Bulletin of the **S**ection of **L**ogic
Volume 5/3 (1976), pp. 87–93

reedition 2011 [original edition, pp. 87–96]

David Miller

ON PROBABILITY MEASURES FOR DEDUCTIVE SYSTEMS I

1. The problem of defining probability measures not simply on the class of sentences of a language, but on the class of its deductive systems (relative to some configuration of logical rules) was first raised by Mazurkiewicz [1]. In the paper – the second part of which seems never to have been published – Mazurkiewicz purported to provide an axiomatic theory of probability suitable for application to deductive systems. His theory, however, is essentially a finitistic one, and cannot as it stands accommodate deductive systems that are not finitely axiomatizable. No progress at all seems to have been made on the problem since.

In this paper I shall not say anything directly pertinent to the axiomatic question; it strikes me as a rather difficult one. I shall instead consider some of the possibilities for providing probabilities for deductive systems by extending measures already defined on the algebra of sentences of a formalized language.

2. A deductive system is any set of sentences closed under the operation Cn of logical consequence. If a is a consequence of the system A we shall write $a \in A$. Deductive systems are obviously partially ordered by the relation of set inclusion. Tarski [3], section 2, shows that the ordering is in fact a lattice ordering. For reasons that will become apparent I prefer to consider the dual lattice generated by the ordering $\supseteq$. The disjunction $A \vee B$ of two systems is defined as their intersection. The conjunction AB is not the union (this is not a system) but the set $Cn(\{ab : a \in A$ and $b \in B\})$. The lattice has a unit, namely L, the set of logical theorems; and it has a zero, namely S, the set of all sentences. As Tarski shows, it is possible to define an operation akin to negation. I shall not need this at present, but will refer to it below.

The Lindenbaum/Tarski algebra of a language is the Boolean algebra obtained by identifying sentences that are interderivable and defining the operations in the natural way. Suppose that p is a probability measure defined on this algebra. Then we will have

(1) $0 = p(contr) \leq p(a) \leq p(taut) = 1,$

where *contr* is the inconsistent element and *taut* the tautological one; and

(2) $p(a) + p(b) = p(a \vee b) + p(ab),$

the general law of addition. Can we extend p to a probability-like measure q defined throughout the lattice of systems?

3. The most natural way of proceeding is to write

(3) $q(A) = inf\{p(a) \ : \ a \in A\}.$

This was suggested to me by Zoltan Domotor (to whom I am indebted for many helpful discussions on these matters). It is a straightforward business to show that

(4) $0 = q(S) \leq q(A) \leq q(L) = 1,$

but distinctly harder to establish the general addition law

(5) $q(A) + q(B) = q(A \vee B) + q(AB).$

It is this result that I wish now to prove. The proof depends on a lemma about deductive systems that has some considerable interest in its own right.

Lemma.

$$A \vee B = \{a \vee b \ : \ a \in B \ and \ b \in B\};$$
$$AB = \{ab \ : \ a \in A \ and \ b \in B\}.$$

Proof. Clearly every sentence of the form $a \vee b$ belongs to both A and B, and so to their intersection. Moreover, since c is the same as $c \vee c$, every element of $A \vee B$ is identical with a disjunction of elements of A and B. This proves the first part.

In the same way it is immediate that $\{ab \ : \ a \in A$ and $b \in B\} \subseteq AB$. For the converse, suppose $c \in AB$. Then it is derivable from from some finite subset of $\{ab \ : \ a \in A$ and $b \in B\}$; which is to say that it is derivable from a single such sentence ab. But then c is identical with $ab \vee c$, which by distributivity is the same as $(a \vee c)(b \vee c)$. Since $a \vee c \in A$ and $b \vee c \in B$, the result is proved.

THEOREM 1. $q(A) + q(B) = q(A \vee B) + q(AB)$.

PROOF. By the definition (3) and the Lemma we have

$$\begin{aligned} q(A \vee B) + q(AB) &= inf\{p(a \vee b) \ : \ a \in A \text{ and } b \in B\}+ \\ &\quad\; inf\{p(ab) \ : \ a \in A \text{ and } b \in B\} \\ &\leq inf\{p(a \vee b) + p(ab) \ : \ a \in A \text{ and } b \in B\} \\ &= inf\{p(a) + p(b) \ : \ a \in A \text{ and } b \in B\}, \text{ by (2)} \\ &= inf\{p(a) \ : \ a \in A\} + inf\{p(b) \ : \ b \in B\} \\ &= q(A) + q(B), \qquad \text{by (3).} \end{aligned}$$

If the inequality is strict there are sentences $a \in A$, $b \in B$, $c \in A \vee B$ such that for all $x \in A$ and $y \in B$

$$p(c) + p(ab) < p(x \vee y) + p(xy).$$

Let x be ac and y be bc. Then

$$p(c) + p(ab) < p((a \vee b)c) + p(abc),$$

thanks to the distributive law again. But by the law of monotony for p we have

$$p((a \vee b)c) \leq p(c) \text{ and } p(abc) \leq p(ab),$$

which gives us a contradiction.

The measure q accordingly satisfies both (4) and (5). If the lattice of systems were Boolean q would be an orthodox probability measure. But in fact it is only a Brouwerian algebra, so that there are several points of divergence. In particular if, following Tarski, we define $\neg A$ as $\bigcap\{Cn(\neg a) : a \in A\}$, we will not have in general

(6) $q(A) + q(\neg A) = 1$,

but only

(7) $q(A) + q(\neg A) \geq 1$.

Of course, for finitely axiomatizable A equation (6) holds, since q is an extension of p. Thus we do not appear to have a fully informative axiom for the probability of a system's negation.

4. If $q(B) \neq 0$ we can of course define $q(A, B)$, the relative (or conditional) probability of A given B, by $q(AB)/q(B)$. But how are we to proceed when $q(B) = 0$?

Let us suppose that we are given a probability measure $p(a, b)$ that satisfies the axioms of Popper's [2], appendix $*v$. Accordingly $p(a, b)$ is defined for every pair of sentences a, b. Even if b is contradictory $p(a, b)$ is defined: in this case it equals 1, whatever a may be.

Is it possible to extend p to a measure $q(A, B)$ that is defined for every pair of systems and behaves at least partly like a probability measure? The problem is immeasurably more difficult than that treated in the previous section. Here I intend to outline one intuitively appealing method of making the extension, and to exhibit a respect in which it is seriously defective.

No simple monotony law holds for the second argument of the function $p(a, b)$. However, it follows at once from the general multiplication law

(8) $p(ab, c) = p(a, bc) \cdot p(b, c)$

that

(9) if $a \vdash b \vdash c$ then $p(a, c) \leq p(a, b)$.

That is, $p(a, b)$ increases monotonically as b becomes a stronger and stronger logical consequence of a. It therefore seems reasonable to write

(10) if $a \vdash B$ then $q(a, B) = sup\{p(a, b) \ : \ b \in B\}$,

and, more generally,

(11) if $A \vdash B$ then $q(A, B) = sup\ q(A, b) \ : \ b \in B\}$.

In combination with the obvious generalization of (3) this becomes

(12) if $A \vdash B$ then $q(A, B) = \underset{b \in B}{sup} \ \underset{a \in A}{inf} \ \ p(a, b)$.

Once this is agreed, we can set generally

(13) $q(A, B) = q(AB, B)$,

and q is everywhere defined.

THEOREM 2. *The following axioms of Popper's system ([2], p. 349) are satisfied by q:*

$A1$ $\exists C \exists D\ q(A,B) \neq q(C,D)$
$A2$ $(\forall C\ q(A,C) = q(B,C)) \rightarrow q(D,A) = q(D,B)$
$B1$ $q(AB,C) \leq q(A,C)$
$B2$ $q(AB,C) = q(A,BC) \cdot q(B,C)$.

PROOF. $A1$ is obvious, and $B1$ is easy to prove from the definition of q. It is possible to establish $B2$ in a fairly straightforward manner, but the proof is here omitted. $A2$ then follows from $B2$, given that $AB = BA$.

Popper's axiom C of complementation not surprisingly fails in general. What is disconcerting is that the one remaining axiom, which in our notation is

(14) $q(A,A) = q(B,B)$,

also does not hold in general. (Note that if it were to hold then axiom C would have to fail. For Popper has shown that his six axioms guarantee that the domain of element is reducible to a Boolean algebra).

If A is axiomatizable then $q(A,A) = 1$. We shall give an example of a system B for which $q(B,B) < 1$. Indeed, we shall show that for each $x \in B$ there is a $y \in B$ such that $p(y,x) \leq 1/2$. That $q(B,B) \leq 1/2$ then follows from (12).

Tarski, Mostowski and Robinson [4], p. 51, give an example of a finitely axiomatizable system Q that is essentially undecidable. The system B we shall construct is a complete extension of Q, and is therefore not finitely axiomatizable ([4], p. 14). Let b_0 be a single axiom for Q, and let $\{a_j : j \in \omega\}$ be an enumeration of all the sentences in the language in which Q is expressed. We suppose there to be a probability measure $p(a_i, a_j)$ defined for every pair of sentence of this language. Define the sequence $\{b_j : j \in \omega\}$ by:

(i) if $b_j \vdash a_j$ or $b_j \vdash \neg a_j$ then $b_{j+1} = b_j$;

otherwise (ii) $\begin{cases} \text{if} \quad p(a_j, b_j) \leq 1/2 \quad \text{then} \quad b_{j+1} = b_j a_j, \\ \text{if} \quad p(\neg a_j, b_j) < 1/2 \quad \text{then} \quad b_{j+1} = b_j \neg a_j. \end{cases}$

Then

(15) $B = Cn(b_j \ / \ j \in \omega)$.

THEOREM 3. *For the system B defined in (15) we have $q(B,B) \leq 1/2$.*

PROOF. If $i \leq j$ then b_j decides a_i (that is, $b_j \vdash a_i$ or $b_j \vdash \neg a_i$). Since no axiomatizable extension of Q is complete, the sentence b_j cannot decide a_k for every $k > j$. Thus for each b_j there is some $k > j$ for which clause (ii) above comes into action; in short, for each j there is a $k > j$ such that $p(b_k, b_j) \leq 1/2$. Now choose $x \in B$. Certainly for some j we have $b_j \vdash x$. Choose $k > j$ for which $p(b_k, b_j) \leq 1/2$. Then by (9) above $p(b_k, x) \leq p(b_k, b_j) \leq 1/2$.

Thus for each $x \in B$ there is a $y \in B$ such that $p(y, x) \leq 1/2$. By (12) above $q(B, B) \leq 1/2$.

In fact something a lot stronger holds.

THEOREM 4. *If* $q(B, B) \neq 1$ *then* $q(B, B) = q(B, L) = 0$.

PROOF. By Theorem 2, axiom $B2$, $q(BB, B) = q(B, BB) \cdot q(B, B)$, which proves that $q(B, B)$ is either 0 or 1. Likewise, $q(BB, L) = q(B, BL) \cdot (B, L)$, so that $q(B, L)$ is 0 unless $q(B, B) = 1$.

THEOREM 5. *If* $q(B, B) \neq 1$ *then* $q(A, B) = 0$.

PROOF. Again by the multiplication law, $q(AB, B) = q(A, BB) \cdot q(B, B)$. Thus by Theorem 4, $q(AB, B) = 0$. So by (13), $q(A, B) = 0$.

These results show, I think conclusively, that the method of extending p that is given in (12) and (13) is unsatisfactory. This is a surprising, and somewhat disappointing, conclusion. But an analysis of the proof of Theorem 3 suggests one source of the trouble: this is that the limiting processes to the two arguments of $q(A, B)$ are done separately, rather then together. If a way can be found of circumventing this, perhaps a more adequate method of extension may be found.

References

[1] S. Mazurkiewicz, *Ueber die Grundlagen der Wahrscheinlichkeitsrechnung I*, **Monatschefte für Mathematik und Physik** 41 (1933), pp. 343–352.

[2] K. R. Popper, **The Logic of Scientific Discovery**, London, 1959.

[3] A. Tarski, *Grundzüge des Systemenkalkül*, Erster Teil, **Fundamenta Mathematicae** 25 (1935), pp. 503–526. Translated in A. Tarski, **Semantics, Metamathematics**, London, 1956.

[4] A. Tarski, in collaboration with A. Mostowski and R. M. Robinson, **Undecidable Theories**, Amsterdam, 1953.

University of Warwick
England