one (if a significant one) of the factors that will influence the assignation in any particular instance. Other contextual factors include the extent and nature of the hearer's prior knowledge, the nature of the discourse of immediately prior utterances, the nature of the hearer's interpretation of that discourse, and so on. Interpretation is also creative in the important sense that the meaning assigned to a heard utterance may be new; the thought triggered by a heard utterance may be one that has not been previously entertained by the hearer. Moreover, the response of the hearer to a particular utterance may not be limited simply to an increase in the quantity of thoughts making up her mental state; her attitude toward other propositions may change. One thought has a habit of leading to another.

In principle, given a speaker, an utterance, a hearer, a discourse of prior utterances, and a community of language users to which the speaker and hearer belong, we may determine that the utterance simultaneously has meanings of at least three different kinds: (i) its *conventional meaning,* i.e., the proposition conventionally attributed by the community to the sentence instantiated by the utterance; (ii) its *speaker meaning,* i.e., the proposition instantiated by the thought expressed by the speaker of the utterance; and (iii) its *hearer meaning,* i.e., the proposition instantiated by the thought entertained by the hearer on interpreting the utterance. These meanings may or may not coincide. If (ii) is not the same as (i), then the likelihood of (iii) being the same as (ii) will depend on the hearer's success in making sense of the discourse that provides the context for the utterance.

Table 2. Some entities commonly defined in philosophy of language.

	Linguistic	Mental	Physical
Tokens	Utterances	Thoughts	Situations
Types	Sentences	Propositions	

A Simple, General Ontology

The system of categories presented in Table 2 is provided both as a summary of the foregoing discussion and as an example of the sort of taxonomy that is typically proposed in philosophy of language. Entities in the "Token" categories are datable particulars; entities in the "Type" categories are nondatable classes of particulars.

The distinctions between categories formalized in this system are one possible result of our determining the necessity of differentiating among entities of the following kinds:

1. particular physical states of affairs: e.g., the *situation* (at time *t*) in which Hank is standing in front of his car;
2. general representations of individual states of affairs: e.g., the *proposition* that Hank is standing in front of his car;
3. particular internal instantiations of such representations: e.g., Lucy's *thought* (at time *t*) that Hank is standing in front of his car;
4. general external expressions of such representations: e.g., the English *sentence* "Hank is standing in front of his car"
5. particular instantiations of such expressions: e.g., Lucy's *utterance* (at time *t*) "Hank is standing in front of his car"

A *situation* is a possible state of affairs in the physical world. At any given time *t*, only some situations "obtain" or are "the case" in actuality.

A *proposition* is an abstract, mental representation of a particular situation. There is a one-to-one correspondence between situations and propositions. Propositions may be evaluated in terms of their truth: a true proposition is one that represents a situation that obtains.

A *thought* is a particular attitude toward a proposition in the mind of a particular person. Different people can have thoughts about the same proposition; a single person can have different thoughts about the same proposition. Thoughts include beliefs; a belief is an acceptance of a given proposition as true.

An *utterance* is a particular expression in symbolic form of a particular thought. The same thought may be expressed by different utterances. Utterances are what can be said to have meaning (i.e., to be meaningful); in this sense, meaning (i.e., meaningfulness) is a property of utterances. Whether an utterance is meaningful or not depends on the occurrence or nonoccurrence of an event in which a human recognizes that utterance to

express a particular thought. In this sense, strictly speaking, an utterance's meaningfulness is a property of the situation consisting not only of utterance *u*, but also agent *a* and time *t*.

A *sentence* is the class of all utterances that share a particular symbolic form. The same form can be used (by a single person or by different people) to express different thoughts. The *meaning* of a sentence is a proposition *p:* we might say that the conventional meaning of the sentence "Hank is fat" is the proposition that Hank is fat. Since we are free to interpret any given sentence in any way we like, however, there is a many-to-many correspondence between sentences and propositions. The same sentence can represent different propositions; the same proposition can be represented by different sentences. In a strict sense, the particular proposition *p* that is assigned at time *t* by agent *a* to a sentence *s* is a property of a situation—is literally *assigned* as the outcome of a human act—and not something that inheres in the sentence itself.

We might say that the *form* of an utterance is the sentence that it instantiates, and the *content* of an utterance is the proposition that is instantiated by the thought that the utterance expresses.

Conceptions of Information

We are now in a position to distinguish among three general kinds of sense in which "information" has historically been used in the IS literature. These three genera may be considered as the top level of a taxonomy of concepts denoted by the term (Table 3). In the first kind of sense, the concept of information is understood to designate *particulars* (i.e., individual objects or events) of certain types; in the second kind of sense, the designata are certain types of human *action;* and in the third kind of sense, the designata are *universals* (i.e., properties) of certain types.

Information-as-Particular

Objects or events of the kinds that are designated by the concept of information-as-particular are not necessarily "physical" or tangible. In this particular context, the term "object" is used to refer both to linguistic entities (such as words) and to mental entities (such as concepts), as well as to physical entities. The distinction between the linguistic, the mental, and the

Table 3. Conceptions of information

Information-as-particular	Utterances Thoughts Situations
Information-as-action	Communication
Information-as-universal	Informativeness Relevance

physical serves as the basis for a division of the information-as-particular category into three subcategories: the *utterance* as information; the *thought* as information; and the *situation* as information.

The utterance as information

In this sense, corresponding roughly to Buckland's category of "information-as-thing,"[10] the concept of information is understood to designate symbols, signs, or signals,[11] i.e., noises or marks (or even aromas or flavors) that are interpreted in some way by the hearer or viewer (or smeller or taster). A generic term commonly used in many different contexts to denote aggregations of such signals is "data." Binary digits are data of one of the simplest kinds; words, images, and sounds are also data. Data may be aggregated in various forms and at various levels, as (to take the example of textual data) alphabetic characters, words, sentences,[12] paragraphs, chapters, and books. Middle-range aggregations of data are typically considered generically as documents (or "document-like objects"). In the sense presented here, then, "utterance," "data," "signal," and "document" are more or less functionally equivalent.

The conception of the utterance as information may be considered as an objectivist view, not simply by virtue of the physical existence of utterances, but primarily by virtue of the supposedly objective nature of the criteria to be used in determining whether something is classifiable as information or not. In effect, this view commits one not only to the proposition that information is anything that is interpretable—i.e., anything that is capable of being interpreted—but also that the interpretability of an entity does not depend on its historically having been interpreted. Entities can thus be classed as information on the basis of their potential to inform.

A distinction is often drawn between natural signs—the forms of physical objects such as clouds, tree stumps, smoke, tracks, and rocks—and conventional signs. This distinction can serve not only to separate naturally occurring signs from human artifacts, but also to highlight a supposed difference between the objective meaning of natural signs and the (at best) intersubjective meaning of conventional signs. Once the decision is taken to consider naturally occurring entities as interpretable, the way is clear for a definition of information that encompasses "everything"—everything that has the potential to be treated as a source of meaning, that is.

The thought as information

In this sense, corresponding to Buckland's category of "information-as-knowledge," the concept of information is understood to designate messages, i.e., the concepts or thoughts that are the product of a hearer's interpretation of signals. The distinction between signal and message is explicitly drawn not only in the philosophy of language (where the distinction may be cast as one between utterance and thought, or between sentence and proposition), but also, famously, in the mathematical theory of

communication[13] and in semiotics (where the contrast is between the signifier and the signified).

A metaphor in widespread illustrative usage is the distinction between physical container and mental content: documents are regarded as the (physical) expressions, representations, or records of conjunctions of (mental) thoughts. Some authors, developing the metaphor further, have promoted the view of documents as vehicles by which messages are transferred, via a channel, conduit, or canal, across space and time. We should be careful, however, not to infer from such an account of the communication process that any given document has a single corresponding message—for example, that intended by its author—the recovery of which is the aim of any hearer; for it is clear both that any individual signal can be interpreted in multiple ways and that any individual message can be expressed in multiple ways. The conduit metaphor seems to serve only to reify the controversial idea that information is something that can somehow "flow" from one place to another. Such flow would be possible only if messages were inherent properties of signals, rather than separate entities assigned to signals by humans. These two different views of the nature of messages may be characterized as, respectively, an objectivist and a subjectivist perspective on information.

Signals formed or recorded on media of certain kinds (e.g., paper, tape) persist over time, with the result that they may be considered at a date later than that of their creation. The set of messages expressed by all signals stored in this way is sometimes referred to as the world of recorded, public, explicit, or objective knowledge, to contrast to the world of private, tacit, or subjective knowledge. The sense in which "objective" and "knowledge" are used in such formulations is sometimes ambiguous, however. If it is actually the set of signals themselves (rather than the set of messages expressed by those signals) that is being referred to, then "objective" is appropriate but "knowledge" is not; if instead the referent is the set of thoughts or meanings that could potentially be assigned to those signals, then "knowledge" (in the everyday sense) seems appropriate, but it might be argued that "objective" is not, since such knowledge exists in people's minds—i.e., in the subjective realm of consciousness that is usually contrasted with the physical world—not in signals.

A common strategy in IS is to define "information" in such a way as to denote that class of messages that share a particular property, or (more accurately) those that stand in a relationship of a particular kind to the context in which their source signals are interpreted. A message might be classified as information if it satisfies any (or any combination) of criteria of the following kinds: (i) truth—i.e., its corresponding with the facts; (ii) utility—its potentially being used to further the goals of the hearer; (iii) novelty—its not having been assigned by the hearer to any previous signal; (iv) unexpectedness—its coming as a surprise to the hearer; (v) uncertainty-

reduction—its increasing the strength of the hearer's current attitude toward any proposition; and so on.[14] Taken together, such criteria may be considered as criteria for the *relevance* of a message to a hearer.

On this account, the effect of the signal that is the source of the message is not simply to trigger some interpretative response on the part of the hearer, but more specifically to cause a transformation of a particular kind in the hearer's mental state. It is the message that is information; the signal that serves as the source of the information is merely "informative," or, as philosophers of language might say, "meaningful."[15] All signals are at least potentially informative; only some—those whose hearer meanings are true and/or useful and/or novel, etc.—are actually so. Thus the determination of the informativeness of a signal at time t is a subjective matter, one that relies on our determining the truth, utility, novelty, etc. of the meaning assigned to it at time t by an individual hearer. This is roughly the position of adherents to the influential "cognitive viewpoint" in IS.[16]

One reading of the cognitive view of information amounts to saying that only some messages—those that fulfill the informative potential of signals—are information. But we might decide that it is more useful to consider (i) not only that all signals are potentially informative, but also that all messages are potentially relevant and (ii) that potential informativeness and potential relevance are matters of degree. If we define information to encompass only actually relevant messages, then any decision as to whether some entity is information or not becomes a wholly subjective matter (in the sense that only individual hearers can determine the actual relevance of messages), and the term "information" becomes of limited application. On the other hand, if we define information to encompass all potentially relevant messages, then the class denoted by "information" becomes the same as that denoted by "message," and the former term becomes redundant. In any case, it would appear that determining the extent to which a message is relevant to hearer a at time t is what is more important. I shall return to this point in a moment.

The situation as information

In this sense, the concept of information is understood to designate situations—i.e., states of affairs, arrangements of physical objects and properties, that may or may not obtain in actuality. Such a conception may be considered as an extended version of the idea (discussed under "The utterance as information," above) that information is everything that is interpretable; in the present case, even the requirement of interpretability becomes irrelevant since it is assumed that every possible entity is necessarily interpretable. On this reading, the physical world is made up of data, or (as is often said) of differences. Any entity, from the simplest to the most complex, may be defined by specifying the ways in which it may be distinguished from another; such a specification is the substance of the entity;

the entity is its own specification. The distinction that, until now, we have sustained between physical and linguistic entities is eradicated. For many, the usefulness of assigning the label "information" to the class of all things remains to be demonstrated.

Information-as-Action

The category of information-as-action corresponds roughly to Buckland's "information-as-process," encompassing conceptions of information that apply the term to sequences of events that involve humans either as agents (subjects) or as patients (objects) or both, and that may thus be treated as acts or actions.

Communication as information

It might be instructive at this point to contrast the common usages of two words of similar morphology. We might talk about having received "a communication" (i.e., a document), but more generally "communication" is used to refer to the process of communicating, whereas "information" is seldom used (in ordinary English language, at least) to refer to the processes of informing or of becoming informed, of expressing thoughts or of interpreting utterances. Yet these latter are the technical senses that are meant here. Instead of saying that documents are information, or that documents contain information, we might wish to say that our very acts of creating, classifying, storing, retrieving, and/or interpreting documents are information. I would suggest, however, that we already have perfectly adequate labels for those acts.

Information-as-Universal

The category of information-as-universal includes conceptions of information that apply the term to certain attributes, or properties, of objects or events. For example, the particular conception of information that is associated with the mathematical theory of communication is one that defines information as a measurable, quantitative property of signals. Thus it makes sense to talk of the *amount* of information in a given signal.[17]

Informativeness as information

Given the everyday usage of "information" in reference to objects rather than properties, it can be helpful, when thinking about information-as-property, to substitute "informativeness" for the more common term, and to consider the degree of informativeness (rather than the amount of information) of signals.

Is there an analogy to be drawn between the informativeness of signals and the meaningfulness of utterances? Given the sense in which meaningfulness was spoken of earlier, it might seem that to talk of the *degree* of meaningfulness of a signal would be to stretch the notion inappropriately. We might decide that it is more sensible to retain the idea that meaning-

fulness is a binary property—that a signal is either meaningful or not. There do seem to be other senses, however, in which we can comprehensibly assess the value both of signals and of propositions.

One of these is the sense in which "information" is used in the mathematical theory of communication. To evaluate the extent to which a signal is informative, we need to determine the size of the set of signals from which the given signal is selected; the larger this set, the greater the uncertainty that a particular signal will be selected, and the more informative the final selection. Strictly speaking, then, informativeness is less an inherent property of a signal per se, and more a property of the decision made to select that signal from a particular set of alternatives.

Relevance as information

Decisions of this kind might well be construed as events rather than objects. Once we realize that we are not necessarily restricted to signals when identifying the kinds of entities that may "have information" to a greater or lesser extent, other conceptions of information in the category of information-as-property suggest themselves. For instance, we might wish to develop a conception of the informativeness not merely of signals, but of messages, or even of situations. Such a conception might involve taking into account not only the size of the set of entities from which the given entity is selected, but also the history of previous such selection-events. In this way we might arrive at conceptions of information that are essentially equivalent to contemporary conceptions of relevance.

Some Concluding Remarks

We have now seen, through an analysis of the categories to which the term "information" is variously applied in IS, how those categories are well-understood in fields such as philosophy of language, communication studies, and semiotics, and how labels other than "information" have been used to effectively distinguish among those categories in those fields. I would like to conclude by suggesting that the treatment of the theory of meaning adopted in this paper serves the purpose of highlighting the precise location where any theory of information becomes truly interesting—and where, incidentally, fields other than IS have made, and continue to make, a greater deal of progress.

The point is that a good theory of meaning should do more than explain what it is to say that a signal is meaningful. It needs to explain *how* a person assigns a particular meaning to a given signal; how one person determines the meaning that has been assigned to a signal by another (in other words, how communication takes place); how certain meanings come to be conventionally associated with certain signals, and so on. However, these seem to be less metaphysical questions than they are psychological—questions, in other words, for the cognitive sciences, or (to the extent that they involve philosophy at all) for the philosophy of mind.

Now, if we wished to maintain one view of the nature of information that is commonly assumed in the IS literature—that of the thought as information—then it would seem that these questions about the mechanics of cognitive processes become crucially important for us to answer . . . or at least for us to find answers for in the literature of cognitive psychology. (I am prompted to inquire what exactly it is that is achieved by attempting to corral such weighty and long-standing questions under the rubric of "information studies" when they already attract wide interest from communities that, on the face of it, are rather better equipped to deal with the kinds of issues that are raised by practical brain research.)

Alternatively, if we are more convinced of the usefulness of another common conception of information—that of relevance as information—then we should recognize, perhaps, that a good theory of information should do more than simply explain what it is to say that a document's content is relevant. Ideally, it would explain how the extent to which a document's content is relevant (to agent *a* at time *t*) may accurately be predicted. Of course, we have several such theories in IS, the most advanced being those that cluster in the area known as probabilistic information retrieval (IR). Relevance research lies at the heart of probabilistic IR and of information studies in general. Relevance researchers are also those IS people who are most keenly aware of the significance for IS of current work being done in pragmatics and the philosophy of language.[18] For those of us who, while sharing a concern for the ways in which definitions of terms shape perceptions, directions, and agendas, do not view ourselves as relevance researchers per se, I suggest that it would be worthwhile to reflect on the coherence, parsimony, and utility of a theory of information that, in its essential conception of information-as-property, also happens to trace its lineage back to Shannon's original "information theory."[19]

Notes

1. Recent reviews in information studies (IS) of the literature on the nature of information include those by Capurro and Hjørland (2002), Case (2002), Cornelius (2002), and Dick (2002).
2. Some might prefer the substitution of "symbolic" for "linguistic" in formulations of this kind.
3. The division into three and the distinctive nature of the categories recall the Popperian conception of worlds 1, 2, and 3. But Popper (1968) wished to emphasize the epistemic qualities of worlds 2 and 3; each was conceived as a world of "knowledge" (subjective knowledge in the case of world 2, and objective knowledge in the case of world 3). A closer analogy would be to any triadic model of the sign in semiotics: see, for example, Nöth (1990) for a review of these models.
4. The explicit identification of a category of physical entities in this model betrays a commitment to realism about the external world, i.e., an assumption of the existence of a world beyond our collected mental states that constrains those mental states in certain ways.
5. There are many good introductions to the main issues in philosophy of language; a recent example is by Lycan (2000). Blair (2002) provides an overview of the applications of philosophy of language to the study of information retrieval.
6. Another account might privilege events rather than objects but treat events similarly as property-bearing entities occurring in relation to one another.

7. Other technical definitions of "sentence" are also variously suggested in the philosophical literature.
8. We might say that such a proposal asks us to commit to the ontological priority of the physical world over thought, and thought over language.
9. It is intended that the term "speaker" denotes any human source of an utterance, in spoken or written form; similarly, "hearer" denotes any human interpreter of an utterance.
10. See Buckland (1991).
11. Semioticians are usually careful to distinguish the senses of these three words. The concept of "sign" is commonly modeled as a dyad of word and object (e.g., signifier and signified) or as a triad (e.g., sign vehicle, significatum, and denotatum). It is important to note that the intention at this point in the present paper is to use "signal" and "utterance" interchangeably, to denote signifier-tokens.
12. "Sentence" is used here in the everyday sense of the term.
13. See Shannon and Weaver (1949).
14. See, for example, Machlup (1983).
15. Cognitivists might even prefer "transformative."
16. See, for example, Belkin (1990).
17. It should be noted that this conception of amount of information is quite different from the notion of quantity that is assumed, for example, in studies that seek to determine how much information there is to be found in the world. In studies of that kind, it is the quantity of data (numbers of documents or numbers of bytes) that is measured.
18. In his paper summarizing the significance for IS of the theory of psychological relevance developed by the pragmaticists Sperber and Wilson (1986), Harter (1992) suggests that "[r]elevance and information-as-process are intimately related; we may not need both ideas (or terms) in information science. Moreover employing two terms may be detrimental to the development of theory, since it suggests that information and relevance are different, when perhaps they can usefully be regarded as one and the same."
19. The implications for IS of the particular vision of information-as-property developed by Dretske (1981) have perhaps not yet been fully recognized; see van Rijsbergen and Lalmas (1996) and Bonnevie (2001) for interpretations from different perspectives.

References

Belkin, N. J. (1990). The cognitive viewpoint in information science. *Journal of Information Science, 16,* 11–15.

Blair, D. C. (2002). Information retrieval and the philosophy of language. In B. Cronin (Ed.), *Annual review of information science and technology: Vol. 37, 2003* (pp. 3–50). Medford, NJ: Information Today.

Bonnevie, E. (2001). Dretske's semantic information theory and meta-theories in library and information science. *Journal of Documentation, 57*(4), 519–534.

Buckland, M. K. (1991). *Information and information systems.* New York: Praeger.

Capurro, R., & Hjørland, B. (2002). The concept of information. In B. Cronin (Ed.), *Annual review of information science and technology: Vol. 37, 2003* (pp. 343–411). Medford, NJ: Information Today.

Case, D. O. (2002). The concept of information. In *Looking for information: A survey of research on information seeking, needs, and behavior* (pp. 39–63). New York: Academic Press.

Cornelius, I. (2002). Theorizing information. In B. Cronin (Ed.), *Annual review of information science and technology: Vol. 36, 2001/02* (pp. 393–425). Medford, NJ: Information Today.

Dick, A. L. (2002). Ontological aspects of information. In *The philosophy, politics and economics of information* (pp. 31–39). Pretoria, South Africa: Unisa Press.

Dretske, F. I. (1981). *Knowledge and the flow of information.* Cambridge, MA: MIT Press.

Harter, S. P. (1992). Psychological relevance and information science. *Journal of the American Society for Information Science, 43*(9), 602–615.

Lycan, W. G. (2000). *Philosophy of language: A contemporary introduction.* London: Routledge.

Machlup, F. (1983). Semantic quirks in studies of information. In F. Machlup & U. Mansfield (Eds.), *The study of information: Interdisciplinary messages* (pp. 641–671). New York: Wiley.

Nöth, W. (1990). Sign. In *Handbook of semiotics* (pp. 79–91). Bloomington: Indiana University Press.

Popper, K. (1968). Epistemology without a knowing subject. In B. van Rootselaar & J. F. Staal (Eds.), *Proceedings of the third international congress for logic, methodology and philosophy of science* (pp. 333–373). Amsterdam: North-Holland.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication.* Urbana: University of Illinois Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition.* Cambridge, MA: Harvard University Press.

van Rijsbergen, C. J., & Lalmas, M. (1996). Information calculus for information retrieval. *Journal of the American Society for Information Science, 47*(5), 385–398.