Dretske

Knwowledge and the flow of information

Dr Greig Russell

14 September, 2019

Contents

1 Chapter 2 - Communication and Information

There is an obvious distinction between the message within the signal and the signal itself. Yet one can tell something of the message from the signal, With the analogy being made that a bucket can not tell you what is in the bucket but does tell you what can be in the bucket.

Dretske (p.41) makes the point that communication theory does not address what specific information the message holds, rather the focus is on the max amount of information the message could hold. In this he quotes Shannon. *note to self some authors disagree with Shannon's interpretation of his own theory.*

Communication theory is focused on the average amount of information contained in a message (entropy). Not the amount of information contained in a particular message.

Dretske (p.42) makes the distinction between semantic meaning and information. The former is derived from the latter, and although often one and the same, not necessarily so. Dretske uses the example of bridge bidding convention, where the semantic meaning is declared but not what information it contains in this particular instance is not. He makes the point that the information contained in a message can exceed the literal semantic meaning of a message.

Information is that commodity capable of yielding knowledge and what information contained in a signal is what we can learn from it (p.44)

The but to that is for Dretske argues that *false* information, as it does not yield verifiable knowledge, is not a true member of the set of true knowledge, *note to self, surely this can only be determined either by hindsight or by reference to prior to learning - very loosely defined.*

1.0.1 Key argument (p.46)

1. Information is a semantic idea.
2. Semantics is the study of meaning.
3. Communication is not a theory of meaning.

Hence

4. Communication theory is not a theory of information

Dretske's counter That is a theory does provide a satisfactory account of meanig, it does not provide a satisfactory account of any semantic concept.

For Dretske, while there is some logic for a communication engineer trying to right size a communication channel to talk about about the entropy of a system, there is no such concept of average information in two messages on two different topics.

He makes the argument for a pre-arranged system where a 1 is transmitted if the chess piece is on square 2, but 0 otherwise. Communication theory says this is a one bit message, but if it is on square 2 it contains 6 bits of information. Of course, if it is not on square 2 then it contains only 0.05 bits of information. So on average our message contains only 0.095 bits of information.

For a communication engineer this is a highly inefficient system, and a better solution is to give up ever knowing where the chess piece ever actually is but reducing the possibilities of where it might be. If you are only interested in knowing when the chess piece is on square 2, then you are actually much worse off with the more efficent communication system as you get less information all of the time in the communication, even if the channel capacity is greater.

Dretske, hwever, argues that Communication theory can be used to estimate the quantity of information contained in a given signal (p. 51)

He argues that key are the surprisal of a given message or the amount of information contained in a given message.

$$I(S_i) = -log_2 \frac{1}{p(s_a)} \tag{1}$$

and the amount of information contained in r_{a} about s_{a} is given by;

$$I_s(r_a) = I(s_a) - E(r_a) \tag{2}$$

where

$$E(r_a) = -\sum P(s_i|r_a).log_2 P(s_i|r_a) \tag{3}$$

Dretske emphasises that this is a non-standard use of the Communication theory to determine the information contained in a given message as opposed to the entropy (average amount of information) that is possible for messages using a particular channel. **Dretske (p.53-54) argues that to determine the absolute amount of information in a given message is challenging, given the need to determine a infinite set of possible circumstances. Instead he argues we can more practically determine the comparitive amount of information between different messages**.

Key is the need to know that r_{a} carries as much information as was contained in s_{a} ie. that equivication was zero. In comparison, if equivication is not zero, it contains r_{a} contains less information than was in s_{a}. It does not matter that we can not define or even know that equivication has occurred, for Dretske information is only about whether the information received was the same or less than the information sent. Belief of the recipent is irrelevant to whether equivication actually occurred.

Note to self p.56 carries an interesting link to Hume and humans being trapped in a circle of one's own perceptions.

1.1 Xerox principle of information transmission.

1. If message A carries the information contained in message B
2. and message B carries the information contained in message C Hence then message A carries the same information as contained in message C

For Dretske, the Xerox principle is fundamental to information flow. He makes the point (p.60) that message C may contain extra noise, so cumulative losses due to equivication can occur at each step but message A could still contain the same total information. Indeed even if the message is not complete any correlation between message C and message A would be enough so that information is still preserved. A message does not have to tell us everything to tell us something (p.61)

The message however muct have sufficient capacity to carry the information. So loss of bits (ie. capacity) degrades the message absolutely while loss of content may not degrade the information.

2 Chapter 3: A semantic theory of information

2.1 Core theorem

For a signal to carry the information that s is F, it must be the case that;

1. The signal carries as much information about s as as would be generated by s's being F.
2. s is F
3. The quantity of information the signal carries about s (or includes) that quantity generated by s's being F (and not s being G).

Where 3) is meant to include that the signal has to have the capacity of carrying that s is F and that it actually does carry that information.

So that the definition of information content is:

A signal r, carrys the information that s is F is equivalent to saying that the conditional probability of s's being F, given r (and k) is 1, (but, given k alone is less than 1).

Where k is all prior knowledge about the possibilities that exist at the source. The example used was if a signal is known to be red or blue, and the message says it is not blue, then you know it is red but only if you know red was the only other choice, if you thought green was a valid option then you would none the wiser. k also descibes the relativisim of informational content. As each receiver of a signal may have different prior knowledge ie differing values for k, so that while one receiver can get a conditional probability of 1 another may fall short.

Note that the informational content of a signal is being expressed in the form "s is F" refers to some item at source (p. 66). It is also meant as a de re sense, namely it is about the fact that this s is F, not that all or another s is F (the de dicto sense). For Dretske, a signal's de re iformation is determined by 1) the individual s about which the signal carries information and 2) The information (determined by the open sentence "... is F" it carries about that individual.

Dretske raises what might be objecton the definition of informational content by discussing what happens in a situation such that;

P(A|r) = .07

P(B|r) = .80

P(C|r) = .07

P(D|r) = .06

So the best that can be said of r is that it is probably B. Dretske argues this is not a bug, rather it is a feature as this is how we usually express the information we receive.

Note to self but this objection that s is B does not have a conditional probability of 1. So I might ageee it is the usage, how does it fulfil Dretske's definition for information content. I think this might be covered by the discussion below, not that Drekste joined the dots.

This is not elaborated upon, instead we move on to the idea of nested information, where;

The information that t is G is entirely nested within s being F, so that s is F also carries the information that t is G. Further a signal may contain many units of information, that have different meanings, either overtly or by being nested wtihin another piece of information. The latter is analytically nested meaning within information. Nomically (following the laws of nature) the statement that s is F, even if this obeys a law of ature does not automatically imply that t is G as in this de re instance it might not (as yet) because the conditional probability is not 1.

Dretske digresses on p.73 to discuss correlation and to distinguish that given correlation does not imply causation, namely there may not be a unity conditional probability that if F is true, therefore G must also be true in either a de dicto or de re sense, that is H might be nested within F but if it is present then G is nested within it, but not because F is true.

This is intentionality is that s is F can not be substitued by s is H without changing the meaning. Underpinning information intentionality is nomic regularities. Namely that s is F is a manifestation of a law of nature.