Is Intelligent Belief Really Beyond Logic?

Paul Snow

P.O. Box 6134
Concord, NH 03303-6134 USA
paulsnow@delphi.com

Abstract

"Impossibility theorems" have recently appeared in the AI literature which have been interpreted as forbidding truth-functional uncertainty calculi. Such "logicist" calculi do in fact exist. For example, case-based reasoning principles entail a truth-functional probability-agreeing scheme whose strengths and weaknesses are not so different from those of the usual belief representation methods. No claim is made that belief modeling should be conducted exclusively along logicist lines, but a non-trivial common ground of intuition does exist beneath logicism and its alternatives.

Introduction

A prominent community in the world of artificial intelligence are the *logicists.* Briefly put, logicists hold that intelligent cognitive states, including belief states, can be faithfully emulated by some kind of logic. As used here, a logic is a formalism which attaches values to sentences so as to be both *truth-functional* and *equivalence preserving.*

"Truth-functional" means that the values which belong to any sentences s and t determine the values of $\neg s$, $\neg t$, $s \vee t$, $s \wedge t$, and of all other compounds, regardless of the identities of s and t. This contrasts, for instance, with some probability distributions in which $p(s \wedge t)$ cannot be determined from $p(s)$ and $p(t)$. "Equivalence preserving" means that if two sentences are equivalent in the usual Boolean sense, such as $\neg(s \wedge t)$ and $\neg s \vee \neg t$, then the two sentences enjoy identical values. All probability distributions are equivalence preserving, but fuzzy set membership grades generally are not, since $s \wedge \neg s$ may have a different value than that of $\varnothing$ (for further discussion and motivation of this aspect of fuzzy logic, see Dubois and Prade 1994).

A second AI community is united by the conviction that something different from a logic is needed to emulate belief. Members of this community differ among themselves about what that "something different" might be, but all espouse some formalism which is not necessarily truth-functional, or not necessarily equivalence preserving. We shall call the members of this community *credalists*, and any of the espoused formalisms a *credo.*

There is consensus among credalists that any belief formalism worthy of the name agrees ordinally with some Sugeno function (dissertation work cited in Prade 1985). The principal feature of a Sugeno function is that if sentence s implies sentence t, then t is no less credible than s. "Implies" is read in the usual Boolean fashion for the equivalence preserving calculi, and only slightly differently for fuzzy membership grades.

The Sugeno criterion is an abstraction of how implications behave in Boolean logic, and many credos include Boolean logic as a special case. Probability does, and so does the possibility calculus. Of course, Boolean logic makes no allowance for uncertainty apart from variable expressions.

It is safe to characterize the logicist-credalist relationship as chilly. Amity reached a local minimum in 1993 with the celebrated publication of a logicist philippic by Elkan (also 1994). Although principally urging against the fuzzy credalists, Elkan in passing denounced the entire Sugeno communion while demanding a calculus in which *red* $\wedge$ *watermelon* would be more credible than *red* alone when dubious melons are at issue, contrary to the counsel of the implication involved. Presumably, what Elkan really wanted was that *red* $\wedge$ *watermelon* be more credible than *red* $\wedge$ $\neg$*watermelon*, which all popular credos can readily provide. (Elkan 1994 uses a different melon story.)

The resulting flap moved Dubois and Prade (1994) to rebuttal. They portray Elkan as merely restating a special case of their own theorem about the impossibility of any calculus with more than two truth values which is both equivalence preserving and truth-functional over all the connectives for arbitrary (or even Sugeno-agreeing) assignments of values to sentences.

Although they affirmed the correctness of Elkan's result, Dubois and Prade disputed Elkan's interpretation of it. Their own interpretation included that "Elkan's trivialization result kills truth-functional uncertainty handling systems," meaning the equivalence preserving ones, since the authors had distinguished fuzzy sets from other credos earlier in the paper.

The purpose of the present paper is to question whether that killing really took place. The theorems are correct, but it is argued that their interpretation has been too sweeping and has failed to distinguish the extent to which different credos conflict with logicism. With some attention to the assignment of values to sentences, logicist uncertainty handling with multiple truth values is possible. At least one such calculus does not appear so different from other established credos. That calculus is motivated here in the axiomatic style prevalent in the credalist community. The class of multivalued logicist calculi is *not* subsumed by probability models nor by other popular credos.

Distinct Truth Values

As is routine, we assume that any domain to be discussed is finite and closed under the usual Boolean connectives, with at least four sentences, $\{ s \vee \neg s, s, \neg s, \varnothing \}$, at least two of which are uncertain. A sentence in the domain which is implied by no non-equivalent sentence besides $\varnothing$ shall be called an *atom*, and distinct atoms are mutually exclusive. The sentences in the domain are completely ordered and no sentence is ranked equally with any Boolean contradictions except for those contradictions themselves. Complete ordering and the absence of false sentences ease exposition, but can be relaxed.

If multivalued logicist credos exist, then every sentence must have its own truth value, distinct from the truth value of any non-equivalent sentence. To see this, suppose s and t are not equivalent, but share the same intermediate truth value strictly between that of tautology and that of contradiction. Without loss of generality, we take $s \wedge \neg t$ to be ranked strictly ahead of a contradiction (since s and t are not equivalent, at least one of $s \wedge \neg t$ and $t \wedge \neg s$ must be non-contradictory). Then the expression $(s \vee t) \wedge \neg t$ must have the value of $s \wedge \neg t$, while $(t \vee t) \wedge \neg t$ is a contradiction, and so has a distinct value from $s \wedge \neg t$, but the expressions would have the same value in a truth-functional regime.

That each non-equivalent sentence have its own distinct truth value is also a sufficient condition for the existence of a multivalued logicist calculus on a domain. The reason is obvious: the values would then be an *index* of the sentences, and so truth-functionality and equivalence preservation are assured by the bijection between the sentences and their indices.

One can also force a violation of the logicist criteria without introducing a third truth value. The example of the preceding paragraph also serves for a $\{0, 1\}$ formalism if, as is allowed in possibility theory, s and t share the value of unity, but neither $s \Rightarrow t$ nor $t \Rightarrow s$ obtains in any Boolean sense. This, along with the ties problem for a third truth value, means that possibility is non-logicist on *all* occasions *except* when it strictly emulates Boolean logic. The same is also true of fuzzy membership grades, of course, where the value assigned to a compound fuzzy sentence is necessarily equal to the value of at least one of its constituents.

In the case of fuzzy grades, it is equivalence preservation which is jettisoned. Dubois and Prade (1994) motivate this semantically with a notion of "partial truth" which they take as an intuitive primitive. Lukasiewicz, whose $\aleph_1$ formalism antedates fuzzy theory (Lukasiewicz and Tarski 1930), chose a syntactic response to the same problem, defining equivalence to be the possession of identical truth values within the calculus, and nothing else. Possibility measures retain equivalence preservation, but give up truth-functionality for AND and for negation.

The situation with probability distributions is different. Dubois and Prade's theorem is correct as far it goes, in that probability is not always truth-functional except for negation. But while *any* non-Boolean use of the usual fuzzy-based calculi necessarily violates logicism, that is not so for probability distributions. Some probability distributions violate, others do not.

Existence of Logicist Credos

All finite Boolean algebras of sentences can be ordered so as to admit a Sugeno-agreeing logicist formalism. One construction which furnishes a way to do this uses *atomic bound systems* of algebraic constraints.

> **Definition.** A (strict and complete) *atomic bound system* comprises simultaneous linear constraints upon probabilities over a finite plural set of sentences whose atoms are transitively, completely, and strictly ordered according to any criterion. The system requires that for each atom *a:*
>
> $$p(a) > \Sigma_{\text{atoms b ordered behind a}}\ p(b)$$
>
> or else $p(a) > 0$ for the atom of the lowest order. The system also requires that the sum of the atomic probabilities equal unity (total probability). No other constraints bind.

Atomic bound systems express the rule that the ordering between any two exclusive disjunctions is determined by comparing the highest order atom in each sentence, and only those two atoms. The rule is "possibilist" (at least for exclusive sentences), but the solutions are probabilities.

These constraint systems have been successfully applied to a variety of belief modeling tasks, such as establishing the coherence of a scheme for prior-free statistical inference ("ignorant induction"), furnishing a probabilistic model of Kraus, Lehmann and Magidor (1990) default entailment, and grounding a probabilistic semantics for *possibilistic* strict inequalities. Pointers into the relevant literature appear in Snow (1996); the default and possibilistic applications are also discussed in Benferhat, Dubois and Prade (1997).

Atomic bound systems are always consistent, i.e., there are always probability distributions which satisfy the constraints. This can be shown by considering a particular solution, which also illustrates the logicist character of the probability distributions so described.

One solution of an atomic bound system for n atoms assigns to the i-th lowest atom the probability $2^{i-1} / (2^n - 1)$, for example {1/31, 2/31, 4/31, 8/31, 16/31}. By the familiar summation properties of the powers of two which underlie binary number notation, it is clear that (1) the constraints of the system are satisfied, and (2) the probability of each of the 32 distinct sentences in question differs from that of any other sentence.

In binary notation, the numerator of each fraction represents the presence or absence of each atom in the disjunctive rendering of the sentence. The index of any sentence and the value assigned to it are thus yoked together, and restate one another. If the operations in the logic are the usual bitwise AND, OR, and complement, then equivalence preservation is attained.

One way that belief revision can be modeled is by incorporating potential evidence into the domain (i.e., by constructing a joint probability distribution). Observation of evidence then results in a projection of the sentence set. If the remaining values are simply renormalized, then this scheme is completely consistent with Bayes' Theorem.

Reasonableness

This section explores a set of assumptions which are sufficient for the adoption by a logicist of a calculus which agrees with a solution of an atomic bound system. That is, the assumptions support a function $v()$ which assigns real values to sentences so that there is some atomic bound solution $p()$ for which, if A and B are sentences:

$$v(A) \geq v(B) \text{ just in case that } p(A) \geq p(B)$$

In the language of Cox (1946), $v()$ is an increasing transform of a probability distribution. Naturally, $v()$ assigns the same values to equivalent sentences. Weak inequalities will continue to appear in the discussion to accommodate possible equivalencies. As usual, we assume complete ordering for ease of exposition. This can be relaxed with some adjustments in the assumptions.

We begin by assuming that the believer accepts the Sugeno constraints on a belief representation, and wishes to manage beliefs in logicist fashion (for whatever reason). We shall use the following Sugeno and logicist properties:

[S] Consequents are at least as credible as antecedents in implications.

[L] Non-equivalent sentences are strictly ordered.

In motivating a further assumption, we take it that the believer assents that expressions like "if I learn C, then I would think A is no less credible than B" correspond to the relationship $v(A \wedge C) \geq v(B \wedge C)$. So, "seeing red-fleshed fruit would suggest watermelon more strongly than not" becomes $v(\mathit{red} \wedge \mathit{watermelon}) > v(\mathit{red} \wedge \neg \mathit{watermelon})$.

It is sometimes more convenient to use "given" or "conditional" notation rather than conjunctions to distinguish quickly what was learned from what was concluded. So, the notation "$v(A \mid C) \geq v(B \mid C)$" is shorthand for the relationship $v(A \wedge C) \geq v(B \wedge C)$, where C is not contradictory, and if one inequality is strict, then so is the other. This reflects conventions in both probability and possibility which relate conditional expressions like $p(A \mid C) \geq p(B \mid C)$ with the conjunctive $p(A \wedge C) \geq p(B \wedge C)$. "Unconditional" expressions, e.g. $v(A) \geq v(B)$, coincide with tautological "given" expressions, such as $v(A \mid C \vee \neg C) \geq v(B \mid C \vee \neg C)$.

Using this shorthand, let us assume that the believer agrees to a general form of the OR property of case-based reasoning, which may be written:

$$\text{if } v(A \mid C) \geq v(B \mid C) \text{ and } v(A \mid D) \geq v(B \mid D) \text{, then } v(A \mid C \vee D) \geq v(B \mid C \vee D)$$

In words, if learning C would lead to the same belief ordering as learning D would, then knowing that at least one of C and D is true suffices for that ordering to hold.

We shall use only two specific consequences of this general OR property. The first is

$$\text{[Q] if } v(A \mid C) \geq v(B \mid C) \text{ and } v(A \mid \neg C) \geq v(B \mid \neg C), \text{ then } v(A) \geq v(B)$$

That says: if learning C would favor A over B, and learning $\neg C$ would also favor A over B, then we can skip the lesson: A is favored over B, since both possible cases favor that ordering conclusion, and one of them must be true.

We also use the related but distinct Kraus, Lehmann, and Magidor (1990) version of OR,

$$\text{[K] If } A \mid\!\sim C \text{ and } B \mid\!\sim C \text{, then } A \vee B \mid\!\sim C$$

with their connective interpreted as $A \mid\!\sim C$ *just in case that* $v(C \mid A) > v(\neg C \mid A)$. This interpretation is discussed in Benferhat, *et al.* (1997). The KLM authors motivate [K] on much the same basis as the general OR principle was offered here, that is, as an intuitively appealing feature of case-based reasoning.

Theorem. The valuation $v()$ agrees with an atomic bound solution.

Proof (sketch). By [L], all distinct atoms are strictly ordered. By the interpretation of the KLM connective, for atoms a and b, $v(a) > v(b)$ is equivalent to $a \vee b \mid\!\sim a$. Hence, if $v(a) > v(b)$ and for atom c, $v(a) > v(c)$, then $a \vee b \mid\!\sim a$ and

$a \vee c \mid\!\sim a$. So, by [K], $a \vee b \vee c \mid\!\sim a$, which by Boolean simplification is $v(a) > v(b \vee c)$, which is the typical constraint of an atomic bound system. The argument generalizes to any number of atoms.

Ordinal agreement for all mutually exclusive sentences between $v()$ and a solution of an atomic bound system follows easily. Let A and B be exclusive sentences, with a being the highest valued atom in A, and b the highest valued atom in B. WOLG, suppose $v(a) > v(b)$, then by the result shown in the last paragraph, $v(a) > v(B)$. By [S], $v(A) \geq v(a)$ since $a \Rightarrow A$, and by transitivity, $v(A) > v(B)$. As discussed in the preceding section, $p(A) > p(B)$ would obtain under parallel circumstances.

[Q], [L], and complete ordering imply the property of *quasi-additivity*, $v(A) \geq v(B)$ just in case that $v(A \neg B) \geq v(B \neg A)$, which holds in every probability distribution, and binds the ordering of any pair of sentences to echo that of their mutually exclusive "difference." Thus the ordering of any pair of sentences is the same under $v()$ as under an atomic bound solution. //

Although the theorem depends only on the weaker [Q] and [K] versions of OR, it can be shown that atomic bound solutions do comply with the stronger general version of the principle. Atomic bound solutions also comport with all of the KLM "rational" default entailment principles (Snow 1996; Benferhat, *et al.* 1997), not just [K].

A logicist believer with a taste for case-based reasoning or defaults might therefore be interested in atomic bound systems. In explaining that interest, this logicist would not sound much different from more traditional credalists explaining their own favorites within the Sugeno universe.

Expressiveness

This section addresses possible concerns about the absence of equal degrees of belief for non-equivalent sentences, something that would be a feature of any logicist credo.

The current objective is *not* to build models of indeterminate physical systems, like a fair coin, in which the representation of equipoise is crucial. Presumably, there is no quarrel that the statistical behavior of mass phenomena may be modeled outside of logic, just as temperatures and heat flows might be. While you may *believe* that a head is as likely as a tail with such-and-such a coin, that says more about coins than about your psyche.

It is not at all unusual for credos to experience difficulty with equality. In the possibilist calculus, for instance, equal possibility values are referred to a "tie breaking" function called *necessity*. This is not an elective enrichment. Without it, possibility is unable to distinguish between a tautological certainty and a strictly uncertain sentence which disjoins the highest-ranking atom.

Probability also has its run-ins with an overloaded equality relationship (i.e., the same arithmetic comparative corresponds to more than one belief pattern). The most familiar example is the case of three atoms a, b, c about which one is ignorant. One cannot assign different probabilities to them without asserting an ordering among the atoms. In assigning equal probability to each atom, however, one appears to assert that $a \vee b$ is strictly more credible than c, even while professing ignorance. A necessitarian can make an explanation of this. The intent here is not to disparage the merits of such explanations, but merely to point out that there is something to explain, and that the explanation needs a theory of interpretation which cannot be inferred from the calculus and its values alone.

What can be expressed in logicist calculi is that "neither is a more credible than b, nor b more credible than a." This can be done by representing beliefs by non-singleton sets of valuations, the same maneuver by which probability expresses partial orders (Kyburg and Pittarelli 1996). This would overload the absence of order with the equipoise relationship, but as noted, overload in one way or the other has not heretofore been a fatal problem for credos. One might also note that in the real-world domain of civil litigation in the United States, lack of order and equipoise are interchangeable (overloaded) in determining whether the complaining party has met its burden of proof.

The implementation of a partially ordered set representation can be quite compact for the calculus arising from an atomic bound system, even while preserving a "possibilist" style of ordering. Atoms can simply be removed from the disjunctions on the right side of the atomic bound inequalities for the atoms which they tie.

Other Logicist Calculi

Credos constructed from the atomic bound system are not the only kind of probabilistic logicist calculus, nor are all logicist calculi necessarily probabilist.

Among the probability distributions, infinitely many happen to be "self-avoiding" without being ordered in possibilist fashion, for example {2/11, 4/11, 5/11} for three atoms. Whether there are other families of such distributions with meaningful orderings and simple functions for AND and OR is unknown, but self-avoiding distributions appear throughout probability space, arbitrarily close to any finite probability distribution.

Proof (sketch). Start with any finite distribution $p()$. Each step gives a self-avoiding distribution for some of the atoms in the domain. Choose any two atoms, a and b. Let $q(a \mid a \vee b) =$

$p(a) / [p(a) + p(b)]$ if the atoms have different $p()$, or otherwise as close to one-half as desired; $q(b \mid a \vee b)$ is just $1 - q(a \mid a \vee b)$. Pick a third atom, c. It is easy to show that the only "forbidden" values of $q(c \mid a \vee b \vee c)$ have the form $x / (1 + x)$, where x is the difference in q-probability between two exclusive c-free sentences given $a \vee b$. So, $q(c \mid a \vee b \vee c) = p(c) / [p(a) + p(b) + p(c)]$, or as close as desired, if the ratio happens to "forbidden;" $q(a \mid a \vee b \vee c)$ is $1 - q(c \mid a \vee b \vee c)$ times $q(a \mid a \vee b)$, and similarly for b.

A k-th atom's "forbidden" values are based on the (always finite number of) sentence pairs formed from the previous k-1 atoms, The new atom's $q(\,|\,)$ is the applicable $p()$ ratio, or as close as desired. The new values for the atoms already selected are the product of $1 - q(\,|\,)$ for the new atom times their old values. Continue until all atoms have been incorporated. When all the atoms are brought in, there is a self-avoiding distribution $q()$ arbitrarily close in Euclidean distance to the original $p()$. //

So, while it is true that not all probability distributions are truth-functional (except for negation), it is also true that candidate logicist distributions are found "almost everywhere" throughout any finite-dimension probability space. This is a strikingly different picture from the one suggested by the bare impossibility theorems.

Other candidate logicist calculi are incoherent. E.g.,

$$a = .1,\ b = .2,\ c = .3;\ \ a \vee b = .5,\ a \vee c = .4,\ b \vee c = .6;\ \ a \vee b \vee c = 1$$

is self-avoiding, complies with the Sugeno requirement, and agrees with no probability distribution (since $a \vee b$ and $a \vee c$ violate quasi-additivity).

It is hard to find any popular credo which violates quasi-additivity for strict inequalities (compliance is easily shown from the Benferhat, *et al.*, 1997 discussion for the fuzzy family and the schemes subsumed by possibility, and from Kyburg, 1987 for Dempster-Shafer). Maybe, then, the incoherent logicist credos are simply counterintuitive. On the other hand, their existence shows that a logicist stance does not in itself restrict the inductive reasoner to any unusually small slice of the Sugeno universe.

Conclusions

Results such as those studied by Elkan or Dubois and Prade, while correct, do not "kill" equivalence preserving truth-functional uncertainty handling, contrary to the latter pair's claim. The most conspicuous distinguishing mark of logicist calculi, the lack of a direct expression of equipoise, is not very distinctive. Established credos have their own troubles with equality and need interpretative mechanisms to untangle true equipoise from other belief states.

Although the particular example calculus developed in this paper is a kind of probability distribution and probability is rich in candidate logicist distributions, the overall relationship between existing credos and logicist ones is not a matter of subsumption. Even the inclusion of the two-valued Boolean calculus in existing credos is not so much the *subsumption* of Boole as one element of an *intersection* of "ordinary" and logicist credos.

Viewed as a potential competitor in the credalist arena, a logicist credo like the example developed here can combine computational efficiency similar to that of possibility with the famous normative properties of probability. An innate suitability for default entailment and case-based belief revision comes at no extra charge.

Even if the logicist credos are not to one's taste, one might still rejoice that such things exist. The ultimate point of any credo is to perform successful inductive reasoning, some part of which is to convince others about the merits of the inferences made. That is easier when people's conceptual frameworks about what makes sense in the absence of deductive warrant are at least somewhat similar.

References

Benferhat, S., Dubois, D. and Prade, H. 1997. Possibilistic and standard probabilistic semantics of conditional knowledge. In Proceedings AAAI, 70-75.

Cox, R.T. 1946. Probability, frequency, and reasonable expectation, *American Journal of Physics* 14: 1-13.

Dubois, D. and Prade, H. 1994. Can we enforce full compositionality in uncertainty calculi?, In Proceedings AAAI. 149-154.

Elkan, C., 1993. The paradoxical success of fuzzy logic, In Proceedings AAAI. 698-703.

Elkan, C. 1994. The paradoxical success of fuzzy logic. *IEEE Expert* 9: 3-49 (with commentaries).

Kraus, S., Lehmann, D. and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44: 167-207.

Kyburg, H.E., Jr., 1987. Bayesian and non-Bayesian evidential updating. *Artificial Intelligence* 31: 271-294.

Kyburg, H. and M. Pittarelli, M. 1996. Set-based Bayesianism. *IEEE Trans.SMC* 26-A: 324-339.

Lukasiewicz, J. and Tarski, A. 1930. Investigations into the sentential calculus, In A. Tarski 1956. *Logic, Semantics, Metamathematics*, Oxford: Oxford University Press.

Prade, H. 1985. A computational approach to approximate and plausible reasoning with applications to expert systems. *IEEE Trans. PAMI:* 7, 260-283.

Snow, P. 1996. Inductive reasoning with conditional probabilities. In Proceedings FLAIRS, 300-304.