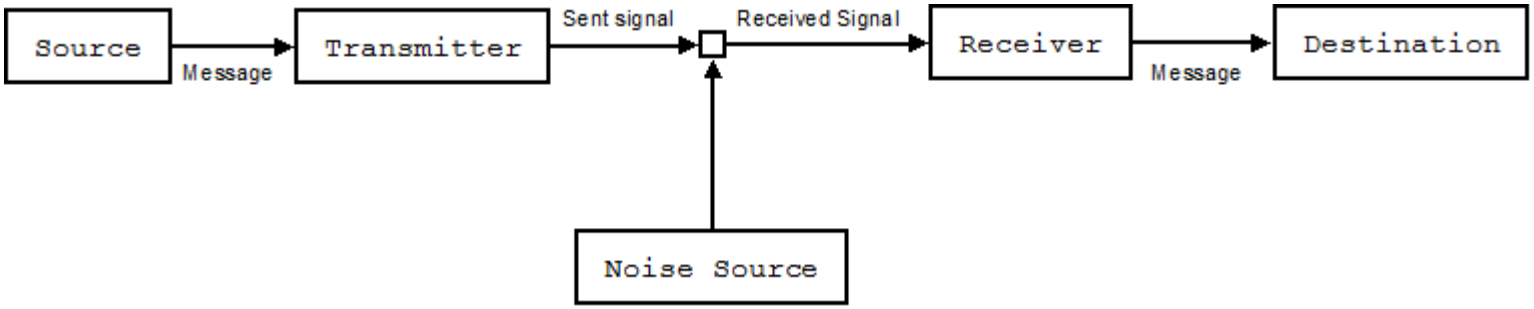
Objectivism and Subjectivism about Information in Dretske's Semantic Theory of Information.

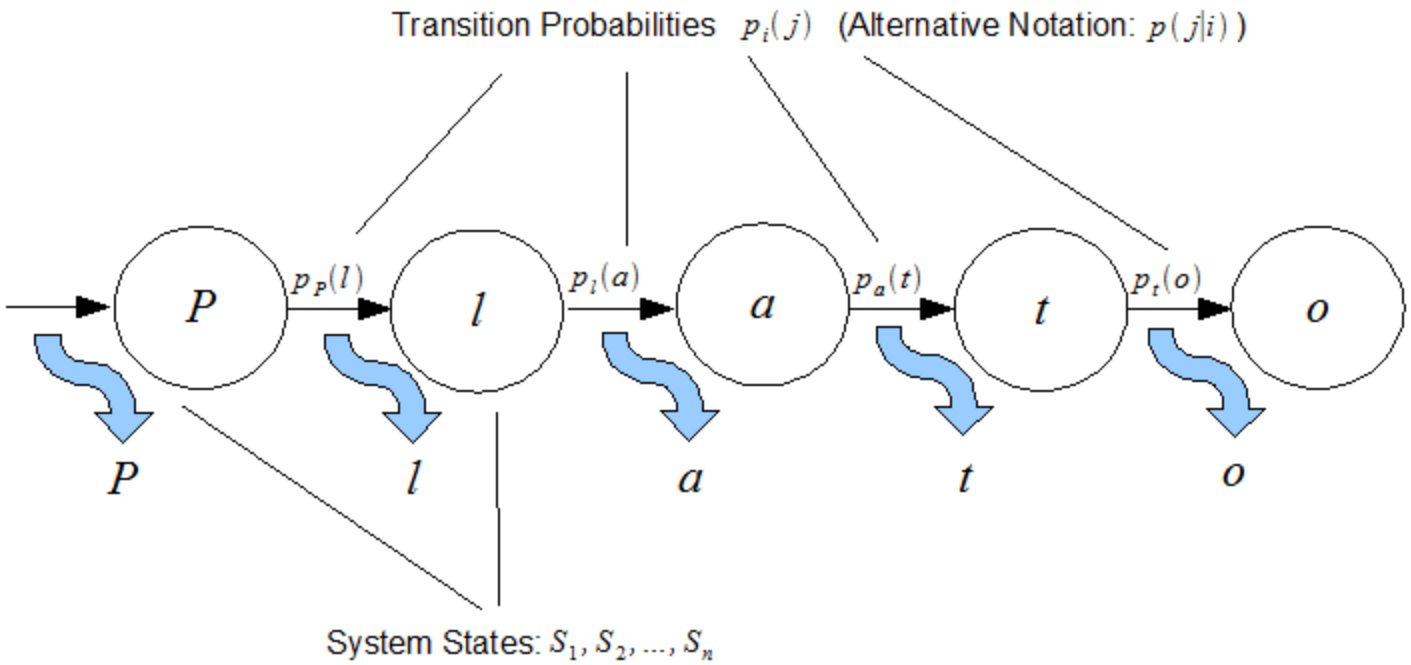
Shannon's Schema



The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have *meaning; that is they refer* to or are correlated according to some system with certain physical or conceptual entities. **These semantic aspects of communication are irrelevant to the engineering problem. The significant aspect is that the actual message is one *selected from a set of possible messages.*** *The system must be designed to operate for each* possible selection, not just the one which will actually be chosen since this is unknown at the time of design.

If the number of messages in the set is finite then this number or any monotonic function of this number can be regarded as a measure of the information produced when one message is chosen from the set, all choices being equally likely. (Shannon, The Mathematical Theory of Communication, 1948.)

Shannon Source State Sequences



$$H = -\sum p_i \log p_i$$

Information is *Objective* Uncertainty...

We have represented a discrete information source as a Markoff process. Can we define a quantity which will measure, in some sense, how much information is "produced" by such a process, or better, at what rate information is produced?

Suppose we have a set of possible events whose probabilities of occurrence are $p_1, \ldots, p_n$. **These probabilities are known but that is all we know concerning which event will occur. Can we find a measure of how much "choice" is involved in the selection of the event or of how uncertain we are of the outcome? [how uncertain the outcome objectively is.]** (Shannon, 1948)

Choice and Uncertainty: Objective and Statistical

- Shannon uses the term “choice” with quotes because he is interested in the *objective* statistical conception of both choice and selection. This conception involves the objective occurrence of a given specific outcome from the possibility space. It is not about any agent cognitive choice in the psychological sense, except where this is part of the source mechanism for state and symbol determination.

Dretske's Semantic Theory of Information: Preparatory Remarks

To get the semantic theory off the ground, Dretske first requires:

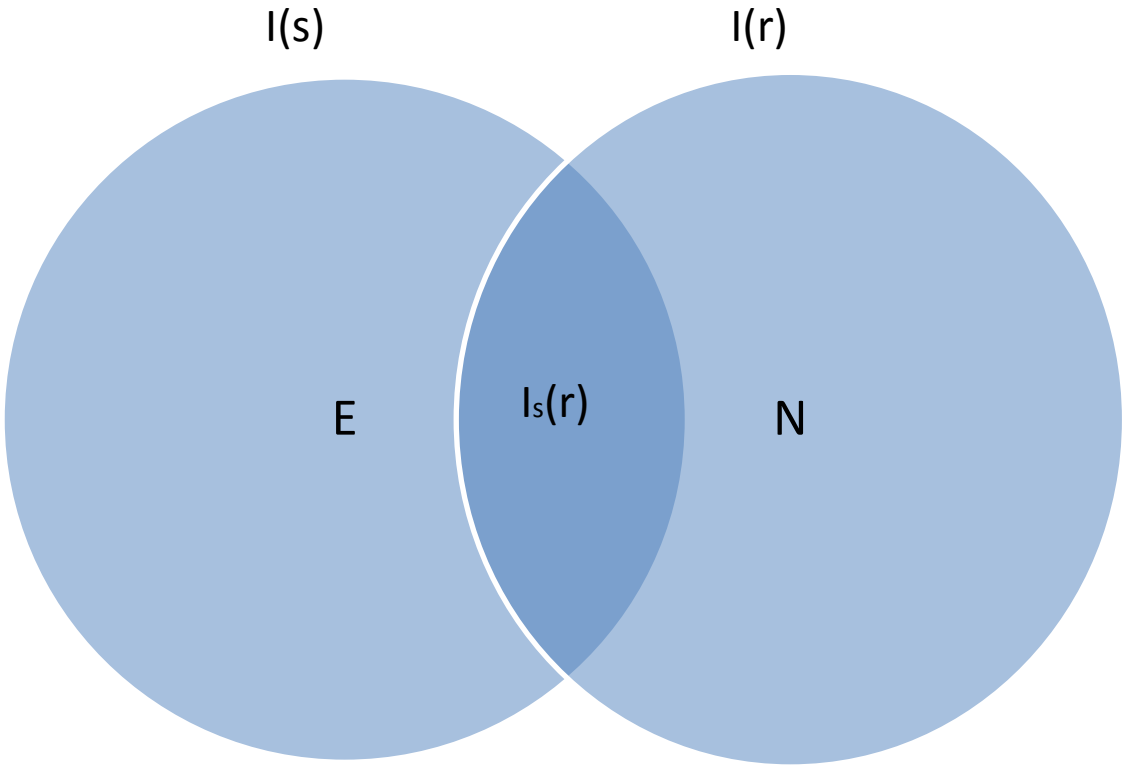
1. A measure of the information in a specific individual source state and corresponding message or symbol.
2. That no absolute specific value of the measure need be determinable or calculable by a receiving agent.
3. This measure can be used to compare signal and message information content on the basis of what the receiving cognitive agent knows about the source state of affairs which received messages are 'about'.

Two conceptions of information: Objectivist and Epistemic/Subjectivist.

1. The conception of uncertainty and probability of possible states that Dretske needs to get the semantic theory working is *objective*.
2. The conception that he needs to make the theory work for cognitive agents is *subjective*.

These are two fundamentally different conceptions of information which form the basis of two completely different conceptions of source and message information content. Moreover, it is simultaneously the case on Dretske's approach that 1. (objective possibility set and probability distribution) is the *exclusive* determiner of message information content, but 2. (subjective **knowledge** of the subset of possibilities) is **also** a determiner.

Quantities of the form $H = -\sum p_i \log p_i$ (the constant *K merely amounts to a choice of a unit of measure)* play a central role in information theory as measures of information, choice and uncertainty. The form of H will be recognized as that of entropy as defined in certain formulations of statistical mechanics where is the probability p_i of a system being in cell *i of its phase space.* H is then, for example, the H in Boltzmann's famous *H theorem. We shall call* $H = -\sum p_i \log p_i$ *the entropy of the set of probabilities* $p_1, \ldots, p_n$ *.If* x is a chance [random] variable then we will write $H(x)$ for its entropy. (Shannon, 1948)



Dretske's explication of Shannon's Theory

Surprisal of particular event: $I(si) = \log 1/p(si) = -\log p(si)$

Average Information of Source: $I(s) = \Sigma p(si).I(si)$

Average Information at Receiver: $I(r) = \Sigma p(ri).I(si)$

Information transmitted: $I_s(r) = I(r) - \text{Noise}$ OR $I_s(r) = I(s) - \text{Equivocation}$

Dretske's equations for individual source states and signal contents:

(Dretske, KFI, 50)

The quantities of interest then, are the amount of information generated by a particular event or state of affairs s_a :

2.1 $$I(s_a) = \log 1/\ p(s_a)$$

and the amount of information carried by a particular signal r_a about s_a :

2.2 $$I_s(r_a) = I(s_a) - E(r_a)$$

The Objective Possibility Space Problem

It may seem as though (2.1) and (2.2) are not very useful tools. Aside from contrived examples involving games (or in applications to telecommunications systems **in which there is some theoretically restricted range of possibilities with their associated probabilities**) **the use of these formulas to compute amounts of information seems hopeless**. How, for example, do we calculate the amount of information generated by Edith's playing tennis? And how much information about her activities is contained in the light reaching Herman-a casual spectator? (Dretske, KFI,)

“Restricted Range of Possibilities”

- In Shannon’s theory this is still objective. It is established in most cases of the discrete case using frequency data to determine prior probability distributions.

In order to answer questions like this [slide II] in accordance with (2.1) and (2.2) one needs to know: (1) the alternative possibilities (e.g. Edith's eating lunch, Edith's showering [**clearly these are infinite**]); (2) the associated probabilities of all these alternative possibilities; (3) the conditional probabilities (given the **configuration of the photons** reaching Herman). Obviously, in most ordinary communication settings **one knows** none of this. It is not even clear whether one **could know it.** What, after all, are the alternative possibilities to Edith's playing tennis? Presumably there are some things that are possible (e.g. Edith going to the hairdresser instead of playing tennis) and some things that are not possible (e.g., Edith turning into a tennis ball), but how does one begin to catalog these possibilities? (Dretske, KFI, 53)

Infinite possibility space with infinite set of possibilities gives, (by Kolmogorov's 2nd and 3rd axioms):

$$\lim_{i \to \infty} p(n_i) = \lim_{N \to \infty} \frac{1}{N} = 0$$

$$N = \left| \{ n_1, \ldots, n_i \} \right|$$

$$H(n_i) = \log \frac{1}{p(n_i)} = -\log p(n_i)$$

These difficulties in applying (2.1) and (2.2) to concrete situations (situations of interest to someone, like myself, trying to formulate a theory of knowledge) may seem formidable indeed, but as we shall see, they present obstacles only to an overambitious use of the formulas. If one seeks an *absolute measure,* a *definite numerical figure,* for the amount of information generated by an event or carried by a signal, one must be in a position to determine the range of possibilities and their associated probabilities. (Dretske, KFI, 53.)

Comparative Information Content

Despite these limitations, formula's (2.1) and (2.2) have an important use. They can be used to make *comparisons*, in particular, comparisons between the amount of information generated by the occurrence of an event and the amount of information a signal carries about an event. Such comparisons can be made without ever determining absolute values for either magnitude. That is, one can use these formulas in the way one would use a piece of string that is not marked off in inches and feet.One can use the string to determine whether A is longer than B without ever determining the length of either A or B. (Drestke, KFI, 53.)

Language Examples and Subjectivity

Say we are using a discrete message which is linguistic and has meaning for a human receiver by way of grammatical rules and other empirically established 'codes' for fixing reference and semantic content. The human receiver might be cognitively aware – on the basis of experience - that a certain message they receive is to be less expected than other messages. They will not, however, have any idea of the objective probability value associated with that message out of the set of possible messages, nor have any concrete idea of the number of possible messages in the *objective* set of possible messages (lexicon, for example.)

Animals

Take the case of animals. A given animal may be psychologically surprised to hear a sound produced by one of its species that is never or very rarely produced by that species. In other words, based on its empirically attained knowledge of what possible sounds can occur at another creature of its kind. There will be an objective probability associated with the sound which will be small compared to the probabilities of other more commonly used sounds in the set of possible animal sounds. However, the uncertainty the animal may 'experience' has nothing to do with them *knowing objective probability values*.

Ineliminable Objective Component

- For both of the previous examples, it is these *objective* probability values and the *objective* statistical uncertainty that is associated with them that are required for the formulas for individual source and signal/message information content that Dretske develops.
- The receiving agent can make no comparison on the basis of these formulae without some knowledge of what can (objectively) possibly obtain at the source: knowledge of the objective possibilities. The objective possibilities are *ineliminable* in the comparative formulation relying on the single source state and signal content measures. Specific knowledge about them is necessarily required for both an absolute specific measure *and* still also for a comparison.

Objective-Subjective Tension

- The information content of a message and of the source state **cannot** be *determined* by BOTH of the following:
  1. **Exclusively** by the objective probability distribution over possible states regardless of what the agent knows AND also
  2. By what the agent receiver knows subjectively about the possibilities that can obtain at the source .
  - The set of possibilities for 1. must be greater than that for 2. in most cases. On Shannon's averaging formulation and Dretske's contraction of it each set would give different information content and values, especially for possibilities with unequal probabilities.

Yet the above is just what is required by Dretske, and also what he states to be the case. However, the epistemic-probability determined conception of information at s source/message cannot be juxtaposed on the objective-probability determined conception of information at a source/message. The former precludes the latter by definition. 1. is statedly exclusive.

Initial Problem statement:

How does one know anything about the comparative possibilities and comparative/relative information content without having a well-enough defined range of possibilities for a source to at least determine which source involves more possibilities? Take p to be the proposition “There are more possible places to live in Wisconsin than on Adams Street in Wisconsin”. That I have been told (sent the message) q, (that) “Denny lives on Adams Street in Wisconsin” only provides me with *more* information than being told that r “Denny lives in Wisconsin” IF it is in fact true that p (in other words, iff p). Why?...

Objective Message Content: Objectivism about Information/Information Content

Dretske points out that even if a cognitive agent either does not believe or misunderstands the message [q], then they have still received the information that [q] via the signal that bears the message [q]. In other words, the receiving agent can determine the state of affairs at the source to be other than it actually is having received the message [q], but information from the source state of affairs will still be encoded into [q]: [q] will still carry/**represent** the information *that q*. The information *that q* is in the message regardless of the agent's psychological disposition to message [q] and the belief state/propositional attitude that they adopt towards [q]. It is also statedly independent of their knowledge of the source state of affairs:

It cannot be stressed too strongly...that the set of conditional probabilities defining the equivocation between source and the receiver are meant to be **objective features** of the communication system. **We may not know what these conditional probabilities are, we may not be able to determine what they are, but this is irrelevant to the actual value of the equivocation and therefore irrelevant to whether the signal carries as much information as the source generates...**How much information a message carries is *not* a function of how much information the recipient thinks it carries [based upon their knowledge]. It is a function, simply, of the actual possibilities that exist at *s* and the conditional probabilities of those various possibilities after the message has been received. (Dretske, KFI, 55)

Enter Epistemic/Subjective Probability: Subjectivism about Information

Up to now we have indulged in the harmless fiction that the number of possibilities existing at the source (and their respective probabilities) was fixed independently of what anyone happened to know. This fiction allowed us to develop the information-theoretic framework without distracting complications. (Dretske, KFI, 80)

For an objectivist (probabilistic) conception of information, based upon objective probabilities, which is what Dretske needs for the comparative formulae to be coherent, the number of possibilities at the source must be fixed *objectively*.

The 'Communal Knowledge' Kludge

Furthermore, the fiction is often rendered harmless by the fact that the assessment of the information contained in a signal (both how much information and what information) is carried out against a background of communally shared knowledge in which individual differences are submerged. That is, what is known about the various possibilities at the source (the *k* in our definition of informational content) is the same for all relevant receivers, and the discussion can proceed (for practical purposes) as if the commodity in question were absolute and fixed...often enough, a common frame of reference is understood. Every relevant party knows the same thing about the possibilities arising at the source. (Dretske, KFI, 80.)

Problems Reconciling Objective and Subjective Information

Dretske's initial statements do not involve a "harmless fiction that the number of possibilities existing at the source (and their respective probabilities) was fixed independently of what anyone happened to know."(Dretske, KFI, 80)

Dretske realises that for 2.1 and 2.2 to work as the basis for comparisons between messages, the receiving agent must be able to distinguish which message has more information on the basis of knowing which source has more possibilities. They must know something about the possibilities at the source. The probability space has become *subjective*. The reduction in uncertainty about what obtains at the source has thus become *subjective.*

Objective and Epistemic Information

To know p, I must have acquired information about the state of affairs/source *s* to the effect that p. There are two possibilities: p and ¬p. To know p, it must be that I have previously received a signal carrying information that p, such that there has been a reduction of my uncertainty about what obtains (p or ¬p). In other words, there is a *logical* probability space associated with p about which I must have received some message in the past and about which I must know something. This logical probability space has possibilities that correspond to a statistical probability space, but the statistical probabilities may be unknown and unknowable. Without knowing p or ¬p I will not be in a position to determine which message - q or r – carries the most information. As Dretske himself points out, however, in most cases it will be hard to know the objective probabilities of any of the possibilities for a source. In order for the subjective percipient-receiver to be able to *compare* the messages q ("Denny lives on Adams Street in Wisconsin") and r ("Denny lives in Wisconsin") and determine that one has greater information content *objectively*, they must have received the information that p. After all, the rest of Wisconsin might be desert, and Adams street the only place to live in Wisconsin, in which case each message would carry the same information.

Subjective Source Possibilities

- Receiving agent versus theorist perspective:
- Neither is the information contained in a signal dependent on the receiver's actually learning something from that signal (KFI, 57)
- ...something that should be perfectly obvious on other grounds; viz., that whether or not a signal is equivocal depends on how **we** carve up the possibilities at the source. (KFI, 61)

As we have seen above, how the *agent* carves up the possibilities is supposed to be irrelevant to the information content of the source state and signal. If the 'we' is the theorist, then that is hardly relevant to a given receiving agent's knowledge of the possibilities at the source. The 'agent on the ground' will know nothing of some theorist's carve-up of source possibilities: they will never have received any information about such (it is not even coherent to consider as valid).

Using Shannon's Schema??

Is Shannon's schema even appropriate in the case of epistemic agents and epistemic probability?

- Q. Does the source include agent observers to describe the state of affairs? i.e. Is the source just Wisconsin, or is it Wisconsin plus whatever agent is formulating and sending a message about Wisconsin plus Wisconsin?
- Q. Is Adams Street in Wisconsin a separate source from Wisconsin? This will affect the possibility space and the objective and epistemic probabilities.

[Dretske does discuss nested information content in messages]

It would seem that these should significantly affect how possibilities are assigned.

Revising the Objectivist basis:

In the case of the communication system, the possibilities at the source are determined by frequency data. Shannon's source possibility space associated with the possible source states was objective. When the receiver of information is a cognitive agent percipient who does not *know* what the source state possibilities are, this does not affect what information a message carries.

Dretske's Conditions for Semantic Information:

(C)The quantity of information the signal carries about *s* is (or includes) that quantity generated by *s*'s being *F* [being a certain state of affairs] (and not, say, by *s*'s being G) (Dretske, KFI, 64.)

Informational Content: A signal *r* carries the information that *s* is *F* = The conditional probability of *s*'s being *F* given *r* (and *k*), is I (but, given *k* alone, less than I.)

(Dretske, KFI, 65.)

k

The parenthetical *k* … Is meant to stand for what the receiver already knows (if anything) about the possibilities that exist at the source. (KFI, 65)

If the receiver knows nothing, there is still information on the objective possibilities/probabilities basis. The information content of *s* – *I(s)* - is fixed regardless of what the receiver knows. However, the receiver must know something about *s* in order to get information about *s* from *s*, BUT – this knowledge then is supposed to affect *I(s)* due to the 'carving up of posssibilities' according to specific agent receiver's *k* ?? [See next slide]

Where Dretske's Theory ends up:

...a receiver's background knowledge is relevant to the information he receives (both how much information and *what* information) only to the extent that it affects the value of I(s) – the amount of information generated at the source by the existence of a specific state of affairs. **(KFI, 81.)**

The 'only' here is deceptive: I(s) is fixed according to the *objective* possibility and probability spaces.

The Metaphysical Picture: Causal Pathways

There is nothing wrong with positing a set of possibilities at a source, nor with a theorist or agent carving up the possibilities at a source: but agent-carved possibilities are not objective possibilities and the probabilities associated with them – if also determined by the knowledge of the agent – are also epistemic. Asserting that *I(s)* - the measure of the information at the source which is determined by the objective probability distribution at the source – is affected by the agent's knowledge of the source possibilities *k,* results in an inconsistency. To illustrate it, we can ask how it is that the states of the source could be affected by the agent's knowledge about anything at all including the source states. This goes to the nature of physical communication systems...

Physical Communication Systems

In Shannon's schema - the underlying mechanisms of information transmission are necessarily physical. Even though Shannon models communication systems using mathematical-statistical abstracta, he never asserts that the systems – including sources – are anything but *physical*. A source on Shannon's theory is any stochastic *physical* process. Correspondingly:

- There is no such thing as abstract information on Shannon's theory. Possible source states must all obtain in a physically-causally interactive way: otherwise nothing can be encoded and transmitted.

Unidirectional Causal Pathway

In Shannon's schema - information only 'flows' from the source to the destination. The idea that the knowledge (*k*) of the receiver about the source possibilities can affect the value of *I(s)* requires on Dretske's model that *k* affect the objective source state possibilities themselves. **It is not just a case of *k* affecting *I(s)* as part of a statistical *model* of the source state, the message the receiver receives, and the receiver's knowledge about *s*.** It involves **changes in the objective source possibility space as determined by *k*.** *This would have to involve a causal pathway from destination to receiver, and a flow of information in that direction*. This can **not** transpire on Shannon's schema, and Dretske's interpretation assertedly retains the schema as involving unidirectional information flow.

Semantic Information?

Our definition not only yields a concept of information that satisfies the quantitative demands of communication theory, it also explains the cognitive importance of information by revealing the nature of the connection between information and truth. **It enables us to understand the source (the intentionality of natural laws) of the semantic character of information.** (KFI, 81)

Dretske's conception of information combines two incompatible conceptions of information at a source: that which is *exclusively* determined by objective possibilities, and that which is somehow affected by *k* the subjective epistemic content of the receiver w.r.t *s. The latter can be seen not to be obtainable by the necessarily causal and unidirectional (from source to receiver) nature of information transmission in Shannon's schema.* The combining of objective and epistemic notions of the possibilities at the source is necessary for Dretske's semantic theory of information to work.