(Pure) Logic out of Probability

TON SALES

Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya
Barcelona, Catalonia 08028 (Spain)
e-mail: sales@lsi.upc.es

Abstract: Today, Logic and Probability are mostly seen as independent fields with a separate history and set of foundations. Against this dominating perception, only a very few people (Laplace, Boole, Peirce) have suspected there was some affinity or relation between them. The truth is they have a considerable common ground which underlies the historical foundation of both disciplines and, in this century, has prompted notable thinkers as Reichenbach [14], Carnap [2] [3] or Popper [12] [13] (and Gaifman [5], Scott & Krauss [21], Fenstad [4], Miller [10] [11], David Lewis [9], Stalnaker [22], Hintikka [7] or Suppes [23]) to consider connection-building treatments of Logic and Probability as desirable. Indeed such a line of thinking can be pursued (this author, for one, attempted it in [15-19]). In so doing, one straightforwardly obtains a logic based on —as the simple unifying concept— an *additive non-functional truth valuation* which, though technically indistinguishable from (axiomatic) Probability, can however be totally "decontaminated" from parasitical probabilistic interpretations (such as the usual readings of "event", "probability" or "conditioning") and be given instead a strictly logical reading and justification (in terms of "sentence", "truth" or "relativity"). Once some deeply-ingrained reading habits are overcome, the required concepts and formulas flow easily, and the resulting assertion-based *sentential calculus* becomes a very natural extension of ordinary two-valued reasoning. Furthermore, in the process we get: (a) intuitive geometrical and information-related interpretations of the concepts, (b) a simple theoretical explanation for some poorly justified formulas (intermittently advanced by various authors, some mentioned above), and (c) a semantics —and a proof theory— for general assertions that is unproblematically derived and also fully consistent with empirical or *ad hoc* approximate-reasoning "Bayesian" formulas found by Artificial Intelligence researchers.

Keywords: Sentential logic, boolean algebra, logical semantics, probabilistic semantics, probability logic, many-valued logics, supervaluations, uncertainty, rational belief, proof theory.

The present paper is a summary of [18] and [19] as delivered orally (in Italian) in the Naples Workshop on Nov. 2nd, 1995, under the announced title *Between Logic and Probability*. (That it should be delivered in Naples is just appropriate: the ideas for [18] came effortlessly out of daily uphill trips —shuttling for Arco Felice— on the Mergellina cable train in the mild December mornings of 1981.)

INTRODUCTION

When we argue, we do not always fully assert what we say. We often make half-hearted assertions of sentences we are not sure about, or we even use as assertions sentences we hardly believe to be the case. And yet we proceed by reasoning from such weak premises. If we admit we do, and want to treat this inside Logic, we apparently need to qualify assertions, or rather quantify their strength, and try to follow and control what effects weak assertions may have in the reasoning process, whether and how they affect its logical validity and how we can tell the strength of the conclusion. All this seems to be indeed a proper *logical* subject that, however, very few logicians (Reichenbach [14], Carnap [2] and Popper [12] are the conspicuous exceptions) did ever set out to confront.

Just to fix ideas, take as a first example a reasoning in Physics wherein one of the premises is the positive result of an experiment. Suppose we may even quantify the error ε of the experiment —meaning that the truth of the assertion 'result is positive' is, say, "$1-\varepsilon$". We then perform the formal reasoning —assumed logically valid— and obtain the conclusion. We now want to know what confidence we may have in it, given ε. That, everybody would agree, is a legitimate logician's concern. (It is what we develop in our *proof theory* sketched below.)

For a second example, take the well-known sorites about bald men: "If a man with i hairs is not bald then a man with $i-1$ hairs is still not bald. Suppose a man has n hairs. Therefore, a man with 0 hairs is still not bald". Formally:

$$\begin{array}{l} A_i \rightarrow A_{i-1} \quad (i:1,\ldots,n) \\ A_n \\ \hline A_0 \end{array}$$

This is a paradox because the reasoning is formally correct (it consists of merely n applications of the Modus Ponens rule), the $n+1$ premises are deemed flawless, but the conclusion is outright false (or, more precisely, a *contradictio in terminis*). Usually, it is the length of the argument that is put to blame. There is, however, a more concrete and satisfactory answer we can offer. The n premises $A_i \rightarrow A_{i-1}$ cannot obviously be asserted with the same assurance whatever the index value. That's why the argument fails: for low values of i the premises simply cannot be asserted, even if the rest can, so we can *never* have all premises asserted, and the reasoning is formally valid but vacuously so. As it will be later seen, we propose instead to provide every premise A with a value $v(A)$ in [0,1] —computed in an unspecified way (statistically, by opinion survey, or whatever)— with the unique requirement that a zero value means that the premise is to be taken as false, 1 means a true —and therefore fully assertable— premise, and $v(A) = 1-\varepsilon$ ($\varepsilon > 0$) means that we can assert A but with some apprehension or risk ε. Obviously, the value $v(A_i \rightarrow A_{i-1})$ decreases with i, so that when i is n (or even, say, around $n/2$ or $n/3$) it is 1 or very near 1, but when i approaches, say, $n/10$ —and surely when it becomes zero— the value of $A_i \rightarrow A_{i-1}$ (= the predisposition we have to assert it —or the willingness to assume the risk) comes down to an exceedingly low number. According to our proof theory (introduced below), the conclusion A_0 has the same truth value, at best, as that lowest of numbers (and, thus, the reasoner would be willing to assert the conclusion just no more than he or she would willing to assert $A_1 \rightarrow A_0$).

That we normally need to qualify/quantify what we assert is implied by all mainstream logicians when they treat assertions as true, believed or merely hypothesized sentences, leaving it to the reasoner who uses them to count actively on these factors, as part of the reasoning process itself. After all, following Tarski's introduction of the (T) schema, it is the user who is supposed to be, through the full mastery of the (T) translation process, the sole referee who can validate the truth of the assertions, their aptness to describe an actual situation, their strength (as beliefs) or the relevance of their use in the current logical context. As the reasoner should normally be capable, when asked to do so, to assign consistently relative strengths to the assertions used, it is just natural to assume as we do, first, that assertions have each an associated, measurable strength, and that, second, this strength has significant and measurable effects on the truth of the sentences, the validity of the conclusion and the soundness of the reasoning. That this strength is computable and that it behaves as an *additive valuation* is accounted for by the condition generally assumed to apply: the reasoner behaves in a rational fashion, as classically characterized by Ramsey and De Finetti (though non-rational behavior can be also modelized [17]).

1. "TRUTH" VALUATIONS OVER SENTENCES

We only need to assume that we have a set $\mathcal{L}$ of sentences that form a Boolean algebra (with respect to the three $\wedge$, $\vee$ and $\neg$ connectives and two special sentences $\perp$ and $\top$). We then have a complete Proof Theory by identifying the " $\vdash$ " order defined by the Boolean algebra with the deductive consequence relation. So the algebra of sentences we started with automatically becomes the Lindenbaum-Tarski algebra of all sentences modulo the interderivability relation " $\dashv\vdash$ " given by the $\vdash$ order (i.e. $A \dashv\vdash B$ iff $A = B$). We then only need to assume, further, that all sentences are valued in $[0, 1]$, which can be done in the standard way of a *normalized measure* $v : \mathcal{L} \longrightarrow [0, 1] : A \mapsto [\![A]\!]$, by just requiring that $\top$ gets a value of 1 and that the valuation v is *additive* (i.e. $[\![A \vee B]\!] = [\![A]\!] + [\![B]\!] - [\![A \wedge B]\!]$). We then have also the whole Model Theory of Sentential Logic.

This "truth" valuation is no more nor less than a (finitely aditive) *probability* in all technical senses, but we definitely want to avoid the usual probabilistic connotations and not be distracted away from pure, plain Logic. Thus, for instance, A will be plainly a *sentence* (in a language $\mathcal{L}$), not an *event* (in a sample space Ω), and $[\![A]\!] = 1$, 0 and 1/2 will respectively mean just *truth* (or, more precisely, that "A is taken as truth"), *falsity* and *undecided belief* when expressly *asserting* the sentence A, clearly not the same as (respectively) probabilistic "certainty", zero-probability or balanced odds when evaluating the *uncertain* outcome of A. The presumed Booleanity of the sentences is either assumed (which is undemanding) or it just arises naturally from a "minimal algebra" of sentences as suggested by Popper [12] or from provably equivalent simple assumptions (e.g. by Cox). Also, the additive character of the valuation amounts to having a 'rational' (Ramsey) or 'coherent' (De Finetti) *belief*, a concept so akin to 'strength of assertion' in Logic as to be all but exchangeable.

From the Booleanity of $\mathcal{L}$ and the above properties of the v valuation we immediately obtain:

$$[\![\neg A]\!] = 1 - [\![A]\!] \qquad (1)$$

$$[\![A \wedge B]\!] \leq \min([\![A]\!], [\![B]\!]) \qquad (2)$$

$$[\![A \vee B]\!] \geq \max([\![A]\!], [\![B]\!])$$

$$[\![A \rightarrow B]\!] = 1 - [\![A]\!] + [\![A \wedge B]\!]$$

$$[\![A \leftrightarrow B]\!] = 1 - [\![A \vee B]\!] + [\![A \wedge B]\!]$$

2. SENTENCES AS SET EXTENSIONS, AND TRUTH AS MEASURE

Any Boolean Algebra has a *representation* on a set structure (a field of sets) as Stone proved long ago in a famous theorem (see e.g.[8]). Thus, given the Boolean sentence algebra $\mathcal{L}$, there exist both a set Θ (whatever the meaning we give to its elements θ) *and* a 'representation' function that can be characterized as an isomorphism of $\mathcal{L}$ into the Boolean subalgebra $\mathcal{B}$ of clopens in $\mathcal{P}(\Theta)$, i.e.

$$\rho : \mathcal{L} \longleftrightarrow \mathcal{B} : A \mapsto \mathbf{A} \quad (\mathcal{B} \subset \mathcal{P}(\Theta), \mathbf{A} \subset \Theta)$$

We may indulge in calling the members of Θ *possible worlds*, or *cases* (as did Laplace or Boole [1]) or *possibilities* (as Shafer) or even *observers*, *states*, etc. Θ is thus the real *universe of* *discourse* or *reference frame* (the set of *possible worlds*). It also coincides with Fenstad's [4] *model space*.

We can establish a general, one-to-one correspondence between the two worlds (the language world $\mathcal{L}$ and the referential universe Θ, both made up of "propositions") and their constituent parts, thus:

$$\begin{array}{ccc}
\mathcal{L} & \Longleftrightarrow & \mathcal{B} \\
A\ (A \in \mathcal{L}) & \Longleftrightarrow & \mathbf{A}\ (\mathbf{A} \subset \Theta) \\
A \wedge B & \Longleftrightarrow & \mathbf{A} \cap \mathbf{B} \\
A \vee B & \Longleftrightarrow & \mathbf{A} \cup \mathbf{B} \\
\neg A & \Longleftrightarrow & \mathbf{A}^c \\
\top & \Longleftrightarrow & \Theta \\
\bot & \Longleftrightarrow & \phi \\
A \vdash B & \Longleftrightarrow & \mathbf{A} \subset \mathbf{B}
\end{array}$$

The valuation v and the representation isomorphism ρ clearly induce a [0,1]-valued measure μ in $\mathcal{B} \subset \mathcal{P}(\Theta)$ in such a way that $\mu = v \circ \rho^{-1}$, i.e. $\mu(\mathbf{A}) = [\![A]\!]$.

Intuitively, the measure $\mu(\{\theta\})$ of each individual θ in a *finite* Θ universe is the relevance or the degree of realizability of the given possible world θ. The μ measure corresponds to the weighing function λ in Fenstad's [4] *model space*. As it is known, μ (or λ) is not only additive but countably so; thus μ is eligible as a standard "probability" measure (in the technical sense).

3. CONNECTIVES AND SENTENTIAL STRUCTURE

We usually demand a way to compute the truth value of composite sentences. For the negation connective this is easy: it is given by formula (1). For the rest we have the additivity formula as well as the definitions of $\rightarrow$ and $\leftrightarrow$:

The problem reduces to finding the numerical expression $[\![A \wedge B]\!]$ of the conjunction $A \wedge B$ as a function of the (numerical) "truth" values $[\![A]\!]$ and $[\![B]\!]$ of the component sentences A and B (and the same for the other three binary connectives). But, besides $[\![A]\!]$ and $[\![B]\!]$, we now need a third parameter that we note " α_{AB} " and that we call "*compatibility* between A and B"; its value is defined by $\alpha_{AB} =_{df} 1 - \beta_{AB}$, where:

$$\beta_{AB} = \frac{\min\,([\![A]\!],[\![B]\!]) \; - \; [\![A \wedge B]\!]}{\min\,(\,[\![A]\!],\ [\![B]\!],\ 1 - [\![A]\!],\ 1 - [\![B]\!]\,)} \quad .$$

We call β_{AB} the "*incompatibility* of sentences A and B" and we rename the denominator by calling it "Δ_{AB}". Then, the formulas for the connectives are:

$$\boxed{\begin{aligned}
[\![A \wedge B]\!] &= \min\,([\![A]\!],[\![B]\!]) \; - \; \beta_{AB} \cdot \Delta_{AB} \\
[\![A \vee B]\!] &= \max\,([\![A]\!],[\![B]\!]) \; + \; \beta_{AB} \cdot \Delta_{AB} \\
[\![A \rightarrow B]\!] &= \min\,(1, 1 - [\![A]\!] + [\![B]\!]) \; - \; \beta_{AB} \cdot \Delta_{AB} \\
[\![A \leftrightarrow B]\!] &= 1 - \; |\, [\![A]\!] - [\![B]\!]\, | \; - 2 \cdot \beta_{AB} \cdot \Delta_{AB}
\end{aligned}}$$

So, by knowing $[\![A]\!]$, $[\![B]\!]$ and a single value (either of $[\![A \wedge B]\!]$, $[\![A \vee B]\!]$, $[\![A \to B]\!]$, $[\![A \leftrightarrow B]\!]$, α_{AB} or β_{AB} —or $[\![A|B]\!]$ or $[\![B|A]\!]$, see below) we can compute, via α_{AB} (or β_{AB}), the other seven. Note that the parameter α_{AB} acts as an indicator or measure of the "relative position" of **A** and **B** inside Θ, and that "Δ_{AB}" is a quadruple minimum that only depends on the values of $[\![A]\!]$ and $[\![B]\!]$. Note too that if we suppose that $\beta_{AB} = 0$ for *all* A and B then the above formulas coincide with those given in ordinary many-valued logics.

4. RELATIVE TRUTH

Now suppose we want to express the conjunction value as a product:

$$[\![A \wedge B]\!] = [\![A]\!] \cdot \tau$$

With the current $\mathcal{L}/\mathcal{P}(\Theta)$ representation in mind, we obtain:

$$\tau = \frac{\mu(\mathbf{A} \cap \mathbf{B})}{\mu(\mathbf{A})} \tag{3}$$

We define τ as the *relative truth* " $[\![B|A]\!]$ " (i.e. the "truth of B relative to A"). Though this definition exactly parallels that of conditional probability, the account we give leaves out any probabilistic interpretation of the concept and retrieves it for exclusively logical contexts.

In particular, if $[\![B|A]\!] = [\![B]\!]$ then we say that A and B are *independent*. In that case, the conjunction can be expressed as the product:

$$[\![A \wedge B]\!] \;=\; [\![A]\!] \cdot [\![B]\!]$$

In any other case we say that A and B are mutually *dependent* and speak of the *relative truth* of one with respect to the other. Note the dependence goes both ways and the two situations are symmetric. We have (assuming $[\![A]\!] \neq 0$):

$$[\![A]\!] \cdot [\![B|A]\!] = [\![B]\!] \cdot [\![A|B]\!] = [\![A \wedge B]\!] \quad \text{(a logical \textit{“Bayes formula”})}$$

$$[\![B|A]\!] = 1 - \frac{1 - [\![A \to B]\!]}{[\![A]\!]} \quad .$$

From the latter, note that, in general,

$$[\![B|A]\!] \;\neq\; [\![A \to B]\!]$$

Particularly, we have *always*

$$[\![B|A]\!] \;<\; [\![A \to B]\!]$$

except when either $[\![A]\!] = 1$ or $[\![A \to B]\!] = 1$, in which cases (and they are the *only* ones) $[\![B|A]\!] = [\![A \to B]\!]$. (These facts have been repeatedly noticed by many people, notably by Reichenbach [14], Stalnaker [22], Lewis [9] and Popper [12].)

When two sentences A and B are *independent* then (and this is a necessary and sufficient condition for that to happen):

$$\begin{aligned} \alpha_{AB} &= \max\left([\![A]\!], [\![B]\!]\right) && \text{if } [\![A]\!] + [\![B]\!] \leq 1 \\ &= \max\left([\![\neg A]\!], [\![\neg B]\!]\right) && \text{if } [\![A]\!] + [\![B]\!] \geq 1 \;. \end{aligned}$$

Now, the statement '$A \to B$' can have, among other readings, one logical ("A is sufficient for B" or "B is necessary for A"), another (loosely) "causal" ("A occurs and B follows"). But because $A \to B$ is valued in [0,1], its value $[\![A \to B]\!]$ (and the values $[\![B|A]\!]$, $[\![B \to A]\!]$ and

$[\![A|B]\!]$) now mean only *degrees*, and so $B \to A$ may be —and usually is— read "evidentially" ("B is evidence for A"). Within such a frame of mind,

— $[\![B|A]\!]$ (or "σ_A") could be termed "degree of *sufficiency* or *causality*" of A (or "*causal support* for B"), to be read as "degree in which A is sufficient for B" or "degree in which A is a cause of B". In view of (3), it is roughly a measure of how much of $\mathbf{A}$ is contained in $\mathbf{B}$.

— $[\![A|B]\!]$ (or "ν_A") could be termed "degree of *necessity*" or "*evidence*" of A (or "*evidential support* for A", to be read as "degree in which A is necessary for B" or "degree in which B is evidence (= support of hypothesis) for A (=the hypothesis)". With (3) in mind, it can be seen as how much of $\mathbf{B}$ overlaps with $\mathbf{A}$.

Such measures may be directly estimated by experts, normally by interpreting the θs frequentially, in terms of *cases*, like Boole. ("Cases" may be statistically-based or simply imagined, presumably on the basis of past experience or sheer plausibility.) Thus, σ_A in a causal reading of "$A \to B$" would be determined by answering the question: "How many times (proportionally) —experience shows— A occurs and B follows?" For ν_A, the question would be: "How many times effect B occurs and A has occurred previously?" (Similarly for the evidential reading of "$A \to B$".) Once σ_A and ν_A have been guessed, they may be adjusted (via the

$$\frac{\sigma_A}{\nu_A} = \frac{[\![B]\!]}{[\![A]\!]}$$

relation) and then lead —by straightforward computation— to $[\![A \to B]\!]$, $[\![B \to A]\!]$ and α_{AB}, which allows one to compute all other values for connectives and also to get a picture of the structural relations linking A and B.

Note first that a similar computing process takes place in "Bayesian reasoning", and in the "approximate reasoning" methods as implemented in marketable expert systems, though none of them satisfactorily explains on what logical grounds are the procedures justified, nor can they avoid supplying a (nominally) probabilistic account of them. Such an account would be here clearly misplaced, since there is usually neither (a) a sample space of *events*, but a language of *sentences* (i.e. linguistic descriptions —not necessarily of any "event"), nor (b) a measure based on *uncertainty* and *outcomes* (the topics probability is supposed to deal with) but rather simple estimated *beliefs*, nor (c) an adequate updatable statistical or probabilistic basis to compute the values of the "events" (and their ongoing, dynamical change).

Note also that any reasoning can proceed here in both directions (from A to B and from B to A), because both conditionals $A \to B$ and $B \to A$ claim a non-zero value and thus a "causal" top-down reasoning can be complemented by an "evidential" bottom-up reasoning on the same set of given sentences.

5. THE GEOMETRY OF LOGIC: DISTANCE, TRUTH LIKELIHOOD, INFORMATION CONTENT AND ENTROPY IN $\mathcal{L}$

The fact that we have:

$$[\![A \leftrightarrow B]\!] = 1 - ([\![A \vee B]\!] - [\![A \wedge B]\!])$$

strongly suggests using $1 - [\![A \leftrightarrow B]\!] = [\![A \vee B]\!] - [\![A \wedge B]\!]$ as a measure of the *distance* $\overline{AB}$ (under a given valuation v). So we do. (We remark that all definitions we give here of distance and related concepts are not only applicable to sentences but to *theories* as well, because for a general lattice $\mathcal{L}$ the lattice $\hat{\mathcal{L}}$ of theories derived from each sentence in $\mathcal{L}$ is isomorphic to $\mathcal{L}$.)

[*Definition 1*:] *Distance* (or *Boolean distance*) between two sentences or theories A and B is:

$$d(A,B) =_{df} 1 - [\![A \leftrightarrow B]\!] \;=\; [\![A \vee B]\!] - [\![A \wedge B]\!] \;=\; |[\![A]\!] - [\![B]\!]| + 2 \cdot \beta_{AB} \cdot \Delta_{AB}$$

[*Definition 2*:] *Compatible distance* between two sentences or theories A and B is:

$$d^+(A,B) =_{df} |[\![A]\!] - [\![B]\!]| \;=\; [\![A]\!] + [\![B]\!] - 2 \ \min\ ([\![A]\!],[\![B]\!])$$

We now define a *truth likelihood* value for A —approximating Popper's (and Miller's) [13] *verisimilitude* measure— by making it to equal the distance from A to falsehood, i.e. $d(A,\bot)$. We obtain, immediately:

$$d(A,\bot) = d(\top,\bot) - d(\top,A) = 1 - d(A,\top) = 1 - d(A \triangle \top,\bot) = 1 - d(\neg A,\bot) = 1 - [\![\neg A]\!] = [\![A]\!]$$

So here we have a further interpretation of our "truth values" $[\![A]\!]$ in terms of *truth likelihood* or Popper's *verisimilitude*. We might as well consider $[\![A]\!]$ as a rough measure of *partial truth* or *truth content* of A. In a similar vein, we may recall that Scott [20] once suggested the "truth value" $[\![A]\!]$ of many-valued logics could be interpreted as *one* (meaning *truth*) less the *error* of A (or rather of a measure settling the truth of A) or the *inexactness* of A (as a theory); in this framework, it comes out that, in our terms, $[\![A]\!] = 1 - \varepsilon_A$ and $\varepsilon_A = 1 - [\![A]\!] = d(A,\top)$.

We further observe that, for any sentential letters P and Q, any uniform truth valuation yields $[\![P]\!] = [\![\neg P]\!] = .50$, $[\![P \wedge Q]\!] = .25$ and $[\![P \vee Q]\!] = .75$, which is like saying that, if all letters are equiprobable, the given values are the probability of the given sentence being true (a number that Bar-Hillel and Hintikka called, appropriately, "truth-table probability" [7]). So this value's complement to one should seemingly correspond to the amount of information —in a loose sense— we have when the sentence is true. This is precisely what Bar-Hillel and Hintikka once defined as "degree of information", "semantical information" or "informativeness" $I(A)$ of a sentence A. (Viewed in our terms, this *information content* $I(A)$ equals $1 - [\![A]\!]$, or $[\![\neg A]\!]$.) These authors used the concept to model the reasoning process, assumed to be driven by an increase both of informativeness and parsimony. What is interesting is that they compute the informativeness of composite sentences by combining them according to non-truth-functional rules which yield values that coincide with those predicted by our formulas.

A new measure we can define, which can also be used as an *entropy* (in the sense of De Luca & Termini), is this one:

[*Definition*:] *Imprecision* (or, perhaps, "fuzziness") of a sentence (or theory) A is the value for A of the function

$$f : \mathcal{L} \longrightarrow [0,1] \quad \text{such that} \quad f(A) = 1 - d^+(A,\neg A)$$

It is immediate that:

$$f(A) = 2 \ \min\ ([\![A]\!], 1 - [\![A]\!]) \ ,$$

which is equivalent to saying that the imprecision of a sentence A equals twice the error we make when we evaluate on A the truth of the law of non-contradiction or of the excluded middle by considering there really *is* a non-zero compatibility between A and $\neg A$ (actually there is none, as seen from the formulas). In connection with this fact, we note in passing that:

– *classical* (two-valued) *logic* is the special case of ours in which *all* sentences in $\mathcal{L}$ have *zero* imprecision.

– ordinary *multi-valued logics* —like Łukasiewicz-Tarski's $Ł_\infty$— are the ones where at least *one* sentence in $\mathcal{L}$ has non-zero imprecision. This measure being —as it is— an error function, imprecision is here just the degree in which these logics fail to distinguish contradictions (their lack of "resolving power") .

6. ELEMENTARY PROOF THEORY FOR GENERAL ASSERTIONS

Once we have valued sentences in $[0, 1]$ we now need a Proof Theory that in the most natural way extends standard logic so as to treat imprecise statements or weak assertions, and measure and control whatever effect they may have on reasoning (as well as to explain some results in approximate reasoning methods from Artificial Intelligence).

To begin with, suppose a valid argument, noted $\Gamma \vdash B$ (where Γ are the premises, or a finite subset of them). Classical logic declares it *valid* if B is derivable from Γ in an appropriate deduction calculus. By the completeness property, this amounts to assert the truth of B whenever the premises in Γ are true. Now, the ultimate judge of the truth of the premises is the reasoner. It is the reasoner who decides that each premise used is true (or to be considered true). To justify such a decision, the reasoner may apply a truth criterion like Tarski's well-known definition:

(T) 'A' is true if and only if A is true

So, the reasoner can verify the sentence and declare it true whenever the translation A of the object-language sentence A is found true. It is the reasoner who is full command of the sentences and the translation process, and thus the only one who can validate their truth. In those cases the reasoner declares A when assured that what A describes is precisely the case. If the reasoner is not sure of the result of his/her validation or does not want to commit him/herself to it, then the reasoner may choose not to make a full assertion by claiming that A's verification does not yield an obvious result. In that case, the reasoner may rather easily "qualify" the assertion by assigning numbers in $[0, 1]$ such as $v(A)$ —that we noted "$[\![A]\!]$"— or $\varepsilon(A)$ [$= 1 - v(A)$] meaning that the reasoner believes or is willing to assert A to the degree $v(A)$ or assume it with a risk or estimated error of $\varepsilon(A)$.

The proof theory we now sketch is a slightly extended version of the standard one. Here we understand by *proof theory* the usual syntactical deduction procedures *plus* the computation of numerical coefficients that we must perform alongside the standard deductive process. We do that because a final value of zero for the conclusion would invalidate the whole argument as thoroughly as though the reasoning were formally —syntactically— invalid. As always, any formally valid *argument* will have, by definition, the following *sequent* form:

$$\Gamma \vdash B$$

where B is the conclusion and Γ stands for the list —or, rather, the conjunction— of the premises (or, by the compactness property, of a finite number of them). We have, elementarily:

$$\Gamma \vdash B \quad \Rightarrow \quad [\![\Gamma]\!] \leq [\![B]\!]. \tag{4}$$

We henceforth assume that we have a *valid argument* (so $\Gamma \vdash B$ will always hold), and that all premises are non-zero (i.e. $\forall i \; [\![A_i]\!] > 0$). We distinguish four possible cases:

1) $[\![\Gamma]\!] = 0$ (i.e. the premises are –materially– *inconsistent*). Here by (4) $[\![B]\!]$ can be anywhere between 0 and 1; this value is in principle undetermined, and uncontrollably so.

2) $[\![B]\!] = 0$. This entails, by (4), $[\![\Gamma]\!] = 0$ and we are in a special instance of the previous case. The reasoning is formally valid, no premise is asserted, and the conclusion is false.

3) $[\![\Gamma]\!] \in (0, 1)$ (i.e. the premises are consistent). Then, by (4), $[\![B]\!] > 0$. We have a formally valid argument, we risk assessing the premises (though with some apprehension) and get

a conclusion which can be effectively asserted though by assuming a –bounded– risk. This will be the case we will set to explore below.

4) $[\![\Gamma]\!] = 1$. This condition means that $[\![A_1]\!] = \ldots = [\![A_n]\!] = 1$ and, by (4), $[\![B]\!] = 1$. So the premises are *all* asserted —with no risk incurred— and the conclusion *holds* inconditionally (remember $\Gamma \vdash B$ is formally valid). This is the classical case studied by ordinary two-valued Logic.

We are interested in examining *case 3* above, i.e. formally valid reasoning *plus* assertable premises (though not risk-free assertions) *plus* assertable conclusion (but at some measurable cost). Cases 3 and 4 characterize in a most general way all *sound* reasoning. Case 2 characterizes *unsound* arguments (since in this case having a formally valid argument $\Gamma \vdash B$ does not preclude getting an irrelevant conclusion ($[\![B]\!] = 0$). As this case is the one to avoid, we have:

[*Definition:*] *Unsoundness* of a valid argument $\Gamma \vdash B$ is having $[\![B]\!] = 0$ though the premises are themselves non-zero. So:

[*Definition:*] *Soundness* of a valid argument $\Gamma \vdash B$ is having $[\![B]\!] > 0$ whenever the premises are non-zero.

With that in mind, we can now turn to the basic inference rule, the Modus Ponens (MP). From a strictly logic point of view, this rule is

$$\begin{array}{ll} A & m \\ A \to B & n \\ \hline B & p \end{array} \tag{5}$$

where m, n and p stand for the strength or force (or "truth value") we are willing to assign each assertion; so, in our terms, m, n and p are just our $[\![A]\!]$, $[\![A \to B]\!]$ and $[\![B]\!]$. They are numbers in [0,1] that take part in a (numerical) computation which parallels and runs along the logical, purely syntactical deduction process. This is well understood and currently exploited by reasoning systems in Artificial Intelligence that must rely on numerical evaluations —given by users— that amount to *credibility* assignments (or "certainty factors"), *belief* coefficients, or even —rather confusingly— *probabilities* (often just *a priori* probability estimates); this is the case of successful *expert systems* such as Prospector or Mycin. The trouble with such systems is that they tend to view Modus Ponens as a probability rule (this is made explicit in systems of the Prospector type). They use it to present the MP rule in this way:

$$\begin{array}{l} A(m) \\ A \to B(\sigma) \\ \hline B(p) \end{array} \tag{6}$$

where m and p are the 'probability' (a rather loose term here) of A and B, and "$A \to B(\sigma)$" means that "whenever A happens, B happens with probability σ". Here σ turns out to be just $v(B|A)$ or "$[\![B|A]\!]$", the "relative truth" of B given A; this concept, modelled on a close analogy —*ceteris paribus*— with that of (probabilistic) conditioning, is what we have defined above (in (3)) and called "degree of sufficiency" σ of A —or of necessity of B—, assumed easily elicitable by experts. So it is just natural, and immediate, to compute the p value thus:

$$p \geq \sigma \cdot m$$

or, in our notation,

$$[\![B]\!] \geq [\![B|A]\!] \cdot [\![A]\!]$$

which is just another version of formula (2).

The problem is that what we have, from our purely logical, probability-rid standpoint, is (5), not (6), and in (5) n is not $[\![B|A]\!]$ but $[\![A \to B]\!]$. Recall that $[\![B|A]\!]$ and $[\![A \to B]\!]$ not only do *not* coincide but mean different things (as repeatedly noticed by logicians, and as explained above). Indeed, $[\![A \to B]\!]$ is the value ("truth" we may call it, or "truth minus risk") we assign to the (logical) assertion $A \to B$. Instead, $[\![B|A]\!]$ is a relative measure linking materially, *factually*, A and B (or, better still, the **A** and **B** sets), with no concern whether a true *logical* relation between them exists; we might even have $[\![B|A]\!] < [\![B]\!]$, thereby indicating there exists an *anti*correlation (thus rather contradicting any —logical or other— reasonable kind of relationship between A and B). So we turn back to our (5) rule; note that $m + n \geq 1$ (this *always* holds) and that $[\![B|A]\!]$ can be obtained from $[\![A \to B]\!]$, or vice versa: $[\![A \to B]\!]$ from $[\![B|A]\!]$ through

$$[\![A \to B]\!] \ = \ 1 - [\![A]\!] \cdot (1 - [\![B|A]\!]) \tag{7}$$

(which is useful, since $[\![B|A]\!]$ is directly obtainable from experts).

The following theorem states the *soundness* condition for the MP rule.

The Modus Ponens rule

$$\begin{array}{ll} A & m \\ A \to B & n \\ \hline B & p \end{array} \quad \text{(we assume } m \text{ and } n \text{ are both non-zero)}$$

is *sound* (and thus $[\![B]\!] \neq 0$) if one of these four equivalent conditions hold:

1) $m + n > 1$
2) $[\![B|A]\!] > 0$
3) $[\![A \wedge B]\!] > 0$
4) Either $[\![A]\!] + [\![B]\!] > 1$ (and thus $[\![B]\!] > 1 - m$) or both A and B are compatible ($\alpha_{AB} > 0$) and not binary-valued.

In both sound and unsound cases we have the following easily computable bounds for the value $[\![B]\!]$ of the MP conclusion:

$$[\![A]\!] + [\![A \to B]\!] - 1 \quad \leq \quad [\![B]\!] \quad \leq \quad [\![A \to B]\!] \tag{8}$$

or equivalently, in shorter notation:

$$m + n - 1 \ \leq \ p \ \leq \ n \ .$$

(Such bounds have been discovered again and again by quite diverse authors; see e.g. Genesereth & Nilsson [6]). The lower bound —which equals $[\![A \wedge B]\!]$ — is reached when $\alpha_{AB} = 1$ and $[\![A]\!] \geq [\![B]\!]$, while the upper bound is reached when $\alpha_{AB} = 0$ and $[\![A]\!] + [\![B]\!] \geq 1$. Naturally we usually know neither $[\![B]\!]$ nor α_{AB} beforehand, so we don't know whether the actual value $[\![B]\!]$ reaches either bound or not, nor which is it; we can merely locate $[\![B]\!]$ inside the $[m + n - 1, n]$ interval.

But this interval can be narrowed. Since a very reasonable constraint a conditional $A \to B$ may be expected to fulfill ([19]) is that A and B be (assumedly) non-independent –and not binary– and *positively* correlated (i.e.: $[\![A \wedge B]\!] > [\![A]\!] \cdot [\![B]\!]$), so we have:

$$[\![A]\!] + [\![A \to B]\!] - 1 \quad \leq \quad [\![B]\!] \quad < \quad [\![B|A]\!]. \tag{9}$$

Here, if A and B are fully or strongly compatible, $[\![B]\!]$ will be nearer the lower bound. Thus, we can only increase our $[\![B]\!]$ if we are assured that A and B are independent (in the sense that $[\![A \wedge B]\!]$ equals $[\![A]\!] \cdot [\![B]\!]$, see above): we then obtain the highest value $[\![B]\!] = [\![B|A]\!]$. On the other hand, the more we confide instead in a strong logical relation between A and B, the more we should lean towards the low value given by

$$[\![B]\!] = [\![A \wedge B]\!] = m + n - 1. \tag{10}$$

(Under very reasonable elementary hypotheses ([19]), we easily get these bounds for $[\![B]\!]$: $[\![A]\!]/2 \ < \ [\![B]\!] \ \leq \ [\![A]\!]$.)

Now, if what we want is not an interval, however narrowed, but a precise value for $[\![B]\!]$ we should favor the lower value, the one given by (10) above. There are lots of reasons ([19]) for this choice of value in the absence of more relevant information, so we will stick to it.

But if we want not merely a pair of bounds —or a favored lower bound— for the conclusion B of an MP but the *exact* value $[\![B]\!]$, the obvious candidate formula for this follows easily: suppose we are given not only $[\![B|A]\!]$ but also $[\![A|B]\!]$ (that we note by σ and ν) and we assume them estimated by experts (see [19]). We then formulate MP as

$$\begin{array}{l} A(m) \\ A \to B(\sigma, \nu) \\ \hline B(p) \end{array}$$

which is exactly (6) except that the conditional has prompted evaluation of relative truths of A and B in both directions. The value is computable at once from the above definition of $[\![A|B]\!]$:

$$[\![B]\!] \ = \ \frac{[\![B|A]\!] \cdot [\![A]\!]}{[\![A|B]\!]} \qquad \text{or} \qquad p \ = \ \frac{\sigma \cdot m}{\nu}$$

Note that the above logical formula coincides formally with Bayes's theorem —whence the adjective ("Bayesian") for any calculus that uses it— except that it deals not with hard-to-compute probabilities of events but with beliefs (or assertion strengths) of sentences. Note also that this value is the one that some *approximate reasoning* systems (e.g. Prospector) unqualifiedly assign to $[\![B]\!]$ supposedly on purely probabilistic grounds —and falsely assuming that $[\![B|A]\!]$ is the same as $[\![A \to B]\!]$—; see, for instance, the Genesereth & Nilsson [6] text, where the logical equation above is said to be Bayes's formula.

If we wanted the MP presented in the more traditional *logical* way (5), first we would directly estimate the truth value $[\![A \to B]\!]$ of the conditional, or compute it from σ through (7) —or both, and use each estimate as a cross-check on the other—, so we would now have, along with the expert guess of ν:

$$\begin{array}{ll} A & m \\ A \to B & n\ (\nu) \\ \hline B & p \end{array}$$

(where $n = 1 - m \cdot (1 - \sigma)$), and so

$$[\![B]\!] \;=\; \frac{[\![A \wedge B]\!]}{[\![A|B]\!]} \;=\; \frac{m + n - 1}{\nu}$$

that naturally fits the (9) bounds (when ν runs along from 1 to $[\![A]\!]$).

References

[1] Boole, G. (1854), *An Investigation of the Laws of Thought*, Dover, N.Y. (1958)

[2] Carnap, R. (1950), *Logical Foundations of Probability*, U. Chicago Press

[3] Carnap, R. (1962), The aim of inductive logic, in: E. Nagel, P. Suppes & A. Tarski, eds., *Logic, Methodology and Philosophy of Science*, Stanford

[4] Fenstad, J.E. (1967), Representations of probabilities defined on first order languages, in: J.N. Crossley, ed., *Sets, Models and Recursion Theory*, North Holland

[5] Gaifman, H. (1964), Concerning measures in first order calculi, *Israel J. of Mathematics*, **2**, 1-18

[6] Genesereth, M. & Nilsson, N. (1987), *Logical Foundations of Artificial Intelligence*, Morgan Kaufmann

[7] Hintikka, J. (1973), *Logic, Language Games and Information*, Oxford

[8] Koppelberg, S., Monk, J.D. & Bonnet, R., eds. (1989), *Handbook of Boolean algebras* (3 vols.), North Holland

[9] Lewis, D. (1976), Probabilities of conditionals and conditional probabilities, *Philosophical Review*, **85**, 297-315

[10] Miller, D.W. (1978), On distance from the truth as a true distance, in: J. Hintikka, I. Niinilnuoto & E. Saarinen, eds., *Essays in Mathematical and Philosophical Logic*, Reidel

[11] Miller, D.W. (1981), A Geometry of Logic, in: H. Skala, S. Termini & E.Trillas, eds., *Aspects of Vagueness*, Reidel

[12] Popper, K.R. (1959), *The Logic of Scientific Discovery*, Hutchinson

[13] Popper, K.R. & and Miller, D.W. (1987), Why probabilistic support is not inductive, *Phil. Trans. R. Soc. Lond.*, A **321**, 569-591

[14] Reichenbach, H. (1935), *Wahrscheinlichkeitslehre* (Engl. transl. *The Theory of Probability*, U. Calif. Press (1949))

[15] Sales, T. (1982), Una lògica multivalent booleana (LMB), in: *Actes 1er Congrés Català de Lògica Matemàtica* (Barcelona, January 1982), 113-116

[16] Sales, T. (1982), *Contribució a l'anàlisi lògica de la imprecisió*, Universitat Politècnica de Catalunya (Ph. D. dissertation)

[17] Sales, T. (1992), *Propositional Logic as Boolean many-valued logic*, TR LSI-92-20-R, Universitat Politècnica de Catalunya

[18] Sales, T. (1994), Between Logic and Probability, *Mathware & Soft Computing*, **1**, 99-138

[19] Sales, T. (1996), Logic of Assertions, *Theoria*, **24** (in the press)

[20] Scott, D. (1973), Does many-valued logic have any use?, in: S. Körner, ed., *Philosophy of Logic*, Blackwell (1976)

[21] Scott, D. & Krauss, P. (1968), Assigning probabilities to logical formulas, in: J. Hintikka & P. Suppes, eds., *Aspects of Inductive Logic*, North Holland

[22] Stalnaker, R. (1970), Probability and conditionals, *Philosophy of Science*, **37**, 64-80

[23] Suppes, P. (1968), Probabilistic inference and the concept of total evidence, in: J. Hintikka & P. Suppes, eds., *Aspects of Inductive Logic*, North Holland