KYBERNETIKA — VOLUME 14 (1978), NUMBER 1

# A Boolean-Valued Probability Theory

IVAN KRAMOSIL

New foundations for probability theory are suggested and investigated, derived from the formalized systems theory and using also some other branches of mathematical logic. The probability measures are supposed to take their values not in the unit interval of reals, but in a special kind of Boolean algebras. Some analogies as well as differences of this approach with respect to the classical Kolmogorov probability theory are investigated.

## 1. INTRODUCTION AND MOTIVATION

A brief and meta-mathematically oriented discussion of some aspects of Kolmogorov probability theory seems to be an adequate introduction to this paper. We suppose the reader to be familiar with the axiomatic foundations of this theory, however, no more detailed knowledge of probability theory is necessary in order to be able to understand what follows.

Kolmogorov conceived the probability theory as a special case of measure theory and measurable functions theory. Probability is nothing else than a set function satisfying certain axioms, random events are defined as measurable subsets of a universe, random variables are reduced to measurable functions. The technical advantages of such an approach are obvious – measure theory and real functions theory are developed and rich theories having at their disposal a powerful mathematical apparatus together with a number of important results and all this can be used in order to develop a powerful and rich probability theory.

Another aspect of Kolmogorov probability theory, usually also presented as its positive feature, its *logical consistency*, deserves a more detailed analysis. A number of older probability theories (geometric probability, von Mises probability etc.) have failed because of violating the consistency meta-principle when trying to define what the probability of a random event *is*. Kolmogorov avoided this problem by leaving it and replacing it by another one: what the probability of a random event *may be*.

Really, Kolmogorov gives a number of conditions which a mapping from the field of random events into the unit interval is to satisfy in order to be admitted as a possible candidate to play the role of probability.

Even if acknowledging the advantages of this approach and the merits of this conception as far as a further development of probability theory is concerned, we are not allowed to hid or neglect some weak points of Kolmogorov probability theory. When applied to a practical problem, the probability theory is confronted again with the question about the actual value of such and such probability and in this context the elimination of this question from the *theory* is of no worth. Hence, this question must be answered by the means, methods or a priori assumptions the correctness, preciseness and justification of which lie outside the probability theory and continue in influencing all the results, no matter how sophisticated they may be, obtained by the means of probability theory from these premises. E.g., considering a sequence of random events consisting in throwing a coin the Kolmogorov probability theory is able to derive the implication "if the probability of any side of the coin is 1/2 and if the different throws are statistically independent, then the probability of a triple occuring of the head-side in three following each other throws is 1/8". Clearly, to derive the unconditioned assertion "the probability of a triple occuring ... is 1/8" is beyond the power of the theory in question, however, just such assertions are requested when probability theory is applied.

The basic idea of this paper consists in the following suggestion: let us abandon the assumption (common to all probability theories known until now) that the value of a probability measure must be a real number and suppose that probability measures take their values in a more general structure. As far as the author knows in probability theory this approach is original, however, in the theory of fuzzy-sets a similar approach exists, see [1], involved by a similar way of reasoning, so the idea to apply this point of view also to probability theory occurs almost immediately.

It is a matter of fact that the assumption ascribing to random events just reals as their probabilities is not based on some principal reasons and is rather a matter of convenience and an analogy with relative frequencies introduced and investigated in statistics. The unit interval of reals is a so rich and powerful structure that the advantages of easy manipulations with probability values offered in this case has been always considered as great enough to dominate the meta-theoretical (von Mises) or identification (Kolmogorov) difficulties. In this paper we shall follow another way of reasoning: to admit worse possibilities of manipulation with probability values in favor of a more easy identification of these evalues.

Supposing that probability is a mapping defined on a certain structure $\mathscr{S}$ of random events and taking its values in another structure $\mathscr{L}$ and leaving the assumption that $\mathscr{L} = \langle 0, 1 \rangle$ immediately the problem arises which structure should be chosen to play the role of $\mathscr{L}$. A discussion concerning this problem can be found in [1], however, it deals with fuzzy-sets and some arguments are only to a degree justified in our field of reasoning. It is not our aim to enrich this discussion here, we should rather suggest and defend one possibility, namely, we propose to take as $\mathscr{L}$ a *Boolean algebra of a special kind.* 

There are many definitions of Boolean algebras, some of them can be found in the second chapter of [3] or in [4]. Here we define a Boolean algebra $\mathscr{A}$ as the structure $\langle A, \equiv, \wedge, \vee, ', 0, 1\rangle$, where $A$ is a nonempty (and usually at least two-elemented) set, $\vee$(supremum) and $\wedge$(infimum) are binary operations on $A$, $'$(complement) is a unary operation on $A$, $\equiv$ is an equivalence relation on $A$ and $0$(zero) and $1$(unit) are two special elements of $A$. Of course, the structure $\mathscr{A}$ is to obey certain axioms to be really a Boolean algebra, we do not write these axioms explicitly referring the reader to [3], [4] or to another source. In $\mathscr{A}$ we define, in the usual way, a partial ordering relation $\leqq$ such that, for any $x, y \in A$, $x \leqq y$ iff $(x \wedge y) \equiv x$, or, equivalently, iff $(x \vee y) \equiv y$.

Besides the most elementary properties of Boolean algebras with which the reader is supposed to be familiar let us consider a possibility of a relativization of the notion of Boolean algebra.

**Lemma 1.** Let $\mathscr{A}$ be a Boolean algebra, let $a \in A$. Denote by $A(a) \subset A$ the subset of all elements of $A$ of the form $a \vee x$, $x \in A$ and define binary operations $\wedge(a)$, $\vee(a)$ on $A$:

$$(a \vee x) \vee(a) (a \vee y) = (\mathrm{df})\, a \vee (x \vee y)\,,$$

$$(a \vee x) \wedge(a) (a \vee y) = (\mathrm{df})\, a \vee (x \wedge y)\,.$$

Define also unary operation $'(a)$ on $A$:

$$(a \vee x)'\,(a) = (\mathrm{df})\, a \vee x'\,.$$

Then $\mathscr{A}(a) = \langle A(a), \equiv, \wedge(a), \vee(a), '(a), a, 1\rangle$ is a Boolean algebra.

Proof. A simple verification of axioms of Boolean algebras. Q.E.D.

Consider, now, a formalized language $\mathscr{B}$ based on the first-order predicate calculus, i.e., roughly speaking, $\mathscr{B}$ is the set of all well-formed formulas of a first-order predicate calculus based theory together with a structure on this set. Suppose that the propositional connectives $\wedge$(conjunction), $\vee$(alternative), $\neg$(negation) and $\Leftrightarrow$(equivalence) are primitive connectives of $\mathscr{B}$ or that they have been defined in $\mathscr{B}$ (the identity of these symbols with those of Boolean algebras will not be misleading). Choose an element $a \in |\mathscr{B}|$ and consider a deducibility relation $\vdash$ defined on $\mathscr{B}$ on the base of the first-order predicate calculus axioms, maybe some other axioms, and the usual predicate calculus deducibility rules. Define an equivalence relation $\equiv$ on $|\mathscr{B}|$ in such a way that $x \equiv y$ iff $\vdash x \Leftrightarrow y$. A well-known theorem (see [3], e.g.) then sounds that

$$\mathscr{L}(\mathscr{B}) = \langle |\mathscr{B}|, \equiv, \wedge, \vee, \neg, a \wedge \neg a, a \vee \neg a\rangle$$

is a Boolean algebra, usually called the *Lindenbaum* or *Lindenbaum-Tarski algebra* over the formalized theory $\langle \mathscr{B}, \vdash\rangle$. And they are just the Lindenbaum algebras over first-order formalized theories which will be taken as the structures in which probability measures should take their values. We are ready to formalize the principal definition of this paper.

**Definition 1.** Let $\Omega$ be a non-empty set, let $\langle \mathscr{B}, \vdash \rangle$ be a formalized theory based on the first-order predicate calculus. Then the triple $\langle \Omega, \mathscr{B}, \vdash \rangle$ is called *Boolean probability space*. Suppose that $\Omega$ is the support of a semantical structure with relations and functions corresponding to the relational and functional constants of $\mathscr{B}$, then $A \subset \Omega$ is called *Boolean random event*, if there is a formula $V_A \in |\mathscr{B}|$ defining $A$ in $\langle \mathscr{B}, \vdash \rangle$ in the usual semantical sense. Let $A \subset \Omega$ be a Boolean random event, then any formula $V_A \in |\mathscr{B}|$ defining $A$ is called the *Boolean probability of A* and will be denoted by $BP(A)$, i.e. Boolean probabilities are formulas defining certain subsets of $\Omega$.

As it can be easily seen the difference between a Boolean random event and its Boolean probability is very narrow and can be neglected on the formalized level. No Boolean random event can be given in the language $\mathscr{B}$ without defining, at the same time, its probability by the same formula.

If compared with the way in which Kolmogorov probability has been constructed a similarity between that and our approaches can be found. We have used, again, another branch of mathematics and try to embed probability theory in this field. Instead of measure and measurable functions theory used by Kolmogorov we try to use mathematical logic and metamathematics, i.e. theory of formalized mathematical theories, to the same goal. The reason for this effort is, roughly speaking, the adequacy of mathematical logic for the purposes of artificial intelligence and robotics. A discussion can be found in the closing part of this paper.

## 2. SOME PROPERTIES OF BOOLEAN PROBABILITIES

A well-known assertion of the Boolean algebras theory sounds that only those Boolean algebras can be complete the cardinality of which equals to $2^n$, i.e. the cardinality of which is either finite or at least continuum. Lindenbaum algebras are, with the exception of the most trivial finite cases, countable, hence, they are not complete. This gives that union and intersection of a sequence of random events are not, in general, random events. It would be possible to embed a Lindenbaum algebra into its minimal complete extension defined on the base of Stone structure associated with this algebra (the construction itself as well as some necessary assertions and their proofs can be found in [3]). However, in such a case random events would become more complex mathematical entities not reducible on definable subsets of the basic space $\Omega$. As the effectivity of random events in the sense of their definability in the given language is the basic principle of our way of handling with probabilities we abandon the demand of completeness and we do not consider these extended algebras.

It is clear, now, that the $\sigma$-aditivity axiom of Kolmogorov probability theory has no reasonable counterpart in our theory. A much more weak assertion can be proved. Consider a Boolean probability space $\langle \Omega, \mathscr{B}, \vdash \rangle$.

**Theorem 1.** Let $A_1, A_2, \ldots$ be a sequence of Boolean random events, then for any $n$

$$BP(\bigcup_{i=1}^{n} A_i) \geqq BP(A_j)\,, \quad j \leqq n\,, \quad BP(\bigcup_{i=1}^{n+1} A_i) \geqq BP(\bigcup_{i=1}^{n} A_i)\,.$$

If, moreover, $A_i \leqq \neg A_j$ for any $i, j$, $i \neq j$, (in such a case the Boolean random events are called *mutually disjoint*) and if $A_i \not\equiv \Lambda$ (i.e. $B \wedge \neg B$ for any $B \in |\mathscr{B}|$), then

$$BP(\bigcup_{i=1}^{n} A_i) > BP(A_j)\,, \quad j \leqq n\,, \quad BP(\bigcup_{i=1}^{n+1} A_i) > BP(\bigcup_{i=1}^{n} A_i)\,.$$

Proof. The first two inequalities are trivial. Let $A_1, A_2, \ldots$ be mutually disjoint and non-zero. The propositional tautology $((B \vee A) \to B) \equiv (A \to B)$ gives that $\bigvee_{i=1}^{n+1} A_i \to \bigvee_{i=1}^{n} A_i$ holds iff $A_{n+1} \to \bigvee_{i=1}^{n} A_i$. However, $A_{n+1} \to A_j$, $j \leqq n$, so $A_{n+1} \to \bigwedge_{i=1}^{n} (\neg A_i)$, $A_{n+1} \to \neg(\bigwedge_{i=1}^{n} A_i)$. Immediately follows that $A_{n+1} \to \bigwedge_{i=1}^{n} (\neg A_i) \wedge \bigvee_{i=1}^{n} A_i$, hence, $A_{n+1} \to \Lambda$. Q.E.D.

**Definition 2.** Let $\langle \Omega, \mathscr{B}, \vdash \rangle$ be a Boolean probability space, let $A$, $B$ be Boolean random events. Then the conditional Boolean probability $BP(A/B)$ of $A$ under the condition $B$ is defined as $B \to A$.

**Theorem 2.** Let $A$, $B$ be Boolean random events, then

(a) conditional Boolean probability $BP(\cdot|B)$ is a mapping of the Lindenbaum algebra $\mathscr{L}(\mathscr{B}, \vdash)$ into the Boolean algebra $\mathscr{L}(\neg B)$ (see Lemma 1),

(b) $BP(A|B) \geqq BP(A)$,

(c) if $A_1 \equiv A_2$, then $BP(B|A_1) \equiv BP(B|A_2)$, $A_1, A_2 \in |\mathscr{B}|$,

(d) $BP(A|\mathrm{V}) \equiv BP(A)$, $\mathrm{V} = (\mathrm{df})\ \neg\Lambda$,

(e) $BP(\Lambda|A) \equiv \neg A$, i.e. the zero element of $\mathscr{L}(\neg A)$,

(f) if $\vdash B \to A$, then $BP(A|B) \equiv \mathrm{V}$, if $\vdash B \to \neg A$, then $BP(A|B) \equiv \neg B$, i.e. the zero element of $\mathscr{L}(\neg B)$.

Proof. All the assertions are either trivial or follow immediately from some well-known propositional tautologies.

We must admit that the assertion (b) contradicts the properties of classical conditional probabilities as $BP(A)$ serves as a lower bound for all $BP(A|B)$. If $A$ and $B$

exclude each other, then $BP(A|B) \equiv BP(A)$, if $B$ implies $A$, then $BP(A|B) \equiv V$. On the other hand take into consideration the fact, that conditional probability $BP(\cdot|B)$ is defined to take its values not in $\mathscr{L}$ itself, but in $\mathscr{L}(\neg B)$ and any comparing of $BP(A)$ and $BP(A/B)$ can be done only after the embedding of $\mathscr{L}(\neg B)$ into $\mathscr{L}$ and in the framework of $\mathscr{L}$. In $\mathscr{L}(\neg B)$ the values of $BP(\cdot|B)$ vary from $\Lambda(\mathscr{L}(\neg B))$ to V according to the classical demands. There is still another reason worth of mentioning. The paradoxal inequality (b) is an immediate and inevitable consequence of the so called "paradoxon of classical implication" and cannot be avoided in the framework of classical propositional calculus as the following assertion shows.

**Theorem 3.** There exists no binary classical propositional functor $f$ such that the conditional Boolean probability $BP(A|B)$ defined by $f(A, B)$ would satisfy simultaneously: if $\vdash B \rightarrow A$, then $BP(A|B) \equiv V$, if $\vdash B \rightarrow \neg A$, then $BP(A|B) \equiv \Lambda$.

Proof. If $B \equiv \Lambda$, then $\vdash B \rightarrow A$, $\vdash B \rightarrow \neg A$. However, if the theory $\langle \mathscr{B}, \vdash \rangle$ is consistent, the simultaneous validity of $f(A, B) \equiv \Lambda$ and $f(A, B) \equiv V$ is excluded. Q.E.D.

Multiple conditional probabilities can be defined by the relation $BP(A|B, C) = = (\mathrm{df})\, BP(A|B)|C)$. As

$$BP(BP(A|B)|C) \equiv BP(B \rightarrow A|C) \equiv C \rightarrow (B \rightarrow A) \equiv (B \wedge C) \rightarrow A \equiv \\ \equiv BP(A|B \wedge C),$$

the possibility of elimination of multiple conditional probabilities is conserved as well as the way in which this elimination is performed.

**Definition 3.** Boolean random events $A_1, A_2, \ldots, A_n$ are *mutually independent*, if no among the meta-assertions

$$\vdash (A_1 \wedge A_2 \wedge \ldots \wedge A_{j-1} \wedge A_{j+1} \wedge \ldots \wedge A_n) \rightarrow A_j\,, \quad j \leqq n\,,$$
$$\vdash \neg(A_1 \wedge A_2 \wedge \ldots \wedge A_n)\,,$$
$$\vdash A_1 \vee A_2 \vee \ldots \vee A_n$$

is valid. The Boolean random events $A_1, A_2, \ldots, A_n$ are *pairwise independent*, if any pair $\langle A_i, A_j \rangle$, $i, j \leqq n$, $i \neq j$, is mutually independent. Boolean random events $A_1, A_2, \ldots$ are mutually independent, if any finite subsequence of $A_1, A_2, \ldots$ is mutually independent.

**Theorem 4.** If Boolean random events $A_1, A_2, \ldots, A_n$ are mutually independent, they are also pairwise independent the inverse implication not being generally valid (agrees with the classical probability theory).

Proof. The non-validity of $\vdash (A_1 \wedge \ldots \wedge A_{j-1} \wedge A_{j+1} \wedge \ldots \wedge A_n) \rightarrow A_j$, $\vdash \neg(A_1 \wedge \ldots \wedge A_n)$, $\vdash A_1 \vee \ldots \vee A_n$ immediately gives the non-validity of $\vdash A_i \rightarrow A_j$, $\vdash \neg(A_i \wedge A_j)$, $\vdash A_i \vee A_j$. As a counter-example showing the non-validity of the inverse implication consider the theory $\langle \mathscr{B}, \vdash \rangle$ resulting from the first order predicate calculus by its enriching by five unary predicate constants $P_1, P_2, \ldots, P_5$ and by axioms 

$$(\forall x)\,(P_1(x) \vee P_2(x) \vee \ldots \vee P_5(x))\,,$$
$$(\exists x)\,P_i(x)\,,\quad i = 1, 2, \ldots, 5\,,$$
$$(\forall x)\,[P_i(x) \rightarrow \neg(P_1(x) \vee P_2(x) \vee \ldots \vee P_{i-1}(x) \vee P_{i+1}(x) \vee \ldots \vee P_5(x))]\,,$$
$$i = 1, 2, \ldots, 5\,.$$

Then the Boolean random events $P_1(x) \vee P_2(x) \vee P_3(x)$, $P_2(x) \vee P_3(x) \vee P_4(x)$, $P_3(x) \vee P_4(x) \vee P_5(x)$ are pairwise independent, but not mutually independent, as $\vdash P_1(x) \vee P_2(x) \vee \ldots \vee P_5(x)$.

The following theorem gives, without proofs, some more and easily verifiable properties of mutually independent Boolean random events.

**Theorem 5.** (a) Any of the pairs $\langle \Lambda, \Lambda \rangle$, $\langle \Lambda, V \rangle$, $\langle V, V \rangle$ of Boolean random events is neither mutually nor pairwise independent.

(b) If Boolean random events $A$, $B$ are independent, then none among the Boolean probabilities $BP(A|B)$, $BP(A|\neg B)$, $BP(B|A)$, $BP(B|\neg A)$ is equivalent to V or $\Lambda$.

(c) If Boolean random $A$, $B$ are not independent, then at least one among the Boolean probabilities given in (b) is equivalent to V.

## 3. CONVERGENCE OF BOOLEAN PROBABILITIES

The assertions of limit kind play an important role in the classical Kolmogorov probability theory. Let us examine a possibility how to express at least some of these assertions in our formalism.

**Definition 4.** Consider a formalized theory $\langle \mathscr{B}, \vdash \rangle$ and its Lindenbaum algebra $\mathscr{L}$. Let $\{A_i\}_{i=1}^{\infty}$ be a sequence of Boolean random events, then this sequence tends to a Boolean random event $A_0$, or $A_0$ is the limit of $\{A_i\}_{i=1}^{\infty}$, in symbols $\{A_i\}_{i=1}^{\infty} \nearrow A_0$, if

(a) there is an $n_0$ such that for all $i > n_0$, $A_i \leqq A_{i+1} \leqq A_0$,

(b) if $A_0^*$ is another Boolean random event satisfying (a), then $A_0 \leqq A_0^*$.

I.e., limit is the same as the least upper bound of a monotonous non-decreasing sequence of Boolean random events. As can be easily seen, if the limit of a sequence exists, it is unique up to equivalence $\equiv$.

Having a formalized theory $\langle \mathscr{B}, \vdash \rangle$, suppose to have also an enumeration $E$ of elementary formulas of this theory ascribing different naturals to different elementary formulas. Denote, for any $A \in |\mathscr{B}|$, by $A^P$ the propositional formula resulting when every elementary formula $e$ in $A$ is replaced by the propositional indeterminate $p_{E(e)}$ and all quantifiers are erased.

**Definition 5.** Let $\langle \mathscr{B}, \vdash \rangle$ be a formalized theory with the Lindenbaum algebra $\mathscr{L}$, let $A \in |\mathscr{B}|$. A propositional variable $p$ occuring in $A^P$ is called *substantial* in $A^P$, if there is no propositional formula $A_1$ containing only the indeterminates occuring in $A^P$ but with the exception of $p$ and such that $\vdash A_1 \equiv A^P$. The set of all indeterminates substantial in $A^P$ is denoted by $Var(A)$. Boolean random events $A_1, A_2, \ldots$ from $\mathscr{L}$ are called *mutually strongly independent*, if $\Lambda \not\equiv A_i^P \not\equiv V$, $i = 1, 2, \ldots$ and the sets $Var(A_i)$ are mutually disjoint.

**Theorem 6.** Let $\langle \mathscr{B}, \vdash \rangle$ be a formalized theory, let Boolean random events $A_1, A_2, \ldots$ be mutually strongly independent. Then $\{\bigvee_{i=1}^{n} A_i^P\}_{n=1}^{\infty} \nearrow$ V.

Proof. The assumption $\Lambda \not\equiv A_i^P$ gives that there is, for any $i$, such a mapping $p_i$ from the set of all indeterminates occuring in $A_i^P$ into $\{V, \Lambda\}$ that $\vdash p_i A_i$, where $p_i A_i$ is the formula resulting from the substitution given by $p_i$ and applied to $A_i^P$. Having at our disposal only $n$ propositional indeterminates we can construct just $2^{2^n}$ non-equivalent formulas. This fact and the strong independence of $A_i$ imply that the set $\bigcup_{i=1}^{i} Var(A_i)$ is not finite. Suppose, to come to a contradiction, that there is $A_0 \in |\mathscr{B}|$, $A_0^P \not\equiv$ V, such that $(\bigvee_{j=1}^{i} A_j^P) \leqq A_0^P$ for all $i \geqq i_0$. Immediately follows that there is an $i_1 \geqq i_0$ such that $Var(A_{i_1}) \cap Var(A_0) = \emptyset$. $A_{i_1}^P \not\equiv \Lambda$, $A_0^P \not\equiv$ V, so there exists a mapping $p_0$ of the type described above such that $p_0 A_{i_1} \equiv p_{i_1} A_{i_1} \equiv$ V, $p_0 A_0 \equiv \Lambda$. This gives that $p_0(\bigvee_{j=1}^{i_1} A_j) \equiv \Lambda$, hence, $\vdash (\bigvee_{j=1}^{i_1} A_j^P) \to A_0^P$ does not hold, so $\{\bigvee_{i=1}^{n} A_i^P\}_{n=1}^{\infty} \nearrow$ V. Q.E.D.

In fact, the strong mutual independence represents a sufficient but not necessary condition for the convergence of $\bigvee_{i=1}^{n} A_i^P$ to V.

## 4. NUMERICAL REPRESENTATION OF BOOLEAN PROBABILITIES

Introducing the idea of Boolean probabilities we have mentioned some arguments in favour of our point of view that the classical numerical-valued probability theory suffers from some disadvantages of applicational character. However, its undoubtable priority consists in a very rich structure of the unit interval, enabling to handle easily the probabilities and to derive rather complicated constructions over them. In this chapter we propose an idea, how to use the reals from the unit interval also in our Boolean-valued probability theory, even if in other way than the classical probability does. 

Consider a finite set $\mathscr{S} \subset \mathscr{L}$ of formulas (i.e. Boolean random events from the Lindenbaum algebra over a formalized theory $\langle\mathscr{B}, \vdash\rangle$), denote by $\|\mathscr{S}\|$ the cardinal number of $\mathscr{S}$. In fact, $\mathscr{S}$ is not precisely a set as we admit the possibility that some formulas are contained more than once in $\mathscr{S}$ (sometimes the term *bag* instead of set is used in this case). Formally this situation can be described by defining $\mathscr{S}$ as a set of pair of the type $\langle A, n\rangle$, $A \in \mathscr{L}$, $n$ being the arity of $A$ in $\mathscr{S}$. However, we think that the more intuitive sense of the following definitions and reasonings justifies our way of considering $\mathscr{S}$ to be immediately a subset of $\mathscr{L}$. Besides the demands $\mathscr{S} \neq \emptyset$, $\|\mathscr{S}\| < \infty$, we suppose, till the end of this paper, that

$$\{x : x \in \mathscr{L},\ x \equiv \Lambda\} \cap \mathscr{S} = \emptyset\,,$$

i.e. $\mathscr{S}$ does not contain a contradictory formula. On the other hand, a "global" inconsistency of $\mathscr{S}$, i.e. the possibility that $\mathscr{S} \vdash \Lambda$, is not excluded.

**Definition 6.** Let $\pi$ be the mapping of the Cartesian product $\mathscr{L} \times \mathscr{L} \times \mathscr{P}_{fin}(\mathscr{L})$ into the real line defined, for any $A, B \in \mathscr{L}$, $\mathscr{S} \in \mathscr{P}_{fin}(\mathscr{L})$, as

$$\pi(A \mid B, \mathscr{S}) = \frac{\|\{x : x \in \mathscr{S},\ \vdash x \to (A \wedge B)\}\|}{\|\{x : x \in \mathscr{S},\ \vdash x \to B\}\|}$$

supposing the set $\{x : x \in \mathscr{S},\ \vdash x \to B\}$ is not empty, $\pi(A/B, \mathscr{S})$ is not defined otherwise. The real number $\pi(A|B, \mathscr{S})$ is then called the *numerical image of the Boolean conditional probability* $BP(A|B)$ *with respect to* $\mathscr{S}$. The numerical values of the unconditioned Boolean probabilities are defined by setting $B \equiv V$, immediately follows, as $\mathscr{S} \neq \emptyset$, that these images are always defined.

**Theorem 7.** For every $A, B \in \mathscr{L}$, $\mathscr{S} \subset \mathscr{L}$, if $\pi(A|B, \mathscr{S})$ is defined, then $0 \leqq \pi(A|B, \mathscr{S}) \leqq 1$. Let $A_1, A_2, \ldots, B \in \mathscr{L}$, $\mathscr{S} \subset \mathscr{L}$, let there be at least one $x \in \mathscr{S}$ such that $\vdash x \to B$, let $\vdash A_i \to \neg A_j$ for any $i, j = 1, 2, \ldots, i \neq j$. Then there exists an $n$ such that

$$\pi\Big(\bigvee_{i=1}^{n} A_i|B, \mathscr{S}\Big) \geqq \sum_{i=1}^{\infty} \pi(A_i|B, \mathscr{S})\,.$$

Proof. The first assertion follows immediately from the fact that

$$\{x : x \in \mathscr{S}, \vdash x \to A \wedge B\} \subset \{x : x \in \mathscr{S}, \vdash x \to B\}\,.$$

Let $\vdash x \to A_i$, then $\vdash x \to \neg A_j$, hence, not $\vdash x \to A_j$. This gives that

$$\{x : x \in \mathscr{S},\ \vdash x \to A_i\} \cap \{x : x \in \mathscr{S},\ \vdash x \to A_j\} = \emptyset\,,\quad i \neq j\,.$$

$\mathscr{S}$ is finite, so there is the maximal index $n$ such that $\{x : x \in \mathscr{S},\ \vdash x \to A_n\} \neq \emptyset$. $\vdash x \to A_i$, $i \leqq n$, implies $\vdash x \to \bigvee_{i=1}^{n} A_i$, hence, $\{x : x \in \mathscr{S},\ \vdash x \to (\bigvee_{i=1}^{n} A_i) \wedge B\} \supset$ $\supset \bigcup_{i=1}^{n} \{x : x \in \mathscr{S},\ \vdash x \to (A_i \wedge B)\}$, $\|\{x : x \in \mathscr{S},\ \vdash x \to (\bigvee_{i=1}^{n} A_i) \wedge B\}\| \geqq \sum_{i=1}^{\infty} \|\{x : x \in$ $\in \mathscr{S},\ \vdash x \to (A_i \wedge B)\}\|$ and the assumption $\{x : x \in \mathscr{S},\ \vdash x \to B\} \neq \emptyset$ gives immediately the assertion. Q.E.D.

The mapping $\pi$ has the same properties as the so called inner measures sometimes studied in measure theory, see [2]. The super-aditivity of $\pi$ can be replaced by the usual $\sigma$-aditivity when the meta-relation $\vdash$ of provability is replaced by the intuitionistic provability $(I) \vdash$.

**Theorem 8.** Let $\pi_1$ be defined in the same way as $h$ just with $\vdash$ replaced by $(\mathrm{I}) \vdash$, let the other condition of Theorem 7 be satisfied. Then, for the same $n$ as in the assertion of Theorem 7,

$$\pi_1(\bigvee_{i=1}^{n} A_i | B, \mathscr{S}) = \sum_{i=1}^{\infty} \pi_1(A_i | B, \mathscr{S}) .$$

Proof. The properties of intuitionistic provability give that $(I) \vdash x \to (\bigvee_{i=1}^{k} A_i)$ implies that $(\mathrm{I}) \vdash x \to A_j$ for at least one $j \leqq k$. The other conditions give, as in the proof of Theorem 7, that there is just one $j \leqq n$ with this property, so

$$\|\{x : x \in \mathscr{S},\ (I) \vdash x \to (\bigvee_{i=1}^{n} A_i) \wedge B\}\| =$$
$$= \sum_{i=1}^{\infty} \|\{x : x \in \mathscr{S},\ (I) \vdash x \to (A_i \wedge B)\}\|$$

which implies the assertion. Q.E.D.

Let us mention some other properties of the mapping $\pi$, rather trivial but interesting enough when compared with the classical probability.

**Theorem 9.** (a) Let $A, B, C \in \mathscr{L}$, $\mathscr{S} \subset \mathscr{L}$, $\vdash A \to C$, let $\{x : x \in \mathscr{S},\ \vdash x \to B\} \neq \emptyset$, then $\pi(A|B, \mathscr{S}) \leqq \pi(C|B, \mathscr{S})$, $\pi(\Lambda|B, \mathscr{S}) = 0$, $\pi(\mathrm{V}|B, \mathscr{S}) = 1$.

(b) Let $A, B \in \mathscr{L}$, $\mathscr{S} \to \mathscr{L}$, $\{x : x \in \mathscr{S},\ \vdash x \to B\} \neq \emptyset$. If $\vdash B \to A$, then $\pi(A|B,|\mathscr{S}) = 1$, if $\vdash B \to \neg A$, then $\pi(A|B, \mathscr{S}) = 0$.

Proof. All the assertions follow from the definition of $\pi$ after an easy calculation. Q.E.D.

Hence, $\pi(A|B, \mathscr{S})$ varies from 0 to 1 in the same way as the classical probability does and the paradoxon of classical implication does not influence the behaviour of $\pi$.

**Theorem 10.** Let $A, B \in \mathscr{L}, \mathscr{S}, \mathscr{S}', \mathscr{S}_1, \mathscr{S}_2, \ldots \subset \mathscr{L}$, $0 < \|\mathscr{S}\|, \|\mathscr{S}'\|, \|\mathscr{S}_i\| < \infty$, let there exist, for any $n$, an $x \in \bigcup_{i=1}^{\infty} \mathscr{S}_i$ such that $\vdash x \to B$, let $\mathscr{S}_i \cap \mathscr{S}_j = \emptyset$, $i \neq j$, then, for $n \to \infty$,

$$\left|\pi(A|B, \mathscr{S} \cup \bigcup_{i=1}^{n} \mathscr{S}_i) - \pi(A|B, \mathscr{S}' \cup \bigcup_{i=1}^{n} \mathscr{S}_i)\right| \to 0 .$$

Proof. The assumptions assure, that for any $n \geqq n_0$ the values $\pi(A|B, \mathscr{S} \cup \bigcup_{i=1}^{n} \mathscr{S}_i)$, $\pi(A|B, \mathscr{S}' \cup \bigcup_{i=1}^{n} \mathscr{S}_i)$ are defined. Moreover, $\|\{x : x \in \bigcup_{i=1}^{n} \mathscr{S}_i, \vdash x \to B\}\| \to \infty$, $n \to \infty$. The fact that

$$\left|\pi(A|B, \mathscr{S} \cup \bigcup_{i=1}^{n} \mathscr{S}_i) - \pi(A|B, \mathscr{S}' \cup \bigcup_{i=1}^{n} \mathscr{S}_i)\}\right| \leqq$$

$$\leqq \frac{\|\{x : x \in \mathscr{S} \cup \mathscr{S}', \vdash x \to (A \wedge B)\}\|}{\|\{x : x \in \bigcup_{i=1}^{n} \mathscr{S}_i, \vdash x \to B\}\|} \to 0 , \quad n \to \infty ,$$

implies the desired assertion. Q.E.D.

**Corollary.** Let $A \in \mathscr{L}$, $\mathscr{S}, \mathscr{S}', \mathscr{S}_1, \mathscr{S}_2, \ldots \subset \mathscr{L}$, $0 < \|\mathscr{S}\|, \|\mathscr{S}'\|, \|\mathscr{S}_i\| < \infty$ $\mathscr{S}_i \cap \mathscr{S}_j = \emptyset$, $i \neq j$, then, for $n \to \infty$,

$$\left|\pi(A, \mathscr{S} \cup \bigcup_{i=1}^{n} \mathscr{S}_i) - \pi(A, \mathscr{S}' \cup \bigcup_{i=1}^{n} \mathscr{S}_i)\right| \to 0 .$$

Proof. Set $\mathrm{V} \equiv B$ in Theorem 10. Q.E.D.

## 5. SOME DISCUSSION AND CONCLUSIVE REMARKS

Having started from the idea of non-numerical-valued probability measure we have come back to the numerical-valued mapping $\pi$ of the field of random events taking its values in the unit interval and similar, to a degree, with a classical probability measure on a finite probability space. However, the ontological level of $\pi$ is different from that taken by a classical probability measure. Formally, the difference is expressed by the presence of the argument $\mathscr{S}$ when defined. To illustrate this matter let us personify the set $\mathscr{S}$ of formulas by calling it as an observer, i.e. $\mathscr{S}$ describe the knowledge and experience of an observer who uses the theory $\langle \mathscr{B}, \vdash \rangle$ in order to describe the environment in which he is situated and about which he wants to derive conclusions of stochastic character. The multiplied occurence of some formulas in $\mathscr{S}$ expresses the multiple experience causing the result of this multiple experience to play more important role, i.e. to influence more significantly the values $\pi(\cdot, \mathscr{S})$, than an individual and never more repeated experience.

Now, consider two observers $\mathscr{S}, \mathscr{S}'$, i.e. $\mathscr{S}, \mathscr{S}' \in \mathscr{P}_{fin}(\mathscr{L})$. Having a Boolean random event $A$, its Boolean probability $BP(A)$ is objective in the sense that it is common for $\mathscr{S}$ and $\mathscr{S}'$ (up to logical equivalence) under the condition that $\mathscr{S}$ and $\mathscr{S}'$ use the same theory $\langle \mathscr{B}, \vdash \rangle$ when describing the environment, in another words, if $\mathscr{S}$ and $\mathscr{S}'$ "speak in the same language". On the other hand, the values $\pi(A, \mathscr{S})$ and $\pi(A, \mathscr{S}')$ may differ, they are subjective, and the question "which of the values $\pi(A, \mathscr{S})$, $\pi(A, \mathscr{S}')$ is correct or which is better?" cannot be answered not because of some technical difficulties, but because this question has no sense in our way of understanding the probability theory. Only if $\vdash A \rightarrow B$ for Boolean random events $A, B$ and, at the same time, $\pi(A, \mathscr{S}') > \pi(B, \mathscr{S}')$, then $\mathscr{S}$ is justified, if speaking in the same language as $\mathscr{S}'$, to say to his colleague that he, i.e. $\mathscr{S}'$, is wrong, that $\mathscr{S}'$ has made an error and, moreover, $\mathscr{S}$ is able to prove to $\mathscr{S}'$ this error and to persuade him about it. And again, if $\vdash A \rightarrow B$ and, say, $\pi(A, \mathscr{S}) = 0{\cdot}21$, $\pi(B, \mathscr{S}) = 0{\cdot}29$, $\pi(A, \mathscr{S}') = 0{\cdot}12$, $\pi(B, \mathscr{S}') = 0{\cdot}74$, then both the images of the Boolean probabilities $BP(A)$, $BP(B)$ are to the same degree "true" or "false" and the question "which is better?" is not answerable as all such questions are beyond any sense.

The probabilitistic approach to many problems of artificial intelligence is forced by the increasing descriptional and computational complexity of various problems the automated solution of which is the desired goal in automatic problem solving-perhaps the key branch of artificial intelligence. Considering robotics as the principal applicational branch for automatic problem solving we can immediately see that "from the robot's point of view" Boolean probabilities are much more useful than the classical ones because of the two following reasons.

First, at least at the present level of automatic problem solving, any robot "intelligent" enough to solve more difficult problems must be able to handle, at least partially, with formulas of an appropriate logical calculus and deducibility relation among them, as the most important problem solving methods are based on such a conversion (situation calculus). So the idea to use this apparatus and these robot's abilities also in order to take some statistically based decisions and to use some statistical estimation or testing methods is quite natural and justified. On the other hand, as classical probabilities of random events are not immediately given by their formalized descriptions, the use of classical probability theory in automatic problem solving would request either to implement the necessary probability values immediately or to implement some computing rules enabling to compute these values using some other data. In the case of Boolean probabilities the situation is much more simple, even if robot must replace some Boolean probabilities by their numerical images, e.g., being forced to compare some probabilities not comparable in the Boolean sense, such a calculation is an easy matter of a rather universal routine. Here we do not consider, of course, the difficulties arising from the possible large cardinality of the set $\mathscr{S}$ and from the fact that the deducibility relation $\vdash$ is, in general, undecidable. There are some possibilities how to make the handling with Boolean probabilities more effective, e.g., to replace $\mathscr{S}$ by some its representative subset (by a random sample, say) or to replace $\vdash$ by an effective immediate consequence relation and such modifications would be of great practical importance. However, these problems are still open and their more detailed investigation would bring us beyond the intended framework of this paper.



(Received April 18, 1977.)

REFERENCES


[1] J. Goguen: L-Fuzzy Sets. Journal of Math. Analysis and Applications *18* (1967), 147—174.

[2] P. R. Halmos: Measure Theory. D. van Nostrand Comp., Inc., Toronto, New York, London 1951.

[3] H. Rasiowa, R. Sikorski: The Mathematics of Metamathematics. Third Edition, PWN, Warszawa 1970.

[4] R. Sikorski: Boolean Algebras. Second Edition, Springer Verlag, Berlin—Göttingen—Heidelberg—New York 1964.


*Dr. Ivan Kramosil, CSc., Ústav teorie informace a automatizace ČSAV (Institute of Information Theory and Automation — Czechoslovak Academy of Sciences), Pod vodárenskou věží 4, 182 08 Praha 8. Czechoslovakia.*