Bar-Hillel and Carnap's Account of Semantic Information

February 19, 2010

Bar-Hillel and Carnap's account of semantic information falls under the probabilistic approach to information. [1] Their idea was to measure the quantity of semantic information associated with a statement within a given language in terms of the set of possible worlds it rules out and an *a priori* logical probability space over those worlds. The general idea is based on the *Inverse Relationship Principle*, according to which the amount of information associated with a proposition is inversely proportional to the probability associated with that proposition. Using some probability measure *pr*, two measures of information are provided, *cont* and *inf*, such that:

$$cont(A) =_{df} 1 - pr(A)$$

and

$$inf(A) =_{df} -log_2(pr(A))$$

pr here refers to logical probability, as developed by Carnap. [2] It is a measure on the space of possible worlds, which can be represented in a given language using what are called state descriptions. Here is a basic example to illustrate these concepts. The scenario, represented using the familiar propositional logic, concerns the weather conditions for a day in the coming week. The three properties being considered are whether the day will be hot (h), rainy (r) and windy (w). Since there are three variable propositions involved, there are eight logically possible ways things can be, there are eight possible worlds. Here is a truth table depicting this:

	h	r	w
1	T	T	T
2	T	T	F
3	T	F	T
4	T	F	F
5	F	T	T
6	F	T	F
7	F	F	T
8	F	F	F

Now, a state description is a conjunction of atomic statements, consisting of each atomic statement or its negation, but never both. Each of these eight possible worlds can be represented with a state description. For example, the state description for state 1 is $h \wedge r \wedge w$ and the state description for 8 is $\neg h \wedge \neg r \wedge \neg w$. For the sake of example, the probability distribution I use here will be a simple uniform one. That is, since there are eight possible worlds, each world has a probability of $\frac{1}{8}$ that it will obtain or be actual.[1]

Now we can perform some calculations. Each world or state description has a probability of $\frac{1}{8}$, so the probability of it being hot, rainy and windy $(h \wedge r \wedge w)$ is $\frac{1}{8}$. Given this

$$cont(h \wedge r \wedge w) = 1 - \tfrac{1}{8} = 0.875$$

and

$$inf(h \wedge r \wedge w) = -log_2(\tfrac{1}{8}) = 3$$

State descriptions are highly informative; a statement which narrows things down to just one of the possible worlds contains the most information. By contrast, if a statement tells us that it will be hot and rainy but tells us nothing about whether or not it will be windy $(h \wedge r)$, it would be less informative, as the calculations agree. $(h \wedge r)$ is true in two of the possible states, so:

$$cont(h \wedge r) = 1 - \tfrac{2}{8} = 0.75$$

and

$$inf(h \wedge r) = -log_2(\tfrac{2}{8}) = 2$$

Further still, a statement which only says that it will be hot or windy $(h \vee w)$ is contains relatively little information. $(h \wedge r)$ is true in all but two of the possible states, so:

$$cont(h \vee w) = 1 - \tfrac{6}{8} = 0.25$$

and

$$inf(h \vee w) = -log_2(\tfrac{6}{8}) = 0.415$$

Finally, a tautological statement, which gives us no reduction of the original 8 possibilities, contains no information:

$$cont(h \vee \neg h) = 1 - 1 = 0$$

[1] Carnap discusses various types of more sophisticated logical probability measures. See http://plato.stanford.edu/entries/probability-interpret/#LogPro for an example.

and

$$inf(h \vee \neg h) = -log_2(1) = 0$$

It is also worth noting that according to these measures, contradictions contain the most information:

$$cont(h \wedge \neg h) = 1 - 0 = 1$$

and

$$inf(h \wedge \neg h) = -log_2(0) = \infty$$

The counter intuitiveness of associating maximal informativeness with necessarily false contradictions will be discussed shortly.

Further to these core measurement formulas, Bar-Hillel and Carnap added formulas to measure conditional informativeness, that is the informativeness of a statement already given another statement.

Given the following definitions:

$$cont(A|B) =_{df} cont(A \wedge B) - cont(B)$$

and

$$inf(A|B) =_{df} inf(A \wedge B) - inf(B)$$

it so happens that $cont(A|B) = cont(B \supset A)$ and $inf(A|B) = -log_2(pr(A|B))$.

Does the Bar-Hillel/Carnap account provide an acceptable measure of semantic information? To a certain extent its results accord with our intuitions and are expected. For example, the statement $A \wedge B$ is simply more informative than the statement $A \vee B$. From this perspective, the more states or possible worlds a statement rules out, the more informative it is. Inevitably though, this approach suffers from some problematic results. The most prominent concerns its wayward calculative assignment of maximal informativeness to contradictions, a result which indicates the discord in basing a quantitative measure of our intuitive notions of information and informativeness on logical content. Bar-Hillel and Carnap provide the following commentary on the matter:

> It might perhaps, at first, seem strange that a self-contradictory sentence, hence one which no ideal receiver would accept, is regarded as carrying with it the most inclusive information. It should, however, be emphasised that semantic information is here not meant as implying truth. A false sentence which happens to say much is thereby highly informative in our sense. Whether the information it carries is true or false, scientifically valuable or not, and so forth, does not concern us. A self-contradictory sentence asserts too much; it is too informative to be true. [1, pg. 229]

So the more a sentence says the more informative it is. But when does a sentence 'say much'? We intuitively judge the statement $A \wedge B$ to say more than the statement $A \vee B$ not just because it is less probable or excludes more possible worlds, but also because it does a better, more detailed job of describing how presumably things are. $A \wedge \neg A$ on the other hand does not at all do a good job of describing how things presumably are. It does not discriminate and selectively narrow down on the actual state of affairs (unless a contradiction does actually occur). These considerations suggest that perhaps semantic information is not just about content and should be meant as implying truth.

Also, another issue related to contradictions concerns conditional information formulas. For example, it is the case that $cont(A|\bot) = 0$:

Proof. □

$$
\begin{aligned}
cont(A|\bot) &= cont(\bot \supset A) \\
&= cont(\neg\bot \vee A) \\
&= cont(\top) \\
&= 0
\end{aligned}
$$

Either some contradiction tolerance (this is another example of an application for paraconsistent logic) or an alternative method of conditional information calculation would be welcome here. Hypothetically speaking, if, given the information of a somehow true contradiction, then 'all hell breaks loose' and statements will arguably no longer be informative. But given false contradictory information (the natural alethic value of all contradictory information), surely other statements in general should not lose their conditional informativeness. A paradigmatic example is a database. If the database does contain contradictory data, this should not mean that other data cease to be informative, for they can still very well be informative in the sense of saying something veridical or being useful, independently of the false contradictory data.

References

[1] Yehoshua Bar-Hillel and Rudolf Carnap, 'Semantic Information'. The British Journal for the Philosophy of Science, Vol. 4. No. 14, 1953, 147157.

[2] Carnap, R., 1950, Logical Foundations of Probability, Chicago: University of Chicago Press.