Propositional logic as Boolean many-valued logic

Ton Sales

Report LSI-92-20-R

Propositional Logic

as

Boolean Many-Valued Logic

TON SALES

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

Abstract: It is commonly assumed that *Boolean logic* is synonymous with *two-valued logic*. However, this needs not be the case. As Boole suspected [7], and Probability theorists know, one can value propositions in, say, the real unit interval —whatever the purpose of doing this— and yet preserve the Boolean-algebra structure of propositions *as well as* the common laws of ordinary logic. This report (which is a barely-updated summary of a Ph.D. Thesis [38] written in 1981-82 [see the disclaimer at the end]) is an exploration of some of the consequences —some interesting, some a bit unexpected— of valuing propositions in [0,1]. What emerges is a *many-valued logic* which is, perhaps surprisingly, also *Boolean*. All in a neat and natural way. And though the analysis suggests some parallels with Probability Theory, the development falls strictly within the *logical theory* of Propositional Calculus, to whose classical and well-known version this report aspires to add new insights. Moreover, the expounded theory, initially aimed at a better understanding of logical concepts, turns out to admit a proof theory that has a direct application to hypothetical and approximate reasoning (forced by *imprecision* or other causes).

General remark: Most proofs in this report are both elementary and straightforward, and all are based on well-known results from elementary Logic and Lattice Theory. They are, therefore, generally omitted and left to the reader.

Index:

0. INTRODUCTION: THE STRUCTURE WITHIN PROPOSITIONAL LOGIC

Ordinary Propositional Logic is commonly considered satisfactory for many purposes. Also, it has a widely known and much studied lot of convenient properties, good-natured proof and model theories as well as a simple and adequate algebraic structure (the Boole algebra). The whole theory of Propositional Logic is built on the basis that every proposition —an interpreted sentence— is always assigned only one of two values: *true* (or 1) and *false* (or 0); so much, that this *two*-valuedness is commonly felt as inseparable from the theory. However, as we argue below, this *needs not* be so. Two-valuedness is, for all purposes, not the *only* case to consider. Suppose we confront any of the following typical situations —among many possible—, where we use (and so we must value) e.g. (1) propositions that are imprecise, vague or uncertain (and even paradoxical or meaningless), or (2) unsure or merely believed propositions (e.g. as premises) or (3) rules and general laws that have exceptions, or (4) propositional theories which are only partially true —i.e. that have some defective 'truth content' in Popper's sense [32]—, or (5) propositions with an experimental error affecting their truth value (Scott [39]). Also suppose that because of some of the above reasons (to name a few), and because we dare not or cannot give it a plain *false* or *true*, we willingly assign *one* single proposition a truth value other than 0 or 1 (e.g.: "undetermined" or non-existent —as in three-valued logics [5][24]—, or a number in the [0,1] interval —as in Probability or Łukasiewicz-Tarski's many-valued Logic [27][35]). In that case, ordinary [two-valued] Propositional Logic is at odds, it has never even suggested such possibilities were tractable, and gives no hint anyway about how to proceed, get a valid argument out of it and control or parameterize the conclusion's logical validity.

And, nevertheless, Propositional Logic *can* accommodate —we contend— such eventualities in a perfectly simple and natural way. In the sequel we explore how this can be done, with minimum strain and maximum generality. The resulting [Boolean] Propositional Logic —that we have called **BML** (an obvious acronym)— is like the well-known one in virtually all important aspects —only conveniently augmented— except in having *more than two* values available, so that ordinary —that is, *two-valued*— Propositional Logic becomes the *particular case* of **BML** when all propositions are *precise*, two-valued propositions. Thus, we get a natural extension of all Propositional Logic properties and, moreover, puzzling new insights into the inner workings of Propositional Logic's *propositions* and their inherent mutual relationships, which had been disguised and concealed within the disarming simplicity of traditional two-valued analyses of Propositional Logic. But not only the inner algebraic propositional structure is revealed: the analysis seems to shed new light on the apparently simple *conditional* statement and the seemingly innocuous relations between that statement and event-conditioning (as studied in Probability). Also, from the syntactic (proof-theoretic) standpoint, what we further obtain is an intriguing new possibility of having formally *valid* but *unsound* reasoning instances. We will explore these as well as their necessary and sufficient conditions, and will try to measure and control the degree of relevance of such arguments.

Note therefore that everything that will be said in the sequel is just a full-thrusted *complement* of —and fully compatible with— *all* laws and presentations of classical Propositional Logic. It will show a rich texture of relations and extensional structure linking the plain-looking and (apparently) dull propositions of Propositional Logic. The starting point, as the whole theory that follows, cannot be more elementary.

1. BASIC PROPERTIES OF PROPOSITIONS

Suppose, first, we have a propositional language $\mathcal{L}$, whose members (henceforth denoted $A, B, \ldots$) are sentences. We shall assume that $\mathcal{L}$ is closed by the propositional operators $\wedge$, $\vee$ and $\neg$ and, moreover, that it has the structure of a Boolean algebra. Thus, the structure $\langle \mathcal{L}, \wedge, \vee, \neg, \bot, \top \rangle$ (with signature $\langle 2, 2, 1, 0, 0 \rangle$) has the following properties (for arbitrary A, B and C):

1. $A \wedge A = A$ and $A \vee A = A$ *(Idempotency)* (1)
2. $A \wedge B = B \wedge A$ and $A \vee B = B \vee A$ *(Commutativity)* (2)
3. $A \wedge (B \wedge C) = (A \wedge B) \wedge C$ and $A \vee (B \vee C) = (A \vee B) \vee C$ *(Associativity)* (3)
4. $A \wedge (B \vee A) = A$ and $A \vee (B \wedge A) = A$ *(Absorption)* (4)
5. $A \wedge (B \vee C) = (A \wedge B) \vee (A \wedge C)$ and $A \vee (B \wedge C) = (A \vee B) \wedge (A \vee C)$ *(Distributivity)* (5)

6. $A \wedge \perp = \perp$ and $A \vee \perp = A$ *(Infimum)*, and $A \wedge \top = A$ and $A \vee \top = \top$ *(Supremum)* (6)

7. $A \wedge \neg A = \perp$ and $A \vee \neg A = \top$ *(Complementation)* (7)

The resulting structure is, of course, the well-known Lindenbaum-Tarski algebra of propositions. It is, if *finite*, of cardinality 2^N (for some N) and *atomic*, with exactly N atoms. If there are n propositional letters in the language —acting as generators of the resulting *free* Boolean algebra— the structure has 2^{2^n} elements, with 2^n atoms. In case there are infinitely many letters (generators) the algebra has the same cardinality as the set of generators, and is *not* atomic. (Nevertheless, we shall restrict our attention generally, for clarity's sake, to the finite case.)

As in all Boolean algebras, a partial order (that we denote by $\vdash$) is definable in the structure —which is then characterized as a Boolean lattice, with bounds $\perp$ and $\top$— through the equivalence:

8. $A \vdash B \iff A \wedge B = A \iff A \vee B = B$ *(Partial order)* (8)

Suppose, secondly, that we impose a valuation on this structure, taking values from the real $[0,1]$ interval. The valuation takes any sentence A in the language and assigns it a real number $[\![A]\!]$ (that we call "truth value of A", whatever sense we make of this). In particular, it assigns the interval bounds 0 and 1 to the lattice bounds $\perp$ and $\top$, respectively. As usual for valuations, this one is monotonic and additive. Thus, we have:

9. There is a valuation $v : \mathcal{L} \longrightarrow [0,1] : A \mapsto [\![A]\!]$ such that:

   9a. $[\![\perp]\!] = 0$, and $[\![\top]\!] = 1$. (9)

   And also, as it is generally understood for any valuation:

   9b. If $A \vdash B$ then $[\![A]\!] \leq [\![B]\!]$ *(Monotonicity)* (10)

   9c. For any A and B, $[\![A \wedge B]\!] + [\![A \vee B]\!] = [\![A]\!] + [\![B]\!]$ *(Additivity)* (11)

Once we are given the nine conditions 1-9 above —admittedly over-redundant (actually, 2 and 5-7 plus 9a and 9c would suffice)— we have all we need to proceed.

Observe that if the valuation described in 9 had been defined in $\{0,1\}$ instead of [0,1], conditions 1-9 would have completely characterized the standard [two-valued] Propositional Logic. In particular, from 1-8 arises all of its proof theory, while conditions 9a-c —together with the [immediately derivable] facts (12-13) (see below)— describe completely its model theory (its truth tables, in fact).

It is to be specially emphasized that condition 9 as stated above —that the v valuation has [0,1] as its range of values rather than merely $\{0,1\}$— is the *single* differing point between our presentation of Propositional Logic and the ordinary two-valued view of it. The first naturally subsumes the second as a special case. It also provides the starting point for our theory.

Several remarks about point 9 are in order here: As to the set of values (the "truth set"), the [0,1] interval could have as well been replaced for greater generality by any bounded linearly-ordered cancellative monoid. Moreover, $[\![\top]\!] = 1$ in 9a could have been verbalized by saying that v, as a measure, is *normalized*. Also, 9b could have been omitted or made implicit, since it can be derived from the rest of 9. Still, 9c could have been omitted since any *valuation* on a Boolean algebra is usually understood —by definition— to have the additive property (thus making the explicit statement of 9c redundant). Finally, note that in the *infinite* case 9c should have been replaced by an appropriate additivity formula (e.g. *"Countable additivity"*), as done in Probability Theory.

(A further remark, on notation: $[\![A]\!] = 1$ will be sometimes abbreviated as "$\models_v A$", which has the advantage of showing explicitly that a particular valuation v is involved.)

From conditions (1-11) the six formulas below follow immediately:

$$[\![\neg A]\!] = 1 - [\![A]\!] \tag{12}$$

$$[\![A \wedge B]\!] \leq [\![A]\!] \leq [\![A \vee B]\!] \tag{13}$$

$$[\![A \wedge B]\!] \leq \min([\![A]\!], [\![B]\!]) \tag{14}$$

$$[\![A \vee B]\!] \geq \max([\![A]\!], [\![B]\!]) \tag{15}$$

$$[\![\bigvee_{i=1}^{i=n} A_i]\!] = \sum_{i=1}^{i=n}[\![A_i]\!] - \sum_{i<j}[\![A_i \wedge A_j]\!] + \ldots + (-1)^{n+1}[\![\bigwedge_{i=1}^{i=n} A_i]\!] \tag{16}$$

$$\text{If for all } i \text{ and } j\ (i \neq j)\quad A_i \wedge A_j = \perp \quad \text{then} \quad [\![\bigvee_{i=1}^{i=n} A_i]\!] = \sum_{i=1}^{i=n} [\![A_i]\!] \tag{17}$$

(*Remarks*: The slightly informal (16) is the classic Poincaré formula for sets [now applying to propositions]. A special case of the former, (17) is the propositional analogue of the classic formulation of *additivity* (e.g. in Probability or Measure theory); the condition it imposes on propositions is typically met by *atoms*.)

As $\mathcal{L}$ here is finite, and therefore atomic, any A is decomposable into [the disjunction of] its atoms (i.e. those it covers, under the lattice $\vdash$ order). Combining this fact with (17) (see the last remark in the previous paragraph), we have the following decompositions (here rather loosely formalized):

$$\forall A \in \mathcal{L} \qquad A = \bigvee_{a_i \vdash A} a_i \qquad \text{and} \qquad [\![A]\!] = \sum_{a_i \vdash A} [\![a_i]\!]\,. \tag{18}$$

So, the truth valuation v is completely defined by the truth values it assigns to the atoms of $\mathcal{L}$; from these basic values the truth value of any proposition can be computed. (More on (18) will be said in later sections.) Also, by direct application of (17) we have, if $a_1, \ldots, a_N$ are the atoms in $\mathcal{L}$:

$$\sum_{i=1}^{i=N} [\![a_i]\!] = 1 \tag{19}$$

i.e. the truth values of all the atoms —for any valuation— add up to one.

If we now define the *conditional* or *if then* connective in the usual —classical— manner:

$$A \rightarrow B \ =_{def}\ \neg A \vee B \tag{20}$$

then the following formulas immediately obtain:

$$[\![A \rightarrow B]\!] = 1 - [\![A]\!] + [\![A \wedge B]\!] \tag{21}$$

$$[\![A \rightarrow B]\!] = 1 + [\![B]\!] - [\![A \vee B]\!] \tag{22}$$

$$[\![A \rightarrow B]\!] = 1 - [\![A \wedge \neg B]\!] \tag{23}$$

and, from these, also:

$$[\![A \rightarrow B]\!] - [\![B \rightarrow A]\!] = [\![B]\!] - [\![A]\!] \tag{24}$$

$$[\![A \rightarrow B]\!] + [\![\neg A \rightarrow B]\!] = 1 + [\![B]\!] \tag{25}$$

$$[\![A \rightarrow B]\!] + [\![A \rightarrow \neg B]\!] = 2 - [\![A]\!] \tag{26}$$

which interrelate the truth values of A, B, $A \rightarrow B$, $B \rightarrow A$ and so on, as well as:

$$[\![A \rightarrow B]\!] = [\![A \wedge B]\!] + [\![\neg A]\!] \tag{27}$$

$$[\![A \rightarrow \neg B]\!] = 1 - [\![A \wedge B]\!] \tag{28}$$

the first of which decomposes the value of the conditional into a sum of two terms: one due to the logical conjunction of the conditional operands, plus a second due to the negation of the antecedent. Also, we get:

$$[\![A \wedge B]\!] = [\![A \rightarrow B]\!] - [\![A \rightarrow \neg A]\!] = [\![B \rightarrow A]\!] - [\![B \rightarrow \neg B]\!] \tag{29}$$

$$[\![A]\!] = [\![A \wedge B]\!] + [\![A \wedge \neg B]\!] = [\![A \vee B]\!] + [\![A \vee \neg B]\!] - 1. \tag{30}$$

The first, rather intriguingly, allows us to consider the value of any logical conjunction of two propositions as the value of the conditioning that one exerts on the other *less* the value of the conditional of the first on its own negation. The second is a *complementation* law (notice B is arbitrary).

As is to be expected in a Boolean algebra,

$$A \vdash B \iff \neg A \vee B = \top. \tag{31}$$

If we denote $A = \top$ by "$\vdash A$", this can be written as:

$$A \vdash B \iff\ \vdash A \rightarrow B, \tag{32}$$

more in line with the usual formulation of the Deduction Theorem of elementary logic. (As a matter of fact, A could actually be a conjunction $A_1 \wedge \ldots \wedge A_n$ of propositions, called *premises*; in that case, the first A would rather be written as a list, so the left-hand part would read "$A_1, \ldots, A_n \vdash B$".)

Now, we define the relation between A and B given by $[\![A \to B]\!] = 1$ —that we note by "$A \models_v B$" (notice it depends on the particular valuation v chosen)—:

$$[Definition:]\quad \text{“}A \models_v B\text{”} \text{ if and only if } [\![A \to B]\!] = 1. \tag{33}$$

This relation is a *quasiordering* (see formulas (36) and (38) below) that contains the lattice order $A \vdash B$ (i.e. $A \vdash B \Rightarrow A \models_v B$) and satisfies the following five conditions:

$$\text{If } A \models_v B \text{ then } [\![A]\!] \leq [\![B]\!] \text{ (but note the converse is } \textit{not} \text{ true in general)} \tag{34}$$

$$A \models_v B \iff [\![A]\!] = [\![A \wedge B]\!] \iff [\![B]\!] = [\![A \vee B]\!] \tag{35}$$

$$A \models_v A \quad (\textit{Reflexivity} \text{ of } \models_v) \tag{36}$$

$$\text{If } A \models_v B \text{ and } B \models_v A, \text{ then } [\![A]\!] = [\![B]\!] \tag{37}$$

$$\text{If } A \models_v B \text{ and } B \models_v C, \text{ then } A \models_v C \quad (\textit{Transitivity} \text{ of } \models_v). \tag{38}$$

If we finally define the *biconditional* or *equivalence* connective in the usual manner:

$$A \leftrightarrow B =_{def} (A \to B) \wedge (B \to A) \tag{39}$$

then the following formulas are immediately obtainable:

$$[\![A \leftrightarrow B]\!] = [\![A \to B]\!] + [\![B \to A]\!] - 1 \tag{40}$$

$$[\![A \leftrightarrow B]\!] = 1 - [\![A \vee B]\!] + [\![A \wedge B]\!] \tag{41}$$

$$[\![A \leftrightarrow B]\!] = 1 - [\![A]\!] - [\![B]\!] + 2 \cdot [\![A \wedge B]\!] \tag{42}$$

$$[\![A \leftrightarrow B]\!] = 1 + [\![A]\!] + [\![B]\!] - 2 \cdot [\![A \vee B]\!] \tag{43}$$

as well as these —rather puzzling— properties:

$$[\![A \leftrightarrow B]\!] = 2 \cdot [\![A \to B]\!] + ([\![A]\!] - [\![B]\!]) - 1 = 2 \cdot [\![B \to A]\!] + ([\![B]\!] - [\![A]\!]) - 1 \tag{44}$$

$$[\![A \to B]\!] + [\![B \to A]\!] = [\![A \leftrightarrow B]\!] + [\![A \leftrightarrow A]\!] = [\![B \leftrightarrow A]\!] + [\![B \leftrightarrow B]\!]. \tag{45}$$

We now define the relation between A and B given by $[\![A \leftrightarrow B]\!] = 1$ and we note it by "$A =_v B$" (notice the dependence on the particular valuation v chosen):

$$[Definition:]\quad \text{“}A =_v B\text{”} \text{ if and only if } [\![A \leftrightarrow B]\!] = 1. \tag{46}$$

Note that "$=_v$" (which we can call "equality under a valuation") is an equivalence relation and contains the ordinary propositional [Lindenbaum-Tarski] identity (i.e. $A = B \Rightarrow A =_v B$). It yields the following two properties:

$$\text{If } A =_v B \text{ then } [\![A]\!] = [\![B]\!] \quad (\textit{Truth-value equality}). \text{ (The converse is } \textit{not} \text{ true in general)} \tag{47}$$

$$A =_v B \text{ if and only if } A \models_v B \text{ and } B \models_v A. \tag{48}$$

Now, the definition below follows the usual line:

$$[Definition:]\quad \text{“}\models A\text{”} \text{ if and only if } [\![A]\!] = 1 \textit{ for all valuations} \tag{49}$$

(*Remark*: Here "$[\![A]\!] = 1$ *for all valuations*" means "$v(A) = 1$ for all [0,1]-valuations v of A". From this point onwards, "*for all valuations*" will be sometimes shortened —quite informally— to "$(\forall v)$".)

As is obvious, $\models$ contains $\models_v$ (i.e. $\models A \Rightarrow \models_v A$), and the given definition can be written in this way: "$\models A$ iff $(\forall v) \models_v A$".

Naturally,

$$\text{If } A = B \text{ then } [\![A]\!] = [\![B]\!] \textit{ for all valuations} \tag{50}$$

(because the valuation is meant to be a function in the mathematical sense).

This has a corollary:

$$\text{If } \vdash A \text{ then } \models A \quad (\textit{Soundness}) \tag{51}$$

(because $A = \top$ yields $[\![A]\!] = [\![\top]\!] = 1$ for any valuation).

Conversely, we are forced to admit —by convention— that:

$$\text{If } [\![A]\!] = [\![B]\!] \textit{ for all valuations}, \text{ then } A = B \tag{52}$$

because we have no other way to distinguish any two propositions through the semantic means available (i.e. the [0,1]-valuations). Actually, we know that the ordinary $\{0,1\}$-valuations suffice to distinguish and identify all the elements of a Boolean algebra (this is a consequence of the Prime Ideal Theorem). So we have a considerably softened version of (52):

$$\text{If } [\![A]\!] = [\![B]\!] \textit{ for all } \textbf{binary} \textit{ valuations}, \text{ then } A = B \tag{53}$$

where "binary" means that the only values allowed are 0 and 1.

As a corollary of (52) we get:

$$\text{If } \models A \text{ then } \vdash A \ \ (\textit{Completeness}) \tag{54}$$

or, equivalently,

$$\vdash A \text{ if } [\![A]\!] = 1 \textit{ for all valuations.} \tag{55}$$

Note this is quite a strong requirement: recall that we are dealing with [0,1]-valuations, and there are inordinately more of them than classical $\{0,1\}$-valuations. Nevertheless, as a consequence of (53), this is no major drawback, and the weaker phrase "for all *binary* valuations" can be substituted in (55). Moreover, as it will be seen, a considerably less stringent condition can be imposed on v, and we shall have as a consequence that *not all* valuations v of A —binary or otherwise— will need to yield 1 so that we have $\vdash A$: actually, the existence of *just one* that does can be proved to suffice. In fact, if we define:

[*Definition:*] A *positive valuation* p is the particular case of the ordinary v valuation in which all atoms of $\mathcal{L}$ are assigned non-zero values. Or else, in a more general, straightforward and traditional manner:

$$p \text{ is a } v : \mathcal{L} \longrightarrow [0,1] : A \mapsto [\![A]\!] \text{ such that, for any } A, \quad [\![A]\!] = 0 \text{ if and only if } A = \bot \tag{56}$$

then we get as a result the rather surprising fact that the process of verifying "$\models A$" reduces to verifying that this very simple condition holds:

$$[\textit{Theorem:}] \models A \quad \text{if and only if} \quad \models_p A \textit{ for just } \textbf{one} \textit{ positive valuation } p. \tag{57}$$

because —for the *if* part (the converse is trivial)— the existence of a single *positive* valuation p of $\mathcal{L}$ that yields $[\![A]\!] = 1$ (i.e. $p(A) = 1$) suffices for guaranteeing that A covers all atoms of $\mathcal{L}$, and so $A = \top$.

The two relationships (50) and (52) shown above between propositions and values can be combined to yield this [informally stated] *semantical characterization* of propositional identity:

$$A = B \text{ if and only if } (\forall v)\ A =_v B \tag{58}$$

where "$\forall v$" can be here —and subsequently— taken to mean "for all *binary* valuations v" (see (53)).

Also, the *soundness* and *completeness* conditions, taken together, yield this equivalence:

$$[\textit{Completeness theorem} \text{ (first form):}] \vdash A \text{ if and only if } \models A \tag{59}$$

Now we put forward this (that we state informally):

$$[\textit{Definition:}] \text{ "}A \models B\text{" if and only if } (\forall v)\ \{[\![A]\!] = 1 \Rightarrow [\![B]\!] = 1\}. \tag{60}$$

It is easy to show that

$$A \models B \text{ if and only if } \models A \rightarrow B \tag{61}$$

Thus, $\models$ contains $\models_v$ (i.e. $A \models B \Rightarrow A \models_v B$), and the last definition can also be written: "$A \models B$ iff $(\forall v)\ A \models_v B$".

Recalling the above semantical characterization of propositional identity ("$A = B$ iff $(\forall v)$ $[\![A]\!] = [\![B]\!]$"), we would like to have a reasonable characterization of "$A \models B$" along the same line: something as "$A \models B$ iff $(\forall v)$ $[\![A]\!] \leq [\![B]\!]$". Unfortunately, this does *not* hold. The most we can have is just the rightward half:

If $A \models B$ then $[\![A]\!] \leq [\![B]\!]$ *for all valuations* (and the converse is *not* true in general). (62)

Now, again from the given semantical characterization of propositional identity (58) we have:

$$A = A \wedge B \quad \text{if and only if} \quad [\![A]\!] = [\![A \wedge B]\!] \textit{ for all valuations.} \tag{63}$$

Note the the left-hand side is equivalent to writing "$A \vdash B$", while the right-hand part amounts to saying "$[\![A \rightarrow B]\!] = 1$ for all valuations" (or else, by definition, "$\models A \rightarrow B$", that we have shown to be equivalent to "$A \models B$"). So we are led to this new characterization of the soundness and completeness condition:

[*Completeness theorem* (second form):] $A \vdash B$ if and only if $A \models B$. (64)

A could represent a list (or, better, a conjunction) of propositions $A_1, \ldots, A_n$ —the *premises*—. In that case, it would read thus:

[*Completeness theorem* (third form):] $A_1, \ldots, A_n \vdash B$ if and only if $A_1, \ldots, A_n \models B$ (65)

where the left-hand A's are the conjuncts of $A = A_1 \wedge \ldots \wedge A_n$, while the right-hand term is [demonstrably] equivalent to stating "$\{ [\![A_1]\!] = \ldots = [\![A_n]\!] = 1 \Rightarrow [\![B]\!] = 1 \}$ *for all valuations*".

2. REPRESENTATIONS: PROPOSITIONS AS EXTENSIONS

Two widely-known and yet under-exploited results in Boolean Algebra Theory (see [3]) have to do with *representations* of Boolean algebras on set structures. They can be stated thus:

A. Every *finite* Boolean algebra has a representation in the algebra of *all* subsets of a [finite] set.

B. [*Stone's Representation Theorem:*] Every Boolean algebra is *representable* on —isomorphic to— a field of sets.

In particular, the algebra $\mathcal{L}$ of propositions defined above (with 2^N elements) has a representation in —is isomorphic to— :

(a) The set of all subsets (the *power set*) of a set of N arbitrary elements.

(b) The set of binary valuations (two-valued epimorphisms) of $\mathcal{L}$.

(c) The lattice $\hat{\mathcal{L}}$ of theories of $\mathcal{L}$. (A *theory* is, algebraically, a *filter* (dual ideal); for lattices such as $\mathcal{L}$, filters are always *principal*).

(d) A discrete [Stone] topology.

(e) A direct product of N copies of $\{0, 1\}$ (or, equivalently, the set of strings of N 0's and 1's).

Whenever the algebra $\mathcal{L}$ is *finite*, the representations are just those mentioned (a-e), and the associated theorems are relatively trivial. In the general, *infinite* case, the corresponding representation theorems are described by slightly different, more open characterizations; and they follow from deep [and beautiful] results in Boolean Algebra Theory —basically, the classical Birkhoff and Stone *Representation Theorems* of the nineteen-thirties.

As we said before, we shall concentrate on the *finite* case for clarity's sake, since the infinite case differs mainly in the lack of the more intuitive features of the finite counterpart, and results are easily extrapolable. Indeed, the infinite case implies $\mathcal{L}$ is atomless and its homomorphic image (set-representation) is a field of sets that is properly contained in the power set; it lacks therefore the [very intuitive] *atomicity* of $\mathcal{L}$ as well as the simpler and more natural *powerset* structure on which a finite $\mathcal{L}$ maps.

So we begin by applying the Stone Theorem (or, rather, its finite-instance trivialization) to our case. We have:

[*Representation theorem:*] For any given finite algebra $\mathcal{L}$ of propositions (of order 2^N) there exist:

a) a finite set Θ (of order N), and

b) an isomorphism of $\mathcal{L}$ into the power set $\mathcal{P}(\Theta)$ of Θ, i.e.

$$\rho : \mathcal{L} \longleftrightarrow \mathcal{P}(\Theta) : A \mapsto \mathbf{A} \quad (\mathbf{A} \subset \Theta) \tag{66}$$

So, every time we have a *Propositional Logic* we have also an inherent accompanying structure or *universe* that we make here *explicit* and name Θ; it is explicitly definable from its sentences $A \in \mathcal{L}$. (This happens *always*, even in strictly two-valued logics; the interplay of propositions inside Θ for that case is studied below, in section 10.)

Though clearly there is no need to name or qualify the members of Θ , we may indulge in calling them *possibilities* (following Shafer [40]) or *eventualities* (to distinguish them from Probability's elementary *events*) or, metaphorically, even *states* or *possible worlds* (or, for that matter, *observers*, [observational] *set-ups*, *observations*, *instants* of time or *stages* of development, elementary *situations* or *contexts* in which things happen, basic *considerations* or *judgments* we have on things, discernible basic expression *modes* or *registers* we adopt to describe them, and so on). Θ is thus configured —in a most general and open-ended way, so as to fit any interpretation— as the real *universe of discourse* or *reference frame* (the set of *possibilities*). Each $\mathbf{A} \in \mathcal{P}(\Theta)$ is a *proposition* (now a rather ambiguous and perhaps devious terminology: $\mathbf{A}$ is, to speak properly, the set-theoretic, extensional counterpart of the proposition A ; in contrast, $A \in \mathcal{L}$ is clearly a linguistic, intensional item). The ρ function is what we call *representation* function. Thus, $\rho(A) = \mathbf{A}$ is the [set] *representation* of the [linguistic] proposition A . Speaking freely, we could say that $\mathbf{A}$ is the *extension* of A, and A the *intension* of $\mathbf{A}$; $\mathbf{A}$ could even be called as well *meaning* or *reference* of A —innocently enough, were these not such loaded words. Since ρ is an isomorphism, there is an inverse function $\rho^{-1} = \delta$ we can call *description* function. Thus, $\delta(\mathbf{A}) = A$ is the [linguistic] proposition that "describes" the [set] proposition $\mathbf{A}$ associated to it. The existence of an isomorphism guarantees there is a *representation* for any [Boolean-structured] *description*, and vice versa.

The set Θ has N elements (where N is the binary logarithm of the order of the $\mathcal{L}$ lattice). Thus, it has as many elements (*possibilities*) as there are *atoms* in $\mathcal{L}$. 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:

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

(The arrows are meant to be read "corresponds isomorphically to"; also, "$\mathbf{A}^c$ " is the set-complement of $\mathbf{A}$, a is an *atom* of $\mathcal{L}$, and θ is an element of Θ.)

The one-to-one correspondence between linguistic, *propositional atoms* (noted a) and *possibilities* (set elements θ) is particularly suggestive. Boolean structure imposes (see (18)) that for any proposition A in $\mathcal{L}$, and for any proposition $\mathbf{A}$ of $\mathcal{P}(\Theta)$, both A and $\mathbf{A}$ are decomposable into their constituent atoms (i.e. those they cover, under their respective lattice order $\vdash$ and $\subset$):

$$\forall A \in \mathcal{L} \quad A = \bigvee_{a_i \vdash A} a_i \ , \qquad \text{and} \qquad \forall \mathbf{A} \subset \Theta \quad \mathbf{A} = \bigcup_{\theta_i \in \mathbf{A}} \{\theta_i\} \ . \tag{67}$$

This could be stated thus: " any proposition A is a disjunction of [mutually incompatible] *elementary propositions*, whose isomorphic counterparts are [in principle mutually independent] *possibilities* within a certain universe of discourse; the union of these gives the extensional, set-theoretical counterpart $\mathbf{A}$ of the original proposition A ". It could be called "*Principle of Analycity*" and it would correspond to the metaprinciple put forward by Wittgenstein in the Tractatus (1921) [45]. Now this classical [and informally stated] thesis is quite naturally formalized and explained in terms of $\mathcal{L}$ (and the universe Θ): it is just an immediate consequence of the *finite* —and thus *atomic*— character of $\mathcal{L}$. The "second Wittgenstein" —explicitly rejecting analycity— would naturally fit into the limit, *infinite* case: whenever $\mathcal{L}$ is [countable but] no longer finite or generable by a finite vocabulary of propositional [so-called] "atoms" (i.e. letters), $\mathcal{L}$ ceases to be atomic, and thus every proposition A ceases to be expressible in terms of propositional atoms —that no longer exist in the structure (in fact infinite downward chains appear)— and the current proposition/set-of-possibilities isomorphism is destroyed.

Dually, in a *finite* $\mathcal{L}$ (and Θ), any proposition A, and particularly any elementary proposition a (or possibility θ), can be analyzed as —"decomposed" into— the conjunction (or intersection) of more complex propositions —*clauses*, in fact— covering a under $\vdash$ (or θ under $\subset$). Thus (again loosely stated, as (18) or (67) above):

$$\forall a \in \mathcal{L} \quad a = \bigwedge\nolimits_{a \vdash A_j} A_j \,, \qquad \text{and} \qquad \forall \theta \in \Theta \quad \{\theta\} = \bigcap\nolimits_{\theta \in \mathbf{A}_j} \mathbf{A}_j \ . \tag{68}$$

This formulation, valid only for finite languages (and universes), fits nicely —better still if we describe the θ as "states" or "possible worlds"— with Carnap's view [8] that a *possible world* —whatever this is intended to mean— is [definable as] "the conjunction of all propositions that are true in it". As before, this would no longer hold for the limiting case of an *infinite* world Θ of possibilities.

Incidentally, the $\mathcal{L}$ / $\mathcal{P}(\Theta)$ isomorphism —in the finite case— lends theoretical legitimacy to the intuitive practice of representing propositions as Venn-diagram domains in an extensional universe, where propositional operations translate into set operations and logical implication is viewed as set inclusion.

We give an elementary example of what this isomorphism can do to trigger an intuitive, extensional visualization of logical propositions, operations and relations. Suppose we have a language $\mathcal{L}$ consisting of 32 propositions and —therefore— a referential frame Θ consisting of five elements (of an unspecified nature) $\theta_0 - \theta_4$. We display here, graphically, four different two-proposition situations, obtained by varying the first (A), with the second (B) fixed. Shown also in each configuration are the five most common composite propositions (naturally, "$A \leftarrow B$" means $B \rightarrow A$). To get a quick extensional identification of propositions, each is shown as the set of characteristic values of its set-representation (a subset of Θ). Furthermore, the traditional 0/1 values of characteristic functions have been replaced by nil/box signs, for enhanced effect.

1) $\rho(A) = \{\theta_0, \theta_1\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 1):

Θ		A	B		$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftarrow B$	$A \leftrightarrow B$
θ_0		■				■		■	
θ_1		■				■		■	
θ_2			■			■	■		
θ_3			■			■	■		
θ_4			■			■	■		

2) $\rho(A) = \{\theta_1, \theta_2\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 2):

Θ		A	B		$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftarrow B$	$A \leftrightarrow B$
θ_0							■	■	■
θ_1		■				■		■	
θ_2		■	■		■	■	■	■	■
θ_3			■			■	■		
θ_4			■			■	■		

3) $\rho(A) = \{\theta_2, \theta_3\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 3):

Θ		A	B		$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftarrow B$	$A \leftrightarrow B$
θ_0							■	■	■
θ_1							■	■	■
θ_2		■	■		■	■	■	■	■
θ_3		■	■		■	■	■	■	■
θ_4			■			■	■		

4) $\rho(A) = \{\theta_3, \theta_4\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 4):

Θ		A	B		$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftarrow B$	$A \leftrightarrow B$
θ_0							■	■	■
θ_1							■	■	■
θ_2			■			■	■		
θ_3		■	■		■	■	■	■	■
θ_4		■	■		■	■	■	■	■

Remarks:

- Observe that, in all, 14 out of the possible 32 propositions that could have been set up in this universe Θ are on display here. (And notice this is *not* a free Boolean algebra.)

- In the first case, since $\rho(A \vee B) = \mathbf{A} \cup \mathbf{B} = \Theta$ (a fact which is visually obvious), we have $\vdash A \vee B$. In the third and fourth we have $\vdash A \rightarrow B$ for similar reasons.

- In the third and fourth cases, clearly, $A \vdash B$ (since $\mathbf{A} \subset \mathbf{B}$, a fact that can be quickly recollected after a simple, casual glance at the tables).

As to the precise nature of the "possibilities" $\theta \in \Theta$, nothing needs to be specified about them except, maybe, that each θ is just a *discernible unit*, merely something that can be isolated and told apart precisely from all other θs ; so Θ is —just— a *frame of discernment* (in Shafer's sense). In such an undemanding and abstract setting we can analyze in detail the behavior and mutual relationship of *propositions* (or, more to the point, *sentences*), that we here define merely as linguistic entities that are members of a [Boolean] structure $\mathcal{L}$ and that have, by the representation theorem, a set-theoretic counterpart. This analytic procedure will give us new insights into the interior workings and structure of what we ordinarily call "propositions", entities we use to consider utterly simple —by its very nature— and generally dismiss without further analysis. As the example above shows, to speak now of negation, conjunction, disjunction, conditional, deductive consequence, etc. means to be able to give an accompanying set-theoretic operation or relation that may be used to illustrate the [initially] purely syntactic interplay among propositions of $\mathcal{L}$.

It is remarkable that when Boole, just after studying classes, set out to analyze *propositions* (in 1847), he conceived them by means of an alternative interpretation of his elective symbol x (already introduced for classes) which —he said— now stood for the *cases* (or "conceivable sets of circumstances") —out of a given hypothetical "universe"— in which the proposition was true. This is stunningly close to our set-extensional representation of propositions in the set Θ of *possibilities*, to which the Stone theorem gives rigorous legitimacy. And, since set propositions are measurable —an idea we shall immediately turn to—, an easy step carries the picture into probability, a step Boole inevitably made when he, in the last chapters of his 1854 book [7], somewhat obscurely likened the product $x \cdot y$ of two propositions to the probability of simultaneaously having both and the sum $x + y$ to the probability of having either [provided both were mutually exclusive].

Something similar occurred later to MacColl [28] when he distinguished between propositions that were *certain*, *variable* or *impossible* (meaning they were always, sometimes or never the case), an idea Peirce [31] a bit later more or less borrowed when he spoke of the need of telling apart *necessary* from *contingent* propositions (the first making up a fraction of them, the second being the general case, see [35]). Logicians' traditional reaction to this approach has always been plainly to deny that what is currently being described here is a proposition at all, but a *propositional function*. This is what Russell did when he reviewed MacColl's work in 1906 —to an unconvinced MacColl's strong protestations when arguing back at Russell—. Such is the conventional answer today, but the Representation Theorem bridges both views by converting a proposition A (a piece of the language $\mathcal{L}$) into a propositional function that takes values in $\{0, 1\}$, the values of a characteristic function, so defining a set $\mathbf{A}$ in a world of *cases* or *circumstances* (Boole's terms, corresponding to what we have called *possibilities*) and distinguishing MacColl's *certain* (or Peirce's *necessary*) propositions (our case $\mathbf{A} = \Theta$) from *variable* or *contingent* ones (the general $\mathbf{A} \subset \Theta$ case).

Though we have just stated that no specification needs to be given about the precise nature of the possibilities θ, one particular —and yet totally general— view of what each θ stands for is of paramount importance here. Suppose $\mathcal{L}$ is *free*. $\mathcal{L}$ has then —provided it is finite— 2^{2^n} elements, is atomic and has 2^n atoms. So $\rho(\mathcal{L}) = \mathcal{P}(\Theta)$ has 2^{2^n} elements and Θ has 2^n. We now assign to each θ the meaning of *interpretation*, in the standard [Propositional Logic] model-theoretic usage of the word. As is known, an interpretation —in this technical, traditional sense— sets the conditions under which a proposition is to be evaluated according to precise truth-functional rules. So, interpretations allow us to distinguish and identify each and every one of the 2^{2^n} propositions in $\mathcal{L}$ (recall that the $\{0, 1\}$-valuations of a Boolean algebra suffice to identify all its members, and note that, given n propositional letters in the [generated] lattice $\mathcal{L}$, there are exactly 2^n atoms in it, just as many as interpretations required for the n letters). This fits nicely with the intuitive idea that the θs are just "discernible units"; it is the case that, from the algebraic standpoint, they really are. The view just expounded of the θs as full-blown logical *interpretations* makes considerably more sense than may appear at first sight, and will be explored immediately (in section 3). We shall stick to it as our standard view of Θ.

However, with an eye to gaining still newer insights into the mechanism and applicability of *propositions*, we give below twelve examples —separately described but not necessarily disjoint— of various interpretation instances of "proposition" and "possibility" that are clearly superfluous but may give further clues to the functioning and inner structure of the members of $\mathcal{L}$. We review these examples summarily and regroup them into five distinct types or "views" (later revisited, in section 4). Given a proposition A, each example gives a distinct meaning to each elementary possibility θ belonging to the extension $\mathbf{A}$ of A.

A. Statistical view of propositions. Examples:

A1. $\theta =$ Affirmative (or negative) response about proposition A from someone being questioned. (Or $\theta =$ Person being questioned: box/nil means a yes/no answer.)

A2. $\theta =$ Observer or information source about fact A.

A3. $\theta =$ Test or observation some experimenter can or does make of proposition A (or validation instance of a [proposed] general law A).

A4. $\theta =$ Context in which proposition A does (or does not) happen.

A5. $\theta =$ Formal consideration or judgment on the proposition A one person (or a plurality) can hold or formulate. (For instance, from a point of view θ, the person judges A to be the case [but not from $\theta' \neq \theta$].)

A6. $\theta =$ Mode or instance of application —in a linguistic context— of A. (A is here taken more as a term or utterance than as the proposition this utterance expresses, and θ is the specific circumstance or linguistic context in which the utterance A is brought into play.) This approach is especially suitable to model *vagueness*, an inherent feature of natural language. Thus, a vague "John is tall" can be "precisified" (Carnap's term [9]) into $\theta =$ "John is 6 ft. tall", $\theta' =$ "John is 6 ft. 4 in. tall" and so on, each to be valued according to the degree the speaker thinks the particular θ describes —or applies to— the original vague sentence. The θs here are Carnap's *precisifications* or Fine's *specifications* (defined as "ways of making a vague expression precise" [15]), or the elementary propositions making up a *communication class* which —as Cresswell considers [12]— any vague expression ultimately reduces to.

B. Probabilistic or Evidential view of propositions. Examples:

B1. $\theta =$ Elementary event (ω) in a [complex] event A. This view, which likens Θ to the sample space Ω of Probability is consistent with several authors' view of Logic as something related in certain senses to Probability, especially Boole (in his propositional characterization of the x variable, in 1847), MacColl (in 1897) and Peirce (c. 1902). Later this view was considerably taken to extremes by Reichenbach [34] and Carnap [9], who at a certain point came to consider "truth", roughly, as probability equaling one.

B2. $\theta =$ A [Shafer's] *possibility* θ in a proposition A inside an evidential *frame of discernment* Θ. All considerations and caveats given by Shafer [40] apply here.

C. "Partial truth" or "Truth content" view of propositions:

C1. θ = Elementary component of a proposition or theory A (in the sense of Haack [18]: A is [partially] true depending on how many of its two-valued components —conjuncts, in fact— are true). Such a treatment may be made compatible with Popper's truth-content and verisimilitude ideas [32], and be subject to related metric considerations (e.g. distance between theories).

D. Modal view of propositions. Examples:

D1. θ = Instant of time (or stage of development) in which A happens. This case is subsumable under the next heading (D2). It is also consistent with Kripke's modeling of Intuitionist Logic [25].

D2. θ = "Possible world" in which A is supposed to happen (or, perhaps, *state* in which A is or occurs). We can incorporate all usual modal descriptions into this model; we can speak, for instance, of *necessity*, *belief* (or *knowledge*), *obligation*, *accessibility* —through program execution—, or *provability* —in Gödel's [modal] sense [17]—, if we use the conventions of the alethic, epistemic, deontic, dynamic or [Gödel's] provability version, respectively, of Modal Logic [11].

E. Properly standard-Logic view of propositions:

E1. θ = *Interpretation* (in the ordinary [Propositional Logic] model-theoretic usage of the word) under which the proposition A is to be evaluated according to precise truth-functional rules. As we said previously, such an apparently outrageous view makes considerably more sense than appears at first sight (especially when $\mathcal{L}$ is *free*), and is the one we shall make our standard (since, indeed, any view we can give about the nature of the θs can be embedded naturally into the 'θs-as-interpretations' frame). It will be explored immediately below (in section 3).

These or other interpretations of the θs are not necessary, but may be useful sometimes. In the context of the views we have just listed, Tables 1-4 above (which we shall reuse and expand later) are perhaps best understood intuitively when interpreted either in Boole's original terms or else in terms of type A. Each θ is e.g. a context or situation (Boole), a test (A3), an observer (A2) or a person answering a questionnaire (A1). In the first case, A "happens" (or "it is the case that A") when a box is present, and the whole A column represents A as a general proposition (or rather as a [propositional] function of its diverse cases of application). In the latter case, a box in its row on column A means the individual has answered yes to the specific request, the A column summarizes the global [i.e. the sample population's] answers to question A, and the five composite propositions represent collective opinions (from those subjects polled); these are logically deduced —or rather [truth-functionally] constructed— from each individual's answers (note there is effectively a standard truth-value computation for each row or observer θ, and note also that this same fact applies to Boole's view of θ as "cases" or "circumstances"). A simple look-up at the tables also allows to capture such [global] facts as $\vdash A \vee B$ or $A \vdash B$ (which now become immediately apparent).

Other views can also explain the tables satisfactorily or supply additional and enriching hints. For instance, the B1 approach can be used to understand Boole's description (in terms of x) of propositions —which, he implies, are ordinarily two-valued but sometimes may fluctuate in-between— a description which insensibly took him into probability. B1 can similarly be used to model MacColl's and Peirce's treatment of propositions as contingent entities —in fact propositional functions (that these authors, like Boole, tended to interpret as "probabilities").

3. POSSIBILITIES AS INTERPRETATIONS

From now on, we shall take $\mathcal{L}$ to be a Boolean algebra *freely generated* by n propositional letters. So, $\mathcal{L}$ will have cardinality 2^{2^n} and, as before (see section 2), it will admit representations in (among others)

(a) $\mathcal{P}(\Theta)$, where Θ has 2^n arbitrary elements.

(b) The set of binary valuations (two-valued epimorphisms) of $\mathcal{L}$.

For the infinite case (i.e. infinitely many propositional letters available), $\mathcal{L}$ turns out to have the same cardinality as the set of letters. If this is countable, $\mathcal{L}$ also is. The same general-case classical [Birkhoff-Stone] theorems apply, but here we have two particular remarks to add:

- First: there is indeed an isomorphism between $\mathcal{L}$ and a [Boolean] algebra of sets $\mathcal{B}$. Here, $\rho(\mathcal{L}) = \mathcal{B}$ and $\mathcal{B} \subset \mathcal{P}(\Theta)$; and Θ (and $\mathcal{B}$) is countably infinite if the set of generators of $\mathcal{L}$ (and $\mathcal{L}$) is. *Nevertheless*, $\mathcal{B}$ does *not* ever equal $\mathcal{P}(\Theta)$.

- Second: the new situation is beautifully characterized by an additional, later (i.e. post-Stone) result by Mostowski and Tarski [30] (and Hanf [21]): "A propositional algebra freely generated by a countable vocabulary of infinitely many propositional [so-called] "atoms" (i.e. letters) is (1) *countably infinite*, (2) *atomless*, and (3) isomorphic to an *interval algebra* (i.e. the [Boolean] algebra generated by the semiopen intervals of a linear order) and also to the clopen algebra of the Cantor continuum $\mathbf{2}^\omega$ ".

(However, as before, we shall not further explore this line of enquiry.)

Now, turning to the finite case, all we said above about Θ holds, but this set now has 2^n elements (where n is the number of propositional letters available in the description language $\mathcal{L}$). Also, Wittgenstein's "analycity" of propositions (67) and Carnap's definition of "possible world" (68) are here valid *only* because the propositional language $\mathcal{L}$ used has a finite number of basic, "atomic" expressions available: they no longer hold when the language we need to describe reality requires a non-finite amount of resources.

Clearly, especially in view of the representation theorem (the (a) and (b) above), the 2^n elements of Θ turn out to be —or can be properly identified with— the model-theoretic standard *interpretations* of Propositional Logic: there are exactly 2^n of them (given n propositional letters), and each interpretation corresponds to one particular string of n 0's and 1's that identifies each one of the 2^n atoms (minterms) of $\mathcal{L}$ (which are one-to-one associates of the 2^n members of Θ).

This view of the θ as the familiar truth-value *interpretations* of elementary logic is what we shall take as *standard* from now on. This is not to exclude all other views of Θ as a set of general, abstract *possibilities* (or, as we mentioned, different entities metaphorically describable as eventualities, states, possible worlds, observers, set-ups, observations, instants, considerations, judgments, modes, expression registers, frames of mind, etc.). In fact, as will be seen later —when we turn our attention back to the valuations of $\mathcal{L}$—, the view that the 2^n elements of Θ are just the common Propositional Logic *interpretations* will subsume naturally all other views of what the elements of Θ can be thought to be, and every *possibility* will just be a particular instance of a *logical interpretation.*

4. TRUTH AS [ADDITIVE] MEASURE

Given that there is a valuation $v : \mathcal{L} \longrightarrow [0,1] : A \mapsto [\![A]\!]$ (defined in section 1, condition number 9) and that there is, also, an isomorphism $\rho : \mathcal{L} \longleftrightarrow \mathcal{P}(\Theta) : A \mapsto \mathbf{A}$ (see the *Representation Theorem* in section 2), clearly v and ρ induce a [0,1]-valued measure (a secondary valuation, in fact) on $\mathcal{P}(\Theta)$:

[Definition:] $\mu : \mathcal{P}(\Theta) \longrightarrow [0,1]$ is the valuation on $\mathcal{P}(\Theta)$ induced by the isomorphism $\rho : \mathcal{L} \longleftrightarrow \mathcal{P}(\Theta)$ in such a way that $\mu = v \circ \rho^{-1}$, i.e. $\mu(\mathbf{A}) = [\![\delta(\mathbf{A})]\!] = [\![A]\!]$. (Note "$[\![\ ]\!]$" is $v(\ )$ and "δ" is ρ^{-1}). (69)

The properties of v (see section 1) are automatically passed on to μ, so that we have, for instance:

$$\mu(\phi) = 0, \quad \text{and} \quad \mu(\Theta) = 1. \tag{70}$$

$$\text{If } \mathbf{A} \subset \mathbf{B} \text{ then } \mu(\mathbf{A}) \leq \mu(\mathbf{B}) \ \textit{(Monotonicity).} \tag{71}$$

$$\text{For any } \mathbf{A} \text{ and } \mathbf{B},\ \mu(\mathbf{A} \cap \mathbf{B}) + \mu(\mathbf{A} \cup \mathbf{B}) = \mu(\mathbf{A}) + \mu(\mathbf{B}) \ \textit{(Additivity).} \tag{72}$$

$$\mathbf{A} \cap \mathbf{B} = \mathbf{A} \quad \text{if and only if} \quad \mu(\mathbf{A} \cap \mathbf{B}) = \mu(\mathbf{A}) \ \textit{for all valuations } \mu. \tag{73}$$

which are the counterpart of formulas (9-11) and (63), respectively, of section 1.

As the counterpart of (12-17) we have, trivially:

$$\mu(\mathbf{A}^c) = 1 - \mu(\mathbf{A}) \tag{74}$$

$$\mu(\mathbf{A} \cap \mathbf{B}) \ \leq \ \mu(\mathbf{A}) \ \leq \ \mu(\mathbf{A} \cup \mathbf{B}) \tag{75}$$

$$\mu(\mathbf{A} \cap \mathbf{B}) \leq \min(\mu(\mathbf{A}), \mu(\mathbf{B})) \tag{76}$$

$$\mu(\mathbf{A} \cup \mathbf{B}) \geq \max(\mu(\mathbf{A}), \mu(\mathbf{B})) \tag{77}$$

$$\mu(\textstyle\bigcup_{i=1}^{i=n} \mathbf{A}_i) = \sum_{i=1}^{i=n} \mu(\mathbf{A}_i) - \sum_{i<j} \mu(\mathbf{A}_i \cap \mathbf{A}_j) + \ldots + (-1)^{n+1}\mu(\bigcap_{i=1}^{i=n} \mathbf{A}_i) \tag{78}$$

$$\text{If for all } i \text{ and } j \ (i \neq j) \ \ \mathbf{A}_i \cap \mathbf{A}_j = \phi \ \text{ then } \ \mu(\textstyle\bigcup_{i=1}^{i=n} \mathbf{A}_i) = \sum_{i=1}^{i=n} \mu(\mathbf{A}_i) \tag{79}$$

Conditions (70) and (79), together with the fact that $\mathcal{P}(\Theta)$ is a [finite] Boolean algebra (and, *a fortiori*, a σ-algebra), clearly characterize the $\mu : \mathcal{P}(\Theta) \longrightarrow [0, 1]$ valuation as an *additive measure* of the sets (propositions) $\mathbf{A}$ of $\mathcal{P}(\Theta)$. It also allows us to see the "truth"-value $[\![A]\!]$ of a proposition A as the *measure* $\mu(\mathbf{A})$ of the set of its corresponding *possibilities*.

The counterpart of (18) is (again loosely formalized):

$$\forall \mathbf{A} \subset \Theta \qquad \mu(\mathbf{A}) = \textstyle\sum_{\theta_i \in \mathbf{A}} \mu(\{\theta_i\}) \, . \tag{80}$$

which tells us that μ is completely defined by the values it assigns to the *possibilities* of Θ (as v by those assigned to the *atoms* of $\mathcal{L}$), and that the measure of any proposition $\mathbf{A} \subset \Theta$ can be computed from them. Also (as (19)):

$$\textstyle\sum_{i=1}^{i=\mathrm{card}\Theta} \mu(\{\theta_i\}) = 1 \, . \tag{81}$$

Four particular cases of the μ and v valuations may be interesting:

1) *Positive measure*: Here, like we said in definition (56), *all* values attributed to the atoms are non-zero. If $\mathcal{L}$ is free —as we have assumed since section 3— with n generators, then for every proposition A and every corresponding set-representation $\mathbf{A}$ all 2^n *interpretations* (elements of Θ) in $\mathbf{A}$ add up their contributions to the numerical value $\mu(\mathbf{A}) = [\![A]\!]$. Thus, we have as an immediate consequence:

$$\models A \quad \text{if and only if} \quad \models_p A \textit{ for all } \textbf{positive} \textit{ valuations } p, \tag{82}$$

a condition which is reducible (see (57)) to this quite undemanding one:

$$\models A \quad \text{if and only if} \quad \models_p A \textit{ for just } \textbf{one} \textit{ positive valuation } p. \tag{83}$$

It can easily be seen that the definition of "$\models A$" given in section 1 is now fully consistent with the traditional definition of "$\models A$", namely: "$\models A$ iff A is true in all [2^n] interpretations".

2) *Uniform measure*: It is a special case of the former. Here all θ in Θ are represented with equal weight —presumably obeying to our ignorance, or perhaps translating the intuitive idea of "equiprobability"—, so that the truth value or measure of a proposition coincides with the relative cardinal of its extension:

$$[\![A]\!] = \frac{\mathrm{card}(\mathbf{A})}{\mathrm{card}(\Theta)}. \tag{84}$$

3) *Non-positive measure*: By this we mean any valuation that *may* have some zero-valued θ. If this does happen, then *not* all interpretations θ necessarily contribute to the overall value of the proposition they are in, so this value lacks the logical significance it had in positive valuations (when non-zero values for all θ were mandatory). In this case only the positive-valued interpretations count (because they contribute something to the truth value), and so we are now free to give the elements of Θ (that we here denote by the more general and uncommitted term *possibilities* rather than by the Logic-biased "interpretations") any meaning we choose. For instance, if we try the *possible world* approach, then those θ that are valued above zero are precisely the worlds (or states) that "count" [in "our" world] (i.e. that determine that a given proposition A is to be considered *necessary*, *believed* (or *known*), *compulsory*, *accessible* [through program execution], *provable* and so on [according to whether our approach is, respectively, alethic, epistemic, deontic, dynamic, proof-theoretic, etc.]). In this way, we can effectively give non-zero values to all worlds accessible (in Kripke's sense) from our given world, and zero values to all others. Thus, we shall have $[\![A]\!] = 1$ if and only if A is true in all accessible worlds (i.e. is "necessary"). Conversely, if we are given a modal operator $\Box$ with any precisely-defined intended meaning, we can give a semantic characterization of it through the appropriate v or μ valuation (that will then prescribe which atoms or worlds precisely are to be assigned non-zero values), so that:

$$\Box A \quad \text{if and only if} \quad [\![A]\!] = 1 \text{ [for that particular valuation]} \tag{85}$$

If we vary the valuation, the meaning of $\Box A$ varies accordingly. Naturally, if the valuation is what we have called *positive* we fall back on the "$\models A$" case (as it should be: logical truth, after all, is an obvious instance of modal necessity, and the latter should include the former as a special case).

4) *Selective measure*: It is a special instance of the latter case. It consists of a zero-valuation for all possibilities in Θ except for *one* particular θ (say θ_r), which for obvious reasons (i.e. additivity) is assigned 1 as value. Such a valuation allows us to select a *possibility* (possible world) θ_r among those in the universe Θ. This case, also, corresponds to the simple *two-valued* case of *ordinary propositional logic*, as here "$[\![A]\!] = 1$" simply expresses the assertion "A is true" (and "$[\![A]\!] = 0$" "A is false"). It also means that we live in —precisely— θ_r (the "real world"), where everything happens (and nothing happens elsewhere). We shall have more of this later (see section 10).

We now work and expand the example we gave in section 2 (note however that, for simplicity reasons —and against our own advice in section 3—, the Boolean algebra used there has only five atoms and is therefore *not* free). Suppose that, for illustration purposes, we value the elements of Θ in four different ways (μ_1-μ_4), and that we make the measure μ_1 *uniform*, μ_2 *positive*, μ_3 *non-positive*, and μ_4 *selective*. So we get:

(Table 5)

Θ	μ_1	μ_2	μ_3	μ_4
θ_0	.2	.10	0	0
θ_1	.2	.15	0	0
θ_2	.2	.20	.6	1
θ_3	.2	.25	.3	0
θ_4	.2	.30	.1	0
$\mu(\Theta) =$	1	1	1	1

1) $\rho(A) = \{\theta_0, \theta_1\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 6):

A	μ_1	μ_2	μ_3	μ_4	B	μ_1	μ_2	μ_3	μ_4
■	.2	.10	0	0					
■	.2	.15	0	0					
					■	.2	.20	.6	1
					■	.2	.25	.3	0
					■	.2	.30	.1	0
$[\![A]\!] =$	.4	.25	0	0	$[\![B]\!] =$	.6	.75	1	1

$A \wedge B$	μ_1	μ_2	μ_3	μ_4	$A \vee B$	μ_1	μ_2	μ_3	μ_4
					■	.2	.10	0	0
					■	.2	.15	0	0
					■	.2	.20	.6	1
					■	.2	.25	.3	0
					■	.2	.30	.1	0
$[\![A \wedge B]\!] =$	0	0	0	0	$[\![A \vee B]\!] =$	1	1	1	1

$A \to B$	μ_1	μ_2	μ_3	μ_4	$A \leftrightarrow B$	μ_1	μ_2	μ_3	μ_4
[redacted]	.2	.20	.6	1					
[redacted]	.2	.25	.3	0					
[redacted]	.2	.30	.1	0					
$[\![A \to B]\!] =$	.6	.75	1	1	$[\![A \leftrightarrow B]\!] =$	0	0	0	0

2) $\rho(A) = \{\theta_1, \theta_2\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 7):

A	μ_1	μ_2	μ_3	μ_4	B	μ_1	μ_2	μ_3	μ_4
[redacted]	.2	.15	0	0					
[redacted]	.2	.20	.6	1	[redacted]	.2	.20	.6	1
					[redacted]	.2	.25	.3	0
					[redacted]	.2	.30	.1	0
$[\![A]\!] =$	.4	.35	.6	1	$[\![B]\!] =$	.6	.75	1	1

$A \wedge B$	μ_1	μ_2	μ_3	μ_4	$A \vee B$	μ_1	μ_2	μ_3	μ_4
					[redacted]	.2	.15	0	0
[redacted]	.2	.20	.6	1	[redacted]	.2	.20	.6	1
					[redacted]	.2	.25	.3	0
					[redacted]	.2	.30	.1	0
$[\![A \wedge B]\!] =$	.2	.20	.6	1	$[\![A \vee B]\!] =$	.8	.90	1	1

$A \to B$	μ_1	μ_2	μ_3	μ_4	$A \leftrightarrow B$	μ_1	μ_2	μ_3	μ_4
[redacted]	.2	.10	0	0	[redacted]	.2	.10	0	0
[redacted]	.2	.20	.6	1	[redacted]	.2	.20	.6	1
[redacted]	.2	.25	.3	0					
[redacted]	.2	.30	.1	0					
$[\![A \to B]\!] =$	.8	.85	1	1	$[\![A \leftrightarrow B]\!] =$	.4	.30	.6	1

3) $\rho(A) = \{\theta_2, \theta_3\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 8):

A	μ_1	μ_2	μ_3	μ_4	B	μ_1	μ_2	μ_3	μ_4
[redacted]	.2	.20	.6	1	[redacted]	.2	.20	.6	1
[redacted]	.2	.25	.3	0	[redacted]	.2	.25	.3	0
					[redacted]	.2	.30	.1	0
$[\![A]\!] =$	.4	.45	.9	1	$[\![B]\!] =$	.6	.75	1	1

$A \wedge B$	μ_1	μ_2	μ_3	μ_4
■	.2	.20	.6	1
■	.2	.25	.3	0
$[\![A \wedge B]\!] =$	.4	.45	.9	1

$A \vee B$	μ_1	μ_2	μ_3	μ_4
■	.2	.20	.6	1
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A \vee B]\!] =$	.6	.75	1	1

$A \rightarrow B$	μ_1	μ_2	μ_3	μ_4
■	.2	.10	0	0
■	.2	.15	0	0
■	.2	.20	.6	1
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A \rightarrow B]\!] =$	1	1	1	1

$A \leftrightarrow B$	μ_1	μ_2	μ_3	μ_4
■	.2	.10	0	0
■	.2	.15	0	0
■	.2	.20	.6	1
■	.2	.25	.3	0
$[\![A \leftrightarrow B]\!] =$	.8	.70	.9	1

4) $\rho(A) = \{\theta_3, \theta_4\}$, $\rho(B) = \{\theta_2, \theta_3, \theta_4\}$ (Table 9):

A	μ_1	μ_2	μ_3	μ_4
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A]\!] =$	.4	.55	.4	0

B	μ_1	μ_2	μ_3	μ_4
■	.2	.20	.6	1
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![B]\!] =$	.6	.75	1	1

$A \wedge B$	μ_1	μ_2	μ_3	μ_4
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A \wedge B]\!] =$	.4	.55	.4	0

$A \vee B$	μ_1	μ_2	μ_3	μ_4
■	.2	.20	.6	1
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A \vee B]\!] =$	.6	.75	1	1

$A \rightarrow B$	μ_1	μ_2	μ_3	μ_4
■	.2	.10	0	0
■	.2	.15	0	0
■	.2	.20	.6	1
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A \rightarrow B]\!] =$	1	1	1	1

$A \leftrightarrow B$	μ_1	μ_2	μ_3	μ_4
■	.2	.10	0	0
■	.2	.15	0	0
■	.2	.25	.3	0
■	.2	.30	.1	0
$[\![A \leftrightarrow B]\!] =$	.8	.80	.4	0

(Note that, due to lack of table space, we have done without the "$A \leftarrow B$" proposition.)

It is easy to verify that *all* relevant formulas of section 1 (i.e. (9-24) and (40-45)) are satisfied in the four cases 1-4. Moreover, observe that in case 1,

$$[\![A \wedge B]\!] = \max(0, [\![A]\!] + [\![B]\!] - 1) \qquad \text{and} \qquad [\![A \vee B]\!] = \min(1, [\![A]\!] + [\![B]\!]) \tag{86}$$

for all four measures μ_1-μ_4; in cases 3 and 4 we have:

$$[\![A \wedge B]\!] = \min([\![A]\!], [\![B]\!]) \qquad \text{and} \qquad [\![A \vee B]\!] = \max([\![A]\!], [\![B]\!]) \tag{87}$$

also for all four measures. These results, that will be rigorously justified in section 6, already suggest that the classical logical connectives (here $\wedge$ and $\vee$) are *non-functional*, and that their values (i.e. the truth values of the composite propositions constructed with them) depend on the "relative position" —a concept that we shall later parameterize accurately— of **A** and **B** (inside Θ). We also have (note the left-hand condition is satisfied in cases 3 and 4) that:

$$A \vdash B \quad \Rightarrow \quad [\![A \rightarrow B]\!] = 1 \tag{88}$$

for the four —actually for all— valuations (as it should be).

Intuitively, the measure $\mu(\{\theta\})$ of each individual θ in the Θ universe should seemingly correspond to the relative importance or the relevance this individual has in that universe. Thus, in a view of Θ of the A2 type (see section 2) where the θ are *observers*, $\mu(\{\theta\})$ would represent the importance a "superobserver" assigns to each particular θ. In a *tests* (A3) or *possible worlds* (D2) view, $\mu(\{\theta\})$ would be the relevance attributed to test θ or the degree of realizability of the given possible world. And so on.

In our standard view of the θs as *interpretations* (E1), we would have:

E1. $\mu(\theta)$ = Relevance of interpretation. $\mu(\mathbf{A}) = [\![A]\!]$ = Degree of [logical] validity of A. (Note measure μ should be *positive* here.) We have:

$$\text{“} \models A \quad \text{if and only if} \quad [\![A]\!] = 1 \text{”} .$$

If the valuation were not positive but *selective* then we would have:

$$\text{“ } A \text{ is true} \quad \text{if and only if} \quad [\![A]\!] = 1 \text{”} .$$

For general, non-positive valuations we would have three situations, characterized by:

I) $[\![A]\!] = 0$

II) $[\![A]\!] \in (0,1)$

III) $[\![A]\!] = 1$

that would correspond to what, historically, all formalizations of *Three-valued Logic* have characterized as:

I) 'A is false'

II) 'A is undetermined' —either because the truth value of A is [existent but] unavailable or because it does not exist (i.e. there is a "truth-value gap")

III) 'A is true'.

We shall stick to this reading of "truth values" in three-valued Logics later, when we turn back to analyze their truth tables (in section 8).

Now we turn back to our list of restricted views A-D. With views of type A in mind, Table 5 (that lists the twenty values assigned to the five possibilities θ_0-θ_4 by the four valuations μ_1-μ_4) allows, through the relatively smooth transition it suggests between the uniform μ_1 and the selective μ_4, the modeling of the process of scientific *explication* (Carnap's term). We begin in a situation with poor evaluation criteria —so we give each possibility θ (observation, possible explanation, etc.) equal weight—, then we progressively discard some possibilities made irrelevant by the enquiry until we focus on a unique θ (here θ_2). Under this perspective, A would be a proposition or theory under elaboration or discussion, Θ would be the set of, say, experiments, contexts or polemizers, each with its own independent verdict on the theory being discussed, and each being itself evaluated (according to likelihood, seriousness, relevance, ...) by those who globally appraise the theory in its state of progress. Such a process —being eliminative— would refine the theory, as usual in science (and also in public opinion: different observers can, even maintaining their views, be weighed equally at first by the public and then progressively one or several opinion-leaders can arise and consolidate); but the inverse process is also possible: a single explanation or observer is given utmost importance at first (probably because of lack of alternatives), then more explanations are raised or more observers have own opinions.

To give an idea of the range of applicability of the $\mathcal{L}/\mathcal{P}(\Theta)$ isomorphism, we next reexamine the list of the twelve different examples we suggested in section 2.

A1. $\mu(\theta)$ = Weight given to the person being questioned. $\mu(\mathbf{A}) = [\![A]\!]$ = State of opinion (of those being questioned) about A. This is an intuitive and trivial view of Θ, with a direct application to questionnaire theory. Each distinct μ or v valuation corresponds to a different aggregation function (and so μ_1 is an arithmetic mean while μ_4 is a selection function). Note that our conception of Θ has been based so far on an implicit assumption: that opinions are mutually independent; if individual opinions influence each other or there are collective/individual interferences then the additivity assumption (72) should be relaxed. (We shall do that in section 11.)

A2. $\mu(\theta)$ = Observer's credibility. $\mu(\mathbf{A}) = [\![A]\!]$ = Consensus over A. This is a variant of A1 (with a possible application to Science Theory) according to which the "truth" of a proposition or theory A is the result of consensus among (e.g.) scientists at a given moment. This can be seen in agreement with a certain simplified theorization of scientific truth in the Kuhnian tradition.

A3. $\mu(\theta)$ = Evaluation or reliability of test. $\mu(\mathbf{A}) = [\![A]\!]$ = "Truth" of A. Here $[\![A]\!] = 1$ amounts to assert that A is true *by induction* on the N (or 2^n) observed cases. Obviously, it is *limited* induction, because it ranges just over a finite number of cases, over cases that were previously selected (by choosing Θ) and, moreover, with an attached weighing function presumably reflecting the relevance of each observation. With these constraints in mind, this adequately formalizes *induction*. Nevertheless, when possible observations are potentially infinite (the most frequent situation), we are again in an infinite $\mathcal{L}$, and all the appropriate caveats apply.

A4. $\mu(\theta)$ = Relative importance of given context. $\mu(\mathbf{A}) = [\![A]\!]$ = Global value of A considering all the relevant contexts. (This is consistent with Boole's presentation.)

A5. $\mu(\theta)$ = Weight or relevance of given consideration. $\mu(\mathbf{A}) = [\![A]\!]$ = Result of a complex consideration or judgment on A. (A model, maybe, for "subjective probability".)

A6. $\mu(\theta)$ = Likelihood of application. $\mu(\mathbf{A}) = [\![A]\!]$ = Degree of applicability (or frequency of application) of the utterance A in some given circumstance or linguistic context. As we said, this approach is especially suitable to model the *vagueness* of natural language. Thus, a vague "John is tall" can be "precisified" or "specified" into —or [implicitly] communicated as— θ = "John is 6 ft. tall", θ' = "John is 6 ft. 4 in. tall" and so on, each to be valued according to the degree the speaker thinks the particular θ describes —or applies to— the original vague sentence. "Fuzzy-set"-theory practitioners would no doubt say that $[\![A]\!]$ is the "truth value" of the [vague] sentence A.

B1. $\mu(\theta)$ = Additive valuation of event θ (i.e. Probability, in the current mathematical sense). $\mu(\mathbf{A}) = [\![A]\!]$ = Probability of A. (Nothing to add here; we are in the mainstream of an intensively explored subject, and we can take advantage of all its methods and insights.)

B2. $\mu(\theta)$ = Bayesian (i.e. additive) valuation of *possibility* θ. $\mu(\mathbf{A}) = [\![A]\!]$ = *Belief* in A (in Shafer's sense). (Also nothing to add, except for the fact that the additivity (72) we currently assume assimilates our case to what Shafer calls *Bayesian* beliefs; section 11 will loosen additivity, so that our analysis will be coextensional with Shafer's.)

C1. $\mu(\theta)$ = Value or relevance of the given component of the theory. $\mu(\mathbf{A}) = [\![A]\!]$ = Partial truth (in Haack's sense) of A, or "Truth content" (in Popper's). (Nothing to add to what was said on this view before.)

D1. $\mu(\theta)$ = Relevance of the given instant (or stage). $\mu(\mathbf{A}) = [\![A]\!]$ = Value of A in time (or "mathematical truth" in the intuitionists' sense).

D2. $\mu(\theta)$ = Relevance of world. $\mu(\mathbf{A}) = [\![A]\!]$ = Degree of contingency of A. Here we have that $[\![A]\!] = 1$ means "A is necessary", that $[\![A]\!] \in (0,1)$ means "A is contingent", and that $[\![A]\!] = 0$ means that "A is impossible" (thus abusing the alethic terminology; actually, proper terms should be used for non-alethic modal contexts). Dually, we would have $[\![A]\!] \neq 0$ corresponding to "A is possible" or "$\diamond A$". (It is noticeable that P. Henle gave this same meaning for $\diamond A$ when interpreting Lewis's $S5$ in the Boole-Schröder algebra.)

E1. (See above, before A1).

Given that our standard view is that $\mathcal{L}$ is *free*, generated by n propositional letters, any other view (like e.g. A1) is subsumable into it by considering that any possible proposition is generated by n letters, so that there are 2^{2^n} such propositions. If there are 2^N possible answers to question A by N persons being polled, there must be n letters and 2^n possibilities θ such that $N \leq 2^n$ (this is necessary to be able to distinguish all N individuals). At least $2^n - N$ possibilities not representing persons will be valued zero. If only one person is questioned, his or her answer will automatically yield the truth value $[\![A]\!]$ of A. If there are N (> 1) of them, they may disagree (and then $[\![A]\!]$ will be a weighed mean of their opinions). Or else they may all agree; in that case $[\![A]\!] = 1$ will signal not only truth but solid unanimity. Similarly for the rest of views (A-D).

It is to be remarked that, by the representation theorem (66), we *always* have a set structure $\mathcal{P}(\Theta)$ when we use a propositional language $\mathcal{L}$. If $\mathcal{L}$ is freely generated by n letters then $\mathcal{L}$ contains 2^{2^n} distinct elements (propositions) and automatically we have as their set counterparts 2^{2^n} subsets of Θ, where Θ is the set of 2^n elements (of unspecified nature) that correspond to the 2^n atomic propositions in $\mathcal{L}$. Though we are generally uninterested in the meaning of this universe Θ and its "possibilities" or "eventualities" θ, they are there to take, and possibly to help us in figuring out or interpreting better how the [linguistic] propositions of $\mathcal{L}$ behave. For instance, suppose the generators of $\mathcal{L}$ (a lattice of theories) are just a vocabulary of n elementary theories (posing as basic explanations for observed facts [and also as prediction components for yet unobserved ones]). There *is* —by the representation theorem— a universe Θ of 2^n possibilities (of still unknown nature) whose 2^{2^n} different combinations (sets) correspond to all 2^{2^n} theories in $\mathcal{L}$. An easy interpretation of Θ is this: each θ is a corroboration instance of each atomic theory of $\mathcal{L}$. Given a theory A of $\mathcal{L}$, A is expressed by a syntactical combination of the elementary theories that generate $\mathcal{L}$ and it simultaneously has a set counterpart: the set $\mathbf{A}$ of eventualities θ where A is found to hold. Thus, it is just natural to value the "truth" of A by measuring the θs in $\mathbf{A}$, so we have $[\![A]\!] = \mu(\mathbf{A})$. Moreover, we have that if, for instance, for a given pair of theories A and B we find that $A \vdash B$, then we also find that $\mathbf{A} \subset \mathbf{B}$, so that when we follow a reasoning down the $\vdash$ direction, we also can be sure that, as it concerns possibilities, the ones contained in $\mathbf{A}$ are also in $\mathbf{B}$, so that the reasoning process monotonically maintains possibilities (if the θ were distinct respondents to the A question, $A \vdash B$ would mean they are faithful to their own past affirmative responses); and vice versa: if all possibilities (or yes answers) were not preserved when going from A to B, then valid reasoning could not proceed in that direction. Such considerations may help to illustrate an underlying process going on whenever we —apparently only syntactically— use propositions.

5. RELATIVE TRUTH

Suppose we want to express the conjunction value as a product:

$$[\![A \wedge B]\!] = [\![A]\!] \cdot \tau \qquad \text{(or } [\![A \wedge B]\!] = \tau' \cdot [\![B]\!]) \tag{89}$$

We have (provided $[\![A]\!] \neq 0$):

$$\tau = \frac{[\![A \wedge B]\!]}{[\![A]\!]} = \frac{v(A \wedge B)}{v(A)} = v_A(B) \tag{90}$$

which yields on $\mathcal{L}$ a new valuation $v_A : \mathcal{L} \longrightarrow [0,1]$ with the same properties as the original valuation v (indeed v_A satisfies equations (9) and (11), as is easy to prove).

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

$$\tau = \frac{\mu(\mathbf{A} \cap \mathbf{B})}{\mu(\mathbf{A})} \tag{91}$$

Thus, the new valuation takes in account, out of every subset of Θ, only the part contained in $\mathbf{A}$, and it gives it a value related only to that part. So we define τ as the *relative truth* " $[\![B|A]\!]$ " (i.e. the "truth of B relative to [or conditioned on] A"):

[*Definition:*] *Relative truth* of B with respect to A is the quotient

$$[\![B|A]\!] = \frac{[\![A \wedge B]\!]}{[\![A]\!]} \qquad ([\![A]\!] \neq 0) \tag{92}$$

[*Remark*: Here we break with our own implicit notational convention according to which double brackets separate the syntax (inside) from the outer world of semantics; now, what is inside is not syntax —clearly it is not a wff (the stroke there is not Sheffer's)— but a bungled expression, a half-hearted concession to the time-honored 'conditional probability' notation.]

If $[\![B|A]\!] = [\![B]\!]$ (which, incidentally, implies that also $[\![B|\neg A]\!] = [\![B]\!]$), then we say that A and B are *independent* (because the valuation $v_A(B) = v(B)$ remains unaffected by A). In that case, the conjunction can be expressed as the product:

$$[\![A \wedge B]\!] = [\![A]\!] \cdot [\![B]\!] \tag{93}$$

(and then we have also:

$$[\![A \rightarrow B]\!] = 1 - [\![A]\!] + [\![A]\!] \cdot [\![B]\!] = 1 - [\![A]\!] \cdot [\![\neg B]\!] \quad). \tag{94}$$

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 symmetrical. We have (always assuming $[\![A]\!] \neq 0$):

$$[\![A]\!] \cdot [\![B|A]\!] = [\![B]\!] \cdot [\![A|B]\!] = [\![A \wedge B]\!] \tag{95}$$

$$[\![A|A]\!] = 1 \tag{96}$$

$$[\![A|\neg A]\!] = [\![\neg A|A]\!] = 0 \tag{97}$$

$$[\![A \wedge B|A]\!] = [\![B|A]\!] = [\![A \rightarrow B|A]\!] \tag{98}$$

$$[\![A \vee B|A]\!] = 1 = [\![A \rightarrow B|B]\!] \tag{99}$$

$$[\![B|A]\!] + [\![\neg B|A]\!] = 1 \quad \textit{(Complementation)} \tag{100}$$

$$[\![B]\!] = [\![A]\!] \cdot [\![B|A]\!] + [\![\neg A]\!] \cdot [\![B|\neg A]\!] \quad \textit{(Distribution)} \tag{101}$$

$$[\![A \rightarrow B]\!] = 1 - [\![A]\!] \cdot [\![\neg B|A]\!] \tag{102}$$

$$[\![B|A]\!] = \frac{[\![A \rightarrow B]\!]}{[\![A]\!]} - \frac{[\![A \rightarrow \neg A]\!]}{[\![A]\!]} = 1 - \frac{1 - [\![A \rightarrow B]\!]}{[\![A]\!]} \quad . \tag{103}$$

Note that, in general,

$$[\![B|A]\!] \neq [\![A \rightarrow B]\!] \tag{104}$$

Particularly, we have *always*

$$[\![B|A]\!] \ < \ [\![A \rightarrow B]\!] \tag{105}$$

except when either $[\![A]\!] = 1$ or $[\![A \rightarrow B]\!] = 1$, in which cases (and they are the *only* ones):

$$[\![B|A]\!] = [\![A \rightarrow B]\!] \ .$$

We have also:

$$[\![B|A]\!] = 1 \quad \text{if and only if} \quad A \models_v B \tag{106}$$

$$[\![B|A]\!] = 0 \quad \text{if and only if} \quad A \models_v \neg B$$

Naturally, from (106) we get: $A \models B$ iff $(\forall v)\ [\![B|A]\!] = 1$, and so: $A \models B \Rightarrow [\![B|A]\!] = 1$ (107)
(Also, as (57) or (83): $A \models B$ iff $(\exists p)\ [\![B|A]\!] = 1$, where p is a *positive* valuation.)

Two remarks are in order. First: formula (95) is the logical equivalent of *Bayes' formula*, but note there is no privileged direction here, i.e. neither A nor B can be thought of as "prior" or "posterior" to the other (actually, such terms would make no special sense in this exclusively logic context), and the formula is —contrasting with its interpretation problems in Probability— unproblematically symmetric. Second: the (104) inequality together with formula (95) suggest a reason for the current $P(A \rightarrow B)\ /\ P(B|A)$ confusion incurred when trying to measure [probabilistically] the [logical] $A \rightarrow B$ statement. Such cavalier identification may plausibly arise from a straightforward translation from $A \wedge (A \rightarrow B)$ —which occurs in Modus Ponens and is logically equivalent to $A \wedge B$— into $[\![A]\!] \cdot [\![B|A]\!]$ —which equals the value of $A \wedge B$—, thus turning a [logical] *meet* into a [numerical] *product* along the way, and —on the whole— unwittingly assimilating $A \rightarrow B$ to its value $[\![A \rightarrow B]\!]$ and then to $[\![B|A]\!]$, which is then multiplied by [the value of] A. To see that, just compare the following two valid decompositions (the first of which is an elementary Boolean property, while the second is formula (95) again):

$$[\![A \wedge B]\!] \;=\; [\![A \wedge (A \rightarrow B)]\!] \text{ (note that } [\![A \wedge (A \rightarrow B)]\!] \;\neq\; [\![A]\!] \cdot [\![A \rightarrow B]\!] \text{, see below)}$$
$$[\![A \wedge B]\!] \;=\; [\![A]\!] \cdot [\![B|A]\!]$$

where it can be noticed that, to get $[\![A \rightarrow B]\!]$ identified with $[\![B|A]\!]$, a [logical] meet between propositions must be transmuted into a [numerical] product between values, and this can be done in a value-preserving way only if A and $A \rightarrow B$ are independent —which is emphatically *not* the case (see below, after table 12) except when either proposition is binary-valued; also, the $[\![A \rightarrow B]\!]$ must be ultimately assimilated to $[\![B|A]\!]$. But "$B|A$" neither is the same thing as "$A \rightarrow B$" nor has the same value or meaning, as we saw above.

An example, taken from typical probability scenarios, will give the feeling of what we mean by *relative truth* (or what we could as well call "conditional truth", after the model of "conditional probability"). Suppose we toss a coin three times; if we represent 'heads' by '$\odot$' and 'tails' by '$\oplus$' then we have eight possible outcomes:

$$\theta_0 \equiv \odot\odot\odot\,,\ \theta_1 \equiv \odot\odot\oplus\,,\ \theta_2 \equiv \odot\oplus\odot\,,\ \theta_3 \equiv \odot\oplus\oplus\,,$$
$$\theta_4 \equiv \oplus\odot\odot\,,\ \theta_5 \equiv \oplus\odot\oplus\,,\ \theta_6 \equiv \oplus\oplus\odot\,,\ \theta_7 \equiv \oplus\oplus\oplus\,.$$

(So we have an eight-element frame Θ.) We now express the following facts:

$A \equiv$ 'First toss is heads'.

$B \equiv$ 'Second toss is heads'.

$C \equiv$ 'Exactly two consecutive heads are obtained'.

Naturally,

$$\begin{aligned} \mathbf{A} &= \{\theta_0, \theta_1, \theta_2, \theta_3\} \\ \mathbf{B} &= \{\theta_0, \theta_1, \theta_4, \theta_5\} \\ \mathbf{C} &= \{\theta_1, \theta_4\}\ . \end{aligned} \tag{108}$$

So we have (assume μ is uniform):

(Table 10)

Θ	A	B	C	$A \wedge B$	$A \wedge C$	$B \wedge C$
θ_0	■	■		■		
θ_1	■	■	■	■	■	■
θ_2	■					
θ_3	■					
θ_4		■	■			■
θ_5		■				
θ_6						
θ_7						
$\mu =$	1/2	1/2	1/4	1/4	1/8	1/4

where we note that:

$$[\![A \wedge B]\!] = 1/4 = [\![A]\!] \cdot [\![B]\!]$$
$$[\![A \wedge C]\!] = 1/8 = [\![A]\!] \cdot [\![C]\!]$$
$$[\![B \wedge C]\!] = 1/4 \;\neq\; 1/8 = [\![B]\!] \cdot [\![C]\!]\,.$$

According to this, A and B are mutually independent, and so are A and C, but not B and C, because C is conditioned by B.

Not that this example is anything new: on the contrary, it is a most trivial and commonplace exercise in Probability. Yet, suppose that we now vary slightly the problem and we say, not that three consecutive tosses are made (which involves a *time* sequence), but that we have a three-coin array (so now, e.g. ⊙⊕⊙ means there is 'tails' in the *center* coin) and that three possible coin configurations particularly interest us:

$A \equiv$ 'Leftmost coin is heads'.

$B \equiv$ 'Center coin is heads'.

$C \equiv$ 'Exactly two coins in contact have heads in them'.

These are obviously the same A, B, C as before, but they now refer to a situation in space, and there is no time or action involved at all. Thus, Probability's traditional way of describing the relationship between B and C as a "*conditioning* (or *chaining*) of one [previous] event (B) on another [subsequent] event (C)" is rather at odds with the absence of the time dimension, because here the *dependence* of B and C is due strictly to the spatial configuration of the coins (and not to any "previous" event). Moreover, this dependence is mutual: C's truth also influences B's (notice that $C \vdash B$, as clearly shown in the table); labeling this reverse influence as "conditioning" would give the description, even in the tossed-coins interpretation, a rather contrived unnaturality (because the fact that C happens presupposes that B and A have already happened).

And note that the inherent symmetry we have pointed out in logical *dependence* —so natural and easy to accept (see e.g. (95), a logical "Bayes formula")— is mirrored, in Probability, in the unease probabilists feel at the definition, distinction and interplay of *a priori* and *a posteriori* probabilities, two symmetric concepts that —because Probability so much subliminally relies on *time* (and because time is a one-way arrow)— prove so hard to accept as such (i.e. as indeed *symmetric*).

Actually, we suggest that the terms "*logical dependence* between two propositions A and B" or else "*truth* of some proposition A *relative* to some other proposition B" is a more accurate —and general— description of fact than "*conditioning* of event B on event A". Particularly, the proposed formulation avoids the usual time-dependent and frequentist connotations so often associated to probabilistic terminology —especially when described phenomena are of a purely logical nature. We also claim that this assimilation of the 'conditional probability' concept into Logic is, aside from practical applications, enriching for both Logic and Probability theories. From the conceptual side, it allows to recognize and establish the clear-cut *distinction* between the [generally diverging] values " $[\![B|A]\!]$ " and " $[\![A \to B]\!]$ " (see (104)). It also helps to clarify the sharp-edged *difference* between the " $[\![B|A]\!] = 1$" and " $A \models B$ " situations (which —though they always go hand in hand, see (106-107)— are conceptually *distinct*).

All that was said in this section about truth relativization of B by a "conditioning" proposition A is extensible to the general case where there is not one but n propositions $A_1, \ldots, A_n$ that bear on B's truth. We call these propositions —which could easily be the axioms of a certain theory (and B a potential theorem)— *context*. Then definition (92) generalizes into:

[*Definition:*] *Relative truth* of B with respect to a *context* $A_1, \ldots, A_n$ is the quotient

$$[\![B|A_1, \ldots, A_n]\!] = \frac{[\![A_1 \wedge \ldots \wedge A_n \wedge B]\!]}{[\![A_1 \wedge \ldots \wedge A_n]\!]} \qquad ([\![A_1 \wedge \ldots \wedge A_n]\!] \neq 0) \tag{109}$$

6. CONNECTIVES AND PROPOSITIONAL STRUCTURE

The goal now is to find the truth value of composite propositions in $\mathcal{L}$. For the negation connective this is easy: it is given by formula (12). For the rest we have the three following formulas that are a direct spin-off of additivity (11) and the definitions (20) and (39):

$$[\![A \vee B]\!] = [\![A]\!] + [\![B]\!] - [\![A \wedge B]\!] \tag{110}$$

$$[\![A \to B]\!] = 1 - [\![A]\!] + [\![A \wedge B]\!] \tag{111}$$

$$[\![A \leftrightarrow B]\!] = 1 - [\![A]\!] - [\![B]\!] + 2 \cdot [\![A \wedge B]\!] \tag{112}$$

(formulas (111) and (112) are (21) and (42) again).

So the problem now 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 propositions A and B.

For any $\mathbf{A} = \rho(A)$ and $\mathbf{B} = \rho(B)$ we have, obviously:

$$\phi \subset \mathbf{A} \cap \mathbf{B} \subset \mathbf{A} \subset \mathbf{A} \cup \mathbf{B} \subset \Theta \tag{113}$$

and, because of the monotonicity (71) of μ:

$$0 \leq \mu(\mathbf{A} \cap \mathbf{B}) \leq \mu(\mathbf{A}) \leq \mu(\mathbf{A} \cup \mathbf{B}) \leq 1 \tag{114}$$

In (113) there is a smooth, comprehensive gradation of possible cases. We distinguish four extreme cases; the first two concern intersection, the third and the fourth union:

- Case 1: $\mathbf{A} \cap \mathbf{B} = \phi$. We have:

$$\mu(\mathbf{A} \cap \mathbf{B}) = 0 \tag{115}$$

- Case 2: Subcase (a): $\mathbf{A} \cap \mathbf{B} = \mathbf{A}$ (i.e. $\mathbf{A} \subset \mathbf{B}$). We have:

$$\mu(\mathbf{A} \cap \mathbf{B}) = \mu(\mathbf{A}) \leq \mu(\mathbf{B})$$

or Subcase (b): $\mathbf{A} \cap \mathbf{B} = \mathbf{B}$ (i.e. $\mathbf{B} \subset \mathbf{A}$). We have:

$$\mu(\mathbf{A} \cap \mathbf{B}) = \mu(\mathbf{B}) \leq \mu(\mathbf{A})$$

In either subcase we have:

$$\mu(\mathbf{A} \cap \mathbf{B}) = \min\,[\mu(\mathbf{A}), \mu(\mathbf{B})] \tag{116}$$

- Case 3: Subcase (a): $\mathbf{A} \cup \mathbf{B} = \mathbf{A}$ (i.e. $\mathbf{B} \subset \mathbf{A}$). We have:

$$\mu(\mathbf{A} \cup \mathbf{B}) = \mu(\mathbf{A}) \geq \mu(\mathbf{B})$$

or Subcase (b): $\mathbf{A} \cup \mathbf{B} = \mathbf{B}$ (i.e. $\mathbf{A} \subset \mathbf{B}$). We have:

$$\mu(\mathbf{A} \cup \mathbf{B}) = \mu(\mathbf{B}) \geq \mu(\mathbf{A})$$

In either subcase we have:

$$\mu(\mathbf{A} \cup \mathbf{B}) = \max\,[\mu(\mathbf{A}), \mu(\mathbf{B})] \tag{117}$$

- Case 4: $\mathbf{A} \cup \mathbf{B} = \Theta$. We have:

$$\mu(\mathbf{A} \cup \mathbf{B}) = 1 \tag{118}$$

Cases 2 and 3 describe a single situation —that we shall call 'case $\oplus$'— because, thanks to additivity (72) of μ:

$$\begin{aligned}\mu(\mathbf{A} \cup \mathbf{B}) &= \mu(\mathbf{A}) + \mu(\mathbf{B}) - \mu(\mathbf{A} \cap \mathbf{B}) = \\ &= \mu(\mathbf{A}) + \mu(\mathbf{B}) - \min\,[\mu(\mathbf{A}), \mu(\mathbf{B})] = \\ &= \max\,[\mu(\mathbf{A}), \mu(\mathbf{B})]\end{aligned}$$

and vice versa. On the other hand, cases 1 and 4 can be treated as one, because we have, respectively:

$$\mu(\mathbf{A} \cap \mathbf{B}) = 0 \qquad (119) \qquad \text{(for case 1), and}$$

$$\mu(\mathbf{A} \cup \mathbf{B}) = 1 \qquad (120) \qquad \text{(for case 4),}$$

that by (72) yield, respectively:

$$\mu(\mathbf{A} \cup \mathbf{B}) = \mu(\mathbf{A}) + \mu(\mathbf{B}) \qquad (121) \qquad \text{(for case 1), and}$$

$$\mu(\mathbf{A} \cap \mathbf{B}) = \mu(\mathbf{A}) + \mu(\mathbf{B}) - 1 \qquad (122) \qquad \text{(for case 4).}$$

We can cover the situations described by (119-122) with a single name —we shall call it 'case $\ominus$'— and a pair of formulas:

$$\mu(\mathbf{A} \cap \mathbf{B}) = \max\,[0, \mu(\mathbf{A}) + \mu(\mathbf{B}) - 1], \tag{123}$$

that summarizes in a single statement formulas (119) and (122), and

$$\mu(\mathbf{A} \cup \mathbf{B}) = \min\,[1, \mu(\mathbf{A}) + \mu(\mathbf{B})], \tag{124}$$

that does the same for formulas (120) and (121).

To justify results (123) and (124) it suffices to consider that:

- If $\mu(\mathbf{A} \cap \mathbf{B}) = 0$ (119) then $\mu(\mathbf{A} \cup \mathbf{B}) = \mu(\mathbf{A}) + \mu(\mathbf{B}) \leq 1$, so $\mu(\mathbf{A}) + \mu(\mathbf{B}) - 1 \leq 0$ and $\mu(\mathbf{A} \cap \mathbf{B}) = \max\,[0, \mu(\mathbf{A}) + \mu(\mathbf{B}) - 1] = 0$, as it should be.

- If $\mu(\mathbf{A}\cup\mathbf{B}) = 1$ (120) then $\mu(\mathbf{A}\cap\mathbf{B}) = \mu(\mathbf{A})+\mu(\mathbf{B})-1 \geq 0$, so $\mu(\mathbf{A}\cap\mathbf{B}) = \max\,[0, \mu(\mathbf{A})+\mu(\mathbf{B})-1] = \mu(\mathbf{A})+\mu(\mathbf{B})-1$, also as it should be.

So the formula (123) is right for the two cases (119) and (120) it so compactly summarizes. And identically, *coeteris paribus*, for (124) as a summarizing formula for the union in both same cases.

Thus, by tracking what happens with measure μ when $\mathbf{A}$ (as $\mathbf{B}$) goes all the way —in smooth gradation— from ϕ to Θ (see (113)), it is easy to see that not one but *many* values are possible for $\mu(\mathbf{A}\cap\mathbf{B})$ and $\mu(\mathbf{A}\cup\mathbf{B})$, and that those values are strictly *bounded* as prescribed by formulas (116-117) in one extreme case, and by (123-124) in the other. This has a straightforward translation into *truth values* and *composite propositions*. The first thing we learn is that the binary connectives —as propositional functions— are *not* functional, i.e. they yield different values for a proposition despite the fact that the operands may have stable values. (We shall see below, however, that the binary connectives are actually *functional*, but in *three* —not two— variables.) The second is that the range of *values* of composite propositions has, nevertheless, strict and prescribable *bounds*. We analyze that, and distinguish the two extreme cases we mentioned:

A) **Case** $\oplus$: It summarizes cases 2 and 3 above. This situation is what we call *maximum compatibility* between two propositions A and B. The value of the connectives is given by:

$$[\![A \wedge B]\!] = \min\,([\![A]\!], [\![B]\!]) \tag{125}$$

$$[\![A \vee B]\!] = \max\,([\![A]\!], [\![B]\!]) \tag{126}$$

(which are (116-117) in truth-value notation).

We shall often abbreviate the right-hand members as " $[\![A \wedge B]\!]^{+}$" and " $[\![A \vee B]\!]^{+}$ ", respectively.

Case $\oplus$ obviously corresponds to any of the situations described next:

$A \quad B$

$\mathbf{A} \subset \mathbf{B}$, which is equivalent to $A \vdash B$ and, graphically,

$\mathbf{B} \subset \mathbf{A}$, which is equivalent to $B \vdash A$ and, graphically,

all of which justifies our speaking of "*maximum compatibility*". We could have called this case also simply *compatibility* or *coherence* (because of lack of incoherence, see case $\ominus$) or *mutual implication* (because here either $A \vdash B$ or $B \vdash A$). The situation here is one of [mutual] *dependence*, as $[\![B|A]\!]$ —or $[\![A|B]\!]$— equals one. (Later we shall speak also of *correlation*.)

From $[\![A \wedge B]\!]$ and (111-112) we obtain at once:

$$[\![A \rightarrow B]\!] = \min\,(1, 1 - [\![A]\!] + [\![B]\!]) \tag{127}$$

$$[\![A \leftrightarrow B]\!] = 1 - |[\![A]\!] - [\![B]\!]| \tag{128}$$

We shall abbreviate the right-hand members as " $[\![A \rightarrow B]\!]^{+}$ " and " $[\![A \leftrightarrow B]\!]^{+}$ ".

B) **Case** $\ominus$: It summarizes cases 1 and 4 above. There is what we call *minimum compatibility* between two propositions A and B. The value of the connectives is given by:

$$[\![A \wedge B]\!] = \max\,(0, [\![A]\!] + [\![B]\!] - 1) \tag{129}$$

$$[\![A \vee B]\!] = \min\,(1, [\![A]\!] + [\![B]\!]) \tag{130}$$

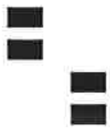
(which are (123-124) in truth-value notation).

We shall often abbreviate the right-hand members as " $[\![A \wedge B]\!]^{-}$ " and " $[\![A \vee B]\!]^{-}$ ", respectively.

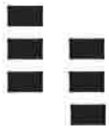
Case $\ominus$ obviously corresponds to any of the situations described next:

$A \quad B$

$\mathbf{A} \cap \mathbf{B} = \phi$, which is equivalent to $A \wedge B = \perp$ and, graphically,



$\mathbf{A} \cup \mathbf{B} = \Theta$, which is equivalent to $\vdash A \vee B$ and, graphically,



all of which justifies our speaking of "*minimum compatibility*". We could have called this case also simply *incompatibility* or *incoherence* (because either $A \wedge B = \perp$ or $\neg A \wedge \neg B = \perp$), or *mutual contradiction* (because here either $A \vdash \neg B$ or $\neg A \vdash B$).

From $[\![A \wedge B]\!]$ and (111-112) we obtain at once:

$$[\![A \rightarrow B]\!] = \max\,(1 - [\![A]\!], [\![B]\!]) \tag{131}$$

$$[\![A \leftrightarrow B]\!] = |[\![A]\!] + [\![B]\!] - 1| \tag{132}$$

We shall abbreviate the right-hand members as " $[\![A \rightarrow B]\!]^-$ " and " $[\![A \leftrightarrow B]\!]^-$ ".

Note: As corollaries to (125-132) we have that:

- If $[\![A]\!] = 0$ then: $[\![A \wedge B]\!] = 0$, $[\![A \vee B]\!] = [\![B]\!]$, $[\![A \rightarrow B]\!] = 1$ and $[\![A \leftrightarrow B]\!] = 1 - [\![B]\!]$. (133)
- If $[\![B]\!] = 0$ then: $[\![A \wedge B]\!] = 0$, $[\![A \vee B]\!] = [\![A]\!]$ and $[\![A \rightarrow B]\!] = [\![A \leftrightarrow B]\!] = 1 - [\![A]\!]$. (134)

So, in summary, the value of the connectives is always inside a slack interval, with bounds $\ominus$ and $\oplus$:

Connective	Case	Minimum value	Actual value	Maximum value	Case
$\wedge$	$\ominus$	$\max\,(0, [\![A]\!] + [\![B]\!] - 1)$	$[\![A \wedge B]\!]$	$\min\,([\![A]\!], [\![B]\!])$	$\oplus$
$\vee$	$\oplus$	$\max\,([\![A]\!], [\![B]\!])$	$[\![A \vee B]\!]$	$\min\,(1, [\![A]\!] + [\![B]\!])$	$\ominus$
$\rightarrow$	$\ominus$	$\max\,(1 - [\![A]\!], [\![B]\!])$	$[\![A \rightarrow B]\!]$	$\min\,(1, 1 - [\![A]\!] + [\![B]\!])$	$\oplus$
$\leftrightarrow$	$\ominus$	$\|[\![A]\!] + [\![B]\!] - 1\|$	$[\![A \leftrightarrow B]\!]$	$1 - \|[\![A]\!] - [\![B]\!]\|$	$\oplus$

We have, perhaps more graphically:

$$\begin{array}{ccccccc}
\ominus & [\![A \wedge B]\!]^- & \leq & [\![A \wedge B]\!] & \leq & [\![A \wedge B]\!]^+ & \oplus \\
\oplus & [\![A \vee B]\!]^+ & \leq & [\![A \vee B]\!] & \leq & [\![A \vee B]\!]^- & \ominus \\
\ominus & [\![A \rightarrow B]\!]^- & \leq & [\![A \rightarrow B]\!] & \leq & [\![A \rightarrow B]\!]^+ & \oplus \\
\ominus & [\![A \leftrightarrow B]\!]^- & \leq & [\![A \leftrightarrow B]\!] & \leq & [\![A \leftrightarrow B]\!]^+ & \oplus \\
 & \downarrow & & \downarrow & & \downarrow &
\end{array}$$

0 ———————— c ———— a ———— d ———————— 1

The above diagram is actually oversimplified: each connective would have a different projection on the [0,1] line, and each would have a different triple of values c, a and d; in each triple, c and d are the minimum (c) and maximum (d) values, respectively, that the actual value (a) of the corresponding composite proposition can reach, as the swing along [0,1] takes a from the $\ominus$ to the $\oplus$ case.

A rather stunning fact about the above graph is that the width $d - c$ is *constant* for the three first connectives (and exactly double that length for the biconditional). Indeed,

$$[\![A \wedge B]\!]^+ - [\![A \wedge B]\!]^- = [\![A \vee B]\!]^+ - [\![A \vee B]\!]^- = [\![A \rightarrow B]\!]^+ - [\![A \rightarrow B]\!]^- = ([\![A \leftrightarrow B]\!]^+ - [\![A \leftrightarrow B]\!]^-)/2 =$$
$$= \min\,(\, [\![A]\!],\ [\![B]\!],\ 1 - [\![A]\!],\ 1 - [\![B]\!]\,)\ ,$$

a quadruple minimum that only depends on the values of $[\![A]\!]$ and $[\![B]\!]$. We note this value by " $\Delta(A,B)$" or " Δ_{AB}" and notice that $\Delta(A,B) = \Delta(\neg A, B) = \Delta(A, \neg B) = \Delta(\neg A, \neg B)$.

Another striking fact about connectives is that we can parameterize their values through a unique parameter we name " $\alpha(A,B)$ " or " α_{AB} " and we call "*degree of compatibility* [between propositions A and B]" or "relative position [of propositions **A** and **B** (inside Θ)]". Its value is:

$$\alpha_{AB} = \frac{[\![A \wedge B]\!] - [\![A \wedge B]\!]^-}{[\![A \wedge B]\!]^+ - [\![A \wedge B]\!]^-} \tag{135}$$

Symmetrically we define a second parameter we name " $\beta(A,B)$ " or " β_{AB} " —that we call "*degree of incompatibility* [between propositions A and B]"— through the formula

$$\beta_{AB} =_{df} 1 - \alpha_{AB} \tag{136}$$

Naturally, $0 \leq \alpha_{AB} \leq 1$ and, simultaneously, $1 \geq \beta_{AB} \geq 0$ —where the leftmost and rightmost bounds refer to cases $\ominus$ and $\oplus$, respectively, so that both cases are completely determined by one parameter (or both of them):

Case $\oplus$ (*Maximum compatibility*): $\alpha_{AB} = 1$ (or $\beta_{AB} = 0$).
(Note that then —and only then— $[\![A \wedge B]\!] = [\![A \wedge B]\!]^+$.)

Case $\ominus$ (*Minimum compatibility*): $\alpha_{AB} = 0$ (or $\beta_{AB} = 1$).
(Note that then —and only then— $[\![A \wedge B]\!] = [\![A \wedge B]\!]^-$.)

Both cases coincide if —and only if— at least one of the propositions A or B is valued binarily. In this situation —which is equally well described by both case profiles— α_{AB} and β_{AB} are *undetermined*, and the actual value of the connectives is given by *any* of the formulas above.

In the general case, the parameter α_{AB} acts as an indicator or measure of "relative position" of propositions **A** and **B** inside Θ, and also as a cursor ranging inside the [fixed] interval between bounds, pointing to the actual value of the connective. We can formulate each connective as a linear function (a convex combination of case $\oplus$ and case $\ominus$ values) "interpolating" between bounds (=the extreme $\oplus$ and $\ominus$ values), so that its effective value is given by the values $[\![A]\!]$ and $[\![B]\!]$ *and* the parameter α. (Thus each connective is functional in *three* variables, the third being α.)

Indeed we can, and get the following set of formulas (where (137) derives directly from (135) while (138-140) are obtained from (137) via (110-112)):

$$\boxed{\begin{aligned}
[\![A \wedge B]\!] &= \alpha_{AB} \cdot [\![A \wedge B]\!]^+ + \beta_{AB} \cdot [\![A \wedge B]\!]^- && (137)\\
[\![A \vee B]\!] &= \alpha_{AB} \cdot [\![A \vee B]\!]^+ + \beta_{AB} \cdot [\![A \vee B]\!]^- && (138)\\
[\![A \rightarrow B]\!] &= \alpha_{AB} \cdot [\![A \rightarrow B]\!]^+ + \beta_{AB} \cdot [\![A \rightarrow B]\!]^- && (139)\\
[\![A \leftrightarrow B]\!] &= \alpha_{AB} \cdot [\![A \leftrightarrow B]\!]^+ + \beta_{AB} \cdot [\![A \leftrightarrow B]\!]^- && (140)
\end{aligned}}$$

So by knowing a single value (either of $[\![A \wedge B]\!]$, $[\![A \vee B]\!]$, $[\![A \rightarrow B]\!]$, $[\![A \leftrightarrow B]\!]$, α_{AB} or β_{AB}) we can compute the other five.

Alternatively, formulas (137-140) can be replaced by this set:

$$\llbracket A \wedge B \rrbracket = \min\,(\llbracket A \rrbracket, \llbracket B \rrbracket) - \beta_{AB} \cdot \Delta_{AB} \tag{141}$$

$$\llbracket A \vee B \rrbracket = \max\,(\llbracket A \rrbracket, \llbracket B \rrbracket) + \beta_{AB} \cdot \Delta_{AB} \tag{142}$$

$$\llbracket A \rightarrow B \rrbracket = \min\,(1, 1 - \llbracket A \rrbracket + \llbracket B \rrbracket) - \beta_{AB} \cdot \Delta_{AB} \tag{143}$$

$$\llbracket A \leftrightarrow B \rrbracket = 1 - \mid \llbracket A \rrbracket - \llbracket B \rrbracket \mid - 2 \cdot \beta_{AB} \cdot \Delta_{AB} \tag{144}$$

where it is prominent that the value of the connectives is the value for *case* $\oplus$ plus a *negative* correction (except for $\vee$, whose correction is *positive*) of size proportional to the *incompatibility* β_{AB} and the [constant] *interval length* Δ_{AB} (which is a function of $\llbracket A \rrbracket$ and $\llbracket B \rrbracket$ only).

Incidentally, connectives for case $\oplus$ —in fact [generally upper] *bounds* for actual values— coincide with values [functionally] assigned by Łukasiewicz-Tarski to the connectives in their $Ł_\infty$ logic. On the other hand, connectives for case $\ominus$ (in fact [generally lower] bounds for actual values) coincide with values [functionally] assigned to them by *threshold logic*. The difference here is that those connectives are no longer *functional* in the truth values of the operands, but act as mere bounds for actual values. These depend not only on the truth-values of the component propositions but also on a third term indicating their relative position as well. (Nothing new in this to Probability, long accustomed to producing different values after the relative disposition of component events.)

The α parameter satisfies the following relations:

$$\alpha(A, A) = 1 \qquad \text{and} \qquad \alpha(A, \neg A) = 0 \tag{145}$$

$$\alpha(A, B) = \alpha(B, A) \tag{146}$$

$$\alpha(A, B) = \alpha(\neg A, \neg B) = 1 - \alpha(A, \neg B) = 1 - \alpha(\neg A, B) \quad [\ = \beta(\neg A, B)\ ,\ \text{etc.}\ ] \tag{147}$$

(and analogously for β). With those relations in mind, numerous properties of our many-valued logic can be easily proved. For instance:

$$\llbracket A \wedge \neg A \rrbracket = \alpha(A, \neg A) \cdot \llbracket A \wedge \neg A \rrbracket^+ + \beta(A, \neg A) \cdot \llbracket A \wedge \neg A \rrbracket^- = \llbracket A \wedge \neg A \rrbracket^- = 0$$

$$\llbracket A \vee \neg A \rrbracket = \alpha(A, \neg A) \cdot \llbracket A \vee \neg A \rrbracket^+ + \beta(A, \neg A) \cdot \llbracket A \vee \neg A \rrbracket^- = \llbracket A \vee \neg A \rrbracket^- = 1$$

$$\begin{aligned}\llbracket A \rightarrow B \rrbracket &= \alpha(A, B) \cdot \min\,(1, 1 - \llbracket A \rrbracket + \llbracket B \rrbracket) + \beta(A, B) \cdot \max\,(1 - \llbracket A \rrbracket, \llbracket B \rrbracket) = \\ &= \beta(\neg A, B) \cdot \min\,(1, \llbracket \neg A \rrbracket + \llbracket B \rrbracket) + \alpha(\neg A, B) \cdot \max\,(\llbracket \neg A \rrbracket, \llbracket B \rrbracket) = \llbracket \neg A \vee B \rrbracket\end{aligned}$$

We continue elaborating on the example of section 4 (tables 5-9), but now we only use, for simplicity, the μ_2 valuation (see table 5). As before, we shall consider four cases 1-4 (corresponding to tables 6-9).

1) We have:

$\llbracket A \rrbracket = .25$, $\llbracket B \rrbracket = .75$ and, e.g., $\llbracket A \rightarrow B \rrbracket = .75$ (from table 6)

$\Delta_{AB} = .25$ and $\beta_{AB} = (\llbracket A \rightarrow B \rrbracket^+ - \llbracket A \rightarrow B \rrbracket)/\Delta_{AB} = 1$ (through simple arithmetic)

so that we obtain:

$\alpha_{AB} = 0$

$\llbracket A \wedge B \rrbracket = .25 - 1 \times .25 = 0$ (because of (141))

$\llbracket A \vee B \rrbracket = .75 + 1 \times .25 = 1$ (because of (142))

$\llbracket A \rightarrow B \rrbracket = .75$ (given)

$\llbracket A \leftrightarrow B \rrbracket = .5 - 2 \times 1 \times .25 = 0$ (because of (144))

$\llbracket A|B \rrbracket = \llbracket A \wedge B \rrbracket \ / \ \llbracket B \rrbracket = 0/.75 = 0$

$\llbracket B|A \rrbracket = \llbracket A \wedge B \rrbracket \ / \ \llbracket A \rrbracket = 0/.25 = 0$

(notice the above values for connectives are in perfect agreement with those computed for μ_2 in table 6).

2) We have:

$[\![A]\!] = .35$, $[\![B]\!] = .75$ and $[\![A \rightarrow B]\!] = .85$ (from table 7)

$\Delta_{AB} = .25$ and $\beta_{AB} = (1 - .85)/.25 = 3/5$ (straightforward computation)

so that we obtain:

$\alpha_{AB} = 2/5$

$[\![A \wedge B]\!] = .35 - 3/5 \times .25 = .20$ (because of (141))

$[\![A \vee B]\!] = .75 + 3/5 \times .25 = .90$ (because of (142))

$[\![A \rightarrow B]\!] = .85$ (given)

$[\![A \leftrightarrow B]\!] = 1 - |.35 - .75| - 2 \times 3/5 \times .25 = .60 - .30 = .30$ (because of (144))

$[\![A|B]\!] = .20/.75 = 4/15$

$[\![B|A]\!] = .20/.35 = 4/7$

(note that values agree with those computed in table 7 for μ_2).

3) We have:

$[\![A]\!] = .45$, $[\![B]\!] = .75$ and $[\![A \rightarrow B]\!] = 1$ (from table 8)

$\Delta_{AB} = .25$ and $\beta_{AB} = (1 - 1)/.25 = 0$ (straightforward computation)

so that we obtain:

$\alpha_{AB} = 1$

$[\![A \wedge B]\!] = .45 - 0 \times .25 = .45$ (because of (141))

$[\![A \vee B]\!] = .75 + 0 \times .25 = .75$ (because of (142))

$[\![A \rightarrow B]\!] = 1$ (given)

$[\![A \leftrightarrow B]\!] = 1 - |.45 - .75| - 2 \times 0 \times .25 = .70$ (because of (144))

$[\![A|B]\!] = .45/.75 = 3/5$

$[\![B|A]\!] = .45/.45 = 1$

(note that values agree with those computed in table 8 for μ_2).

4) We have:

$[\![A]\!] = .55$, $[\![B]\!] = .75$ and $[\![A \rightarrow B]\!] = 1$ (from table 9)

$\Delta_{AB} = .25$ and $\beta_{AB} = (1 - 1)/.25 = 0$ (straightforward computation)

so that we obtain:

$\alpha_{AB} = 1$

$[\![A \wedge B]\!] = .55 - 0 \times .25 = .55$ (because of (141))

$[\![A \vee B]\!] = .75 + 0 \times .25 = .75$ (because of (142))

$[\![A \rightarrow B]\!] = 1$ (given)

$[\![A \leftrightarrow B]\!] = 1 - |.55 - .75| - 2 \times 0 \times .25 = .80$ (because of (144))

$[\![A|B]\!] = .55/.75 = 11/15$

$[\![B|A]\!] = .75/.75 = 1$

(note that values agree with those computed in table 9 for μ_2).

The example shows the agreement between computed and table values, and it follows α as it slids along from 0 to 1 as A and B cease to be *incompatible* (i.e. as the degree of *superposition* between **A** and **B** *increases*); in the process, the dependence of A on B (and vice versa) also increases, until $[\![B|A]\!]$ finally reaches one in cases 3-4. More graphically:

(Table 11)

Θ	A	B	A	B	A	B	A	B
θ_0	■							
θ_1	■		■					
θ_2		■	■	■	■	■		■
θ_3		■		■	■	■	■	■
θ_4		■		■		■	■	■
	$\alpha_{AB} = 0$		$\alpha_{AB} = .4$		$\alpha_{AB} = 1$		$\alpha_{AB} = 1$	
	$\llbracket A\|B \rrbracket = 0$		$\llbracket A\|B \rrbracket \simeq .27$		$\llbracket A\|B \rrbracket = .60$		$\llbracket A\|B \rrbracket \simeq .71$	
	$\llbracket B\|A \rrbracket = 0$		$\llbracket B\|A \rrbracket \simeq .57$		$\llbracket B\|A \rrbracket = 1$		$\llbracket B\|A \rrbracket = 1$	

In section 5 we said that two propositions A and B were *independent* when their conjunction could be expressed —in value— as the product of $\llbracket A \rrbracket$ and $\llbracket B \rrbracket$:

$$\llbracket A \wedge B \rrbracket = \llbracket A \rrbracket \cdot \llbracket B \rrbracket \tag{148}$$

It is easily shown that the necessary and sufficient condition for that to happen is:

$$\begin{aligned} \alpha_{AB} &= \max\,(\llbracket A \rrbracket, \llbracket B \rrbracket) \quad \text{if } \llbracket A \rrbracket + \llbracket B \rrbracket \le 1 \\ &= \max\,(\llbracket \neg A \rrbracket, \llbracket \neg B \rrbracket) \;\; \text{if } \llbracket A \rrbracket + \llbracket B \rrbracket \ge 1 \end{aligned} \tag{149}$$

(the expression in the second row is equivalent to $1-\min\,(\llbracket A \rrbracket, \llbracket B \rrbracket)$). Analogously,

$$\begin{aligned} \beta_{AB} &= \min\,(\llbracket \neg A \rrbracket, \llbracket \neg B \rrbracket) \;\; \text{if } \llbracket A \rrbracket + \llbracket B \rrbracket \le 1 \\ &= \min\,(\llbracket A \rrbracket, \llbracket B \rrbracket) \quad \text{if } \llbracket A \rrbracket + \llbracket B \rrbracket \ge 1 \end{aligned} \tag{150}$$

(the expression in the first row is equivalent to $1-\max\,(\llbracket A \rrbracket, \llbracket B \rrbracket)$). Note that A —always compatible with itself— is never independent from itself (except when its truth value is binary). Similarly, A and $\neg A$ —always incompatible— are never mutually independent (except when their truth value is binary).

When two propositions are *independent*, the connectives can be expressed —in value— in this way:

$$\llbracket A \wedge B \rrbracket = \llbracket A \rrbracket \cdot \llbracket B \rrbracket \tag{151}$$
$$\llbracket A \vee B \rrbracket = \llbracket A \rrbracket + \llbracket B \rrbracket - \llbracket A \rrbracket \cdot \llbracket B \rrbracket \tag{152}$$
$$\llbracket A \rightarrow B \rrbracket = 1 - \llbracket A \rrbracket + \llbracket A \rrbracket \cdot \llbracket B \rrbracket \tag{153}$$
$$\llbracket A \leftrightarrow B \rrbracket = 1 - \llbracket A \rrbracket - \llbracket B \rrbracket + 2 \cdot \llbracket A \rrbracket \cdot \llbracket B \rrbracket \tag{154}$$
$$\llbracket A|B \rrbracket = \llbracket A \rrbracket \quad \text{and} \quad \llbracket B|A \rrbracket = \llbracket B \rrbracket \tag{155}$$

Turning back to our example (with μ_2), we now change B and make:

$\rho(A) = \{\theta_0, \theta_1\}$, $\rho(B) = \{\theta_1, \theta_2, \theta_3\}$ (Table 12):

Θ	A		B		$A \rightarrow B$	
θ_0	■	.10				
θ_1	■	.15	■	.15	■	.15
θ_2			■	.20	■	.20
θ_3			■	.25	■	.25
θ_4					■	.30
	$\llbracket A \rrbracket = .25$		$\llbracket B \rrbracket = .60$		$\llbracket A \rightarrow B \rrbracket = .90$	

So we have: $\Delta_{AB} = .25$ and $\beta_{AB} = .40$ and, thus: $\alpha_{AB} = .60$,
which satisfies condition (149), so that A and B are *independent*, and, indeed:

$$[\![A \wedge B]\!] = .15 = [\![A]\!] \cdot [\![B]\!]$$
$$[\![A \vee B]\!] = .70 = [\![A]\!] + [\![B]\!] - [\![A]\!] \cdot [\![B]\!]$$

Summarizing: for any pair of propositions A and B of $\mathcal{L}$, the *truth value* of any composite proposition involving A and B —as well the *relative truth* of one with respect to the other— is computable once the following values are known:

(a) the truth values $[\![A]\!]$ and $[\![B]\!]$ of A and B, and

(b) the value of parameter α_{AB} (or β_{AB}).

(Note that from values $[\![A]\!]$ and $[\![B]\!]$ —that by (a) we are required to know— the value of Δ_{AB} can be immediately computed.)

The values of α_{AB} (or β_{AB}) are usually not known, but two considerations stand out: first, all computations can proceed if we know just *one* of these eight values: $[\![A \wedge B]\!]$, $[\![A \vee B]\!]$, $[\![A \rightarrow B]\!]$, $[\![A \leftrightarrow B]\!]$, $[\![A|B]\!]$, $[\![B|A]\!]$, α_{AB} or β_{AB} —from which all others are derivable at once—; second, in certain cases the value of α is either irrelevant (e.g. when one of the propositions is binary-valued) or immediate. This last happens, for instance, when $A \vdash B$ or $B \vdash A$ (so $\alpha_{AB} = 1$) or when $A \vdash \neg B$ or $\neg B \vdash A$ (so $\alpha_{AB} = 0$). And also, in particular, when $B = A$ or when $B = \neg A$, because then $\alpha(A, A) = 1$ and $\alpha(A, \neg A) = 0$, whence we can easily derive, e.g., the principles of identity ($\vdash A \rightarrow A$), non-contradiction ($\vdash \neg(A \wedge \neg A)$) and excluded middle ($\vdash A \vee \neg A$) —by verifying that, indeed, $[\![A \rightarrow A]\!] = 1$, $[\![A \wedge \neg A]\!] = 0$ and $[\![A \vee \neg A]\!] = 1$ for any valuation. So our logic —that we shall call henceforth '**BML**' (for **B**oolean and **M**any-valued)— is, like ordinary (i.e. two-valued) Boolean logic, consistent with classical Aristotelian principles (which is peculiarly *not* the case of traditional multi-valued logics such as Kleene's [24] or Bochvar's [5] three-valued ones or Łukasiewicz-Tarski's $\mathcal{L}_\infty$ [27]). Also, we have now that **BML** incorporates the whole of classical [two-valued] Boolean Logic as a special case. One further interesting result here is that, for any A and B, propositions A and $A \rightarrow B$ are always incompatible, i.e. $\alpha(A,\ A \rightarrow B) = 0$ (a good thing to remember when we turn to modus ponens and proof theory in section 12).

Requisite (a) is the only one exacted by logics based on *functional* connectives such as Łukasiewicz's $\mathcal{L}_\infty$, where the values of composite propositions are given directly by our case$\oplus$-formulas (and so $\alpha_{AB} = 1$ for every A and B), which means that —viewed from the **BML** perspective— in $\mathcal{L}_\infty$ *all propositions are compatible*. In **BML** —incidentally— assuming such a thing would be clearly unrealistic or utterly impossible, because in $\mathcal{L}$ every proposition A has a negation, and a negation $\neg A$ automatically stands as *incompatible* with A, thus contradicting the assumption [that all propositions are compatible].

Requisite (b) is new and peculiar to the present Boolean multi-valued logic we have called **BML**. The α parameter corresponds to what Reichenbach in his probability logic called '*Kopplungsgrad* ' [34][35]; it acts as a sort of descriptor or "memory" of the [relative] structural disposition of propositions A and B "inside" $\mathcal{L}$ (or of $\mathbf{A}$ and $\mathbf{B}$ inside Θ). This situation parallels the one familiar in Probability, where [only] three commonplace cases are usually considered:

1) One event A is conditioned by the other. Given the probability of A and the degree of conditioning $P(B|A)$, the probability of B can be computed. This is the probabilistic counterpart of our *dependence* (or *relative truth*) relation. If $P(B|A) = 1$ $(\forall P)$, we are in [the subcase $\mathbf{A} \subset \mathbf{B}$ of] our *case* $\oplus$.

2) The two events are independent. Probability of event-conjunction is then the product of probabilities of isolated events. This corresponds to what we called "[propositional] *independence*".

3) The two events are totally disjoint (mutually exclusive). Compound probability of the events is the sum of probabilities of isolated events (as event-conjunction has zero probability). This is the probabilistic counterpart of [the subcase $\mathbf{A} \cap \mathbf{B} = \phi$ of] *case* $\ominus$ (in our terminology).

Note that cases $\oplus$ and $\ominus$ imply that the values of the connectives are those affected with the $+$ sign ($[\![A \wedge B]\!]^+$, etc.) or with the $-$ sign ($[\![A \wedge B]\!]^-$, etc.), respectively, but the converse is not true: the connectives may have $\oplus$ values and yet we may not be in the $\oplus$ case ($A \vdash B$ or $B \vdash A$), and we may have $\ominus$ values with no $\ominus$ conditions ($A \wedge B = \perp$ or $\vdash A \vee B$); such situations arise in the $\oplus$ case when $\mathbf{A} \not\subset \mathbf{B}$ but the "outer" *possibilities* (= the θs in $\mathbf{A}$ that are outside $\mathbf{B}$) are all zero-valued, and in the $\ominus$ case when $\mathbf{A} \cap \mathbf{B} \neq \phi$ but the common *possibilities* (= the θs in $\mathbf{A} \cap \mathbf{B}$) are zero-valued. Such is often the case the case with *selective* valuations (see for example Tables 6-9). Nevertheless, if the truth valuation is *positive* then that converse is true: a $+$ or $-$ value corresponds to a $+$ or $-$ situation —so then by simply looking at the values we can be sure what conditions hold.

7. CONDITIONALS AS REVERSIBLE PROPOSITIONS

Suppose p is a positive valuation. Classically (i.e. in two-valued interpretations of Boolean logic), $[\![A \to B]\!] = 1$ or $\models_p A \to B$ means —by (57), (59) and the representation theorem— $\mathbf{A} \subset \mathbf{B}$, which is a rigid relation that can only be asserted when an inclusion is really present and this fact can be effectively assessed. This happens at least in these three groups of instances:

1) $[\![A \to B]\!] = 1$ or $\models_p A \to B$ (or $\mathbf{A} \subset \mathbf{B}$) represents either a (a) *logical* relation ("A implies B"), or a (b) *semantical* relation (e.g. of the 'semantic postulate' type: "the concept B is entailed by the definition of A"), or a (c) general *mathematical* relation ("A suffices for B", "B is necessary for A"), or a (d) *statistical* relation ("every time one finds that A one also finds that B"), or simply a (e) *heuristic* relation (where rule "if A then B" is to be understood or executed as "every time condition A arises, action or consequence B should follow").

2) $[\![A \to B]\!] = 1$ or $\models_p A \to B$ (or $\mathbf{A} \subset \mathbf{B}$) represents a *causal* relation. This can mean either: (a) A is the cause and B the effect; clearly A needs not be the only alternative cause (i.e. it may be that $\mathbf{A}' \subset \mathbf{B}$ too, with $\mathbf{A}' \neq \mathbf{A}$) but B is clearly the only effect of A (a univocal cause): so $A \to B$ is *causal* and *predictive* (anticipatory). Or (b) A is the effect and B the cause; the effect may not be the only one due to B, but B is the only cause of A (a univocal effect): so $A \to B$ is *causal* yet *retrospective* (diagnostic).

3) $[\![A \to B]\!] = 1$ or $\models_p A \to B$ (or $\mathbf{A} \subset \mathbf{B}$) represents an *evidential* relation: "A is [concluding] evidence to assert B", and so "B is an explanation hypothesis for —or an expected consequence from— the evidence A". (B may be —most often— the real cause of the observed A, it may sometimes be its effect, or it may be neither.) $A \to B$ is thus *evidential*; it is also *retrospective* (or diagnostic) if B causes A, or *predictive* (or anticipatory) if B is the effect.

(We will be interested only in the latter two cases).

In **BML**, the simple fact of being able to value a conditional $A \to B$ in [0,1] turns things around, so we may now have a new perspective on the conditional statement and how it works. For now the arrow no longer signals a rigid condition corresponding to set inclusion $\mathbf{A} \subset \mathbf{B}$ but a *relation* (in the mathematical sense i.e. with possibly overlapping or even disjoint $\mathbf{A}$ and $\mathbf{B}$). Also, the relation between A and B is now completely general; and if, e. g., $A \to B$ is *causal* (but note the causality is now totally general: m causes may be associated to n effects) then $B \to A$ is *evidential*. Further, we have:

a) A and B are *reversible*, and each direction will yield an independent truth value ($[\![A \to B]\!] \neq [\![B \to A]\!]$, generally). If $A \to B$ is meant to be read as 'cause $\to$ effect' (or 'effect $\to$ cause') then $B \to A$ is read 'evidence $\to$ hypothesis'.

b) $A \to B$, *when asserted* (for instance, in a premise), may now simply be a signal (given by the utterer) that the arrow is to be followed in this order (because it flows in the direction of time or evidence) or, more plausibly, that $A \to B$ is more likely the case than $B \to A$, i.e. $[\![A \to B]\!] > [\![B \to A]\!]$ (a relation which, incidentally, amounts to say that $[\![B]\!] > [\![A]\!]$ and, hence, also that $[\![B|A]\!] > [\![A|B]\!]$). Or else asserting $A \to B$ may just signal that $[\![A \to B]\!] > [\![A \to \neg B]\!]$ (and so $[\![B|A]\!] > [\![\neg B|A]\!]$, and $[\![B|A]\!] > 1/2$).

b') Yet, more to the point, asserting "$A \to B$" reasonably seems to imply that *there is* a relation between propositions A and B, so we would expect $[\![A \wedge B]\!] > 0$ (which amounts to $[\![A \to B]\!] > 1 - [\![A]\!]$). Nevertheless, on closer look, this appears *still* to be insufficient. *Asserting* $A \to B$ seems to imply, more restrictively, that $[\![A \wedge B]\!] > [\![A]\!] \cdot [\![B]\!]$ (which amounts to $[\![A \to B]\!] > 1 - [\![A]\!] \cdot [\![B]\!]$, a higher value than $1 - [\![A]\!]$) —because if the sign here were $\leq$ we would have *independence* or even *anticorrelation* (see above), a rather perverse result (contradicting any reasonable expectation raised by the "$A \to B$" assertion). (Remember that in "$A \to B$" the relation between A and B is *logical*, not merely frequential or probabilistic.)

In view of the preceding considerations, we could reinterpret some concepts and values, like $[\![B|A]\!]$ and $[\![A|B]\!]$, when given the $A \to B$ assertion (that we here interpret, for simplicity, as a 'cause $\to$ effect' flow —and so $B \to A$ as 'evidence $\to$ hypothesis'— and note we no longer require that the valuation should be positive):

— $[\![B|A]\!]$ (or "σ_A" —or even "ν_B", see next paragraph—) could be termed "degree of *sufficiency* or *causality*" of A (or "*causal support* for B", or even "*prediction* or *anticipation* capability"), to be read as "degree in which A is sufficient for B" or "degree in which A is [in general non-unique] cause of B". In view of (91), it is roughly a measure of how much of $\mathbf{A}$ is contained in $\mathbf{B}$.

— $[\![A|B]\!]$ (or "ν_A" or "σ_B") could be termed "degree of *necessity* or *evidence*" of A (or "*evidential support* for A", or even "*retrospective* or *diagnostic* capability"), to be read as "degree in which A is necessary for B" or "degree in which B is [in general non-compelling] evidence (= support of hypothesis) for A (=the hypothesis)". With (91) 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 "$A \to B$" 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 occurrred previously as a cause?" (Similarly for the evidential reading of "$A \to B$".) Once σ and ν have been guessed, they lead —by straightforward computation— to $[\![A \to B]\!]$, $[\![B \to A]\!]$ and the α_{AB} value, which allows one to compute all other values for connectives and also to get a picture of the structural relations linking A and B (see also section 12(a).) (Incidentally, the expert can immediately detect there is no [anomalous] anticorrelation just by checking that neither $[\![A|B]\!] < [\![A]\!]$ nor $[\![B|A]\!] < [\![B]\!]$.)

We would thus arrive at an "implication-as-causality" scheme revolving around three basic cases (note we are always following the 'cause $\to$ effect' downstream, predictive direction, but might as well have considered the 'effect $\to$ cause' upward, retrospective direction):

1) $\mathbf{A} \subset \mathbf{B}$: We have $\alpha_{AB} = 1$, $[\![A \to B]\!] = 1$, $[\![A]\!] \leq [\![B]\!]$ and $[\![B \to A]\!] < 1$. The second equation means that A is one of the [possibly several] alternative causes of B, while the second inequality signals B as [partial] evidence in favor of A.

2) $\mathbf{B} \subset \mathbf{A}$: We have $\alpha_{AB} = 1$, $[\![A \to B]\!] < 1$, $[\![A]\!] \geq [\![B]\!]$ and $[\![B \to A]\!] = 1$. The first inequality means that B is one of the [possibly several] effects of cause A, while the second equation suggests [total, determinating] evidence in favor of A (once B occurs).

3) $\mathbf{A}$ and $\mathbf{B}$ overlap: This is the general case. A is one of the [possibly several] alternative causes of effect B [that A causes among possibly several other effects]. We have not only $\alpha > 0$ but reasonably we should have —more restrictedly— $\alpha > \alpha^0$, where α^0 is the compatibility of A and B when they are *independent* (this value is given by formula (149)); this is because we expect causality to imply *positive* correlation. Indeed, $\alpha = \alpha^0$ means independence and, therefore, propositional indifference: nothing to suggest causality; $\alpha < \alpha^0$ means $[\![A|B]\!] < [\![A]\!]$, actually a *negative* correlation, plainly contradicting any reasonable notion of causality or linkage. So, we conclude, one should have $\alpha > \alpha^0$; moreover, both $[\![A \to B]\!]$ and $[\![B \to A]\!]$ are < 1, we have —as always— $[\![A \to B]\!] - [\![B \to A]\!] = [\![B]\!] - [\![A]\!]$, and:

$$[\![A \to B]\!] \geq 1 - [\![A]\!] + [\![A]\!] \cdot [\![B]\!] \qquad \text{(equal sign if and only if}$$

$$[\![B \to A]\!] \geq 1 - [\![B]\!] + [\![A]\!] \cdot [\![B]\!] \qquad A \text{ and } B \text{ are independent)}$$

If these two conditions are met, the conditional $A \to B$ —and the whole chain of reasoning— can be *reversed*. Once that done, two different, non-interfering lines of reasoning appear, one *causal* ($A \to B$), the other *evidential* ($B \to A$). In each —naturally— we compute and keep track of the truth values of the intervening statements, so that the conclusion can be appropriately *qualified* —or, rather, *quantified*— (see section 12). The $B \to A$ reasoning direction happens to make possible —and controllable— what has traditionally been termed 'inductive' or 'plausible reasoning' (or more recently, after Peirce, 'abduction'), which is merely the downward [deductive] process inherently associated to —and simultaneous with— any ordinary [upward] deduction along a Boolean structure. (Here "upward" and "downward" are conventional directions, due to the $\vdash$-operator progress towards the top $\top$.)

We can easily define a *correlation coefficient* κ between A and B as a measure of how far α_{AB} is from α^0_{AB}, ranging from $+1$ (corresponding to $\alpha_{AB} = 1$) through 0 ($\alpha_{AB} = \alpha^0_{AB}$) to -1 ($\alpha_{AB} = 0$). So we do. To obtain the correlation formula, we just fit a curve through the three points $\alpha = 1$, $\alpha = \alpha^0$ and $\alpha = 0$ (quadratically, so we have smooth behavior) and get:

$$\kappa = \frac{\alpha}{\alpha^0(1-\alpha^0)}[(1 - 2(\alpha^0)^2) - (1 - 2\alpha^0)\alpha] - 1 \qquad (156)$$

In particular, any proposition turns out to have, indeed, maximum *negative correlation* (-1) with its own *negation* (as it is natural to expect, anyway).

8. THREE-VALUED LOGIC AS A SPECIAL CASE

Classical three-valued logics, especially Kleene's system of strong connectives [24] and Łukasiewicz's $Ł_3$ [26], give the following tables for the values of connectives (where U stands for "undetermined"):

(Table 13)

$\wedge$	0 U 1		$\vee$	0 U 1		$\rightarrow$	0 U 1
0	0 0 0		0	0 U 1		0	1 1 1
U	0 X U		U	U Y 1		U	U Z 1
1	0 U 1		1	1 1 1		1	0 U 1

where $X = Y = U$ in both Kleene's and Łukasiewicz's tables, and $Z = U$ in Kleene's but $Z = 1$ in Łukasiewicz's. Clearly, Aristotelian principles do not hold in these logics (except that $\vdash A \rightarrow A$ does in Łukasiewicz's), for if A [and so $\neg A$ too] is "undetermined" (i.e. its value is U) there is no way that $\neg(A \wedge \neg A)$ and $A \vee \neg A$ can be assigned in them the 1 value they should have (and, so, these expressions can never be theorems in those logics —unless, that is, theoremhood is extended to designated values other than 1).

Now, if we —as we had already suggested in section 4— abbreviate "$[\![A]\!] \in (0,1)$" (of **BML**) by "$[\![A]\!] = U$" (as in Kleene o Łukasiewicz) the values given in the above tables coincide exactly with those that would have been computed by the **BML** formulas, *except* that X, Y and Z would remain undetermined until we knew α_{AB}. In general, **BML** values would match Kleene's, but in certain cases they would yield differing results:

1) If $[\![A]\!] + [\![B]\!] \leq 1$ and A and B are *incompatible* (typically because $\mathbf{A} \cap \mathbf{B} = \phi$) then $X = 0$.

2) If $[\![A]\!] + [\![B]\!] \geq 1$ and A and B are *incompatible* (typically because $\mathbf{A} \cup \mathbf{B} = \Theta$) then $Y = 1$.

3) If $[\![A]\!] \leq [\![B]\!]$ and A and B are *compatible* (typically because $\mathbf{A} \subset \mathbf{B}$) then $Z = 1$.

Note that in the particular case in which B is $\neg A$ we have *always* $[\![A \wedge \neg A]\!] = 0$ (from 1) and $[\![A \vee \neg A]\!] = 1$ (from 2). We have also $[\![A \rightarrow A]\!] = 1$ (from 3). These three results are perfectly Aristotelian and in full agreement with what is to be expected from a Boolean logic; indeed, the three *must* yield the predicted values just for simple structural Boolean-algebra reasons, as the first two should follow from (7), (9) and (50), and the third from (20) and (7).

The above analysis strongly suggests that if, for some reason, upholding Aristotelian principles is worthwhile, then any three-valued logic should *not* be strictly functional, but allow some *non*-functionally-determined values available in predictable cases. Something in this spirit is what Van Fraassen suggested when he advocated '*supervaluations*' [42]. Now, we contend, the same results are obtained *automatically* through the **BML** formulas.

The same applies to Łukasiewicz-Tarski $Ł_\infty$. This logic, by its very structure, never gets values consistent with Aristotelian principles and, therefore, these are implicitly *negated* in it. Moreover, as we explain in sections 10 and 12, the choice of connectives in this logic is neither very well based on algebraic motives nor very helpful for constructing a reasonable proof theory. And when it comes to *practical* uses of $Ł_\infty$ such as its [extensive] application in Fuzzy Set Theory [2] —where it stands as the base logic—, the 'min' and 'max' values prescribed for the $\wedge$ and $\vee$ connectives by this logic not always make everybody happy: to judge from the current literature, authors often are at pains discussing the relative merits of alternative formulas for the values of composite propositions (see for example [1][16][20][36][41] [46] or [47]). Many intuitive or heuristic reasons given are, viewed from **BML**, relatively easy to understand. Indeed, when propositions are strongly correlated (for instance, because many imply others and few negative propositions are present) then the connectives tend to have our $\oplus$ values, and authors tend to favor them. On the other hand, when propositions are negatively correlated (or many negations intervene), $\ominus$ formulas seem preferable. Standing in middle ground, large systems (like industrial control processes) or long reasoning chains would tend to use many diversely-correlated propositions, and their relative compatibilities would in the long run cancel each other out into a statistical semblance of independence, so it comes as no surprise then that authors are rather prone to suggest products for conjunctions. Elementary statistics show that in a universe Θ with five *possibilities* and uniform measure μ (amounting to "equiprobability") barely a tenth of proposition pairs have α undefined (because one of the propositions is binary-valued), half that proportion have $\alpha = 1$, only one fiftieth have have $\alpha = 0$, and the rest (more than four fifths) have α in the (0,1) interval; if we increase the size of Θ to eight or more elements (*possibilities*) then the proportions of binary, fully compatible and fully incompatible pairs become negligible, and virtually 100 % of proposition pairs have α in (0,1), clustering around independence. No wonder, then, products tend to be suggested when system complexities grow.

9. THE GEOMETRY OF LOGIC: ORDER, DISTANCE, TRUTH LIKELIHOOD, INFORMATIVENESS, IMPRECISION AND ENTROPY IN $\mathcal{L}$

Suppose we are given *any* two propositions A and B, whose truth values are $[\![A]\!]$ and $[\![B]\!]$ and whose conjunction C ($\equiv A \wedge B$) and disjunction D ($\equiv A \vee B$) have values $[\![C]\!]$ and $[\![D]\!]$. Assume —without loss of generality— that $[\![A]\!] \leq [\![B]\!]$ and that both values are in the open interval (0,1). Given the monotonicity (10) of any valuation v, if we project each lattice point A of $\mathcal{L}$ on the point $[\![A]\!]$ of the real interval [0,1], the lattice order $\vdash$ will be preserved in the $\leq$ order of the real line. Any valuation can be conceived as a distinct orientation of the real projection line. Naturally, conditions (64) and (62) require that if $A \vdash B$ then all projections on all possible orientations of the real line keep A projected just *under* B (i.e. $[\![A]\!] \leq [\![B]\!]$).

Now suppose that A precedes B in the lattice order of $\mathcal{L}$, so $A \vdash B$. We have: (a) $\alpha_{AB} = 1$ (so A and B have maximum compatibility), (b) $[\![A \rightarrow B]\!] = 1$ and $[\![B \rightarrow A]\!] = 1 - ([\![B]\!] - [\![A]\!])$, (c) there is a line (representing $A \vdash B$) that links point A to point B, and (d) $A = C$ and $B = D$ (so conjunction C and disjunction D get projected on the real line at the same points as A and B). If A and B keep their truth values but no longer have maximum compatibility (which automatically entails that $A \vdash B$ is no longer the case), then the line joining A and B disappears —as now $A \not\vdash B$— replaced by a quadrilateral (actually a rectangle, see below) joining A, C, B and D, the projection points $[\![C]\!]$ and $[\![D]\!]$ of the conjuction and disjunction become *distinct* from $[\![A]\!]$ and $[\![B]\!]$, and the four projections now follow this order: $[\![C]\!] < [\![A]\!] \leq [\![B]\!] < [\![D]\!]$. The exact amount of the separations is the same for both pairs of points (we call it δ_{AB}, or simply δ):

$$\delta \;=\; [\![A]\!] - [\![C]\!] \;=\; [\![D]\!] - [\![B]\!]$$

(because additivity (11) now reads $[\![A]\!] + [\![B]\!] \;=\; [\![C]\!] + [\![D]\!]$). If we now replace $[\![A \wedge B]\!]$ by its value —as given by (141)— and recall that $[\![A]\!] \leq [\![B]\!]$ we get:

$$\delta \;=\; \beta_{AB} \cdot \Delta_{AB} \tag{157}$$

So, lack or loss of compatibility between A and B means that —though their projected values remain the same— the projected values of their connectives outdistance themselves from $[\![A]\!]$ and $[\![B]\!]$ proportionally to the degree of incompatibility β_{AB}, up to a maximum distance —corresponding to maximum incompatibility, i.e. $\beta_{AB} = 1$— given by the minimum of $[\![A]\!]$ and $1 - [\![B]\!]$ (as this is the value of Δ_{AB} for $[\![A]\!] \leq [\![B]\!]$). In any case we have:

$$[\![A \wedge B]\!] \;=\; [\![A]\!] - \delta \tag{158}$$

$$[\![A \vee B]\!] \;=\; [\![B]\!] + \delta \tag{159}$$

$$[\![A \rightarrow B]\!] \;=\; 1 - \delta \tag{160}$$

$$[\![A \leftrightarrow B]\!] \;=\; 1 - |[\![A]\!] - [\![B]\!]| - 2 \cdot \delta \tag{161}$$

So the values of all the connectives have now a correction factor added that depends on the compatibility α (through δ). The third formula shows the loss of the precedence order $A \vdash B$ and so the deletion of the line formerly joining points A and B (now a simple [untraced] diagonal of the $ACBD$ rectangle). Along with the fourth formula, we can recall (41) (now expressable as "$[\![A \leftrightarrow B]\!] = 1 - ([\![D]\!] - [\![C]\!])$"), which states that "$A =_v B$ iff $[\![C]\!] = [\![D]\!]$", a fact that strongly suggests using $1 - [\![A \leftrightarrow B]\!] = [\![D]\!] - [\![C]\!]$ as a measure of the *distance* $\overline{AB}$ [under a given valuation v]. So we do. (We remark that all definitions we give from now on of distance and related concepts are not only applicable to propositions but to *theories* as well, because for a finite lattice like $\mathcal{L}$ the lattice $\hat{\mathcal{L}}$ of its theories is isomorphic to $\mathcal{L}$.)

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

$$d(A,B) =_{df} 1 - [\![A \leftrightarrow B]\!] = [\![A \vee B]\!] - [\![A \wedge B]\!] = |[\![A]\!] - [\![B]\!]| + 2 \cdot \beta_{AB} \cdot \Delta_{AB} \tag{162}$$

(Naturally, if $A \vdash B$ then $d(A,B) = [\![B]\!] - [\![A]\!]$.)

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

$$d^{+}(A,B) = |[\![A]\!] - [\![B]\!]| = 1 - [\![A \leftrightarrow B]\!]^{+} \tag{163}$$

This distance can also be expressed in this way:

$$d^{+}(A,B) = [\![A]\!] + [\![B]\!] - 2 \min ([\![A]\!], [\![B]\!]) \tag{164}$$

It is interesting to notice that:

- both d and d^{+} are really distances (in the mathematical sense), as indeed, e.g.:

$$d(A,B) \geq 0 \quad , \quad d(A,A) = 0 \qquad \text{and} \qquad d(A,B) = d(B,A) \tag{165}$$

$$d(A,B) + d(A,C) \geq d(B,C) \qquad (\textit{Triangle inequality}) \tag{166}$$

$$A \vdash B \vdash C \;\Rightarrow\; d(A,B) + d(B,C) = d(A,C) \quad (\textit{Linear additivity}) \tag{167}$$

- the distance between two propositions or theories is the same as the distance between their negations or antitheses (i.e. $d(A,B) = d(\neg A, \neg B)$, and the same holds for d^{+})

- the Boolean distance $d(A,B)$ equals the value of the *symmetric difference* between A and B [defined by $A \triangle B =_{df} (A \wedge \neg B) \vee (\neg A \wedge B)$] so we have:

$$d(A,B) = [\![A \triangle B]\!] \qquad \text{and} \qquad d(A,\perp) = [\![A]\!] \tag{168}$$

Moreover, we have the following miscellaneous properties —that hold in any metric Boolean algebra and were comprehensively studied by David Miller in his 'Geometry of Logic' paper [29]:

$$d(A,B) = d(A \triangle B, \perp)$$

$$d(A,B) + d(\neg A, B) = d(A, \neg A) = 1 \tag{169}$$

$$d(A \wedge B, A) = d(B, A \vee B) \tag{170}$$

$$d(A,B) = d(A \wedge B, A \vee B) \tag{171}$$

(Incidentally, (170-1) show the quadrilateral ACBD described above is a rectangle, as we claimed there.)

In view of the previous relations, we could define a *truth likelihood* value for A —approximating Popper's [and Miller's] *verisimilitude* measure— by making it to equal the distance between A and falsehood, i.e. $d(A,\perp)$. We obtain, immediately:

$$d(A,\perp) = d(\top,\perp) - d(\top,A) = 1 - d(A,\top) = 1 - d(A \triangle \top, \perp) = 1 - d(\neg A, \perp) = 1 - [\![\neg A]\!] = [\![A]\!] \tag{172}$$

So here we have a further interpretation of our "truth values" $[\![A]\!]$ in terms of *truth likelihood* or Popper's *verisimilitude* (a promising concept —only partially developed— which initially led this author to suggest it as a means to measure the *truth content* of a theory and to evaluate the distance to other theories). We recall that we already considered $[\![A]\!]$ (in sections 2 and 4, in our C1 interpretation of Θ) as a rough measure of *partial truth* or "truth content" of A. In a similar spirit, we are reminded that a puzzled Scott —and Katz after him— once suggested the "truth value" $[\![A]\!]$ of many-valued logics could be interpreted as one (meaning truth) less the *error* 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 $\varepsilon_A = 1 - [\![A]\!] = d(A,\top)$.

Further along the line of exploring related ideas, we now observe that, for any propositional 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 proposition being true (a number that Bar-Hillel and Hintikka call, appropriately, "truth-table probability" [23]). So this value's complement to one should seemingly correspond to the amount of information —in a loose sense— we have when the proposition is true. This is precisely what Bar-Hillel and Hintikka define as "degree of information", *semantical information* or *informativeness* $I(A)$ of a proposition A. (Viewed in our terms, $I(A)$ equals $1 - [\![A]\!]$, or $[\![\neg A]\!]$.) These authors use 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 propositions by combining them according to rules which amount to a multi-valued truth-value non-functional calculus of the functional-cum-supervaluation type. So, they suggest that equalities

$$I(A \vee B) = I(A) \cdot I(B) \qquad \text{and}$$
$$I(A \wedge B) = I(A) + I(B) - I(A) \cdot I(B)$$

are appropriate, but that $I(A \wedge \neg A)$ and $I(A \vee \neg A)$ should be assigned, respectively, the 1 and 0 values (supervalues, in fact, in Van Fraassen's sense [42]). All these values coincide with those predicted by the **BML** formulas: indeed, if A and B are assumed *independent* the above equations follow automatically from our formulas, while if $B = \neg A$ the 1 and 0 are a consequence of the zero compatibility between A and $\neg A$.

We now define a new measure in $\mathcal{L}$:

[*Definition:*] *Imprecision* of a proposition (or theory) A in $\mathcal{L}$ (or $\hat{\mathcal{L}}$) is the value for A of the function

$$f : \mathcal{L} \longrightarrow [0,1] \quad \text{such that} \quad f(A) = [\![A \leftrightarrow \neg A]\!]^+ \tag{173}$$

(We could as well have used the word "fuzziness" instead of *imprecision*, but the former term —aside from its technical adequacy— is so much tied to Fuzzy Set Theory that it is better to leave it there; moreover, "imprecision" is a rather neutral descriptor, aptly covering *vagueness* —a formal imprecision built *inside* the language— as well as *uncertainty* —of a more epistemological nature.) It is immediate that:

$$f(A) = 1 - d^+(A, \neg A) \tag{174}$$

(and note that, had we used $[\![A \leftrightarrow \neg A]\!]$ in (173) or $d(A, \neg A)$ in (174), we would have had an identically zero function f —and a useless definition—).

Some authors, like Zawirski and Łukasiewicz in the thirties, have justified the deviation from classical principles arising from many-valued logics (especially the non-validity of (7)) not as a drawback of these logics but, on the contrary, as due to the propositional *imprecision* these logics for the first time allow to acknowledge and formalize [35]. They subsequently tend to define imprecision as a measure of how blurred the distinction between a proposition and its negation can appear to be or, equivalently, as a measure of the logical compatibility between them. Or, still, as the inability of the particular logic we use, given an imprecise proposition, to distinguish between positive and negative (so that the more imprecision present the less "separating" or "resolving power"of that logic). In **BML**, as in any other Boolean logic, there is no deviation from classical principles, a proposition and its negation are *always* clearly *distinct*, and they are characterized just by having *zero* compatibility. However, Zawirski's idea can be translated —more usefully than apologetically here— into **BML** (which, as any other many-valued logic, also recognizes and accommodates imprecise statements —i.e. those valued in (0,1)— naturally). In fact, propositional distinctiveness or compatibility are not merely intrinsic properties of propositions; they can be affected as well by our analysis of them (by means of the v valuation): thus, we could as well treat two propositions A and B as possibly indistinct, or independent, or fully compatible, though it may happen that the second is just $\neg A$. So imprecision, rather than being an excuse for ruining laws in the receiving logic, can be redefined as the observer's failure to tell a proposition from its negation (or, equivalently, his inability to identify a contradiction) *in the absence* of the relevant information to make such a distinction possible. More precisely, what $f(A)$ measures is the degree of error we make when in **BML** we find a proposition indistinguishable from its negation, i.e. when we consider —because we lack information or for other reasons— a proposition *compatible* with its own negation. (This is what definition (173) formalizes.) Under this perspective, it is clear that imprecision is an error *we* make. It is more an effect of our measures and considerations over $\mathcal{L}$ (i.e. of the v valuation and the previous distinctions made in $\mathcal{L}$) than a property of the logic or the propositions themselves. Though we as users attribute imprecision to propositions, it does not originate in them but in the way we view or treat them. Indeed, a statement like "the temperature is high" can be precise (or not) depending on the user's point of view: use of (un)tolerably vague language —to match physical perceptions—, or (un)certainty about fact —due to, e.g., (un)availability of a thermometer. It is this standpoint what determines the value he or she assigns to the current proposition —and then a binary value identifies a precise proposition, and an imprecise one gets in **BML** a value in (0,1). All the same, a sentence like "the temperature is low" may be felt as contradictory —and so logically incompatible ($\alpha = 0$)— with the above statement, or it may be felt as not necessarily incompatible with it ($\alpha > 0$) —so their conjunction, for instance, may not (and does not) equal zero. The source of the problem, we admit, is formally indifferent to us. Sometimes it derives from an imprecise use of the language to which the proposition belongs (and then it is *vagueness*), sometimes it is due to an imperfect knowledge of the fact that the proposition is (or is not) actually the case in the real world (and then it is *uncertainty*). [In the context of theories we would rather speak of *ambiguity* and *partial truth* in a Popperian sense].

As is easily deducible [from (164)], we have:

$$f(A) = 2 \min ([\![A]\!], 1 - [\![A]\!]) \tag{175}$$

which is equivalent to saying that the imprecision of a proposition A equals twice the error we make when in **BML** we evaluate on A the truth of the law of non-contradiction or of the excluded middle by considering there really *is* maximum compatibility between A and $\neg A$ (This equivalence is trivial to verify, since the right-hand expression in (175) equals $2 \times [\![A \wedge \neg A]\!]^+$ or $2 \times (1 - [\![A \vee \neg A]\!]^+$.)

Following this line of reasoning, we observe:

- Classical [two-valued] logic ("**B2L**") is the special case of **BML** in which *all* propositions in $\mathcal{L}$ have *zero* imprecision; if there exists just one imprecise proposition then we are in **BML**. (In Gaines's peculiar terminology [–see next section–], **B2L** and **BML** would apparently correspond to what this author somewhat bizarrely calls "probability logic"and "fuzzy logic", respectively.)

- Measure f fulfills a number of conditions that suggest it can be used as an *entropy* measure (in the sense of De Luca & Termini [14]); these conditions are:

  - $f(A) = f(\neg A)$ (i.e. affirmation and negation of A are both equally imprecise)
  - There are in $\mathcal{L}$ 'minimal' propositions (i.e. A such that $f(A) = 0$) as well as 'maximal' propositions (i.e. A such that $f(A) \in (0, 1]$)
  - If $[\![A]\!] < [\![B]\!] < 1/2$ or $[\![A]\!] > [\![B]\!] > 1/2$ then $f(A) < f(B)$

- Naturally, a proposition reaches its highest imprecision value when its truth value is 1/2.

Note, however, that the f measure sometimes plays —often forcedly— the role of a measure of *ignorance*. But, if so, assimilating ignorance to maximum imprecision seems too simplistic (ignorance would then reduce to the $[\![A]\!] = 1/2$ case). Instead, a situation of ignorance seems to be better and more realistically modeled not just inside a proposition, with no concern for alternatives, but rather in the form of a certainty function (plausibly of the sub-additive type) like the one we shall mention in section 11 and we shall call —after Shafer— 'vacuous'.

10. SPECIAL CASES

BML as a logic is general and powerful enough to include familiar, well-known logics as particular instances of itself. This is the case of (a) standard *classical* logic —normally considered to be necessarily two-valued— and of (b) standard [but *non-classical*] many-valued logics like Łukasiewicz-Tarski's $Ł_\infty$ infinite-valued logic [27] (or, more simply, the three-valued formalisms by Kleene [24], Bochvar [5], Łukasiewicz [26] [and others] we already treated in section 8 as a special case of **BML**). But **BML** also helps to explain certain scattered results or ideas that are not normally part of the logic mainstream. We have mentioned already Boole's [7], MacColl's [28] and Peirce's [31] views on [what turned out to be] propositional functions, Carnap's view on [limited] induction [9], or Scott's view of error [39], among others. All of them can be naturally explained inside the **BML** frame. We now turn to other less-known past ideas —more for illustration than straight theorizing— like: (c) the attempt by Gaines at deep-rooted generalization covering both classical and many-valued logic [16][46] (an effort rather akin to ours, at least in its starting point if not in methodology or results), (d) past attempts at formalization of [linguistic] vagueness [4][15][37], (e) attempts to axiomatize partial belief or rational expectations [10][13][43], or even (f) the logic of Quantum Mechanics [6][44]. (Also, Dempster-Shafer theory [40] will be mentioned in the next section.)

What, instead, *cannot* be considered a special case of **BML** is either *intuitionistic logic* (since it operates on a propositional structure which is emphatically *not* Boolean) or traditional *modal logic*. We can integrate modal aspects into **BML** (as we did in view D of sections 2 and 4) but any logic built on this interpretation would be too simple, would have no provisions to accommodate complex Kripkean structures, and would give no clear idea of what is to be understood by "truth" [in the present world] (so, e.g., $[\![A]\!] = 1$ may be translated into "A is necessary [or whatever]", but what is to be done with "A is true" or "$\Box A \rightarrow A$" remains unclear) or of what sense is to be made of operator iteration like in "$\Box\,\Box A$". Moreover, such a logic would block any modal formalization that would not include the necessitation rule. Yet some insights are easily obtainable: the $\mathcal{L}\,/\,\mathcal{P}(\Theta)$ isomorphism allows us to see the $\Box$ operator as a *set* operator that takes $\mathbf{A}$ into $\Box\mathbf{A}$ (where $\Box\mathbf{A}$ is clearly taken to mean "the set of eventualities where A is necessary, believed —or known—, compulsory, program-reachable, provable, etc."). Under such conditions, in modal systems $\mathcal{M}$ having the T axiom (the "necessity" or "knowledge" condition $\vdash_{\mathcal{M}} \Box A \rightarrow A$), this axiom just characterizes the $\Box$ set operator as antitonic with respect to the set-inclusion order, because it merely says that $\Box\mathbf{A} \subset \mathbf{A}$, so that when we state in such systems "$\Box A$" (e.g. "A is necessary" or "known") we just mean that the $\Box$ set-operator restricts $\mathbf{A}$ into a subset of itself —thus making the "necessary" or "known" instances a subset of the merely true ones—. (Naturally, pure *belief* systems would not generally satisfy the condition, and in them $\Box$ would take any $\mathbf{A}$ —representing the believed proposition— into any other set —representing that $\mathbf{A}$ is believed—, possibly quite unrelated to the original.) Also, we could in turn value modal propositions, so $[\![\Box A]\!]$ would have sense as the measure μ of the $\Box\mathbf{A}$ set, and $[\![\diamond A]\!]$ as $1 - \mu(\Box(\mathbf{A}^c))$.

a. Classical two-valued logic (B2L)

B2L is the special case of **BML** in which *every* proposition is *binary* (i.e. $\forall A \in \mathcal{L}\ \ [\![A]\!] \in \{0,1\}$) or, equivalently (by (175)), in which *every* proposition has *zero* imprecision (i.e. $\forall A \in \mathcal{L}\ \ f(A) = 0$). Thanks to the representation isomorphism, we know that this corresponds either to a *selective* valuation or, more generally, to a situation where all non-zero-valued *possibilities* in Θ behave in the same way (note that if there a unique such non-zero-valued possibility then we are back in the case of a selective valuation). Thus, **B2L** corresponds either to a universe where there is only one eventuality to consider (= the "real world" or "the" [only] case we consider) or to a universe where all eventualities θ considered behave solidly in the same manner (e.g., responses make packed unanimity). We can say, then, that "a proposition is true or false in **B2L** if and only if it is true or false simultaneously in all the relevant worlds considered (i.e. all the θs such that $\mu(\theta) \neq 0$)". **B2L** is clearly both *Boolean* and *functional*. Having the first property means that the propositions in the logic can be assumed to make up a Boolean algebra, and no property of the valuations the logic allows on its propositions is inconsistent with this assumption. Having the second (= *functionality*) means that values of composite propositions are functions of -only- the values of component propositions, so $[\![A * B]\!] = [\![A]\!] *' [\![B]\!]$ for any [binary] connective '$*$' (in other words, the valuation is a homomorphism that preserves logical operations).

b. Łukasiewicz-Tarski's L_∞ logic

Łukasiewicz & Tarski's L_∞ logic [27], as **BML**, generalizes **B2L** in the sense that it allows the members of the lattice of propositions to take values in [0,1] other than 0 or 1. Both systems of many-valued logic include classical **B2L** as a special case. Nevertheless, while **BML** gives up *functionality*, L_∞ gives up *Booleanity*. In fact, if the $\{0,1\}$ truth set is extended to [0,1] those two properties of **B2L** cannot be both upheld, and either must be given up.

Indeed, if we wanted to extend **B2L** to the [0,1] truth set case, a reasonable and near-minimal way of doing it is that proposed by Bellman & Gierz [1], which is reducible to postulating (a) that two dual logical operations $\wedge$ and $\vee$ can be defined in the set of propositions that have the commutativity, associativity and distributivity (and so also idempotency) properties, and that this set is *bounded* [with respect to the order $\leq$ defined by those properties] (so that the set is a bounded *distributive lattice*), (b) that there is a [0,1]-valuation of the propositions in this lattice which is a monotonic [increasing] homomorphism (a condition we could call *principle of functionality*). From such elementary, simple and undemanding conditions we deduce immediately (if we note by "$[\![A]\!]$" the value of proposition A)

properties (13-15). But (14) and (15), together with idempotency and the *functionality* condition, yield at once:

$$[\![A \wedge B]\!] = \min([\![A]\!], [\![B]\!]) \tag{176}$$

$$[\![A \vee B]\!] = \max([\![A]\!], [\![B]\!]) \tag{177}$$

and this has some serious consequences we mention below. [But there are other consequences —of proof-theoretic nature— of having the (176-7) constraints built *inside* the logic; they will be mentioned in section 12]. The first consequence is that we happen now to be fully in $Ł_\infty$. The second is that once we get (176-7) both $\neg(A \wedge \neg A)$ and $A \vee \neg A$ can no longer be theorems of the logic (because they yield the 1 value *only* when A is assigned a binary value, so it does not in general equal 1 —a necessary condition). But this is a serious snag: we are compelled to give up *classical* Aristotelian principles —warranted by Booleanity (= conditions (1-7) but especially the last one)— *not* because we want it so but because the algebraic conditions we —or Bellman & Gierz— impose are such that the *non-classical* outcome is unavoidable. The main culprit here is, demonstrably, what we termed "principle of functionality". It breaks Booleanity so narrowly that, by not postulating it, we can preserve the Boolean character of the system and, with it, classical principles (this is just what **BML** does). From a purely algebraic point of view, the giving up of the two traditional laws (7) of Logic means that propositions now form a De Morgan algebra —a fairly interesting structure but considerably *weaker* than Boolean algebra and structurally rather poor (especially as compared to the commonplace generality and simplicity of the latter —otherwise so widely known and universally mastered). What is more surprising is that we give up Booleanity *not* by invoking some metaprinciple (like, e.g., not believing in the validity of the law of the excluded middle —that the intuitionists deny— or doubting the logical value of contradictions constructed through –e.g.– vague or nonsense sentences). Instead, we give up simple and basic theorems only because of technical reasons: we just do not have the laws because we do not seem to be capable of guaranteeing them in the system. Łukasiewicz's comforting thoughts —that many-valued logic allows us creatively to break bivalency, as he was preaching an unconvinced audience in the nineteen-thirties— sound rather like a poor excuse, since the problem is not that many-valued logic may allow law-breaking if we wish it —and moreover, of course, have it all under control— but that in $Ł_\infty$ we are *compelled* to it, with no way out: the fact of valuing *one* single proposition in (0,1) automatically makes the principles of non-contradiction and the excluded middle *invalid*.

BML gets over all this. Both principles are explicitly recognized —in (7)— as theorems (in fact, axioms) of the system, and so we do not walk into non-classical logic, a costly step to make —particularly if there is no compelling need to make it. So, Booleanity is preserved —at the expense of giving up truth-functionality (except for binary values). Added advantages are:

- **BML** maintains a formal symmetry —though with a different interpretation— with Probability Theory (or, more generally, with Measure Theory) whose valuations —also in [0,1]— are unarguably *non-functional* and, moreover, made on a lattice that is [nearly always] assumed Boolean.

- **BML** allows $\mathcal{L}$ to be represented on a Boolean set-algebra, which is both intuitive and easy (to be compared with $Ł_\infty$, that allows to represent the original lattice of propositions only on comparatively awkward and unexplored De Morgan algebra structures, with no clear intuitive meaning).

- **BML** admits $Ł_\infty$ as a special case since, indeed, $Ł_\infty$ behaves exactly as **BML** would do if *no* proposition in $\mathcal{L}$ could be recognized as a *negation* of some other and, then, it would be assigned systematically the $\oplus$-case connective formulas. Naturally that would give an error in the values of composite propositions involving non-fully-compatible propositions, but it would also restore the lost truth-functionality of two-valued logic. Recalling what the $\oplus$-case stands for, $Ł_\infty$ amounts —from the **BML** perspective— to viewing *all* propositions as having *always* maximal mutual compatibility. That means that $Ł_\infty$ conceives *all* propositions as *nested* (i.e. for every A and A', either $\mathbf{A} \subset \mathbf{A'}$ or $\mathbf{A'} \subset \mathbf{A}$ where $\mathbf{A}$ is $\rho(A)$). (Such a picture is strongly reminiscent of Shafer's description of conditions present in what he calls *consonant* valuations, see section 11.)

Thus, in **BML** terms, $Ł_\infty$ is a special case of **BML** (or, better, an approximation to it under special assumptions) characterized by systematic usage of the $\oplus$-formulas. This approximation (a simplification, actually) implies maximum compatibility between all pairs of propositions of $\mathcal{L}$, which entails the fiction of a *total*, linear order $\vdash$ in $\mathcal{L}$ —and $\subset$ in $\mathcal{P}(\Theta)$— (that means a coherent, negationless universe) and which is probably the best assumption we can make when information on propositions is lacking and negations are not involved —or cannot be identified as such. (Such an option is as legitimate as that of assuming, in the absence of information on propositions, that these are independent.) In $Ł_\infty$ obvious negations may be a problem, but it can be solved by applying [error-correcting] *supervaluations* —that **BML** supplies automatically. (And note that, in particular, the error incurred in by $Ł_\infty$ when failing to distinguish between a proposition and its negation —thus not being able to recognize a contradiction— is just the quantity we called *imprecision* in section 9.)

c. Gaines's 'Standard Uncertainty Logic'

In an interesting effort to discriminate between algebraic structures and truth valuations on them —somewhat paralleling and anticipating the aim of the present work— Gaines set out to define what he termed a 'Standard Uncertainty Logic' [16][46], that covered and formalized two common subcases. This logic postulates (a) a proposition lattice with an algebraic structure that is initially assumed Boolean but —by technical reasons— finally admitted to be merely distributive, and (b) a monotonic valuation of the lattice on the [0,1] interval, with the same conditions (9-11) set here for **BML**. Gaines distinguishes two special cases of his logic: (1) what he calls 'probability logic' and defines to be the particular case of his logic where the law of the excluded middle holds (that turns out to be actually coextensional with classical logic, our **B2L**), and (2) what he terms 'fuzzy logic', defined (in one among several alternative characterizations) to be his logic when that law does not hold. Apparently, Gaines's general logic is our **BML** (that covers **B2L** as a special case, just as Gaines's logic becomes his 'probability logic' when "all propositions are binary"), but closer examination reveals that Gaines's 'fuzzy logic' is, simply, $Ł_\infty$. This fact is spelled out by the property he mentions of propositional equivalence (here in our notation):

$$[\![A \leftrightarrow B]\!] \; = \min\,(1 - [\![A]\!] + [\![B]\!], 1 - [\![B]\!] + [\![A]\!])$$

in which the right hand is clearly $1 - |[\![A]\!] - [\![B]\!]|$. This is an expression that is deduced from Gaines's postulates *only* if, necessarily, either $A \vdash B$ or $B \vdash A$ (note this either/or situation is explicitly mentioned by Gaines as a characteristic property in his 'fuzzy logic'). So Gaines makes the implicit assumption that the base lattice is *linearly* ordered (perhaps induced to it by the $\leq$ symbol used for the [partial] propositional order in the lattice). A confirmation for this comes from the fact that classical principles —that hold, by definition, in his 'probability logic'— do not hold in this one. No wonder, then, the base lattice cannot be but merely distributive.

d. Black's and Fine's formalizations of vagueness

After Russell's [37], Black's approach to vagueness [4] has an essentially historical interest today, but we mention it because some of his ideas are curiously reminiscent of particular items of **BML** (especially when interpreted under the A6 viewpoint of section 2 above). Thus, Black suggests valuing the 'consistency' of a "vague entity" by means of the formula:

$$C \; = \; \lim_{n \to \infty} \frac{p}{n}$$

(where p is the number of cases in which the entity is applicable, n the number of cases considered), which, aside from the limit, fits pretty well with our "truth" —or rather applicability— value $[\![A]\!]$ in the case where μ is a uniform measure on $\mathcal{P}(\Theta)$ (so $[\![A]\!]$ is the relative cardinal of **A** in Θ) and the θs are interpreted as application instances, as in A6 (see section 2). Another component of Black's conception that recalls ours is what he calls 'z value' or 'deviation of consistency', that he —obscurely— defines as:

$$z \; = \; \frac{2\,m}{n}$$

where the [undefined] m is called '[number of] most favoured judgments'. Considering the author wanted this measure to capture the degree of vagueness associated to a given vague entity, Black's idea seems to be that of having a measure roughly like our *imprecision* because, to judge from the "z curve" illustrating the paper,

$$z \; = \; 2 \times (1 - \; \max\,(C, \bar{C})) \; = \; \frac{2 \min\,(p, 1 - p)}{n}$$

where the right-hand part of the formula is equivalent to the expression defining $f(A)$ in (175).

Fine [15] also approaches vagueness in a way that **BML** concepts can perhaps illuminate. He essentially advocates for supervaluations on vague propositions that can reveal the truth of $\neg A$ when A is [vague and] false, or the truth of $A \vee \neg A$ (always, no matter A is precise or vague). He calls this mechanism 'supertruth' and says it reveals truth-related 'penumbral connections' existing between vague propositions. But these are automatically resolved in **BML**, so there is no need to supervalue to get the 'supertruth' of vague sentences: vagueness is naturally assimilated there to general imprecision, gets no specific treatment, and all the logic machinery applies.

e. Vicker's formalization of partial belief

Consideration of psychological or epistemic factors in determining the probability function is a time-honored tradition that derives from illustrious precedents like Keynes, Ramsey [33], de Finetti [13] or Carnap [10], who in their formalization of 'subjective probability' (or Carnap's theory of 'partial belief') view probability as an [additive] valuation satisfying specific axioms and "rationality" conditions. Carnap based his theory on [two-valued] predicate logic, but Vickers [43] later asked if an axiomatization of probability could be developed on non-binary logics. He suggested a minimal axiom set he called 'system S':

1. If A is necessary, then $p(A) = 1$.

2. If A and B necessarily exclude each other, then $p(A \vee B) = p(A) + p(B)$.

3. If p is "rational" then necessarily $p(A \wedge B) \leq p(A)$.

5. If a is an atomic proposition, then a and $\neg a$ are inconsistent.

6. Every proposition can be expressed as a disjunction of atoms (= logically independent elementary propositions).

7. [*Principle of Transparency*:] If A' is a proposition such that believing A entails believing A', then $p(A) \leq p(A')$.

Vicker's axioms —with number 4 (involving first-order logic) absent— are all satisfied in **BML**. Any reasonable interpretation of axiom 1 —where 'necessary' presumably means something like "$[\![A]\!] = 1$ for any valuation"— holds there, and axioms 5 and 6 are a consequence of the fact that $\mathcal{L}$ is a Boolean atomic lattice, while axioms 2 and 7 are included in the definition of the v valuation (9-11) and 3 is a consequence of it. Thus, not only **BML** supplies a logical infrastructure on which to build a theory of 'rational belief', but **BML** itself can be thus considered "rational" —in the Carnap-Vickers sense.

f. Quantum logic

Though the logic of Quantum Mechanics arises in Birkhoff and Von Neumann's pace-setting 1936 article, the version we now discuss is the relatively sophisticated one by Bodiou [6] or Watanabe [44]. The starting point is the postulate according to which the base propositional lattice is *not* distributive: one of the current reasons for that is Hilbert space structure (about which we shall not argue); a further reason is the particular interpretation that this logic forces on certain experimental results. Bodiou suggests the following equivalences between probability and "truth":

1. For every law $p(b|a) = 1 \iff_{df} a \leq b$
2. For every law $p(b|a) = 0 \iff_{df} a \leq \bar{b}$ (a and b are "*contradictory*")
3. For every law $p(b|a) \neq 1 \iff_{df} a \wedge b = \phi$ (a and b are "*incoherent*")
4. $a \leq b \iff_{df} \vdash a \rightarrow b$

In our notation this reads

1. $(\forall v)\ [\![B|A]\!] = 1 \iff A \vdash B$
2. $(\forall v)\ [\![B|A]\!] = 0 \iff A \vdash \neg B$

3. $(\forall v)\ [\![B|A]\!] \neq 1 \iff A \wedge B = \perp$

4. $A \vdash B \iff\ \vdash A \rightarrow B$

All relations are **BML** theorems *except* 3, since in **BML** we have:

$$(\forall v)\ [\![B|A]\!] = 0 \iff A \vdash \neg B \iff A \wedge B = \perp\ \Rightarrow\ (\forall v)\ [\![B|A]\!] \neq 1$$

and the last implication cannot be reversed (but see point (3), towards the end of this subsection). Notice that what Bodiou calls "incoherence" ($A \wedge B = \perp$) is a special case of our *incompatibility* ($\alpha_{AB} = 0$, implying either $A \wedge B = \perp$ or $\neg A \wedge \neg B = \perp$).

A reason often given by 'quantum logic' proponents (e.g. Reichenbach in 1944) to explain why distributivity fails in this logic is the interpretation of certain experimental results where:

$$p(a) \cdot p(b|a) + p(\bar{a}) \cdot p(b|\bar{a}) < p(b) \tag{178}$$

a relationship we write in this way:

$$[\![A]\!] \cdot [\![B|A]\!] + [\![\neg A]\!] \cdot [\![B|\neg A]\!] < [\![B]\!] \tag{179}$$

which is the negation of (101) (a theorem we called *distribution*). But this is normally interpreted [by quantum logicians (e.g. Watanabe)] as:

$$(A \wedge B) \vee (\neg A \wedge B) \neq B \tag{180}$$

which obviously signals the breaking of distributivity. However, the even-handed transcription of (178) into (180) is rather cavalier and easy-going. Actually, the reading (178) of experimental results translates immediately, in **BML** notation, into (179). From there we conclude that:

$$[\![A \wedge B]\!] + [\![\neg A \wedge B]\!] < [\![B]\!] \tag{181}$$

which is in clear violation of additivity (11) [but not of distributivity (5)]. The above transcription process illustrates two facts: (a) going directly from (178) into (180) not only transforms *probabilities* [of propositions] into plain *propositions* but unmotivatedly trades an arithmetic *plus* for a logical *join*, and (b) if the (178) experiment has any logical relevance whatsoever, it is for other reasons: (181) implies that the v valuation defined in $\mathcal{L}$ is *subadditive* (making $B = \top$ in (181) one has $[\![A]\!] + [\![\neg A]\!] < 1$), but this quite a different matter. Actually, (181) does not imply the failure of any logical principle, and $\mathcal{L}$ is still Boolean. The only condition that fails to hold in the cases (178) we examine is additivity (11), but this failure is otherwise very reasonable if we admit, with Heisenberg, that in precisely these cases the observer can hardly (or not at all) evaluate experiment conditions and results (and, thus, it is not unnatural that paying the price for the (178) anomaly falls on [our] valuation v of propositions and definitely *not* on their properties or structure).

Bodiou for one defends quantum logic as a useful step towards a more general and ambitious logic he calls "dialectic" [6]. He bases his claim on three reasons —all of which, he contends, support the breaking of distributivity:

1) Non-binary valuation of propositions implies non-distributivity.

2) Certain experimental Quantum Physics contexts are only interpretable without distributivity —or explicitly denying it.

3) Distinguishing incoherence ($A \wedge B = \perp$) from contradiction ($A \vdash \neg B$) is not only reasonable but the contrary (denying the distinction) would be unacceptable; moreover, it allows mutually incoherent —yet non-contradicting— theories to coexist in a same lattice and develop dialectically into a more general and coherent new synthesis theory.

Now, Bodiou's aim, attractive as it may seem, either strains Logic too much (when banning Boolean or merely distributive structures from it) or else it demands too much from it (e.g. distinguishing two ordinarily indistinguishable situations). And, we contend, one requirement is simply unnecessary and misconceived, while the the other is superfluous and can be achieved by other means. As for the point 1 (above), Bodiou's argument is just false: **BML** itself is a counterexample. As for point 2, we have already seen that —aside from the transcription problem from probabilities to propositions— empirical result (178) admits a simple explanation amounting to no more than the presence of subadditive valuations.

And for point 3, **BML** cannot make the proposed distinction —as, in **BML**, incoherence is always equivalent to mutual contradiction (i.e. $A \wedge B = \perp \iff A \vdash \neg B$). But, first, that distiction is not what it seems: the **BML** equivalence is after all what everyone would accept as reasonable; secondly, part of the urge for the distinction derives from a connotational, parasitic sense of the phrase "mutual contradiction" as "invalidation" [of one theory by the other], so it may sound plausible to think of two coexisting incompatible theories yet not contradicting (="invalidating") each other. Nevertheless, if this sense-splitting is just what is wanted, subadditive valuations may, again, be an acceptable and more simple solution.

11. IGNORANCE AS SUBADDITIVITY

Suppose now that the propositions in $\mathcal{L}$ satisfy (1-8) —so we are still in a Boolean algebra— but the v valuation we impose on $\mathcal{L}$ is *subadditive*, i.e. monotonicity (10) and (9) hold but instead of (11) we get:

$$\{\![A \wedge B]\!\} + \{\![A \vee B]\!\} \geq \{\![A]\!\} + \{\![B]\!\} \quad \textit{(Subdditivity)} \tag{182}$$

where we note explicitly by the $\{\![\ ]\!\}$ brackets that v is subadditive. **BML** with (182) instead of (11) will be denoted by $\mathbf{BML}_<$. Instead of (12-17) now he have:

$$\{\![\neg A]\!\} \leq 1 - \{\![A]\!\} \tag{183}$$

$$\{\![A \wedge B]\!\} \leq \{\![A]\!\} \leq \{\![A \vee B]\!\} \tag{184}$$

$$\{\![A \wedge B]\!\} \leq \min(\{\![A]\!\}, \{\![B]\!\}) \tag{185}$$

$$\{\![A \vee B]\!\} \geq \max(\{\![A]\!\}, \{\![B]\!\}) \tag{186}$$

$$\{\![\textstyle\bigvee_{i=1}^{i=n} A_i]\!\} \geq \sum_{i=1}^{i=n}\{\![A_i]\!\} - \sum_{i<j}\{\![A_i \wedge A_j]\!\} + \ldots + (-1)^{n+1}\{\![\bigwedge_{i=1}^{i=n} A_i]\!\} \tag{187}$$

$$\text{If for all } i \text{ and } j\ (i \neq j)\ \ A_i \wedge A_j = \perp \ \text{ then } \ \{\![\textstyle\bigvee_{i=1}^{i=n} A_i]\!\} \geq \sum_{i=1}^{i=n} \{\![A_i]\!\} \tag{188}$$

Condition (187) characterizes the subadditive valuation v as a *lower probability* or as a *belief* $Bel(A)$ [in Shafer's sense] (see [40]). If the inequality signs were inverted, the valuation would become an *upper probability* or [Shafer's] *plausibility* $Pl(A)$. The defective value —the part of the value $v(A)$ attributed neither to A nor to $\neg A$— can be expressed as:

$$\partial(A) = 1 - (\{\![A]\!\} + \{\![\neg A]\!\}) = (1 - \{\![\neg A]\!\}) - \{\![A]\!\} = Pl(A) - Bel(A) \tag{189}$$

We know that a subadditive valuation like $v : \mathcal{L} \longrightarrow [0,1] : A \mapsto \{\![A]\!\}$ —or, better, the measure $\mu : \mathcal{P}(\Theta) \longrightarrow [0,1] : \mathbf{A} \mapsto \mu(\mathbf{A})$ induced on $\mathcal{P}(\Theta)$ by that valuation— defines a function $m : \mathcal{P}(\Theta) \longrightarrow [0,1]$ (called "*basic assignment*" by Shafer) that satisfies:

$$1.\ m(\phi) = 0 \tag{190}$$

$$2.\ \textstyle\sum_{\mathbf{A} \subset \Theta} m(\mathbf{A}) = 1 \tag{191}$$

so that the $\mu(\mathbf{A})$ (or $\{\![A]\!\}$) values are computed from this measure through

$$\{\![A]\!\} = \textstyle\sum_{\mathbf{B} \subset \mathbf{A}} m(\mathbf{B}) \tag{192}$$

and, conversely, the $m(\mathbf{A})$ values can be obtained from $\{\![A]\!\}$ (or $\mu(\mathbf{A})$) through

$$m(\mathbf{A}) = \textstyle\sum_{\mathbf{B} \subset \mathbf{A}} (-1)^{|\mathbf{A}-\mathbf{B}|}\, \mu(\mathbf{B}) \quad \text{for any } \mathbf{A} \subset \Theta \tag{193}$$

Incidentally, it is obvious that $Bel(A)$ and $Pl(A)$ coincide with the traditional concept (in Measure Theory) of *inner measure* (P_*) and *outer measure* (P^*), so that the following chain of equivalences has a transparent meaning (notice Shafer calls $Bel(\neg A)$ "degree of doubt of A"):

$$Pl(A) = P^*(\mathbf{A}) = \textstyle\sum_{\mathbf{B} \cap \mathbf{A} \neq \phi} m(\mathbf{B}) = \sum_{\mathbf{B} \subset \Theta} m(\mathbf{B}) - \sum_{\mathbf{B} \subset \mathbf{A}^c} m(\mathbf{B}) = 1 - P_*(\mathbf{A}^c) = 1 - Bel(\neg A) \tag{194}$$

We know, also, that an *additive* valuation μ is just a basic assignment m such that $m(\mathbf{A}) = 0$ for all $\mathbf{A} \subset \Theta$ *except* for the singletons $\{\theta\}$ of Θ. This is what Shafer calls '*Bayesian* belief'; it has a very intuitive general interpretation here: if a valuation μ is additive then each eventuality θ of Θ has a basic assignment value $m(\{\theta\})$ and the value $\mu(\mathbf{A})$ associated to each set of $\mathcal{P}(\Theta)$ is obtained *strictly* by adding the basic assignments of their eventualities. By contrast, if the valuation is subadditive then not only each eventuality has an assignment value but each subset of Θ has an assignment as well, and the value $\mu(\mathbf{A})$ associated to each set $\mathbf{A}$ of $\mathcal{P}(\Theta)$ is obtained by adding the basic assignments of their eventualities *plus* those of all sets contained in $\mathbf{A}$. In other words, an *additive* valuation implicitly assumes all eventualities in Θ are *independent* (in the following sense: they alone allow us to determine the value $\mu(\mathbf{A})$ of any set of which they are part); on the other hand, a *subadditive* valuation assumes the eventualities are *interdependent* (in the sense that the cumulative value of their basic assignments do not suffice for computing the $\mu(\mathbf{A})$ value: to obtain it, one has still to add them the assignments corresponding to the different sets they may form inside $\mathbf{A}$). This has a transparent meaning under an interpretation of the A1 type (=the θs as question-answering individuals): a subadditive valuation means simply that opinions are due not only to the individual θs' views of the proposition A (they may even have null weight) but *also* to those of opinion-forming collectives, whose emerging opinion "influences" —adds to— that of the individuals in them.

But there are other examples, like this classic result from opinion surveys. Confronted with the question "Should the Government allow public speeches against democracy?", 25% of those polled assented. Substituting the word "prohibit" for "allow" elicited 54% of assenting responses. Since both words are antonyms (the contrary of prohibiting is allowing), it is clear that people had an unattributed gap left between those two complementary concepts, thus:

$$\{\![A]\!\} + \{\![\neg A]\!\} = .25 + .54 \leq 1 \tag{195}$$

which reveals that people value proposition A (=speeches allowed) *subadditively*, and also that "allow" and "fail to prohibit" are analyzable, in Shafer's terms, as $Bel(A)$ and $Pl(A)$, respectively, with values .25 (for belief) and .46 (for plausibility).

One family of subadditive valuations Shafer considers is that of *consonant valuations*, defined through

$$\{\![A \wedge B]\!\} = \min(\{\![A]\!\}, \{\![B]\!\}). \tag{196}$$

One property of such valuations is that all focal elements are nested (see [40]), or, in our language, for every $\mathbf{A}$ and $\mathbf{B}$ with non-zero asignment value m, necessarily $\mathbf{A} \subset \mathbf{B}$ or $\mathbf{B} \subset \mathbf{A}$ (where, as always, $\mathbf{A} = \rho(A)$), so all "focal" sets in $\mathcal{P}(\Theta)$ are linearly ordered by set inclusion (this is, by the way, what we used to call *maximum compatibility* or "case $\oplus$"). Another property of consonant valuations —which Shafer portrays as based on absolutely non-conflicting evidence— is that, for an arbitrary proposition A, we are never in a position to grant positive degrees of belief to both sides of a dichotomy involving A, i.e.

$$\{\![A]\!\} = 0 \quad \text{or} \quad \{\![\neg A]\!\} = 0 \text{ (necessarily one of both).} \tag{197}$$

The particular instance of subadditive valuation in which all non-*true* propositions get the zero value, i.e.

$$\{\![A]\!\} = 1 \quad \text{iff } A = \top\,, \qquad = 0 \text{ otherwise} \tag{198}$$

is what Shafer calls '*vacuous* belief function' and that we prefer to call *ignorance* (because it models ignorance more acceptably than assigning maximum imprecision $f(A)$ to a proposition, see section 9). It is just the particular instance of valuation in which all non-*true* propositions get the zero value. Indeed, we have, for a given A (or for *any* A if there is *total* ignorance):

$$\{\![A]\!\} = \{\![\neg A]\!\} = 0. \tag{199}$$

This is to be compared —if we restrict ourselves to additivity— with being able to obtain just

$$[\![A]\!] = [\![\neg A]\!],$$

a condition also fulfilled in (199), wich can *never* give way —if v is to be additive— to both values A and $\neg A$ being zero (as above), but to the mere and more prosaic

$$[\![A]\!] = [\![\neg A]\!] = 1/2, \tag{200}$$

which implies —through (174) or (175)— also $f(A) = f(\neg A) = 1$, so banally identifying *ignorance* with maximum *imprecision* and at the same time trivializing and impoverishing both concepts. By (198) or

(199), *ignorance* is a byproduct of our inability to value propositions (maybe just one) in the appropriate manner —i.e. additively—. This is distinct from *imprecision*, i.e. our inability to assign —additively— a $\{0,1\}$ value to a given proposition —and thus to *decide* on its truth or falsity; in that case all we can say is that $[\![A]\!]$ (and $[\![\neg A]\!]$) is in the open $(0,1)$ interval and that has non-zero *imprecision* $f(A)$.

Subadditivity enables us to analyze other interesting situations related to Logic. For instance, suppose we have a subadditive valuation assigning A the value $\mu(\mathbf{A}) = \{\![A]\!\} = Bel(A)$, and that a set $\mathbf{A}'$ (not necessarily a subset of $\mathbf{A}$) can be found in $\mathcal{P}(\Theta)$ such that there is an *additive* valuation which assigns $\mathbf{A}'$ precisely the same value. We denote $\mathbf{A}'$ by "$\Box\mathbf{A}$" (where $\Box$ is a set-operator of the sort mentioned in section 10), and $A' = \delta(\mathbf{A}')$ by "$\Box A$". So we have:

$$Bel(A) \;=\; \{\![A]\!\} \;=\; [\![\Box A]\!] \tag{201}$$

The $\Box$ is here a linguistic operator that acts on a sentence A and transforms it into another whose value is the belief (=subadditive truth-value) one can assign non-additively to A, so it seems proper to interpret $\Box$ syntactically as "it is believed that" or "α (= the name of a subject or rational agent) believes that" (in the latter case $\Box$ would rather have been indexed as "$\Box_\alpha$" or "$[\alpha]$"). Further, we would have

$$Pl(A) \;=\; 1 - Bel(\neg A) \;=\; 1 - \{\![\neg A]\!\} \;=\; 1 - [\![\Box\neg A]\!] \;=\; [\![\neg\Box\neg A]\!] \;=\; [\![\diamond A]\!] \tag{202}$$

where we have defined a new operator "$\diamond$" as an abbreviation for "$\neg\Box\neg$". In view of the (202) equality, it could be advisable to interpret $\diamond$ syntactically as "it is *plausible* that" or "α admits as credible" (so that "$\diamond_\alpha A$" —or $\langle\alpha\rangle A$— would read "α finds that A can be believed"), because α just does not believe the contrary. Naturally, if the *Bel* valuation happens to be additive ("Bayesian", in Shafer's terms), then the assigned values of $\Box A$, A and $\diamond A$ collapse into a single number and our truth-valuation of A amounts to our belief of A and also to its plausibility. If it does not, we can distinguish those three values and operate syntactically inside the language $\mathcal{L}$ of propositions. Since for any A we have $\delta = 1 - \{\![A]\!\} - \{\![\neg A]\!\} = Pl(A) - Bel(A) \geq 0$ and so $Pl(A) \geq Bel(A)$, we also have that the $\Box$ operator is consistent with axiom D of normal modal logics:

$$\Box A \rightarrow \diamond A \tag{203}$$

because, by monotonicity, (203) gives

$$Bel(A) \leq Pl(A)\ . \tag{204}$$

(But (203) is not necessary: (204) holds anyway, even when (203) does not.)

Likewise, if the $\Box$ set operator happens to shrink any $\mathbf{A}$ into $\Box\mathbf{A}$ so that $\Box\mathbf{A} \subset \mathbf{A}$ for any $\mathbf{A}$ then this condition translates into axiom T of modal logic:

$$\Box A \rightarrow A\ . \tag{205}$$

Now, this is an interesting case. If axiom (205) —variously known as "principle of knowledge" or "necessity" in *normal* versions of Modal Logic— does hold, then (203) and the usual reading of "$\Box A$" as "A is known [by someone]" or "A is necessary" —and of "$\diamond A$" as "A is admissible (= not known to be otherwise)" or "A is possible"— prompts by (202) and (203) a reinterpretation of our $\{\![A]\!\}$ and $1 - \{\![\neg A]\!\}$, or Shafer's *Bel* and *Pl* (or, equivalently, lower and upper probabilities P_* and P^*) as, respectively, "degree of knowledge" or "necessity" $Nec(A) =_{df} \{\![A]\!\}$ (= $[\![\Box A]\!]$) and "degree of admissibility" or "possibility" $Poss(A) =_{df} 1 - \{\![\neg A]\!\}$ (= $[\![\diamond A]\!]$); naturally, $Nec(A) \leq [\![A]\!] \leq Poss(A)$ for any A (and also $Nec(A) + Poss(\neg A) = Poss(A) + Nec(\neg A) = 1$). If necessity (or knowledge) of A and $\neg A$ are totally incompatible (in the sense that $Nec(\neg A) = 0$ whenever $Nec(A) > 0$) then Nec becomes a 'consonant valuation' of the type mentioned above, and (196) applies, so $Nec(A \vee B) = \max(Nec(A), Nec(B))$ and $Poss(A \wedge B) = \min(Poss(A), Poss(B))$.

Subadditive valuations on a given propositional lattice can be combined by obtaining their basic assignment functions m and computing from them the new valuation through the well-known Dempster rule, which is a conceptually simple operation.

It is interesting to notice that a subadditive valuation may be superimposed on a propositional lattice without breaking its Boolean character, so that $A \wedge \neg A = \perp$ and $A \vee \neg A = \top$ still hold, while $\{\![A]\!\} + \{\![\neg A]\!\} \leq 1$ at the same time. This slightly paradoxical fact may explain that many often-encountered situations —where mere subadditivity ($\{\![A]\!\} + \{\![\neg A]\!\} \leq 1$) was probably the case— have been analyzed historically as invalidating the law of the *excluded middle*, because it was felt that there was a "third possibility" between A and $\neg A$ making for the unattributed value $1 - \{\![A]\!\} - \{\![\neg A]\!\}$, covered neither by A nor $\neg A$. If our analysis is correct, such situations are analyzable in terms of incomplete valuations as in (183), but this does not imply the breaking of bivalence (= the (7) complementarity formula) of any Boolean algebra. Also, it allows analyzing the concept of *ignorance* (see end section 9) and the paradox of Quantum Logic (see (178) in previous section) —as well as Bodiou's claim that a pair of theories may be incoherent yet non-contradicting— as simply subadditive problems not entailing the breakdown of Booleanity or of ordinary logic.

Indeed, as for the quantum paradox, the transcription of (178) not into (179) but into (180) explains why the apparent anomaly can be blamed on the breaking of additivity, not of distributivity, as argued in section 10(f) above. As concerns Bodiou's claim, it can also be seen in $\mathbf{BML}_<$ (the subadditive version of **BML**) that 'incoherence' between theories does not entail their mutual contradiction —thus meeting Bodiou's condition— just by noting that now $\{\![A \rightarrow \neg B]\!\} \leq 1 - \{\![A \wedge B]\!\}$ and that under a given valuation we can have $\{\![A \wedge B]\!\} = 0$ and not have $\{\![A \rightarrow \neg B]\!\} = 1$ (though the converse still holds).

Subadditive valuations can also adequately formalize *certainty* (meaning that $\{\![A]\!\} = 1$ while at the same time $\{\![B]\!\} = 0$ for all $B \vdash A$) as they did —and along parallel lines— with *ignorance* (above). They can as well report satisfactorily how a scientific explanation frame (i.e. the appropriate lattice of theories that cover any observed or predictable true fact) is dynamically replaced by another once the first can no longer account for observed facts: as the valuation of explanations turns subadditive —reflecting that they no longer cover all predictable facts— one is naturally forced to replace the original propositional structure of theories by a new one (still a Boolean lattice) provided with a new —now again additive— valuation on it that restores the balance; the augmented lattice generators are the new vocabulary, and the elementary components of the new structure are the required new explanatory elements (atomic theories) for the presently observed facts. Such a simple valuation-revision and lattice-replacement mechanism may serve to illustrate the basic dynamics of theory change in scientific explanation.

(From now on, we revert to ordinary —i.e. additive— valuations that we shall be using in the rest of the report.)

12. PROOF THEORY

Let us begin by showing 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 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. We propose instead to provide every premise A with a value $[\![A]\!]$ in [0,1] —computed in an unspecified way (statistically, by opinion survey, or whatever)— with the unique requirement that a zero value means a false premise, 1 means a true —and therefore assertable— premise, and $1 - \varepsilon$ ($\varepsilon > 0$) means that we can assert it but with some apprehension or risk ϵ. Obviously, the value of $[\![A_i \rightarrow A_{i-1}]\!]$ 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 —or the risk we are willing to assume— to assert it) comes down to an exceedingly low number. According to our proof theory (developed at some length 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 $A_1 \rightarrow A_0$).

The underlying explanation is this. As always, all elementary propositions used in the above argument generate a Boolean lattice $\mathcal{L}$, and have as automatic counterpart a powerset structure on a set Θ of 2^n eventualities. We can picture these as responses given by 2^n independendent reasoners to the question "Do you think that A is the case?" (where A is one of the 2^{2^n} possible sentences that are expressible in $\mathcal{L}$). We can reasonably assume that at each decrement of i (= number of hairs left) more and more respondents will stop answering yes to the $A_i \rightarrow A_{i-1}$ question, so monotonically shrinking to leave a hard-core set of respondents who still say yes to $A_1 \rightarrow A_0$. It is only this hard set of persisting possibilities (respondents) θ that will validate the whole sorites (actually they are the only ones able to complete it, as they have all its premises asserted), and so it is just natural to get as "truth value" of the conclusion the proportion (or, in general, the weighted mean or measure) of people who have been able to reach the conclusion all the way through.

We now elaborate on the general proof-theoretical aspects of **BML**. First we need to have a formula for the value of the connectives (especially conjuction [of premises]), no matter how large the number of operands (conjuncts) in them. We take the semi-formal (16) and cast it in a slightly more precise form:

For every positive integer n and every collection $A_1, \ldots, A_n$ of propositions of $\mathcal{L}$,

$$[\![\bigvee_{i=1}^{i=n} A_i]\!] = \sum_I (-1)^{|I|+1}[\![\bigwedge_{i\in I} A_i]\!] \tag{206}$$

(where I ranges through all non-empty subsets of $\{1,\ldots,n\}$).

We abbreviate the first $n-1$ terms of (206) as S:

For every positive integer n and every collection $A_1, \ldots, A_n$ of propositions of $\mathcal{L}$,

$$S = \sum_I (-1)^{|I|+1}[\![\bigwedge_{i\in I} A_i]\!] \tag{207}$$

(where I ranges through all non-empty subsets of $\{1,\ldots,n-1\}$).

(Naturally, $S = [\![\bigvee_{i=1}^{i=n} A_i]\!] - [\![\bigwedge_{i=1}^{i=n} A_i]\!](-1)^n$). We then get:

$$\max(0, S-1) \leq [\![\bigwedge_{i=1}^{i=n} A_i]\!] \leq \min_i([\![A_i]\!]) \tag{208}$$

$$\max_i([\![A_i]\!]) \leq [\![\bigvee_{i=1}^{i=n} A_i]\!] \leq \min(1, S) \tag{209}$$

The lower value for $\wedge$ and the upper value for $\vee$ can be replaced by these coarser but often more useful bounds:

$$\max_i(0, \sum_i[\![A_i]\!] - \sum_{i<j}[\![A_i \wedge A_j]\!] - 1) \leq [\![\bigwedge_{i=1}^{i=n} A_i]\!] \leq \min_i([\![A_i]\!]) \tag{210}$$

$$\max_i([\![A_i]\!]) \leq [\![\bigvee_{i=1}^{i=n} A_i]\!] \leq \min(1, \sum_i[\![A_i]\!]). \tag{211}$$

We immediately collect that $[\![\bigwedge_{i=1}^{i=n} A_i]\!]$ reaches its upper bound if there is a set representation $\mathbf{A}_k$ (corresponding to a proposition A_k) that is a subset of all other set representations. Analogously, $[\![\bigvee_{i=1}^{i=n} A_i]\!]$ reaches its lower bound if there is a set representation $\mathbf{A}_k$ that is a superset of all other set representations. Those are sufficient conditions for [the values of] conjunction and disjunction to coincide with the minimum and maximum, respectively. (And notice both conditions are guaranteed by having $\mathcal{P}(\Theta)$ linearly ordered by set inclusion, though there is no need at all to have so strict a requirement.) In the general case, $\wedge$ and $\vee$ will have a value between indicated bounds. If all set representations are pairwise mutually disjoint then the conjunction will have $\max(0, \sum[\![A_i]\!] - 1)$ as its lower bound. And if all the A_is were simply atoms then $[\![\wedge A_i]\!] = 0$ and $[\![\vee A_i]\!] = \sum[\![A_i]\!]$.

We can now proceed to our proof theory (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:

$$\Gamma \vdash B \tag{212}$$

where B is the conclusion and Γ stands for a list of premises. Given an infinite $\mathcal{L}$, the list could be infinite too, but it would always be reducible by the compactness property to a finite list $A_1, \ldots, A_n$. Ambiguously, Γ will also —and most often— stand for the conjunction $\bigwedge A_i$ of the premises $A_1, \ldots, A_n$. We have, elementarily:

$$\Gamma \vdash B \quad \Leftrightarrow \quad \vdash \Gamma \dot{\rightarrow} B \tag{32}$$

$$\Gamma \vdash B \quad \Leftrightarrow \quad \Gamma \models B \tag{65}$$

$$\Gamma \models B \quad \Leftrightarrow \quad \models \Gamma \rightarrow B \quad \Leftrightarrow \quad (\forall v)[\![\Gamma \rightarrow B]\!] = 1 \quad \Rightarrow \quad [\![\Gamma \rightarrow B]\!] = 1 \quad \Rightarrow \quad [\![\Gamma]\!] \leq [\![B]\!].$$

Summing it all up we have, for any arbitrary argument:

$$\Gamma \vdash B \quad \Rightarrow \quad [\![\Gamma]\!] \leq [\![B]\!]. \tag{213}$$

We henceforth assume that we have a *valid argument* (so $\Gamma \vdash B$ will always hold), and distinguish four possible cases:

1) $[\![\Gamma]\!] = 0$ (i.e. the premises are [materially] *inconsistent*). Here by (213) $[\![B]\!]$ can be anywhere between 0 and 1; this value is in principle undetermined, and uncontrollably so (though a limiting condition —an upper bound— will sometime appear in the formulas). This is a case no logician would be interested in, since if one has a formally valid argument but one is in no way risking to assert any of its premises, it is only natural that the value of the conclusion turns out to be anything. (Yet there are cases —when contradictions are involved— in which logicians can and do get interested, see the QS rule below).

2) $[\![B]\!] = 0$. This entails, by (213), $[\![\Gamma]\!] = 0$ 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 (213), $[\![B]\!] > 0$. 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 (213), $[\![B]\!] = 1$. 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. We must first find out the conditions for cases 1 and 2 —so as to exclude them— that we label ***unsound*** arguments (since in those cases having a formally valid argument $\Gamma \vdash B$ does not preclude getting an irrelevant ($[\![B]\!] = 0$) or —worse— uncontrollably-valued conclusion B). As those cases are the ones to avoid and both are characterized by condition $[\![\Gamma]\!] = 0$, we have, summing up the two:

[*Definition 1:*] *Unsoundness* of a valid argument $\Gamma \vdash B$ is having either $[\![B]\!] = 0$ (no matter the value of the premises) or else $[\![B]\!] > 0$ with all premises making up a null conjunction (i.e. $[\![\Gamma]\!] = 0$).

[*Definition 2:*] ***Soundness*** of a valid argument $\Gamma \vdash B$ is having $[\![B]\!] > 0$ whenever the value of the premises is non-zero (and $[\![\Gamma]\!] > 0$).

From this, the next theorem —we could term ***Soundness condition***— immediately follows:

[*Theorem:*] If $[\![\Gamma]\!] \neq 0$ then the argument $\Gamma \vdash B$ is ***sound*** (thus implying $[\![B]\!] \neq 0$). (214)

[*Remark 1*: Reconsidering definition 2, we could venture a more stringent definition of soundness along the following lines: the value of the conclusion should not only be non-zero but also not less than the value of the least asserted premise, so that $[\![B]\!] \geq \min[\![A_i]\!]$ (where $A_i \in \Gamma$). But this condition —apparently reasonable enough— would immediately entail (if only we kept the very natural $\wedge$-introduction rule, see below) that the value of the conjunction connective is rigidly the ***minimum*** function (i.e. $[\![\wedge A_i]\!] = \min([\![A_i]\!])$ in all cases). So, incidentally, we would be back in the inavoidably truth-functional Łukasiewicz-Tarski $Ł_\infty$ logic. This, as we have previously seen, seems neither acceptable nor realistic (it presupposes a linearly ordered negation-free universe, as we found and stated in our $Ł_\infty$ criticism, see section 6 —concluding paragraphs— and 10(*b*)).]

[*Remark 2*: A new and also strict version of definition 2 —more reasonable and less destructive than the one suggested in the previous remark— would be this one: ***Soundness*** of a valid argument $\Gamma \vdash B$ is having $[\![B]\!] \geq 1/2$ (instead of merely > 0) whenever the value of each premise is $\geq 1/2$, as we assume it to be. (We will use this criterion somewhere below.)]

a. Modus Ponens

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

$$\begin{array}{ll} A & m \\ A \to B & n \\ \hline B & p \end{array} \tag{215}$$

where m, n and p stand for the 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 used 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 Mycin or Prospector. 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:

$$\begin{array}{l} A(m) \\ A \to B(\sigma) \\ \hline B(p) \end{array} \tag{216}$$

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 the conditional probability of B given A. (It is what we called —in section 7— "degree of sufficiency" σ of A —or of necessity of B— and 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 none other than (13) followed by (95).

The problem is that what we have, from our logical, probability-rid standpoint, is (215), not (216), and in (215) n is not $[\![B|A]\!]$ but $[\![A \to B]\!]$. Remember $[\![B|A]\!]$ and $[\![A \to B]\!]$ not only do *not* coincide (as we know from (104) already) but mean different things. $[\![A \rightarrow B]\!]$ is the value ("truth" we may call it, or "truth minus risk") we assign to the [logical] assertion $A \rightarrow B$. 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 (215) rule; note that $m+n \geq 1$ (this *always* holds), and that $[\![B|A]\!]$ can be obtained from $[\![A \rightarrow B]\!]$ through (103) or —more usefully (because $[\![B|A]\!]$ is directly obtainable from experts, see section 7)— $[\![A \rightarrow B]\!]$ from $[\![B|A]\!]$ through

$$[\![A \rightarrow B]\!] \;=\; 1 - [\![A]\!] \cdot (1 - [\![B|A]\!]). \tag{217}$$

As an application of all considerations above we now have the following two easy propositions (where, as can be noted, the soundness condition (214) translates into four equivalent conditions):

[*Theorem 1*:] The Modus Ponens rule

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

is *unsound* (including $[\![B]\!] = 0$) if one of these four equivalent conditions hold:

1) $m+n=1$ (218)
2) $[\![B|A]\!] = 0$
3) $[\![A \wedge B]\!] = 0$
4) A and B are incompatible ($\alpha_{AB} = 0$) and $[\![A]\!] + [\![B]\!] \leq 1$.

Moreover, $[\![B]\!] \leq n = 1 - m$ (so $[\![B]\!]$ is either zero or unpredictably somewhere between 0 and n).

Conversely, the following theorem states the *soundness* condition for the MP rule.

[*Theorem 2*:] The Modus Ponens rule

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

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

1) $m+n > 1$ (219)
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 the sound MP case (the only interesting one in practice), we have —by applying formulas (127) and (131)— the following bounds for the value $[\![B]\!]$ of the MP conclusion:

$$[\![A]\!] + [\![A \rightarrow B]\!] - 1 \;\leq\; [\![B]\!] \;\leq\; [\![A \rightarrow B]\!] \tag{220}$$

(or equivalently, in shorter notation: $m+n-1 \leq p \leq n$); 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 know neither $[\![B]\!]$ nor α_{AB} beforehand usually, 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. Admittedly, this is not a tremendous result; in particular, it is not very helpful in pinpointing $[\![B]\!]$ except when either $m = 1$ (then $[\![B]\!] = n = [\![A \wedge B]\!]$) or $n = 1$ (then $[\![B]\!]$ is undetermined, and merely $\geq m$). So we must take a closer look at Modus Ponens as an *argument* and at the conditional statement $A \rightarrow B$ as [the expressive counterpart of] a *relation* (see section 7). It may seem reasonable to assume that the following two provisos apply. *First*, soundness apparently demands not only that $[\![B]\!] \neq 0$ but $[\![B]\!] \geq 1/2$ provided the value of the premises A and $A \rightarrow B$ are also —as we take for granted— above 1/2 (this, that we anticipated in remark 2, makes sure at least that $[\![A]\!] + [\![B]\!] \geq 1$). *Second*, the upper subinterval of $[m+n-1, n]$ comprised by values between $[\![B|A]\!]$ and $n = [\![A \rightarrow B]\!]$ is —quite reasonably— never reached, because it corresponds to the case of actually *anti*correlated A and B. So, we have now a narrower interval to locate the unknown $[\![B]\!]$:

$$[\![A]\!] + [\![A \rightarrow B]\!] - 1 \;\leq\; [\![B]\!] \;\leq\; [\![B|A]\!] \tag{221}$$

(or, equivalently,

$$[\![A \wedge B]\!] \leq [\![B]\!] \leq [\![A \wedge B]\!]/[\![A]\!]$$

or, still, in shorter notation,

$$m+n-1 \leq p \leq \frac{m+n-1}{m}),$$

and here the upper bound is reached when A and B are *independent* ($\alpha_{AB} = \alpha^0_{AB}$).

Now, this result (221) still does not quite help, because we don't usually know much about the relationship —and compatibility— between A and B. (Instead, what we can always be sure of is that $\alpha(A, A \rightarrow B) = 0$ no matter which are A and B, but this is hardly helpful anyway.) We don't even know whether $[\![A]\!] \leq [\![B]\!]$ or $[\![A]\!] \geq [\![B]\!]$. (If the latter were assuredly the case then we could at least write

$$[\![B]\!] \;\leq\; \min([\![A]\!], [\![A \rightarrow B]\!]) \quad \text{or, better,} \quad [\![B]\!] \;\leq\; \min([\![A]\!], [\![B|A]\!]) \tag{222}$$

as the upper bound in (220) and (221).) What we *can* say is that if the first unequality holds ($[\![A]\!] \leq [\![B]\!]$) then (remember provisos 1 and 2 apply):

$$[\![B]\!] = \frac{[\![A \rightarrow B]\!] - \alpha_{AB}}{1 - \alpha_{AB}} \qquad \text{and} \qquad 1 - [\![A]\!] \;\leq\; \alpha_{AB} \;\leq\; 2 \cdot [\![A \rightarrow B]\!] - 1 \tag{223}$$

(but for $\alpha_{AB} = 1$ $[\![B]\!]$ is undefined, merely $\geq m$). And if it is the second ($[\![A]\!] \geq [\![B]\!]$), then:

$$[\![B]\!] = [\![A \rightarrow B]\!] - \alpha_{AB}(1 - [\![A]\!]) \qquad \text{and} \qquad 1 - [\![B]\!] \;\leq\; \alpha_{AB} \;\leq\; \frac{2 \cdot [\![A \rightarrow B]\!] - 1}{2 \cdot (1 - [\![A]\!])} \tag{224}$$

(In both (223) and (224) the lower bounds for α_{AB} are reached for independent A and B.) If we further assumed, as we hinted in section 7, that asserting $A \rightarrow B$ means somehow $[\![A \rightarrow B]\!] \geq [\![B \rightarrow A]\!]$, and so $[\![B]\!] \geq [\![A]\!]$ (and $\sigma_A \geq \nu_A$, see sect. 7) then (223) applies, but not (224), nor (222), and we have merely $[\![A]\!] \leq [\![B]\!] \leq [\![B|A]\!]$ as closest bounds. Furthermore, if $A \rightarrow B$, as an *assertion*, is taken to mean rather $[\![A \rightarrow B]\!] \geq [\![A \rightarrow \neg B]\!]$, an interpretation we suggested there (for which there is some evidence in experimental psychology), then —and only then— $[\![B|A]\!] \geq 1/2$.

In (221), if A and B are fully or strongly compatible, $[\![B]\!]$ will be nearer the lower bound; if they are independent, $[\![B]\!]$ will have the highest value. (While this may seem a paradox, it is not: given the [fixed] values n and m of $[\![A \rightarrow B]\!]$ and $[\![A]\!]$, it is considerably easier for a low-valued $[\![B]\!]$ to yield the given $[\![A \rightarrow B]\!]$ if A and B are compatible; conversely, if they are not fully compatible, or even independent, it will take a high value $[\![B]\!]$ to match the given $[\![A \rightarrow B]\!]$. And we could get a still higher value, but only by demanding that A and B are anticorrelated, a rather absurd proposition.)

Thus, we can only increase our $[\![B]\!]$ if we are assured that A and B are independent (in the sense of (93)): we then obtain a higher value $[\![B]\!] = [\![B|A]\!]$ (but we may consider this one as a rather unwanted side case). Or, the more we confide in a strong logical relation between A and B, the more we should lean towards the low value given by

$$[\![B]\!] = [\![A \wedge B]\!] = m + n - 1. \tag{225}$$

In absence of the relevant information, it seems we should reasonably stick to the $[\![A \wedge B]\!]$ value ($= m + n - 1$) as our safest bet. This value is exactly the measure (or weighted mean) of the possibilities (logical interpretations, polled individuals, etc) θ making up A that also make up B. To understand what that may mean, assume the θs are logical interpretations, in the standard sense; then $[\![B]\!] = [\![A \wedge B]\!]$ is just the "truth" of the argument, i.e. the proportion of interpretations in which the argument (the MP) has been effectively performed and yielded *true* as value. Or assume the θs are independent elementary reasoners, each having full reasoning capabilities and completing its own line of argument in view of the premises it has: $[\![B]\!]$ then equals $[\![A \wedge B]\!]$ and so $\mu(\mathbf{A} \cap \mathbf{B})$. $\mathbf{A} \cap \mathbf{B}$ are precisely the θs in B that have both A and $A \rightarrow B$ as premises (this can be easily visualized on tables 6-9 of section 4), so that they —and only they— have been able actually to complete the Modus Ponens. (Equivalently, if we executed a stochastic process tuning the frequency of each θ to its $\mu(\theta)$ value and performing the MP reasoning each time it were possible, the proportion of cases in which the conclusion B would be reached in the long run would just equal $\mu(B)$, i.e. $[\![B]\!]$.)

(As regards tables 6-9, note that a MP relating the A and B in table 6 would be clearly *unsound*, that nobody would reasonably assert "$A \rightarrow B$" with the anticorrelated A and B of table 7, and that only the μ_2 column in table 8 fulfills the first proviso.)

Now imagine we want not merely a pair of bounds for the conclusion B of an MP but the *exact* value $[\![B]\!]$. Two obvious candidate formulas for this follow immediately from (22) and (24):

$$[\![B]\!] = [\![A \vee B]\!] + [\![A \rightarrow B]\!] - 1 \tag{226}$$

$$[\![B]\!] = [\![A]\!] + [\![A \rightarrow B]\!] - [\![B \rightarrow A]\!] \tag{227}$$

To get something useful out of it, let us suppose we are given not only $\sigma_A = [\![B|A]\!]$ but also $\nu_A = [\![A|B]\!]$ that we shorten to σ and ν and assume estimated by experts (see section 7). We then formulate MP as

$$\begin{array}{l} A(m) \\ A \rightarrow B(\sigma, \nu) \\ \hline B(p) \end{array} \tag{228}$$

which is exactly (216) except that the conditional has prompted evaluation of relative truths of A and B in both directions. The value is computable at once from (95):

$$[\![B]\!] = \frac{[\![B|A]\!] \cdot [\![A]\!]}{[\![A|B]\!]} \qquad \text{or} \qquad p = \frac{\sigma \cdot m}{\nu} \tag{229}$$

(Note this value is the one *approximate reasoning* systems (e.g. Prospector) unqualifiedly assign to $[\![B]\!]$ on purely probabilistic grounds —and falsely assuming, as we saw, that $[\![B|A]\!]$ is the same as $[\![A \rightarrow B]\!]$.) If we wanted the MP presented in the more traditional way (215), first we would directly estimate the truth value $[\![A \rightarrow B]\!]$ of the conditional, or compute it from σ through (217) —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:

$$\begin{array}{ll} A & m \\ A \rightarrow B & n\ (\nu) \\ \hline B & p \end{array} \tag{230}$$

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

$$[\![B]\!] = \frac{[\![A \wedge B]\!]}{[\![A|B]\!]} = \frac{m + n - 1}{\nu} \tag{231}$$

that naturally fits the (221) bounds (when ν runs along from 1 to $[\![A]\!]$). Or else we can use (227) directly, if we previously estimate $[\![B \rightarrow A]\!]$, or compute it from ν.

What is interesting here is that no previous condition is needed. If one wants proviso 2 (having a positive correlation between A and B), this fact, or its absence, can be detected at once just by checking whether $\nu \geq m$ (and then $[\![B]\!] \leq \sigma$). If we require that the first proviso holds (or that simply $[\![A]\!]+[\![B]\!] \geq 1$, a weaker condition) then the following formula gives α_{AB}, once the $[\![B]\!]$ value is computed by either (231) or (227):

$$\alpha_{AB} = \frac{[\![A \to B]\!] - [\![B]\!]}{1 - \max([\![A]\!], [\![B]\!])} \tag{232}$$

We illustrate this with three elementary examples:

1) $\mathbf{A} \subset \mathbf{B}$: We have $\sigma = 1$ and $n = 1$. Moreover: $\nu = [\![A]\!]/[\![B]\!]$, $[\![B]\!] = m/\nu$, and $\alpha_{AB} = 1$.

2) $\mathbf{A} \cap \mathbf{B}$ is —in value— half of $\mathbf{A}$ and also half of $\mathbf{B}$: We have $\sigma = \nu = .5$ and $n = 1 - m/2$. Then $[\![B]\!] = m$, and $\alpha = (1 - 3m/2)/(1 - m)$. If A and B were independent then $[\![B]\!] = m = .5$ and $\alpha = \alpha^0 = .5$.

3) A and B are independent: This fact clearly shows, because $m = \nu$. We have $[\![B]\!] = \sigma = 1 - (1 - n)/m = (m + n - 1)/m$, and $\alpha = \alpha^0 = 1 - \min(\sigma, \nu)$.

b. $\wedge$-Introduction, and the Quodlibet Sequitur (QS) rule

The $\wedge$- Introduction rule is:

$$\begin{array}{ll} A & m \\ B & n \\ \hline A \wedge B & p \end{array} \tag{233}$$

(m and n are here not necessarily non-zero)

where, incidentally, it is clearly seen that the above-mentioned criterion that the conclusion should never be less than the value of the least asserted premise (see remark 1 under 'soundness') would immediately entail that the value of the $\wedge$ connective is necessarily the *minimum* function, and so we would be back in $Ł_\infty$. (Indeed, since $[\![A \wedge B]\!] \leq \min([\![A]\!], [\![B]\!])$ —by (13)— and $[\![A \wedge B]\!] \geq \min([\![A]\!], [\![B]\!])$ —by the given criterion—, so $[\![A \wedge B]\!] = \min([\![A]\!], [\![B]\!])$.)

The *unsoundness* condition for this rule gives:

$$[\![\Gamma]\!] = 0 \text{ if and only if } [\![\text{ Conclusion }]\!] = 0,$$

which is exceptional and characteristic of this type of arguments (for the rest only the *if* part applies). The trivial case here is $m = 0$ or $n = 0$: the rule is then (trivially) unsound. For the general case, the $\wedge$-introduction rule is *unsound* if and only if $[\![A \wedge B]\!] = 0$, which amounts to $[\![A]\!] + [\![B]\!] \leq 1$ *and* $\alpha_{AB} = 0$.

A particularly interesting special case of the argument is where B is $\neg A$. Then the argument can be stated in this way:

$$\begin{array}{ll} A & m \\ \neg A & 1 - m \\ \hline \bot & 0 \end{array} \tag{234}$$

(m is not necessarily non-zero)

which is a *valid* but —to us— *unsound* argument. We concede this fact is not relevant for all practical purposes, because the conclusion is warranted by positing the argument in full, Gentzen-style:

$$\begin{array}{l} \Gamma \vdash A \\ \Gamma \vdash \neg A \\ \hline \Gamma \vdash A \wedge \neg A \end{array} \tag{235}$$

and this makes sense, for $[\![\Gamma \to A \wedge \neg A]\!] = [\![(\Gamma \to A) \wedge (\Gamma \to \neg A)]\!]$ (for any valuation), and the conjunction inside the brackets has *one* as value —since $[\![\Gamma \to A]\!] = 1$ (this is implied by the first line) and also $[\![\Gamma \to \neg A]\!] = 1$ (as implied by the second)—.

Note that this oddity (having an *unsound* yet *valid* argument) is also present in ordinary, two-valued logic. An argument such as $A, \neg A \vdash \bot$ would hardly be validated *just* by model-theoretic considerations, because A and $\neg A$ are *never* assertable simultaneously, by definition.

A related type of argument is the QS rule (the medieval *quodlibet sequitur* or, equivalently, the weak intuitionistic $\neg$-elimination), a very important element in Logic, since it allows detection of contradictions, and subsequent action after that. We have:

$$\begin{array}{ll} A & m \\ \neg A & 1-m \\ \hline B & p \end{array} \qquad (236) \qquad (m \text{ is not necessarily non-zero here})$$

The unsoundness condition, when applied to this argument, gives $[\![\Gamma]\!] = 0$, and so $[\![B]\!]$ is either zero or has an (uncontrollably) arbitrary value. Hence, the rule is again *unsound* to us, but we accept it on the same grounds we accepted the previous argument, and we note also that this rarity is already present in —and ignored by— two-valued logic. Nevertheless, it is perhaps appropriate to remind oneself that $[\![B]\!]$ is indeed demonstrably *arbitrary*. This, which can be blamed on the underlying radical *unsoundness* of the argument, allows $[\![B]\!]$ to be 1 —and this is what ordinary logic chooses it to be— or zero (and thus the argument is clearly unconclusive). The latter solution is, we recall, the one allowed by relevance logics —in certain (controlled) circumstances— or by systems tolerating somehow *local* contradictions.

As an aside, we turn back again to the criterion, suggested in remark 1 above and subsequently discussed, that the conclusion should never be less in value than the least asserted premise. This is here clearly seen as counterproductive, for it would make $[\![B]\!]$ *necessarily* zero whenever the Logic involved is the *classical* two-valued version (since, then, $m \in \{0,1\}$ and thus $p = 0$), contrary to assumptions made there. This is, perhaps, a further reason against the (traditional) many-valued L_∞ logic. We have already contended L_∞ is unrealistic (by forcing the universe to be coherent and negation-free) and unable to recognize contradictions (by sheer lack of "resolving power", so to speak, as we noted when we saw that a contradiction $A \wedge \neg A$ is never valued 0 if A is not binary-valued); now we charge, further, its proof-theory does not allow the reasoner to detect contradictions and act accordingly (resolve them and proceed ahead).

c. Resolution

By *resolution* we understand Quine's "consensus" rule or the disjunctive version of Gentzen's *cut:*

$$\begin{array}{ll} A \vee L & m \\ B \vee \neg L & n \\ \hline A \vee B & p \end{array} \qquad (237)$$

where L is a literal (a letter or a negation of a letter). We have —always— $m + n \geq 1$. Moreover, the soundness condition here implies $m + n > 1$, and $[\![\Gamma]\!] \leq [\![\text{Concl.}]\!]$ is now

$$[\![A \vee B]\!] \;\geq\; [\![A \vee L]\!] + [\![B \vee \neg L]\!] - 1 \qquad (238)$$

(or, equivalently, $p \geq m + n - 1$). This is our best approximation. However, if we want to pinpoint the p value precisely we can do better than that, for we have alternative candidate formulas for the exact value of the conclusion (though not necessarily more useful than (238)), e.g.:

$$[\![A \vee B]\!] = [\![A \vee B \vee L]\!] + [\![A \vee B \vee \neg L]\!] - 1 \quad \text{, or}$$
$$[\![A \vee B]\!] = [\![A \vee L]\!] + [\![B \vee \neg L]\!] + [\![A \vee \neg L]\!] + [\![B \vee L]\!] - [\![A \wedge B]\!] - 2$$
$$[\![A \vee B]\!] = [\![A \vee L]\!] + [\![B \vee \neg L]\!] + [\![A \wedge L]\!] + [\![B \wedge \neg L]\!] - [\![A \wedge B]\!] - 1$$

As is notorious, the Modus Ponens rule can be formulated —as indeed certain other natural-deduction rules can (Modus Tollens, the disjunctive or hypothetical syllogisms, or *reductio ad absurdum*)— as a special case of resolution. We then find again, just by applying the (238) bound, the lower bound in formula (220). (Or we could as well rediscover the exact (227) by simple algebraic manipulation of any of the three formulas above.)

SUMMARY, DISCLAIMER AND REFERENCES

Summary: Many-valued (MV) logic, whatever its philosophical justification or its practical uses, has always been conceptually modeled as a *non-classical* logic. It has thus been, by tradition, inevitably characterized as "deviant". First, because some of the theorems that currently obtain (in MV logic) are not the expected familiar ones —or, rather, because some familiar results fail to show as MV theorems. Also, because the MV approach apparently forces sentences into an algebraic structure that is considerably *weaker* than the usual Boolean algebra that so satisfactorily meets the demands of ordinary propositional logic. This applies as much to the finite-valued formalisms (Post, Kleene, Bochvar, etc.) as to the infinite-valued ones (notably the Łukasiewicz-Tarski $Ł_\infty$ logic). This report (which, as stated in the Abstract, is a barely-updated summary of a Ph.D. Thesis written in 1981-82) has set out to explore some consequences of taking the opposite assumption: that a *many-valued logic* can be consistently conceived and developed that is *classical*, i.e. structurally no different —in both proof-theoretic and algebraic senses— from plain ordinary Boolean logic, from which it inherits the usual convenient properties, with only a moderately small price to pay (non-truth-functionality, which is incidentally the same price that Probability has always had to pay). So, the author hopes to have shown, *Many-valued* and *Boolean* need not be contradictory concepts, as *Boolean* does not necessarily mean *two-valued*. And both can be integrated into ordinary Propositional Logic, whose structures can then be more explicitly revealed.

Disclaimer: Shortly after writing this thesis, back in 1982, I found that Hans Reichenbach and Zygmunt Zawirski had already got the basic idea —which is otherwise simple and obvious enough— in the nineteen-thirties. The idea is, roughly, that imposing [0,1]-valuations on a Boolean algebra should be a natural and sensible thing to do (provided one were ready to renounce the habit of thinking of truth valuations as epimorphisms). These authors, though, were not strictly in the Logic business, and had a different purpose in mind: the grounding of *Probability* on [non-truth-functional] *Logic* (an effort, incidentally, they later discontinued). Mine was, instead, just showing that MV logic could be characterized as *classical* —albeit non-functional— *Logic*. Thus, no matter the goal was different, I considered that historically the point had been made (by them) and so I did not bother to translate the dissertation (written in Catalan) into publishable English papers. Moreover, I sensed that the underlying purpose and results, self-evident as they seemed to me, were to be soon (re)discovered and developed by some other more authoritative logician. I waited ten years for that to happen. To my knowledge, it did not. Rather puzzled, I have now finally made up my mind to publish in English what I had long considered trivial, merely local and anticipatory, and all but redundant, written only for self-illumination. In so doing, I have not dared but summarize and emphasize the main points and, for the rest, leave the underlying work and supporting argumentation and references unmodified, for, in my view, updating would have changed the basic flavor and structure of the material. The ten-year delay has not been completely useless, though, because it has improved my own understanding of the ideas I had so clumsily exposed a decade earlier. I only wish this reluctant working piece of mine will at last prompt the authoritative work I had been expecting all these years from a totally qualified author.

References

[1] Bellman, R. & Gierz, M. (1973), On the analytical formalism of the theory of fuzzy sets, *Info. Sci.*, **5**, 149-156.

[2] Bellman, R. & Zadeh, L. (1976), Local and fuzzy logics, in: J. M. Dunn & G. Epstein, eds., *Modern Uses of Multiple-Valued Logic*, Reidel

[3] Birkhoff, G. (1967) *Lattice Theory*, Am. Math. Soc.

[4] Black, M. (1937), Vagueness, *Phil. Sci.*, **4**, 427-455

[5] Bochvar, D. (1939), On a three-valued logical calculus and its application to the analysis of contradictories, *Matématičéskij Sbornik*, **4**

[6] Bodiou, G. (1964), *Théorie dialectique des probabilités*, Gauthier-Villars

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

[8] Carnap, R. (1947), *Meaning and Necessity*, U. Chicago Press

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

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

[11] Chellas, B. F. (1980) *Modal Logic: an Introduction*, Cambridge

[12] Cresswell, M. J. (1973), *Logics and Languages*, Methuen

[13] de Finetti, B. (1937), La prévision, se lois logiques, ses sources subjectives, *Annales de l'Institut Henri Poincaré*, **7**

[14] de Luca, A. & Termini, S. (1972), A definition of a non-probabilistic entropy in the setting of fuzzy sets theory, *Info. and Control*, **20**, 301

[15] Fine, K. (1975), Vagueness, truth and logic, *Synthese*, **30**, 265-300

[16] Gaines, B. R. (1978), Fuzzy and Probability Uncertainty Logics, *Info. and Control*, **38**, 154-169

[17] Gödel, K. (1933), Eine Interpretation des Intuitionistischen Aussagenkalküls, Engl. transl. in: J. Hintikka, ed., *The Philosophy of Mathematics*, Oxford (1969)

[18] Haack, S. (1974), *Deviant Logic*, Cambridge

[19] Halmos, P. (1974), *Lectures on Boolean Algebras*, Springer

[20] Hamacher, H. (1976), On Logical Connectives of Fuzzy Statements and their affiliated Truth Function, *Proc. 3rd Eur. Meet. Cyber. and Systems Res.*, Vienna

[21] Hanf, W. (1974) Primitive Boolean Algebras, *Proc. of the A. Tarski Symp.*, Am. Math. Soc.

[22] Heyting, A. (1956), *Intuitionism. An Introduction*, North-Holland

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

[24] Kleene, S. C. (1938), On notation for ordinal numbers, *J. Symb. Logic*, **3**, 150-155

[25] Kripke, S. (1965), Semantical analysis of Intuitionistic Logic I, in: J. N. Crossley & M. Dummett, eds., *Formal Systems and Recursive Functions*, North-Holland

[26] Łukasiewicz, J. (1920), O logice trójwartościowej, *Ruch Filozoficzny* (Lwów), **5**, 169-171

[27] Łukasiewicz, J. & Tarski, A. (1930), Untersuchungen über den Aussagenkalkül, *Comptes Rendus Soc. Sci. & Lettres Varsovie*

[28] MacColl, H. (1906), *Symbolic Logic and its Applications*, London

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

[30] Mostowski, A. & Tarski, A. (1949), Arithmetical classes & types of well-ordered systems, *Bull. Am. Math. Soc.*, **55**, 65

[31] Peirce, C. S. (1902), *Minute Logic*, unpublished (see ref. in [35]) (see also C.S. Peirce, *Collected Papers*, Harvard (1933))

[32] Popper, K. R. (1972), Two faces of commonsense, in: K. R. Popper, *Objective Knowledge*, Oxford

[33] Ramsey, F. P. (1931), Truth and Probability, *The Foundations of Mathematics*, London

[34] Reichenbach, H. (1935), *Wahrscheinlichkeit* (Engl. transl. Calif. U. Press 1949)

[35] Rescher, N. (1969), *Many-Valued Logic*, McGraw-Hill

[36] Rodder, W. (1975), On 'And' and 'Or' Connectives in Fuzzy Set Theory, T. Report, *I. für Wirtschaftswiss.*, Aachen

[37] Russell, B. (1923), Vagueness, *Australasian J. Phil. & Psych.*, **1**, 84-92

[38] Sales, T. (1982), *Contribució a l'anàlisi lògica de la imprecisió*, Univ. Polit. Catalunya (Ph. D. dissertation)

[39] Scott, D. (1973), Does many-valued logic make sense?, *Br. Soc. for the Phil. of Sci.*

[40] Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton

[41] Trillas, E., Alsina, C. & Valverde, Ll. (1982), Do we need max, min and 1-j in Fuzzy Set Theory?, in: R. Yager, ed., *Recent Developments of Fuzzy Set and Possibility Theory*, Pergamon

[42] van Fraassen, B. C. (1968), Presuppositions, implication and self-reference, *J. Phil.*, **65**

[43] Vickers, J. M. (1970), Probability and Non-Standard Logics, in: K. Lambert, ed., *Philosophical Problems in Logic*, Reidel

[44] Watanabe, S. (1969), Modified Concepts of Logic, Probability and Information based on generalized continuous characteristic function, *Info. and Control*, **15**, 1-21

[45] Wittgenstein, L. (1922), *Tractatus Logico-Philosophicus*, London

[46] Yager, R. (1979), A Note on Fuzziness in a Standard Uncertainty Logic, *IEEE Trans. Systems, Man & Cyb.*, **9.7**, 387-388

[47] Zimmermann, H. J. (1977), Results of Empirical Work on Fuzzy Sets, *Int. Conf. on Applied Gen. Systems Research*, Binghamton