L&PS – Logic & Philosophy of Science
Vol. VI, No. 1, 2008, pp. 25-33

Can Logical Probability Be Viewed as a Measure of Degrees of Partial Entailment?

Alberto Mura
Department of Theories and Research on Cultural Systems
University of Sassari
e-mail: ammura@uniss.it

ABSTRACT. A new account of *partial entailment* is developed. Two meanings of the term 'partial entailment' are distinguished, which generalise two distinct aspects of deductive entailment. By one meaning, a set of propositions A entails a proposition q if, supposed that all the elements of A are true, q must necessarily be true as well. By the other meaning, A entails q inasmuch the informative content of q is encapsulated in the informative content of A: q repeats a part of what the elements of A, taken together, convey. It is shown that while the two ideas are coextensive with respect to deductive inferences, they have not a common proper *explicatum* with respect to the notion of partial entailment. It is argued that epistemic inductive probability is adequate as an *explicatum* of partial entailment with respect to the first meaning while it is at odds with the second one. A new *explicatum* of the latter is proposed and developed in detail. It is shown that it does not satisfy the axioms of probability.

KEYWORDS: partial entailment, partial implication, logical probability, logical independence, logical content, semantic information.

Published by L&PS – Logic and Philosophy of Science
http://www.units.it/~episteme/ – ISSN: 1826-1043

1. Introduction

In this paper I shall examine the idea of *partial entailment* to investigate to what extent a probability sentence can represent a degree of partial entailment. The idea of probability as a measure of partial entailment goes back to Keynes ([1921] 2004) and Carnap ([1950] 1962). I shall distinguish between two senses of this idea and I shall vindicate the idea of probability as partial entailment with respect to one of them, although it appears to be essentially epistemic in character, contrary to Keynes' and Carnap's claims. My analysis will make essential use of the idea of deductive probability as it has been characterised in Mura (2006).

When we say that a set A of propositions entails a proposition q, there are *two* meanings of such an assertion. In a first meaning (which I shall call 'meaning I') A entails q because if the elements of A are true, then q is also necessarily true. In a second meaning (which I shall call 'meaning II'), A entails q inasmuch as the informative content of q is contained in the informative content of A: q repeats a part of the content of A. Meaning I and meaning II of the word 'entailment', although they are coextensive with respect to deductive inferences, need not admit a common *explicatum* with respect to the idea of *partial entailment*. In fact, as I shall argue, the two concepts need two different *explicata*. In order to show this, I shall focus on the differences between the two concepts of entailment.

Applied entailment in meaning I is a relationship between the *assertability* of the premises and the *assertability* of the conclusion: given that the premises are assertable, the conclusion too is assertable. Surely, epistemic probability *is* a generalisation of entailment with respect to this assertability relation. A degree of probability $\mathbf{P}(h \mid e)$ can indeed be viewed as a measure of the assertability of h given the full assertability of e.[1]

Consider now meaning II. I shall write $e \xrightarrow{r} h$ for 'e entails h to the degree r' according to meaning II. Can we say that $\mathbf{P}(h \mid e) = r$ means that a proportion r of the content of h is contained in e, so that we can properly write $e \xrightarrow{r} h$? Can we say – to put it another way – that when $\mathbf{P}(h \mid e) = r$, h has a degree of assertability r because just a proportion r of h repeats what is already contained in e? Surely not. In fact, if e is analytically true, *no part* of the informative content of h is

[1] However, it should be kept in mind that the following property (called *monotonicity*), satisfied by total entailment, couldn't be approximated by partial entailment: if $A \vdash p$ then, for every B such that $A \subseteq B$, it holds that $B \vdash p$. So the assertability of the conclusion of an argument of partial entailment holds only under the proviso that the premises contain all the relevant available evidence.

already contained in e in spite of the fact that in such a case $\mathbf{P}(h \mid e) = r$ need not be 0 while surely $e \overset{0}{\rightarrow} h$ holds.

However, if we accept the standard view that the content of a proposition is just the class of its logical consequences, it can be shown that $\mathbf{P}(h \mid e)$ can be viewed to the degree at which $\neg h$ entails $\neg e$. This can be argued as follows. If the content of a proposition is given by the set of its logical consequences, we should be able to represent the amount of content by an additive measure across those propositions that share no common factual logical consequences. A measure that satisfies this constraint is, as may easily be verified, the function $Ct(p) = 1 - \mathbf{P}(p)$. Moreover, the set of logical consequences common to e and h is just the set of consequences of $e \vee h$. But in this case a proper measure of the proportion of the content of h contained in e should be given by the formula:

$$\frac{Ct(e \vee h)}{Ct(h)} = \frac{1 - \mathbf{P}(e \vee h)}{1 - \mathbf{P}(h)} = \frac{\mathbf{P}(\neg e \wedge \neg h)}{\mathbf{P}(\neg h)} = \mathbf{P}(\neg e \mid \neg h).$$ [2]

To explain the unconscious shift from $\mathbf{P}(h \mid e)$ to $\mathbf{P}(\neg h \mid \neg e)$ it is helpful to observe that in deductive logic e entails h if and only if $\neg h$ entails $\neg e$. However, from this equivalence it does not follow that the degree to which e entails h must coincide with the degree to which $\neg h$ entails $\neg e$. So, if $\mathbf{P}(\neg e \mid \neg h)$ is adequate as *explicatum* of the partial entailment of e on h, there is no reason to believe that $\mathbf{P}(h \mid e)$ too is adequate in the same respect.

In spite of this, some connection between the entailment of e over h and of $\neg h$ over $\neg e$ should be preserved in the context of partial entailment. Surely, a good *explicatum* of partial entailment should satisfy at least the following requirement: if for some $r > 0$ $e \overset{r}{\rightarrow} h$ then for some $s > 0$ $\neg h \overset{s}{\rightarrow} \neg e$. This requirement may be argued as follows. It would be unreasonable to admit that e totally entails h if and only if $\neg h$ totally entails $\neg e$, while it can happen that e partially entails h (maybe to a very high degree) and $\neg h$ has no deductive connection with $\neg e$. In fact, the idea of partial entailment as a generalisation of total entailment is based upon the supposition that total entailment is a limiting case of partial entailment and, as such, it may be approximated by increasing the degree of partial entailment. As a result, when both r and s tend to 1, $e \overset{r}{\rightarrow} h$ should tend to $\neg h \overset{s}{\rightarrow} \neg e$. Unfortunately, the preceding account does not satisfy this requirement. In fact, $\mathbf{P}(\neg t \mid \neg h) = 0$ independently of $\mathbf{P}(h \mid t)$ (where 't' denotes a tautology) which may well be greater than 0.

[2] See Miller and Popper (1986).

2. Content and logical consequences

The preceding approach rests upon the basic dogma that the content of a proposition coincides with the set of its logical consequences. It should be recognised that the theory by which the content of a proposition is given by the set of its logical consequences sounds *prima facie* very natural from a logical point of view. But, taken seriously, this view carries some consequences that are, in my view, completely unacceptable (see also Gemes 1994).

We have already seen in the preceding section a first flaw in that theory (namely that $e \xrightarrow{r} h$ does not tend to $\neg h \xrightarrow{s} \neg e$ when s approaches 1). But other consequences are even more unacceptable. For example, this view entails that two propositions always share a common content, except when they are logically disjunct, i.e. when their disjunction is a logical truth. In fact, only in that case should two propositions have no factual logical consequence in common. So 'Napoleon was defeated at Waterloo' and 'The population of Chicago amounted in 1990 to 2,783,726 inhabitants' would have a common content. This appears to be completely counterintuitive.

Surely, all agree that two propositions have no common content when they are – in some sense – *logically independent*. So at the heart of our concern there is the idea of logical independence. Before considering in this respect the various forms of independence I shall open a parenthetic digression on *amount of content*. The reason for this depends on the fact that the notion of independence which we are looking for should satisfy the following condition referring to amount of content:

(AC) *Additivity Condition*. There exists at least one *absolute content measure function* additive over logical independent sentences.

3. On amount of content

To begin with, I observe that amount of content may be, just like probability, absolute or relative. We can write $Ct(p \mid q)$ to denote the amount of content of p relatively of q (or given q), where q is assumed to be a consistent sentence. $Ct(p \mid q)$ measures how much information is contained in p but not in q (or, to put it another way, the amount of the content of p that *goes beyond* q). And like absolute probability, absolute amount of content may be considered as amount of content relative to a tautology.

Clearly, $Ct(p \mid q)$ should satisfy the following axioms:

1. $Ct(p \mid q) \geq 0$;
2. $Ct(t \mid q) = 0$ (where 't' is a tautology);
3. If $q \wedge p \vdash r$ then $Ct(p \mid q) \geq Ct(r \mid q)$;
4. $Ct(p \wedge r \mid q) = Ct(p \mid q) + Ct(r \mid p \wedge q)$.

Conditions 1-4 are very general and compatible with a wide class of functions. This class may be split into two subclasses, depending on two different ways to conceive relative content-independence, that is the relation which is fulfilled among three sentences p, q and e when p and q share no common content that goes beyond e. Relative content-independence may be formally expressed by the relation $Ct(p \mid q \wedge e) = Ct(p \mid e)$. If we would like axiomatically to characterise a function Ct satisfying AC, the following axiom should be added:

5a. If p and q are logically independent then $Ct(p \mid q \wedge e) = Ct(p \mid e)$.

The second class is based on the idea that p and q share no common content that goes beyond e when p and q are *probabilistically independent* in the presence of e. In such a case we have to add to the axioms 1-4 the following condition instead of 5a:

5b. If $\mathbf{P}(p \mid q \wedge e) = \mathbf{P}(p \mid e)$ then $Ct(p \mid q \wedge e) = Ct(p \mid e)$.

Axiom 5b is equivalent to 5a if in 5a we adopt separability as logical independence[3] and in 5b we adopt deductive probability as probability. However, except for this very special case, 5a and 5b are two very different ways of characterising content-independence. In fact, 5a aims to capture a *deductive* idea of absence of common content, while 5b aims to capture a more general idea. Indeed, by 5b, two sentences are considered to be without common content when from one sentence no inference about the other can be drawn (and vice versa), including *inductive* (i.e. *ampliative*) inferences.

4. Different kinds of logical independence

For every kind of independence there is a corresponding version of condition 5a. If we give an account of deductive content in terms of logical conse-

[3] For definitions of the various kinds of logical independence discussed in this paper see Mura (2006).

quences, and therefore conceive absence of common content as absence of factual logical consequences, the relation

$$Ct(p \mid q \wedge e) = Ct(p \mid e)$$

should be satisfied when p and q are, in the presence of e, maximally independent (i.e. logically disjunct). So, in such a case, the following constraint should be added to axioms 1-4:

5a.1 If p and q are maximally independent in presence of e (i.e. if $e \vdash p \vee q$) then $Ct(p \mid q \wedge e) = Ct(p \mid e)$.

It can be shown that for every probability function **P**, the function 1 – **P** satisfies the axioms 1-5a.1, provided we put

$$Ct(p \mid q) = Ct(q \supset p \mid t),$$

where ‘*t*’ denotes a tautology.

Notice that by accepting maximal independence as an *explicatum* for content-independence we are not considering the function **P** = 1 – *Ct* as ‘the’ probability. In other terms, there is no immediate reason for choosing as proper logical probability function that probability function which is associated to the proper content function. That constraint, if assumed, must be considered as additional. The celebrated Popperian argument, by which probability and content are opposite, requires this additional assumption. I shall not attempt to discuss the problem of its soundness here.

Maximal independence is ruled out by the considerations we have made above. The same can be said about atomic and truth-functional independence. So the corresponding version of 5a also has to be rejected.

The list of candidates is now restricted to mutual independence, complete independence and separability. The corresponding versions of axiom 5a are the following:

5a.2 If p and q are mutually independent then $Ct(p \mid q \wedge e) = Ct(p \mid e)$;
5a.3 If p and q are completely independent then $Ct(p \mid q \wedge e) = Ct(p \mid e)$;
5a.4 If p and q are separable then $Ct(p \mid q \wedge e) = Ct(p \mid e)$.

Since separability entails complete independence and complete independence entails mutual independence, 5a.2 entails 5a.3 and 5a.3 entails 5a.4.

Unfortunately, 5a.3 and, *a fortiori*, 5a.2, cannot be satisfied except in languages that have less than three completely independent sentences of non-null content.[4]

So the only serious available candidate for axiom 5a is 5a.4. It can be shown that for every Boolean algebra there exists (up to a scalar factor) one and only one *Ct*-function satisfying the axioms 1-4 and 5a.4. More exactly, the functions satisfying the axioms 1-4 and 5a.4 are given by $\log_a \mathbf{P}_\mathrm{D} (p \mid q)$ with $0 < a < 1$ (see Mura 1992). Although the choice of *a* is inessential, it is convenient to assume $a = 1/2$, so that sentences of deductive probability 1/2 turn out to convey 1 information unit.

These results show not only that the additive condition may be consistently satisfied but also (*a*) that it is a condition sufficient to give a *single* measure of amount of content and (*b*) that the idea of amount of content relies on the idea of deductive probability.

5. Partial entailment (sense II) explicated

Once we have at hands an *explicatum* for the idea of amount of content, it is easy to achieve an *explicatum* for partial entailment with respect to sense II. In what follows I shall assume that the content-function Ct_D coincides with the function $\log_a \mathbf{P}_\mathrm{D}$ $(0 < a < 1)$.

I begin by proposing the following desiderata:

---

[4] This is shown by the following considerations. Let *p*, *q* and *r* be three completely independent sentences having non-zero content. It soon follows from the definition of complete independence that $p \wedge q$ and $p \wedge r$ are also completely independent. Therefore 5a.3 entails

$$Ct(p \wedge q \mid p \wedge r) = Ct(p \wedge q \mid t).$$

On the other hand, axiom 3 entails

$$Ct(p \wedge q \mid p \wedge r) = Ct(q \mid p \wedge r).$$

(Since $p \wedge q$ and q are logically equivalent in presence of $p \wedge r$). Moreover, also q and $p \wedge r$ are completely independent, so that

$$Ct(q \mid p \wedge r) = Ct(q \mid t).$$

It follows that

$$Ct(p \wedge q \mid t) = Ct(q \mid t).$$

Finally from this and axiom 4 follows that $Ct(p \mid t) = 0$, contradicting the hypothesis that *p* is having non-zero absolute content.

(i) the degree to which e entails h should express the proportion of the common content of h and e with respect to the content of h;
(ii) the common content (if any) of h and e is given by the difference $Ct_{\mathrm{D}}(h) - Ct_{\mathrm{D}}(h \mid e)$ (or equivalently by $Ct_{\mathrm{D}}(e) - Ct_{\mathrm{D}}(e \mid h)$);[5]
(iii) the more e entails h the less e entails $\neg h$ (and vice versa), so that a suitable *explicatum* should be such that if $e \xrightarrow{r} h$ then $e \xrightarrow{-r} \neg h$;
(iv) if e and h are factual separable propositions, then $e \xrightarrow{0} h$.

These requirements are met by the following definition for the notion of *degree of partial entailment*:

$$\text{DPE} \qquad e \xrightarrow{r} h =_{df} \frac{Ct_{\mathrm{D}}(h) \dot{-} Ct_{\mathrm{D}}(h \mid e)}{Ct_{\mathrm{D}}(h)} - \frac{Ct_{\mathrm{D}}(\neg h) \dot{-} Ct_{\mathrm{D}}(\neg h \mid e)}{Ct_{\mathrm{D}}(\neg h)} = r,$$

where $x \dot{-} y$ is the modified difference that equals $x - y$ if $x > y$ and equals 0 otherwise. The cases in which $Ct_{\mathrm{D}}(h)$ or $Ct_{\mathrm{D}}(\neg h)$ is 0 may be dealt by passage to the limit as $e \xrightarrow{0/0} h$.

It can be easily seen that from DPE it follows that if, for some $r > 0$, $e \xrightarrow{r} h$, then there exists $s > 0$ such that $\neg h \xrightarrow{s} \neg e$, so avoiding the flaw that we found in the approach to partial entailment in terms of logical consequences.

REFERENCES

CARNAP, Rudolf ([1950] 1962): *Logical Foundations of Probability*, Chicago (Ill.): Chicago University Press, 2nd edition.

GEMES, Ken (1994): “A New Theory of Content I: Basic Content”, *Journal of Philosophical Logic*, 23, pp. 595-620.

KEYNES, John Maynard ([1921] 2004): *A Treatise on Probability*, Mineola (N.Y.): Dover Publications.

MILLER, David W., and POPPER, Karl Raimund (1986): “Deductive Dependence”, in *Actes IV Congrés Catalá de logica*, Barcelona: Universitat Politécnica de Barcelona and Universitat de Barcelona.

MOORE, Eliakim H. (1910): “Introduction to a Form of General Analysis”, in *The New Haven Mathematical Colloquium*, New Haven (Conn.): Yale University Press.

[5] The identity $Ct(h) - Ct(h \mid e) = Ct(e) - Ct(e \mid h)$ (which is an obvious desideratum of any *explicatum* for amount of content) is an almost immediate consequence of axiom 4.

MURA, Alberto (1992): *La sfida scettica. Il problema logico dell'induzione*, Pisa: ETS.

——— (2006): "Deductive Probability, Physical Probability, and Partial Entailment", in Mario Alai and Gino Tarozzi (eds.), *Karl Popper Philosopher of Science*, Soveria Mannelli (Cz.): Rubbettino, pp. 181-202.