ASSIGNING PROBABILITIES TO LOGICAL FORMULAS

DANA SCOTT
*Stanford University, Stanford, Calif.**
and
PETER KRAUSS
University of California, Berkeley, Calif.

1. Introduction. Probability concepts nowadays are usually presented in the standard framework of the Kolmogorov axioms. A sample space is given together with a σ-field of subsets – the events – and a σ-additive probability measure defined on this σ-field. When the study turns to such topics as stochastic processes, however, the sample space all but disappears from view. Everyone says "*consider the probability that* $X \geq 0$", where X is a random variable, and only the pedant insists on replacing this phrase by "*consider the measure of the set* $\{\omega \in \Omega : X(\omega) \geq 0\}$". Indeed, when a process is specified, only the distribution is of interest, not a particular underlying sample space. In other words, practice shows that it is more natural in many situations to assign probabilities to *statements* rather than *sets*. Now it may be mathematically useful to translate everything into a set-theoretical formulation, but the step is not always necessary or even helpful. In this paper we wish to investigate how probabilities behave on statements, where to be definite we take the word "*statement*" to mean "*formula of a suitable formalized logical calculus*".

It would be fair to say that our position is midway between that of Carnap and that of Kolmogorov. In fact, we hope that this investigation can eventually make clear the relationships between the two approaches. The study is not at all complete, however. For example, Carnap wishes to emphasize the notion of the *degree of confirmation* which is like a conditional probability function. Unfortunately the mathematical theory of general conditional probabilities is not yet in a very good state. We hope in future papers to comment on this problem. Another question concerns the formulation of

* This work was partially supported by grants from the National Science Foundation and the Sloan Foundation.

interesting problems. So many current probability theorems involve *expectations* and *limits* that it is not really clear whether consideration of probabilities of formulas alone really goes to the heart of the subject. We do make one important step in this direction, however, by having our probabilities defined on *infinitary formulas* involving countable conjunctions and disjunctions. In other words, our theory is σ-additive.

The main task we have set ourselves in this paper is to carry over the standard concepts from ordinary logic to what might be called *probability logic*. Indeed ordinary logic is a special case: the assignment of truth values to formulas can be viewed as assigning probabilities that are either 0 (for *false*) or 1 (for *true*). In carrying out this program, we were directly inspired by the work of Gaifman [1964] who developed the theory for finitary formulas. Aside from extending Gaifman's work to the infinitary language, we have simplified certain of his proofs making use of a suggestion of C. Ryll-Nardzewski. Further we have introduced a notion of a *probability theory*, in analogy with theories formalized in ordinary logic, which we think deserves further study.

In Section 2 the logical languages are introduced along with certain syntactical notions. In Section 3 we define *probability systems* which generalize relational systems as pointed out by Gaifman. In Section 4 we show how given a probability system the probabilities of arbitrary formulas are determined. In Section 5 we discuss model-theoretic constructs involving probability systems. In Section 6 the notion of a *probability assertion* is defined which leads to the generalization of the notion of a *theory* to probability logic. In Section 7 we specialize and strengthen results for the case of finitary formulas. In Section 8 examples are given. An appendix (by Peter Krauss) is devoted to the mathematical details of a proof of a measure-theoretic lemma needed in the body of the paper.

2. The languages of probability logic. Throughout this paper we will consider two different first-order languages, a finitary language $\mathscr{L}^{(\omega)}$ and an infinitary language $\mathscr{L}$. To simplify the presentation both languages have an identity symbol $=$ and just one non-logical constant, a binary predicate $\mathbf{R}$. Most definitions and results carry over with rather obvious modifications to the corresponding languages with other non-logical constants, and we will occasionally make use of this observation when we give specific examples.

The language $\mathscr{L}^{(\omega)}$ has a denumerable supply of distinct individual variables v_n, for each $n<\omega$, and $\mathscr{L}$ has distinct individual variables v_ξ, for each $\xi<\omega_1$, where ω_1 is the first uncountable ordinal. Both languages have logical constants $\wedge$, $\vee$, $\neg$, $\forall$, $\exists$, and $=$ standing for (finite) conjunction, disjunction, negation, universal and existential quantification, and identity as mentioned before. In addition the infinitary language $\mathscr{L}$ has logical constants $\bigwedge$ and $\bigvee$ standing for denumerable conjunction and disjunction respectively. The expressions of $\mathscr{L}$ are defined as transfinite concatenations of symbols of length less than ω_1, and the formulas of $\mathscr{L}^{(\omega)}$ and $\mathscr{L}$ are built from atomic formulas of the forms $\mathbf{R}v_\xi v_\eta$ and $v_\xi = v_\eta$ in the normal way by means of the sentential connectives and the quantifiers. Free and bound occurrences of variables in formulas are defined in the well-known way. (For a more explicit description of infinitary languages see the monograph Karp [1964].) A sentence is a formula without free variables.

We will augment the non-logical vocabulary of our languages with various sets T of new individual constants $t \in T$ and denote the resulting languages by $\mathscr{L}^{(\omega)}(T)$ and $\mathscr{L}(T)$ respectively. It is then clear what the formulas and sentences of $\mathscr{L}^{(\omega)}(T)$ and $\mathscr{L}(T)$ are. For any set T of new individual constants let $\mathscr{S}$ and $\mathscr{S}(T)$ be the set of sentences of $\mathscr{L}$ and $\mathscr{L}(T)$ respectively, and let $\mathscr{s}(T)$ be the set of quantifier-free sentences of $\mathscr{L}(T)$. We adopt analogous definitions for the language $\mathscr{L}^{(\omega)}$.

If Σ is a set of sentences and φ is a sentence, then φ is a *consequence* of Σ if φ holds in all models in which all sentences of Σ hold, and we write $\Sigma \models \varphi$. φ is *valid* if it is a consequence of the empty set, and we write $\models \varphi$. For both languages $\mathscr{L}^{(\omega)}$ and $\mathscr{L}$ we choose standard systems of deduction, and we write $\Sigma \vdash \varphi$ if φ is derivable from Σ. φ is *a theorem* if φ is derivable from the empty set, and we write $\vdash \varphi$. (For details concerning the infinitary language we again refer the reader to Karp [1964].) By the well-known Completeness Theorem of finitary first order logic we have for every $\Sigma \subseteq \mathscr{S}^{(\omega)}$ and every $\varphi \in \mathscr{S}^{(\omega)}$, $\Sigma \vdash \varphi$ iff $\Sigma \models \varphi$. This is not true for the infinitary language $\mathscr{L}$; however, we still have "weak" completeness in the sense that for every $\varphi \in \mathscr{S}$, $\vdash \varphi$ iff $\models \varphi$.

We call two sentences φ and ψ *equivalent* if $\vdash \varphi \leftrightarrow \psi$. It is well-known that this is an equivalence relation, and that the equivalence classes form a Boolean algebra, the so-called Lindenbaum-Tarski algebra of sentences, which in the infinitary case is σ-complete. By the "weak" Completeness Theorem this algebra is isomorphic to a field of sets of models, which in the infinitary case is a σ-field. Let $\mathscr{S}/\vdash$, $\mathscr{S}(T)/\vdash$ and $\mathscr{s}(T)/\vdash$ be the Lindenbaum-Tarski algebras of the respective sets of sentences. Clearly $\mathscr{S}/\vdash$ and $\mathscr{s}(T)/\vdash$ are σ-subalgebras of $\mathscr{S}(T)/\vdash$. We adopt analogous definitions for the language $\mathscr{L}^{(\omega)}$. For any sentence φ, let $\varphi/\vdash$ be the equivalence class of φ. Finally we define some relativized notions. If Σ is a set of sentences we call two sentences φ and ψ equivalent with respect to Σ if $\Sigma \vdash \varphi \leftrightarrow \psi$. It is now clear how the relativized Lindenbaum-Tarski algebras $\mathscr{S}/\Sigma\vdash$, $\mathscr{S}(T)/\Sigma\vdash$ and $\mathscr{J}(T)/\Sigma\vdash$ are defined. Since the "strong" Completeness Theorem fails for the infinitary language $\mathscr{L}$, it is not in general the case that $\mathscr{S}/\Sigma\vdash$ is isomorphic to a σ-field of sets of models. On the other hand $\mathscr{S}^{(\omega)}/\Sigma\vdash$ is isomorphic to a field of models. Again, for any sentence φ, $\varphi/\Sigma\vdash$ is the equivalence class of φ with respect to Σ, and the mapping which sends $\varphi/\vdash$ into $\varphi/\Sigma\vdash$ is a σ-homomorphism of $\mathscr{S}/\vdash$ onto $\mathscr{S}/\Sigma\vdash$.

In general we will present definitions and results for the infinitary language $\mathscr{L}$, and we will show in Section 7 how these results can be specialized, and in many cases strengthened, for the finitary language $\mathscr{L}^{(\omega)}$.

3. Probability systems. We start with the definition of a concept which corresponds to the notion of a relational system in ordinary logic. Recall that if $\mathscr{A}$ is a Boolean algebra then a probability on $\mathscr{A}$ is a σ-additive probability measure on $\mathscr{A}$. A *finitely additive probability* on $\mathscr{A}$ is a finitely additive probability measure on $\mathscr{A}$. For a detailed discussion of these concepts see Halmos [1963] and Sikorski [1964].

DEFINITION. *A probability system (or sometimes, a probability model) is a quintuple* $\langle A, R, \mathrm{Id}, \mathscr{A}, m\rangle$, *where*

(i) A *is a nonempty set;*

(ii) $\langle \mathscr{A}, m\rangle$ *is a measure algebra, that is,* $\mathscr{A}$ *is a Boolean* σ*-algebra and* m *is a strictly positive probability on* $\mathscr{A}$;

(iii) R *is a function on* $A\times A$ *into* $\mathscr{A}$;

(iv) Id *is a function on* $A\times A$ *into* $\mathscr{A}$ *with the substitution property, which means that for all* $a, a', b, b' \in A$,

(a) $\mathrm{Id}(a, a)=\mathbf{1}$

(b) $\sim\mathrm{Id}(a, a')\cup\sim\mathrm{Id}(b, b')\cup\sim\mathrm{Id}(a, b)\cup\mathrm{Id}(a', b')=\mathbf{1}$

(c) $\sim\mathrm{Id}(a, a')\cup\sim\mathrm{Id}(b, b')\cup\sim R(a, b)\cup R(a', b')=\mathbf{1}$.

If $\mathfrak{A}=\langle A, R, \mathrm{Id}, \mathscr{A}, m\rangle$ is a probability system, then $\mathscr{A}$ is a complete Boolean algebra (see Halmos [1963] p. 67), and therefore $\langle A, R, \mathrm{Id}, \mathscr{A}\rangle$ is a Boolean-algebraic model in the ordinary sense (see, e.g., Karp [1964] p. 140).

Id is to be interpreted as identity. If $\mathrm{Id}(a, b)=\mathbf{0}$ for all $a, b\in A$ such that $a\neq b$, we follow a suggestion of Gaifman [1964] and call $\mathfrak{A}$ a probability system with *strict identity* and write $\mathfrak{A}=\langle A, R, \mathscr{A}, m\rangle$. If $\mathfrak{A}$ is a probability system with strict identity and m is a two valued probability, then $\mathscr{A}$ is the two element Boolean algebra and $\langle A, R\rangle$ may be identified with an ordinary model. Most concrete examples of probability systems we have encountered have strict identity. However some intuitively very suggestive model constructions, such as the ultraproduct construction and symmetric probability systems which will be discussed in Section 5, lead beyond the realm of probability systems with strict identity. For this reason we thought it advisable to introduce the more general notion of a probability system.

If $\mathfrak{A}=\langle A, R, \mathrm{Id}, \mathscr{A}, m\rangle$ is a probability system, we define for all $a, b \in A$,

$$a \approx b \text{ iff } \mathrm{Id}(a, b) = \mathbf{1}.$$

The substitution property of Id implies that $\approx$ is a congruence relation on $\mathfrak{A}$. The *cardinality* of $\mathfrak{A}$, denoted by $|\mathfrak{A}|$ is defined to be the cardinality of the set of equivalence classes with respect to $\approx$. If $\mathfrak{A}$ is a probability system with strict identity then $|\mathfrak{A}|$ is the cardinality of the set A. More generally, for any subset $A' \subseteq A$ the *system-cardinality* of A, denoted by $|A'|_{\mathfrak{A}}$, is defined to be the cardinality of the set of equivalence classes with respect to $\approx$ which have a non-empty intersection with A'.

4. Probability interpretations. We now interpret the language $\mathscr{L}$ in probability systems and give a definition of the concept "*a sentence* φ *holds in a probability system* $\mathfrak{A}$ *with probability* α", where $0 \leq \alpha \leq 1$ is a real number. The definition could be given in the traditional way using an analogue of Tarski's concept of satisfaction; however, in the context of probability logic it seems to be more appropriate to use the equally well-known device of new individual constants.

Let $\mathfrak{A}=\langle A, R, \mathrm{Id}, \mathscr{A}, m\rangle$ be a probability system and let $T_{\mathfrak{A}}=\{t_a : a \in A\}$ be a set of new individual constants such that $t_a \neq t_b$ whenever $a \neq b$.

We recursively define a *valuation* function h on $\mathscr{S}(T_{\mathfrak{A}})$ into $\mathscr{A}$:

(i) $h(t_a = t_b) = \mathrm{Id}(a, b)$,

(ii) $h(\mathbf{R}t_a t_b) = R(a, b)$,

(iii) $h(\neg \varphi) = \mathbf{1} \sim h(\varphi)$,

(iv) $h(\bigvee_{i<\xi} \varphi_i) = \bigvee_{i<\xi} h(\varphi_i)$,

(v) $h(\bigwedge_{i<\xi} \varphi_i) = \bigwedge_{i<\xi} h(\varphi_i)$,

(vi) $h(\exists v \varphi) = \bigvee_{a \in A} h(\varphi(t_a))$,

(vii) $h(\forall v \varphi) = \bigwedge_{a \in A} h(\varphi(t_a))$.

The following lemma is well-known and easy to prove:

LEMMA 4.1. (i) *For all* $\varphi, \psi \in \mathscr{S}(T_{\mathfrak{A}})$, *if* $\vdash \varphi \leftrightarrow \psi$, *then* $h(\varphi)=h(\psi)$;

(ii) *h induces a σ-homomorphism from* $\mathscr{S}(T_{\mathfrak{A}})/\vdash$ *into* $\mathscr{A}$.

Proof: By the definition of h it suffices to prove that for all $\varphi \in \mathscr{S}(T_{\mathfrak{A}})$, if $\varphi \vdash$ then $h(\varphi)=\mathbf{1}$. This can be done by considering a standard system of deduction for $\mathscr{L}(T_{\mathfrak{A}})$, and showing that h maps all axioms into $\mathbf{1}$ and that the property of being mapped into $\mathbf{1}$ is preserved under all rules of deduction. For a more detailed presentation see Karp [1964]. (ii) is an immediate consequence of (i).

We identify h with the induced homomorphism and define $\mu_{\mathfrak{A}}(\varphi/\vdash) = m(h(\varphi))$ for all $\varphi \in \mathscr{S}(T_{\mathfrak{A}})$. Then $\mu_{\mathfrak{A}}$ is a probability on $\mathscr{S}(T_{\mathfrak{A}})/\vdash$. (This is a well-known fact in measure theory, and a proof can be found in Halmos [1963] p. 66.) Since hardly any confusion could arise we write $\mu_{\mathfrak{A}}(\varphi)$ for $\mu_{\mathfrak{A}}(\varphi/\vdash)$, and we read "$\mu_{\mathfrak{A}}(\varphi)=\alpha$" as "*$\varphi$ holds in the probability system $\mathfrak{A}$ with probability α*".

If $\mathfrak{A}$ is a probability system with strict identity and m is two valued, then for every $\varphi \in \mathscr{L}(T_{\mathfrak{A}})$, $\mu_{\mathfrak{A}}(\varphi)=1$ iff φ holds in the model $\langle A, R \rangle$. Thus the definition of the probability $\mu_{\mathfrak{A}}$ is a canonical extension of the ordinary definition of truth. Moreover, if $\mathfrak{A}$ has strict identity, then $\mu_{\mathfrak{A}}(t_a = t_b)=0$ for $a, b \in A$ where $a \neq b$, and therefore $\mu_{\mathfrak{A}}$ is two valued on the identity sentences (the sentences without the predicate $\mathbf{R}$).

The next lemma introduces the Gaifman Condition (**G**).

LEMMA 4.2. *Whenever* $\exists v\varphi \in \mathscr{S}(T_{\mathfrak{A}})$, *then*

$$\text{(G)} \qquad \mu_{\mathfrak{A}}(\exists v\varphi) = \sup_{F \in A^{(\omega)}} \mu_{\mathfrak{A}}\Big(\bigvee_{a \in F} \varphi(t_a)\Big);$$

where $A^{(\omega)}$ *is the set of all finite subsets of* A.

Proof: $\mu_{\mathfrak{A}}(\exists v\varphi) = m(h(\exists v\varphi)) = m(\bigvee_{a \in A} h(\varphi(t_a)))$. $\mathscr{A}$ is a measure algebra and therefore satisfies the countable chain condition (see Halmos [1963] p. 67). Thus there exists a countable subset $A' \subseteq A$ such that

$$\bigvee_{a \in A} h(\varphi(t_a)) = \bigvee_{a \in A'} h(\varphi(t_a)).$$

Thus

$$\mu_{\mathfrak{A}}(\exists v\varphi) = m\Big(\bigvee_{a \in A'} h(\varphi(t_a))\Big) = m\Big(h\Big(\bigvee_{a \in A'} \varphi(t_a)\Big)\Big) = \mu_{\mathfrak{A}}\Big(\bigvee_{a \in A'} \varphi(t_a)\Big).$$

Since $\mu_{\mathfrak{A}}$ is σ-additive,

$$\mu_{\mathfrak{A}}\Big(\bigvee_{a \in A'} \varphi(t_a)\Big) = \sup_{F \in A'^{(\omega)}} \mu_{\mathfrak{A}}\Big(\bigvee_{a \in F} \varphi(t_a)\Big).$$

Condition (**G**) now follows from the choice of A'.

We now present the preceding ideas from a slightly different point of view to see that the probability system $\mathfrak{A}$ may be identified with the restriction of the probability $\mu_{\mathfrak{A}}$ to $\mathscr{s}(T_{\mathfrak{A}})/\vdash$. We first observe that from the definition of h and the countable chain condition in $\mathscr{A}$ it follows that for the purpose of our probability interpretation we may assume that $\mathscr{A}$ is the σ-algebra generated by the union of the images of $A \times A$ under Id and R respectively. If the definition of h is restricted to clauses (i)–(v), then obviously Lemma 4.1 holds with $\mathscr{S}(T_{\mathfrak{A}})$ replaced by $\mathscr{s}(T_{\mathfrak{A}})$. Since m is strictly positive on $\mathscr{A}$ we have

$$\{\varphi \in \mathscr{s}(T_{\mathfrak{A}}) : \mu_{\mathfrak{A}}(\varphi) = 0\} = \{\varphi \in \mathscr{s}(T_{\mathfrak{A}}) : h(\varphi) = \mathbf{0}\}.$$

Thus the quotient algebra of $\mathscr{s}(T_{\mathfrak{A}})/\vdash$ modulo the σ-ideal $\{\varphi/\vdash \in \mathscr{s}(T_{\mathfrak{A}})/\vdash : \mu_{\mathfrak{A}}(\varphi)=0\}$ is isomorphic to $\mathscr{A}$, and it is a well-known fact that the probability m on $\mathscr{A}$ may be uniquely recaptured from the probability $\mu_{\mathfrak{A}}$ on $\mathscr{s}(T_{\mathfrak{A}})/\vdash$. (See, e.g., Halmos [1963] pp. 64ff.) Thus the probability system $\mathfrak{A}$ is, up to the obvious isomorphism, determined by the ordered pair $\langle T_{\mathfrak{A}}, \mu_{\mathfrak{A}} \rangle$, where $\mu_{\mathfrak{A}}$ is restricted to $\mathscr{s}(T_{\mathfrak{A}})/\vdash$.

In general any ordered pair $\langle T, \mu \rangle$, where T is a set of new individual constants and μ is a probability on $\mathscr{s}(T)/\vdash$, uniquely determines a probability system $\mathfrak{A}$. Indeed let $A=T$, let $\mathscr{A}$ be the quotient algebra of $\mathscr{s}(T)/\vdash$ modulo the σ-ideal $\{\varphi/\vdash : \varphi \in \mathscr{s}(T), \mu(\varphi)=0\}$, let m be the probability on $\mathscr{A}$ induced by μ, and let $\mathrm{Id}(t, t')$ and $R(t, t')$ be the image of $t = t'/\vdash$ and $\boldsymbol{R}(t, t')/\vdash$ under the canonical homomorphism of $\mathscr{s}(T)/\vdash$ onto $\mathscr{A}$. Then $\mathfrak{A} = \langle A, R, \mathrm{Id}, \mathscr{A}, m \rangle$ clearly is a probability system; it is easy to check that the valuation homomorphism h is the canonical homomorphism, and μ is the restriction of $\mu_{\mathfrak{A}}$ to $\mathscr{s}(T)/\vdash$. Moreover, if $\mu(t = t')=0$ for all $t, t' \in T$ where $t \neq t'$, then $\mathfrak{A}$ has strict identity.

Thus we may also regard a probability system as an ordered pair $\langle T, m \rangle$, where T is a set of new individual constants, and m is a probability on $\mathscr{s}(T)/\vdash$. The probability systems with strict identity are then characterized by the condition $m(t = t')=0$ for all $t, t' \in T$ where $t \neq t'$. This is the form in which Gaifman [1964] introduces the concept of a probability model and, whenever convenient, we will also adopt this terminology.

From this new point of view we have the following extension theorem:

THEOREM 4.3. *Let $\langle T, m \rangle$ be a probability system. Then there exists a unique probability m^* on $\mathscr{S}(T)/\vdash$ which extends m and satisfies the Gaifman Condition: whenever $\exists v\varphi \in \mathscr{S}(T)$, then*

(G) $$m^*(\exists v\varphi) = \sup_{F \in T^{(\omega)}} m^*\Big(\bigvee_{t \in F} \varphi(t)\Big);$$

where $T^{(\omega)}$ is the set of all finite subsets of T.

Proof: The existence of m^* is clear from our considerations above. The uniqueness of the extension will be proved by transfinite induction. During the course of our proof we will make use of analogues of Lemma 7.9 which will be established separately, and of course independently, for the finitary language $\mathscr{L}^{(\omega)}(T)$ in Section 7 of this paper.

For every ordinal $\xi<\omega_1$, we shall define sets of sentences $\mathscr{S}_\xi(T)\subseteq\mathscr{S}_\xi(T)\subseteq\mathscr{S}(T)$ by recursion: First let $\mathscr{S}_0(T)=\mathscr{S}(T)$. Then if $\xi>0$, let $\mathscr{S}_\xi(T)$ be the closure of $\bigcup_{\eta<\xi}\mathscr{S}_\eta(T)$ under denumerable propositional combinations. For every $\xi<\omega_1$, let $\mathscr{S}_\xi(T)$ be the closure of $\mathscr{S}_\xi(T)$ under quantification and finite propositional combinations. Then obviously whenever $\eta<\xi<\omega_1$,

$$\mathscr{S}_\eta(T)\subseteq\mathscr{S}_\eta(T)\subseteq\mathscr{S}_\xi(T)\subseteq\mathscr{S}_\xi(T)$$

and

$$\mathscr{S}(T)=\bigcup_{\xi<\omega_1}\mathscr{S}_\xi(T)=\bigcup_{\xi<\omega_1}\mathscr{S}_\xi(T).$$

Now suppose n_1 and n_2 are both σ-additive probability measures on $\mathscr{S}(T)/\vdash$ which extend m and satisfy condition (**G**). We shall prove by transfinite induction that for every $\xi<\omega_1$ and every $\varphi\in\mathscr{S}_\xi(T)$, we have $n_1(\varphi)=n_2(\varphi)$.

In case $\xi=0$ and $\varphi\in\mathscr{S}_0(T)$, then $n_1(\varphi)=n_2(\varphi)$ by hypothesis. If $\varphi\in\mathscr{S}_0(T)$, then φ may be written in prenex normal form **QM**, where **Q** is a string of quantifiers and **M** is an $\mathscr{S}_0(T)$-matrix of φ; that is, every substitution instance of **M** belongs to $\mathscr{S}_0(T)$. By an obvious analogue of Lemma 7.9 we have $n_1(\varphi)=n_2(\varphi)$. In case $\xi>0$, first observe that $(\bigcup_{\eta<\xi}\mathscr{S}_\eta(T))/\vdash$ is a subalgebra of $\mathscr{S}(T)/\vdash$ that σ-generates $\mathscr{S}_\xi(T)/\vdash$. By way of an induction hypothesis, we assume that $n_1(\varphi)=n_2(\varphi)$ for every $\varphi\in\bigcup_{\eta<\xi}\mathscr{S}_\eta(T)$. Since n_1 and n_2 are both σ-additive measures, we conclude by a well-known extension theorem of measure theory (see Halmos [1950] p. 54), that $n_1(\varphi)=n_2(\varphi)$ for all $\varphi\in\mathscr{S}_\xi(T)$. If $\varphi\in\mathscr{S}_\xi(T)$ then φ may be written in prenex normal form **QM**, where **M** is an $\mathscr{S}_\xi(T)$-matrix of φ; that is, every substitution instance of **M** belongs $\mathscr{S}_\xi(T)$. Again by analogue of Lemma 7.9 we have $n_1(\varphi)=n_2(\varphi)$. Thus by transfinite induction $n_1=n_2$.

Remark: Gaifman first formulated condition (**G**) and published a proof of Theorem 4.3 for the finitary language $\mathscr{L}^{(\omega)}(T)$, which of course is an immediate consequence of Theorem 4.3. The authors subsequently proved Theorem 4.3 for the infinitary language along constructive lines suggested by the uniqueness proof given above. The idea underlying the presentation in this paper, which renders the existence part of Theorem 4.3 almost trivial, was suggested to us by Professor C. Ryll-Nardzewski.

It is now clear how to define various probability-model-theoretic concepts in analogy to the standard concepts of ordinary model theory. We will discuss a few examples in the next section.

5. Model-theoretic concepts in the theory of probability systems. If $\langle T, m\rangle$ is a probability system, let m^* be the extension of m to $\mathscr{S}(T)/\vdash$ satisfying **(G)**, and let $\bar{m}$ be the restriction of m^* to $\mathscr{S}/\vdash$.

DEFINITION. *Let $\langle T_1, m_1\rangle$, $\langle T_2, m_2\rangle$ be probability systems, then*

(i) $\langle T_1, m_1\rangle \subseteq \langle T_2, m_2\rangle$ *iff* $T_1 \subseteq T_2$ *and* m_1 *is the restriction of* m_2 *to* $\mathscr{s}(T_1)/\vdash$;

(ii) $\langle T_1, m_1\rangle \preccurlyeq \langle T_2, m_2\rangle$ *iff* $T_1 \subseteq T_2$ *and* m_1^* *is the restriction of* m_2^* *to* $\mathscr{S}(T_1)/\vdash$;

(iii) $\langle T_1, m_1\rangle \equiv \langle T_2, m_2\rangle$ *iff* $\bar{m}_1 = \bar{m}_2$.

Remark: The concepts defined in (i), (ii) and (iii) correspond to the concepts of *subsystem*, *$\mathscr{L}$-subsystem* and *$\mathscr{L}$-equivalence* respectively in ordinary model theory.

Not many interesting results concerning these concepts are known for the infinitary language $\mathscr{L}$, a phenomenon which the probability-model theory of $\mathscr{L}$ seems to share with the ordinary model theory of $\mathscr{L}$. In many cases the authors have been able to establish for probability logic analogies of major results known from ordinary logic; this is particularly true for the finitary language $\mathscr{L}^{(\omega)}$, for which several results have already been published by Gaifman [1964].

We present next a few standard constructions for probability systems.

Independent Unions. Let I be an index set. For each $i \in I$, let $\mathscr{L}_i$ be the infinitary language whose only non-logical constant is the binary predicate $\mathbf{R}_i$, and let $\mathscr{S}_i$ be its set of sentences. Let T be a set of new individual constants. For each $i \in I$, let $\langle T, m_i\rangle$ be a probability system where m_i is a probability on $\mathscr{s}_i(T)/\vdash$. We shall assume these systems have strict identity. Let $\mathscr{L}$ be the infinitary language whose non-logical constants are all the binary predicates $\mathbf{R}_i$, $i \in I$, and let $\mathscr{S}$ be its set of sentences. For every $i \in I$, $\mathscr{s}_i(T)/\vdash$ and $\mathscr{s}(T)/\vdash$ are isomorphic to σ-fields of sets of models, and $\mathscr{s}(T)/\vdash$ is isomorphic to the σ-field product $\prod_{i \in I} \mathscr{s}_i(T)/\vdash$ of the family $\{\mathscr{s}_i(T)/\vdash : i \in I\}$. $m = \prod_{i \in I} m_i$ is the product measure on $\mathscr{s}(T)/\vdash$ induced by the family $\{m_i : i \in I\}$.

Then we define the *independent union* of the family of probability systems $\{\langle T, m_i\rangle : i\in I\}$, to be the probability system $\langle T, m\rangle$, and denote it by $\sum_{i\in I} \langle T, m_i\rangle$.

We note two corollaries of the construction given with this definition:

COROLLARY 5.1. *For every $i\in I$ and $\varphi\in\mathscr{S}_i(T)$, $m^*(\varphi)=m_i^*(\varphi)$.*

Proof: We argue by transfinite induction along the same lines as the uniqueness part of the proof of Theorem 4.3. Let $i\in I$, and for every $\xi<\omega_1$ define sets of sentences $\mathscr{I}_{i\xi}(T)\subseteq\mathscr{S}_{i\xi}(T)\subseteq\mathscr{S}_i(T)$, as in the proof of Theorem 4.3. If $\varphi\in\mathscr{I}_{i0}(T)$ then $m^*(\varphi)=m_i^*(\varphi)$ by the definition of m. The rest of the induction is carried out as in the proof of Theorem 4.3.

We state a simple fact about product measures: Let X, Y be sets, let $\mathscr{A}$, $\mathscr{B}$ be fields of subsets of X, Y respectively; let $\bar{\mathscr{A}}$, $\bar{\mathscr{B}}$ be the σ-fields generated by $\mathscr{A}$, $\mathscr{B}$ respectively; and let μ, ν be probabilities on $\bar{\mathscr{A}}$, $\bar{\mathscr{B}}$ respectively. Let $\bar{\mathscr{A}}\times\bar{\mathscr{B}}$ be the product σ-field of $\bar{\mathscr{A}}$ and $\bar{\mathscr{B}}$. Then we have:

LEMMA 5.2. *If λ is a probability on $\bar{\mathscr{A}}\times\bar{\mathscr{B}}$ such that $\lambda(A\times B)=\mu(A)\cdot\nu(B)$ for all $A\in\mathscr{A}$, $B\in\mathscr{B}$, then $\lambda(A\times B)=\mu(A)\cdot\nu(B)$ for all $A\in\bar{\mathscr{A}}$, $B\in\bar{\mathscr{B}}$.*

Proof: Let $\mathscr{A}\times\mathscr{B}$ be the field of subsets of $X\times Y$ generated by rectangles $A\times B$, where $A\in\mathscr{A}$, $B\in\mathscr{B}$. Then the condition $\lambda(A\times B)=\mu(A)\cdot\nu(B)$ determines the probability λ on $\mathscr{A}\times\mathscr{B}$. $\bar{\mathscr{A}}\times\bar{\mathscr{B}}$ is σ-generated by $\mathscr{A}\times\mathscr{B}$. Thus this condition determines λ on $\bar{\mathscr{A}}\times\bar{\mathscr{B}}$. The product measure $\mu\times\nu$ on $\bar{\mathscr{A}}\times\bar{\mathscr{B}}$ agrees with λ on $\mathscr{A}\times\mathscr{B}$. Thus $\lambda=\mu\times\nu$, which proves the assertion.

COROLLARY 5.3. *For every $n<\omega$, let $i_n\in I$ and let $\varphi_n\in\mathscr{S}_{i_n}(T)$. Then*

$$m^*(\bigwedge_{n<\omega}\varphi_n)=\prod_{n<\omega} m^*(\varphi_n).$$

Proof: The assertion follows from the continuity of m^* if we can establish: *If $n<\omega$, $i_0,\ldots,i_{n-1}\in I$ and $\varphi_k\in\mathscr{S}_{i_k}(T)$ for all $k<n$, then*

$$m^*(\bigwedge_{k<n}\varphi_k)=\prod_{k<n} m^*(\varphi_k).$$

As in the proof of Theorem 4.3, for every $\xi<\omega_1$, $k<n$ we define sets of sentences $\mathscr{I}_{i_k\xi}(T)\subseteq\mathscr{S}_{i_k\xi}(T)\subseteq\mathscr{S}_{i_k}(T)$. If $\varphi_k\in\mathscr{S}_{i_k}(T)$ for all $k<n$, then there exists $\xi<\omega_1$ such that $\varphi_k\in\mathscr{S}_{i_k\xi}(T)$ for all $k<n$. Accordingly we prove by transfinite induction: *For every $\xi<\omega_1$, if $\varphi_k\in\mathscr{S}_{i_k\xi}(T)$ for all $k<n$, then $m^*(\bigwedge_{k<n}\varphi_k)=\prod_{k<n}m^*(\varphi_k)$.* First if $\varphi_k\in\mathscr{I}_{i_k0}(T)$ for all $k<n$, then the assertion holds by the definition of m^*. If $\varphi_k\in\mathscr{S}_{i_k0}(T)$ for all $k<n$, then for every $k<n$, φ_k may be written in prenex normal form $\mathsf{Q}_k\mathsf{M}_k$, where Q_k is a

string of quantifiers and M_k is an $\mathscr{S}_{i_k 0}(T)$-matrix of φ_k, which means every substitution instance of M_k belongs to $\mathscr{S}_{i_k 0}(T)$. It is now easy to see that by an analogue of Lemma 7.10, straightforward computations with sup's and inf's, and the fact that we have established the assertion for $\varphi_k \in \mathscr{S}_{i_k 0}(T)$, $k<n$, we obtain $m^*(\bigwedge_{k<n} \mathsf{Q}_k \mathsf{M}_k) = \prod_{k<n} m^*(\mathsf{Q}_k \mathsf{M}_k)$. We omit the cumbersome details and illustrate the idea with a simple example. By an analogue of Lemma 7.10,

$$\begin{aligned} m^*(\exists v_0 \mathsf{M}_0(v_0) \wedge \forall v_1 \mathsf{M}_1(v_1)) &= \sup_{F_0} \inf_{F_1} m^*[(\bigvee_{t_0 \in F_0} \mathsf{M}_0(t_0)) \wedge (\bigwedge_{t_1 \in F_1} \mathsf{M}_1(t_1))] \\ &= \sup_{F_0} \inf_{F_1} [m^*(\bigvee_{t_0 \in F_0} \mathsf{M}_0(t_0)) \cdot m^*(\bigwedge_{t_1 \in F_1} \mathsf{M}_1(t_1))] \\ &= \sup_{F_0} m^*(\bigvee_{t_0 \in F_0} \mathsf{M}_0(t_0)) \cdot \inf_{F_1} m^*(\bigwedge_{t_1 \in F_1} \mathsf{M}_1(t_1)) \\ &= m^*(\exists v_0 \mathsf{M}_0(v_0)) \cdot m^*(\forall v_1 \mathsf{M}_1(v_1)), \end{aligned}$$

because $\bigvee_{t_0 \in F_0} \mathsf{M}_0(t_0) \in \mathscr{S}_{i_0 0}(T)$, $\bigwedge_{t_1 \in F_1} \mathsf{M}_1(t_1) \in \mathscr{S}_{i_1 0}(T)$, and those are formulas for which the assertion has already been established.

Next assume $\xi > 0$ and that the assertion holds for all ordinals smaller than ξ. First suppose $\varphi_k \in \mathscr{S}_{i_k \xi}(T)$ for all $k<n$. Remember that for every $k<n$, $(\bigcup_{\eta<\xi} \mathscr{S}_{i_k \eta}(T))/\vdash$ is a subalgebra of $\mathscr{S}_{i_k}(T)$ and σ-generates $\mathscr{S}_{i_k \xi}(T)$. Now suppose $\psi_k \in \bigcup_{\eta<\xi} \mathscr{S}_{i_k \eta}(T)$ for all $k<n$. Then for some $\eta<\xi$, $\psi_k \in \mathscr{S}_{i_k \eta}(T)$ for all $k<n$. Thus, by inductive hypothesis, $m^*(\bigwedge_{k<n} \psi_k) = \prod_{k<n} m^*(\psi_k)$. By an n-dimensional version of Lemma 5.2, $m^*(\bigwedge_{k<n} \varphi_k) = \prod_{k<n} m^*(\varphi_k)$. In the general case where $\varphi_k \in \mathscr{S}_{i_k \xi}(T)$ for all $k<n$, we proceed in the familiar fashion using prenex normal forms and Lemma 7.10, as in the second part of the case $\xi = 0$. This completes the proof of Corollary 5.3.

The construction of independent unions is particularly valuable for the introduction into a given probability system of "*a priori* conditions" such as ordinary relational structures. For example, we may consider a probability system $\mathfrak{A}_1 = \langle A, R, \mathscr{A}, m \rangle$, where the set A has a natural ordering $<$. Then we consider the ordinary relational system $\mathfrak{A}_2 = \langle A, < \rangle$ separately, and form the independent union of $\mathfrak{A}_1$ and $\mathfrak{A}_2$. This allows us to make probability statements involving the ordering, and Corollaries 5.1 and 5.3 show how certain values of the probability of the independent union may be computed from the probabilities of the components.

It is natural to ask for the definition of an analogue of the *direct product* of ordinary relational systems; however, a reasonable, natural generalization of this construction for probability systems does not seem to exist. On the other hand we are able to give an intuitively very suggestive definition of an analogue of the ultraproduct construction of relational systems.

Ultraproducts. Consider again our language $\mathscr{L}$ with one binary predicate **R**. Let I be an index set. For each $i \in I$, let T_i be a set of new individual constants, and let $\langle T_i, m_i \rangle$ be a probability system, where m_i is a probability on $\mathscr{S}(T_i)/\vdash$. Let $T = \prod_{i \in I} T_i$ be the Cartesian product of the family of sets $\{T_i : i \in I\}$. For $\varphi \in \mathscr{S}(T)$ and $i \in I$, let $\varphi | i$ be the projection of φ onto the i^{th} coordinate; that is, replace in φ every $t \in T$ by $t_i \in T_i$. Then for every $\varphi \in \mathscr{S}(T)$ and $i \in I$, $\vdash \varphi$ implies $\vdash \varphi | i$. Finally let λ be a probability on the power set of I. Define for all $\varphi \in \mathscr{S}(T)$ a function m by the equation

$$m(\varphi) = \int_I m_i(\varphi \mid i) \, \mathrm{d}\lambda(i).$$

LEMMA 5.4. (i) *For every $\varphi, \psi \in \mathscr{S}(T)$, if $\vdash \varphi \leftrightarrow \psi$, then $m(\varphi) = m(\psi)$.*

(ii) *For every $\varphi \in \mathscr{S}(T)$, if $\vdash \varphi$, then $m(\varphi) = 1$.*

(iii) *m, regarded as a function on $\mathscr{S}(T)/\vdash$, is a probability.*

Proof: (i) and (ii) are trivial. Thus m may indeed be regarded as a function on $\mathscr{S}(T)/\vdash$, and it suffices to prove σ-additivety. Suppose $\varphi_n \in \mathscr{S}(T)$ for all $n < \omega$, and $\vdash \neg (\varphi_m \wedge \varphi_n)$ for all $m \neq n$. Then $\vdash \neg (\varphi_m | i \wedge \varphi_n | i)$ for all $m \neq n$ and all $i \in I$. Thus for all $i \in I$

$$m_i\Big(\bigvee_{n<\omega} \varphi_n \mid i\Big) = \sum_{n<\omega} m_i(\varphi_n \mid i).$$

Therefore by the Dominated Convergence Theorem

$$\begin{aligned} m\Big(\bigvee_{n<\omega} \varphi_n\Big) &= \int_I m_i\Big(\bigvee_{n<\omega} \varphi_n \mid i\Big) \, \mathrm{d}\lambda(i) \\ &= \int_I \sum_{n<\omega} m_i(\varphi_n \mid i) \, \mathrm{d}\lambda(i) \\ &= \sum_{n<\omega} \int_I m_i(\varphi_n \mid i) \, \mathrm{d}\lambda(i) \\ &= \sum_{n<\omega} m(\varphi_n). \end{aligned}$$

We define the *ultraproduct* with respect to λ of the family of probability systems $\{\langle T_i, m_i\rangle : i\in I\}$, to be the probability system $\langle T, m\rangle$, and denote it by $\prod_{i\in I}\langle T_i, m_i\rangle/\lambda$.

COROLLARY 5.5. *For all* $\varphi\in\mathscr{S}(T)$,

$$m^*(\varphi) = \int_I m_i^*(\varphi \mid i)\,\mathrm{d}\lambda(i).$$

Proof: Define $\mu(\varphi)=\int_I m_i^*(\varphi|i)\mathrm{d}\lambda(i)$, for all $\varphi\in\mathscr{S}(T)$. By the same argument as in Lemma 5.4, μ is a probability on $\mathscr{S}(T)/\vdash$. Clearly μ extends m. By Theorem 4.3 it suffices to prove that μ satisfies the condition (G). Let $\exists v\varphi\in\mathscr{S}(T)$. For every $i\in I$,

$$m_i^*(\exists v\varphi \mid i) = \sup_{F\in T_i^{(\omega)}} m_i^*\big(\bigvee_{t\in F}(\varphi \mid i)(t)\big).$$

Therefore for every $i\in I$ and $n<\omega$ there exists $t_{in}\in T_i$ such that

$$m_i^*(\exists v\varphi \mid i) = \lim_{n\to\infty} m_i^*\big(\bigvee_{k<n}(\varphi \mid i)(t_{ik})\big).$$

For $n<\omega$ define $s_n\in T$ by $s_n(i)=t_{in}$ for all $i\in I$. Then for $i\in I$,

$$m_i^*(\exists v\varphi \mid i) = \lim_{n\to\infty} m_i^*\big(\bigvee_{k<n}\varphi(s_k) \mid i\big).$$

Thus, by the Monotone Convergence Theorem,

$$\begin{aligned}\mu(\exists v\varphi) &= \int_I m_i^*(\exists v\varphi \mid i)\mathrm{d}\lambda(i) = \int_I \lim_{n\to\infty} m_i^*\big(\bigvee_{k<n}\varphi(s_k) \mid i\big)\mathrm{d}\lambda(i)\\ &= \lim_{n\to\infty}\int_I m_i^*\big(\bigvee_{k<n}\varphi(s_k) \mid i\big)\mathrm{d}\lambda(i) = \lim_{n\to\infty}\mu\big(\bigvee_{k<n}\varphi(s_k)\big)\\ &= \mu\big(\bigvee_{k<\omega}\varphi(s_k)\big) = \sup_{F\in T^{(\omega)}}\mu\big(\bigvee_{t\in T}\varphi(t)\big).\end{aligned}$$

Remark: Let $t, t'\in T$ and let $J=\{i\in I : t_i=t'\}$. If for every $i\in I$, $\langle T_i, m_i\rangle$ has strict identity, then we have $m(t\!=\!t')=\lambda(J)$. Thus the ultraproduct construction does not preserve strict identity.

Corollary 5.5 shows that the probability of the ultraproduct is in a very suggestive fashion a "weighed average" of the probabilities of the components. This idea was introduced in Łoś [1962] and is further developed in Fenstad (forthcoming), to which the reader is referred. We conclude this section with a brief discussion of symmetric probability systems.

Symmetric probability systems. Let $\langle T, m\rangle$ be a probability system. A function $\pi \in T^T$ is a *permutation* of T if π is one-to-one and onto. A permutation π is *finite* if $\pi(t)=t$ for all but a finite number of $t \in T$. Following a suggestion of Gaifman [1964] we call $\langle T, m \rangle$ *symmetric* if for every finite permutation π of T and every $\varphi \in \mathscr{S}^{(\omega)}(T)$, $m(\varphi)=m(\varphi^\pi)$, where φ^π is obtained from φ by replacing every individual constant t of φ by $\pi(t)$.

Remark: Gaifman's [1964] definition is apparently stronger; however, Lemma 5.7 shows that the two definitions actually coincide.

The following lemma is easy to prove; for a proof we refer the reader to Hewitt-Savage [1955].

LEMMA 5.6. *Let $\langle T, m \rangle$ be a symmetric probability system. Then for every permutation π of T and every $\varphi \in \mathscr{S}(T)$, $m^*(\varphi)=m^*(\varphi^\pi)$.*

LEMMA 5.7. *Let $\langle T, m \rangle$ be a symmetric probability system. Then for every permutation π of T and every $\varphi \in \mathscr{S}(T)$, $m^*(\varphi)=m^*(\varphi^\pi)$.*

Proof: We proceed by transfinite induction using the familiar method of the proof of Theorem 4.3. First, if $\varphi \in \mathscr{S}_0(T)$, then the assertion is Lemma 5.6. If $\varphi \in \mathscr{S}_0(T)$, then φ may be written in prenex normal form QM as explained before. It is again easy to see by an analogue of Lemma 7.9, by the fact that a permutation π of T is a function onto T, and by elementary arithmetical properties of sup's and inf's that we obtain $m^*(\mathsf{QM})=m^*(\mathsf{QM}^\pi)$. We illustrate the argument with a simple example. By an analogue of Lemma 7.9,

$$m^*(\exists v_0 \forall v_1 \mathsf{M}(v_0, v_1)) = \sup_{F_0 \in T^{(\omega)}} \inf_{F_1 \in T^{(\omega)}} m^*\Big(\bigvee_{t_0 \in F_0} \bigwedge_{t_1 \in F_1} \mathsf{M}(t_0, t_1)\Big).$$

Now $\bigvee_{t_0 \in F_0} \bigwedge_{t_1 \in F_1} \mathsf{M}(t_0, t_1) \in \mathscr{S}_0(T)$, and the assertion has already been established for these formulas, so we have for every $F_0, F_1 \in T^{(\omega)}$

$$m^*\Big(\bigvee_{t_0 \in F_0} \bigwedge_{t_1 \in F_1} \mathsf{M}(t_0, t_1)\Big) = m^*\Big(\bigvee_{t_0 \in F_0} \bigwedge_{t_1 \in F_1} \mathsf{M}^\pi(\pi(t_0), \pi(t_1))\Big).$$

Finally, since π is onto,

$$\begin{aligned} m^*(\exists v_0 \forall v_1 \mathsf{M}(v_0, v_1)) &= \sup_{F_0 \in T^{(\omega)}} \inf_{F_1 \in T^{(\omega)}} m^*\Big(\bigvee_{t_0 \in F_0} \bigwedge_{t_1 \in F_1} \mathsf{M}^\pi(t_0, t_1)\Big) \\ &= m^*(\exists v_0 \forall v_1 \mathsf{M}^\pi(v_0, v_1)). \end{aligned}$$

Next, assume $\xi > 0$ and that the assertion holds for all ordinals smaller than ξ. Recall that $\big(\bigcup_{\eta < \xi} \mathscr{S}_\eta(T)\big)/\vdash$ is a subalgebra of $\mathscr{S}(T)/\vdash$ that σ-generates

$\mathscr{S}_\xi(T)/\vdash$. By the inductive hypothesis $m^*(\varphi)=m^*(\varphi^\pi)$ for every $\varphi\in\bigcup_{\eta<\xi}\mathscr{S}_\eta(T)$. Let Σ be the set of $\varphi\in\mathscr{S}_\xi(T)$ for which the assertion holds. If $\varphi_n\in\Sigma$ and $\vdash\varphi_n\rightarrow\varphi_{n+1}$ for all $n<\omega$, then

$$\begin{aligned} m^*\Big(\bigvee_{n<\omega}\varphi_n\Big) &= \lim_{n\to\infty} m^*(\varphi_n) \\ &= \lim_{n\to\infty} m^*(\varphi_n^\pi) \\ &= m^*\Big(\bigvee_{n<\omega}\varphi_n^\pi\Big). \end{aligned}$$

A similar argument for decreasing sequences proves that Σ is monotone, and therefore by a well-known fact about monotone classes (see Halmos [1950], p. 27), $\Sigma=\mathscr{S}_\xi$(T). If $\varphi\in\mathscr{S}_\xi(T)$, we proceed as in the second part of the case of $\xi=0$.

In a sense the symmetric probability systems are diametrically opposite to ordinary relational systems. In ordinary relational systems the probability is as *concentrated* as possible; in symmetric probability systems it is completely *dispersed*. The condition of symmetry is a severe restriction on a probability system, as an example in Section 6 will demonstrate.

This completes our discussion of probability-model-theoretic concepts, and we now turn to the analogue in probability logic of *theories* in ordinary logic.

6. Probability assertions. In ordinary logic a theory of $\mathscr{L}$ is any subset of $\mathscr{S}$ closed under deduction. In probability logic we first have to define the concept of *probability assertions* which play the role of *sentences* (or, better axioms and theorems) in ordinary logic.

For this purpose we introduce a new language $\mathscr{M}$, the first-order language of real algebra. $\mathscr{M}$ has denumerably many distinct individual variables λ_n, $n<\omega$. The non-logical constants of $\mathscr{M}$ are a binary predicate $\leq$, binary function symbols $+$ and $\cdot$, and individual constants $\mathbf{0}$, $+\mathbf{1}$ and $-\mathbf{1}$. The logical constants of $\mathscr{M}$ are $\wedge$, $\vee$, $\neg$, $\forall$ and $\exists$, standing for (finite) conjunction, disjunction, negation, universal and existential quantification respectively. Formulas and sentences of $\mathscr{M}$ are defined as usual. $\mathscr{M}$ is to be interpreted in the real numbers in the standard way with the obvious meaning being given to the symbols. Let Re denote the set of real numbers, and say that $\mathscr{M}$ is interpreted in the relational system $\mathfrak{R}=\langle \text{Re}, \leq, +, \cdot, 0, +1, -1\rangle$.

The set of sentences of $\mathscr{M}$ true in $\mathfrak{R}$ is called the set of *theorems of real algebra*. An *algebraic formula* is a quantifier free formula of $\mathscr{M}$. Every algebraic formula is equivalent in real algebra to a disjunction (conjunction) of conjunctions (disjunctions) of polynomial inequalities of the form $p \geq \mathbf{0}$ or $p > \mathbf{0}$, where p is a polynomial with integral coefficients. We call an algebraic formula *closed* (*open*) if it is equivalent to a disjunction of conjunctions of polynomical inequations of the form $p \geq \mathbf{0} (p > \mathbf{0})$. It is obvious that in this definition we could have used the conjunctive instead of the disjunctive normal form.

We now make several definitions. A *probability assertion* of $\mathscr{L}$ is an $(n+1)$-tuple $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$, where $n < \omega$, Φ is an algebraic formula with exactly n free variables and $\varphi_0, \ldots, \varphi_{n-1} \in \mathscr{S}$. A probability assertion is called *closed* (*open*) if the algebraic formula is closed (open). A probability system $\langle T, m \rangle$ is a *probability model* of $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$ if the n-tuple of real numbers $\langle \bar{m}(\varphi_0), \ldots, \bar{m}(\varphi_{n-1}) \rangle$ satisfies Φ in $\mathfrak{R}$.

If Σ is a set of probability assertions and Ψ is a probability assertion, then Ψ is a *probability consequence* of Σ iff every probability model of all assertions in Σ is also a probability model of Ψ. Ψ is a *probability law* of $\mathscr{L}$ if Ψ is a probability consequence of the empty set of assertions.

Immediately the familiar questions arise: Is there a method of deductively generating the probability consequences from a given set of probability assertions? Is there a method of deductively generating all probability laws? Under which conditions does a set of probability assertions have a probability model? Is there an analogue of the concept of consistency in ordinary logic? Obviously these questions are interrelated. Before we enter into their discussion, we insert some remarks concerning our definition of probability assertions.

The definition of probability assertions depends both on the language $\mathscr{L}$ and the language $\mathscr{M}$, and it is clear how this definition could be generalized by considering languages $\mathscr{M}'$ with stronger means of expression. The underlying idea of our approach is that we want to investigate polynomial inequalities in "variables" $\mu(\varphi_0), \ldots, \mu(\varphi_{n-1})$ with real coefficients, where $\varphi_0, \ldots, \varphi_{n-1} \in \mathscr{S}$ and μ is interpreted as a probability on $\mathscr{S}/\vdash$. Our definition does not quite realize this idea. For this purpose it would have been appropriate to introduce a language $\mathscr{M}'$ which is like $\mathscr{M}$ but has individual constants for *every* real number. We easily see that by the continuity of addition and multiplication every *closed* probability assertion of $\mathscr{M}'$ is equivalent to a denumerable set of probability assertions of $\mathscr{M}'$ with rational coefficients and, after clearing denominators, to a denumerable set of probability assertions of $\mathcal{M}$. This, however, is not true for *open* probability assertions of $\mathcal{M}'$. We thus fall somewhat short of our objectives. Nevertheless since not very much work has yet been done towards the investigation of probability assertions, we chose the present formulation for its simplicity.

If $\Sigma \subseteq \mathcal{S}$, then Σ determines a set of probability assertions $\{\langle \lambda_0 - \mathbf{1} \geq \mathbf{0}, \varphi\rangle : \varphi \in \Sigma\}$. Accordingly we say that $\langle T, m\rangle$ is a *probability model* of Σ if $\bar{m}(\varphi)=1$ for all $\varphi\in\Sigma$. More generally, if μ is a probability on $\mathcal{S}/\vdash$, then μ determines a set Σ of probability assertions as follows. For every $\varphi\in\mathcal{S}$ we choose sequences of rational numbers p_n/q_n, and p_n'/q_n', such that for all $n<\omega$

$$\frac{p_n}{q_n} \leq \frac{p_{n+1}}{q_{n+1}} \leq \mu(\varphi) \leq \frac{p'_{n+1}}{q'_{n+1}} \leq \frac{p'_n}{q'_n}$$

and

$$\lim_{n\to\infty} \frac{p_n}{q_n} = \mu(\varphi) = \lim_{n\to\infty} \frac{p'_n}{q'_n}.$$

Let

$$\Sigma_\varphi = \{\langle \mathbf{q}_n\lambda_0 - \mathbf{p}_n \geq \mathbf{0}, \varphi\rangle : n < \omega\} \cup \{\langle -\mathbf{q}'_n\lambda_0 + \mathbf{p}'_n \geq \mathbf{0}, \varphi\rangle : n < \omega\}$$

and let $\Sigma = \bigcup_{\varphi\in\mathcal{S}} \Sigma_\varphi$. Accordingly we say that $\langle T, m\rangle$ is a *probability model of* μ if $\bar{m}(\varphi)=\mu(\varphi)$ for all $\varphi\in\mathcal{S}$. It then is obvious that $\langle T, m\rangle$ is a probability model of μ iff $\langle T, m\rangle$ is a probability model of Σ. Moreover, if $\Delta\subseteq\mathcal{S}$ is a complete and consistent theory, then Δ uniquely determines a two valued σ-additive probability measure μ on $\mathcal{S}/\vdash$. In this case $\langle T, m\rangle$ is a probability model of Δ iff $\langle T, m\rangle$ is a probability model of μ.

For the infinitary language $\mathcal{L}$ the questions raised above seem to be rather vexing problems, only a few scanty results could be established by the authors. First let us establish the relationship to ordinary logic.

A set $\Sigma\subseteq\mathcal{S}$ is *consistent* if there exists no $\varphi\in\mathcal{S}$ such that $\Sigma\vdash[\varphi\wedge\neg\varphi]$. It is well known that not every consistent set of sentences of $\mathcal{L}$ has a model (see Karp [1964] p. 32). The trouble already occurs with the ordinary propositional logic of $\mathcal{L}$ and carries over to the probability logic of $\mathcal{L}$. Indeed, let $p_{\xi n}$, $\xi<\omega_1$, $n<\omega$, be a doubly indexed set of propositional constants. Let

$$\Sigma = \{\bigvee_{n<\omega} p_{\xi n} : \xi < \omega_1\} \cup \{\neg [p_{\xi n} \wedge p_{\xi' n}] : \xi < \xi' < \omega_1, n < \omega\}.$$

It is easy to see that Σ is consistent. Suppose then there exists a σ-additive probability measure μ on $\mathcal{S}/\vdash$ such that $\mu(\varphi)=1$ for all $\varphi\in\Sigma$. Then for every $\xi<\omega_1$ there exists $n<\omega$ such that $\mu(p_{\xi n})>0$. Thus there are uncountably many p's such that $\mu(p)>0$. Thus for some $n<\omega$ there are uncountably many

$\xi<\omega_1$ such that $\mu(p_{\xi n})>0$. Since $\mu([p_{\xi n}\wedge p_{\xi' n}])=0$ for $\xi\neq\xi'$, this is a contradiction. Obviously every complete and consistent set of sentences of an infinitary *propositional* language has a model. In infinitary propositional logic the trouble therefore arises from the fact that the Prime Ideal Theorem fails for Boolean σ-algebras.

Naturally the question arises: Does every complete and consistent set $\Sigma\subseteq\mathscr{S}$ have a model? The answer is again no, and a counter-example is due to Professor C. Ryll-Nardzewski. Interestingly enough the counter-example produces a probability model of the complete consistent set of sentences under consideration. The question of whether every complete consistent set $\Sigma\subseteq\mathscr{S}$ has a probability model can, however, be settled by a similar counter-example, and we shall discuss both of these examples in a form slightly modified from Ryll-Nardzewski's original suggestion.

Let $\mathscr{L}$ be an infinitary language with countably many one-place predicates $\mathbf{P}_j$ for each $j<\omega$, and define a probability model $\mathfrak{A}=\langle A, R_j, \mathscr{A}, m\rangle_{j<\omega}$ as follows: Let $A=\omega$, and let $\mathscr{A}$ be the Borel sets of the product space $(2^\omega)^\omega$; that is, the σ-field of subsets of $(2^\omega)^\omega$ generated by all sets of the form $\{\xi\in(2^\omega)^\omega:\xi(i)(j)=1\}$, where $i,j<\omega$. Let m be the product measure on $\mathscr{A}$ determined by $m(\{\xi\in(2^\omega)^\omega:\xi(i)(j)=1\})=\frac{1}{2}$ for all $i,j<\omega$. Finally, for $j<\omega$, define $R_j(i)=\{\xi\in(2^\omega)^\omega:\xi(i)(j)=1\}$ for all $i\in A$. (Note: strictly speaking $\mathfrak{A}$ is not a probability model since $\langle\mathscr{A}, m\rangle$ is not a measure algebra. Thus we would have to consider the quotient algebra $\mathscr{A}/I$ of $\mathscr{A}$ modulo the σ-ideal $I=\{x\in\mathscr{A}:m(x)=0\}$, and lift m up to a strictly positive probability on $\mathscr{A}/I$. In this example, however, all sup's and inf's in $\mathscr{A}$ that have to be taken into consideration are countable; clauses (vi) and (vii) of the definition of the valuation function h make sense; and everything comes out just the same. We can omit the tedious details.) Then let $T_{\mathfrak{A}}=\{t_i:j\in A\}$ be a set of new individual constants such that $t_i\neq t_{i'}$, if $i\neq i'$. Now we observe that for every $\varphi\in S$, the element $h(\varphi)\in\mathscr{A}$ is invariant under all finite permutations of the second coordinate in $(2^\omega)^\omega$. By the well-known 0–1 Law (Hewitt and Savage [1955] p. 496) m is two-valued on $h(\varphi)$. Thus the set $\Sigma=\{\varphi\in\mathscr{S}:m(h(\varphi))=1\}$ is a complete and consistent theory of $\mathscr{L}$. We wish to show that Σ has no model. Indeed, suppose $\mathfrak{B}=\langle B, S_j\rangle_{j<\omega}$ is a model of Σ. Since B must be non-empty, let $b\in B$. For $j<\omega$ define formulas $Q_j(v)=\mathbf{P}_j(v)$, if $b\in S_j$; while $Q_j(v)=\neg\mathbf{P}_j(v)$, if $b\notin S_j$. Then $\exists v[\bigwedge_{j<\omega} Q_j(v)]$ holds in $\mathfrak{B}$. However, as a straightforward computation shows, $m(\exists v[\bigwedge_{j<\omega} Q_j(v)])=0$, which is a contradiction. On the other hand, by its very construction $\mathfrak{A}$ is a probability model of Σ; that is, $\mu_{\mathfrak{A}}(\varphi)=1$ for all $\varphi\in\Sigma$.

For our second example we let $\mathfrak{A}'=\langle A, R_j', \mathscr{A}'\rangle_{i<\omega}$ be that Boolean-algebraic model where $A=\omega$, where $\mathscr{A}'=\mathscr{A}/J$, the ideal J being the σ-ideal of all first-category sets in the Borel algebra $\mathscr{A}$, and where $R_j'(i)=R_j(i)/J$ for all $i,j<\omega$. Since A is countable, we note that the valuation h' for $\mathfrak{A}'$ is such that $h'(\varphi)=h(\varphi)/J$ for all $\varphi\in\mathscr{S}$. But there is a 0–1 Law for category just as there is for measure, hence $\Sigma'=\{\varphi\in\mathscr{S}:h'(\varphi)=\mathbf{1}_{\mathfrak{A}'}\}$ is also a complete and consistent theory of $\mathscr{L}$. We wish to show that Σ' has no probability model. First we note that for every Borel set a of 2^{ω} we can find a formula $\varphi_a(v)$ such that $h'(\varphi_a(t_i))=\{\xi\in(2^{\omega})^{\omega}:\xi(i)\in a\}/J$. Therefore the sentence $\forall v[\varphi_a(v)\leftrightarrow\varphi_b(v)]$ belongs to the set Σ' iff $a/J_0=b/J_0$ where J_0 is the ideal of first-category sets of 2^{ω}. This means that if Σ' had a probability model, then the measure algebra of this model would contain a σ-homomorphic image of the algebra of Borel sets modulo first-category sets. But we know that there is *no* non trivial σ-additive probability on this algebra (see Sikorski [1964] p. 77). Thus Σ' has no probability model. It should be noted that both the models $\mathfrak{A}$ and $\mathfrak{A}'$ could be equipped with a strict identity relation, and they thus afford counter-examples for the logic with strict identity.

The property of a set $\Sigma\subseteq\mathscr{S}$ to have a probability model may be given an algebraic interpretation which can be read off directly from our presentation. We say that a Boolean algebra $\mathscr{A}$ has the *Kelley property* if $\mathscr{A}\sim\{\mathbf{0}\}$ is a countable union of sets with positive intersection number. (For the definition of this and other Boolean-algebraic concepts see Sikorski [1964]; in particular, cf. p. 204.) We then have the following:

COROLLARY 6.1. *Let $\Sigma\subseteq\mathscr{S}$. Then Σ has a probability model iff Σ has a Boolean-algebraic model which is complete, weakly distributive and has the Kelley property.*

Proof: This follows immediately from Kelley's Theorem: A complete Boolean algebra has a strictly positive σ-additive probability measure iff it is weakly distributive and has the Kelley property (see Kelley [1959]).

In Karp [1964] we find the theorem that every *countable* consistent set $\Sigma\subseteq\mathscr{S}$ has a countable model in the ordinary sense. The exact analogue of that result also holds for probability logic. We can also treat the case of theories with identity. To help formulate the result we define the formula θ_n for $0<n<\omega$ to be the formula $\exists v_0 \ldots \exists v_{n-1}\forall v_n \bigvee_{i<n} v_i=v_n$. Note that for each probability system $\langle T, m\rangle$ with strict identity we have $\bar{m}(\theta_n)\in\{0, 1\}$ for $0<n<\omega$.

THEOREM 6.2. (i) *Let μ be a probability on $\mathscr{S}/\vdash$, and let $\Sigma \subseteq \mathscr{S}$ be a countable set. Then there exists a countable probability model $\langle T, m\rangle$ such that for every $\varphi \in \Sigma$, $\bar{m}(\varphi) = \mu(\varphi)$.*

(ii) *If for every $0<n<\omega$, $\mu(\theta_n) \in \{0, 1\}$, then the probability model $\langle T, m\rangle$ may be assumed to have strict identity.*

We will give a detailed proof of part (ii) and leave the proof of part (i) to the reader. To prove this result we require a few lemmas. The first is a measure theoretic generalization of the well-known Rasiowa-Sikorski Lemma (see Rasiowa and Sikorski [1950]). The proof is given in full in the Appendix.

LEMMA 6.3. *Let $\mathscr{B}$ be a Boolean σ-algebra and let $\mathscr{A} \subseteq \mathscr{B}$ be a σ-subalgebra. Let μ be a probability on $\mathscr{A}$, and for every $m, n < \omega$ let $b_{mn} \in \mathscr{B}$. Then there exists a finitely additive probability ν on $\mathscr{B}$ such that*

(i) $\nu(x) = \mu(x)$, *for all* $x \in \mathscr{A}$;

(ii) $\nu(\bigwedge_{n<\omega} b_{mn}) = \lim_{n\to\infty} \nu(\bigwedge_{i<n} b_{mi})$, *for all* $m < \omega$.

For every $\varphi \in \mathscr{S}(T)$ we recursively define an ordinal number $\lambda(\varphi) < \omega_1$, called the *length* of φ, by these equations:

(i) *if φ is atomic*, $\lambda(\varphi) = 1$;

(ii) $\lambda(\neg\varphi) = \lambda(\varphi) + 1$;

(iii) $\lambda(\varphi_1 \vee \varphi_2) = \lambda(\varphi_1 \wedge \varphi_2) = \lambda(\varphi_1) + \lambda(\varphi_2) + 1$;

(iv) $\lambda(\bigvee_{n<\omega} \varphi_n) = \lambda(\bigwedge_{n<\omega} \varphi_n) = \sum_{n<\omega} \lambda(\varphi_n)$.

LEMMA 6.4. *If $\varphi \in \mathscr{S}(T)$, $\lambda(\varphi) \geq \omega$, and $\neg$ occurs in φ only in front of atomic formulas, then there exists a sequence $\psi_i \in \mathscr{S}(T)$ such that $\lambda(\psi_i) < \lambda(\varphi)$ for all $i < \omega$, and either $\vdash \varphi \leftrightarrow \bigvee_{i<\omega} \psi_i$ or $\vdash \varphi \leftrightarrow \bigwedge_{i<\omega} \psi_i$.*

Proof: By transfinite induction on $\lambda(\varphi)$. If $\lambda(\varphi) < \omega$, the assertion holds trivially. Thus assume $\lambda(\varphi) \geq \omega$ and that the assertion holds for all φ' such that $\lambda(\varphi') < \lambda(\varphi)$. If $\varphi = \bigvee_{n<\omega} \varphi_n$ or $\varphi = \bigwedge_{n<\omega} \varphi_n$, the assertion is again trivial. Thus for some $n<\omega$, either $\varphi = \bigvee_{i<n} \varphi_i$ or $\varphi = \bigwedge_{i<n} \varphi_i$. Consider the first case. Since $\lambda(\varphi) \geq \omega$, we see that $\lambda(\varphi_i) \geq \omega$ for some $i<n$. Let m be the largest such integer $i<n$. Then $\lambda(\varphi_m) < \lambda(\varphi)$ and $\lambda(\varphi_i) < \omega$ for all $m<i<n$. By inductive hypothesis there exists a sequence $\psi_j \in \mathscr{S}(T)$ such that $\lambda(\psi_j) < \lambda(\varphi_m)$ for all

$j<\omega$, and either $\vdash\varphi_m\leftrightarrow \bigvee_{j<\omega} \psi_j$ or $\vdash\varphi_m\leftrightarrow \bigwedge_{j<\omega} \psi_j$. Again consider the first case. Then

$$\vdash \varphi \leftrightarrow \bigvee_{j<\omega} \Big[\bigvee_{i<m} \varphi_i \vee \psi_j \vee \bigvee_{m<i<n} \varphi_i\Big].$$

It now follows from well-known laws of ordinal addition that for every $j<\omega$,

$$\begin{aligned}\lambda\Big(\bigvee_{i<m} \varphi_i \vee \psi_j \vee \bigvee_{m<i<n} \varphi_i\Big) &= \lambda\Big(\bigvee_{i<m} \varphi_i\Big) + \lambda(\psi_j) + \lambda\Big(\bigvee_{m<i<n} \varphi_i\Big)\\ &< \lambda\Big(\bigvee_{i<m} \varphi_i\Big) + \lambda(\varphi_m) + \lambda\Big(\bigvee_{m<i<n} \varphi_i\Big)\\ &= \lambda(\varphi),\end{aligned}$$

which proves the assertion. All other cases are treated analogously.

If μ is a finitely additive probability on $\mathscr{S}/\vdash$ and $\Sigma\subseteq\mathscr{S}$ is a set of infinite conjunctions and disjunctions, we say μ *preserves* Σ if for every $\bigwedge_{n<\omega} \varphi_n, \bigvee_{n<\omega} \psi_n \in\Sigma$,

$$\mu\Big(\bigwedge_{n<\omega} \varphi_n\Big) = \lim_{n\to\infty} \mu\Big(\bigwedge_{i<n} \varphi_i\Big)$$

and

$$\mu\Big(\bigvee_{n<\omega} \psi_n\Big) = \lim_{n\to\infty} \mu\Big(\bigvee_{i<n} \psi_i\Big).$$

LEMMA 6.5. *For every $\varphi\in\mathscr{s}(T)$ there exists a denumerable set $\Sigma\subseteq\mathscr{s}(T)$ of infinite conjunctions and disjunctions such that for all finitely additive probabilities m_1 and m_2 on $\mathscr{s}(T)/\vdash$, if they agree on the finitary sentences and preserve Σ, then $m_1(\varphi)=m_2(\varphi)$.*

Proof: By transfinite induction on $\lambda(\varphi)$. If $\lambda(\varphi)<\omega$ then φ is finitary, thus by hypothesis we may take Σ to be empty. Thus assume $\lambda(\varphi)\geq\omega$ and the lemma holds for all ψ such that $\lambda(\psi)<\lambda(\varphi)$. We also may assume that $\neg$ occurs in φ only in front of atomic formulas. By Lemma 6.4 there exists a sequence $\psi_i\in\mathscr{s}(T)$ such that $\lambda(\psi_i)<\lambda(\varphi)$ for all $i<\omega$, and either $\vdash\varphi\leftrightarrow \bigvee_{i<\omega} \psi_i$ or $\vdash\varphi\leftrightarrow \bigwedge_{i<\omega} \psi_i$. Consider the first case. Then for every $n<\omega$, we note that $\lambda(\bigvee_{i<n} \psi_i)<\lambda(\varphi)$. By the inductive hypothesis, for every $n<\omega$ there exists a denumerable set Σ_n such that for all finitely additive probabilities m_1, m_2, if they agree on the finitary sentences and preserve Σ_n, then $m_1(\bigvee_{i<n} \psi_i)=m_2(\bigvee_{i<n} \psi_i)$. Let $\Sigma=\bigcup_{n<\omega} \Sigma_n \cup \{\bigvee_{i<\omega} \psi_i\}$. Then clearly $m_1(\varphi)=m_2(\varphi)$ for all finitely additive probabilities m_1, m_2 which agree on the finitary sentences and preserve Σ. The other case is completely analogous.

Now we begin with the proof of part (ii) of Theorem 6.2. Let μ be a probability on $\mathscr{S}/\vdash$ such that for every $0<n<\omega$, $\mu(\theta_n)\in\{0, 1\}$, and let $\Sigma\subseteq\mathscr{S}$ be countable.

We first consider the case $\mu(\neg\theta_n)=1$ for all $0<n<\omega$. Let $T=\{t_i:i<\omega\}$ be a set of new individual constants such that $t_i\neq t_j$ if $i\neq j$. Let $\bar{\Sigma}\subseteq\mathscr{S}(T)$ be the set of all sentences which may be obtained from subformulas of sentences in Σ by substituting individual constants from T for free variables. (Note that every subformula of a sentence of $\mathscr{L}$ has only *finitely* many free variables, thus every sentence in $\bar{\Sigma}$ contains only finitely many individual constants from T.) $\bar{\Sigma}$ is denumerable. Let $\exists v_{i_n}\varphi_n(v_{i_n})$, $n<\omega$, be an enumeration of all existential sentences in $\bar{\Sigma}$. Choose a sequence $\sigma\in\omega^\omega$ such that

(i) $\sigma(n)+1<\sigma(n+1)$, *for all* $n<\omega$;

(ii) *every individual constant in* φ_n *has index* $<\sigma(n)$.

Consider the set of sentences

$$\Gamma = \{t_i \neq t_j : i < j < \omega\} \cup \{\exists v_{i_n}\varphi_n(v_{i_n}) \rightarrow \bigvee_{i\leq\sigma(n)} \varphi_n(t_i) : n < \omega\}.$$

The following lemma is essentially due to Ehrenfeucht and Mostowski [1961].

LEMMA 6.6. *If* $\mathfrak{A}=\langle A, R\rangle$ *is a denumerably infinite model then there exists a sequence* $a\in A^\omega$ *such that* $\langle A, R, a_n\rangle_{n<\omega}$ *is a model of* Γ *and* $A=\{a_n:n<\omega\}$.

Proof: Assume A is well-ordered in type ω. Define a_m for $m<\omega$ by recursion:

(1) *If for some* $n<\omega$, $m=\sigma(n)$ *and if there exists an* $x\in A\sim\{a_i:i<m\}$ *that satisfies* $\varphi_n(v_{i_n})$ *in* $\langle A, R, a_0,\ldots, a_{m-1}\rangle$, *then let* a_m *be the first such* x.

(2) *Otherwise, let* a_m *be the first element of* $A\sim\{a_i:i<m\}$. Then it is easy to check that $\langle A, R, a_n\rangle_{n<\omega}$ is a model of Γ. We will have $A=\{a_n:n<\omega\}$ because case (2) occurs infinitely often by the definition of σ, which means that A will indeed be exhausted.

Let $\Theta=\{\neg\theta_n:n<\omega\}$, let $\Delta=\Gamma\cup\Theta$, and define a mapping f from $\mathscr{S}/\Theta\vdash$ into $\mathscr{S}(T)/\Delta\vdash$ by $f(\varphi/\Theta\vdash)=\varphi/\Delta\vdash$. Since $\Theta\subseteq\Delta$, the mapping is well defined. Clearly f is a σ-homomorphism. Moreover, f is an isomorphism into. Indeed, let $\varphi\in\mathscr{S}$ and suppose $\Delta\vdash\neg\varphi$. If $\Theta\cup\{\varphi\}$ has a model then, by the Löwenheim-Skolem Theorem, it has a denumerably infinite model. Thus by Lemma 6.6, $\Delta\cup\{\varphi\}$ has a model, contrary to the assumption. Thus, $\Theta\cup\{\varphi\}$ has no model; therefore, by the "weak" Completeness Theorom, $\Theta\vdash\neg\varphi$, which proves the assertion. Now let g be the canonical σ-homomorphism of $\mathscr{S}(T)/\vdash$ onto $\mathscr{S}(T)/\Delta\vdash$; that is, $g(\varphi/\vdash) = \varphi/\Delta\vdash$. For orientaiton we draw a diagram

$$\mathscr{S}/\vdash \to \mathscr{S}/\Theta\vdash \xrightarrow{f} \mathscr{S}(T)/\Delta\vdash \xleftarrow{g} \mathscr{S}(T)/\vdash .$$

As is well-known, since $\mu(\neg\theta_n)=1$ for all $0<n<\omega$, μ may be lifted to $\mathscr{S}/\Theta\vdash$ and subsequently transferred to $\{\varphi/\vdash:\varphi\in\mathscr{S}(T), (f^{-1}\bigcirc g)(\varphi/\vdash)\in\mathscr{S}/\Theta\vdash\}$, which is a σ-subalgebra of $\mathscr{S}(T)/\vdash$. For every $\varphi\in\Sigma$, let $\varphi'\in\mathscr{s}(T)$ be obtained from φ by *eliminating quantifiers*; that is, by successively replacing all existential subformulas $\exists v\psi(v)$ by $\bigvee_{i<\omega}\psi(t_i)$. Then $\Gamma\vdash\varphi\leftrightarrow\varphi'$. Since $(f^{-1}\bigcirc g)(\varphi/\vdash)=(f^{-1}\bigcirc g)(\varphi'/\vdash)$, we have $\mu(\varphi)=\mu(\varphi')$. By Lemma 6.5, for each φ' there exists a denumerable set $\Sigma'\subseteq\mathscr{s}(T)$ of infinite conjunctions and disjunctions such that for all probability measures m_1 and m_2 on $\mathscr{s}(T)/\vdash$ if they agree on the finitary sentences and preserve Σ', then $m_1(\varphi')=m_2(\varphi')$. Finally, let $\bar{\Sigma}'$ be the union of all these sets Σ'. $\bar{\Sigma}'$ is denumerable. By Lemma 6.3, there exists a finitely additive probability v on $\mathscr{S}(T)/\vdash$ which extends μ and preserves $\bar{\Sigma}'$. Let n be the restriction of v to the finitary quantifier-free sentences. Then n is σ-additive. (This is a consequence of Lemma 7.1 which will be established in Section 7.) As is well-known, n may be extended to a probability m on $\mathscr{s}(T)/\vdash$. We claim that $\langle T, m\rangle$ is the desired probability model; that is $\bar{m}(\varphi)=\mu(\varphi)$ for all $\varphi\in\Sigma$. First observe that $m(t_i=t_j)=n(t_i=t_j)=v(t_i=t_j)=\mu(t_i=t_j)=0$ for all $i<j<\omega$. Thus $\langle T, m\rangle$ is indeed a probability model with strict identity. Now let $\varphi\in\Sigma$. Clearly m^* and v agree on the finitary quantifier-free sentences. m^* is σ-additive and therefore preserves $\bar{\Sigma}'$, and v preserves $\bar{\Sigma}'$ by construction. By the definition of $\bar{\Sigma}'$ and Lemma 6.5, $v(\varphi')=m^*(\varphi')$. However $\mu(\varphi)=\mu(\varphi')=v(\varphi')$, as noted above and by the construction of v. Finally we see $m^*(\varphi')=\bar{m}(\varphi)$, by the construction of φ' and condition (G). Thus $\bar{m}(\varphi)=\mu(\varphi)$, which completes the proof in case $\mu(\neg\theta_n)=1$ for all $0<n<\omega$.

In the other case, assume that $N<\omega$ and $\mu(\theta_N)=1$ and that N is the smallest such. Then $\mu(\neg\theta_n)=1$ for all $n<N$ and $\mu(\theta_n)=1$ for all $n\geq N$. Let $T=\{t_i:i<N\}$ be a set of new individual constants such that $t_i\neq t_j$ if $i\neq j$. Let

$$\begin{aligned}\Gamma &= \{t_i \neq t_j : i<j<N\},\\ \Theta &= \{\neg\,\theta_n : n<N\} \cup \{\theta_n : N<n<\omega\},\\ \Delta &= \Gamma \cup \Theta .\end{aligned}$$

As before, define a mapping f from $\mathscr{S}/\Theta\vdash$ into $\mathscr{S}(T)/\Delta\vdash$ such that $f(\varphi/\Theta\vdash)=\varphi/\Delta\vdash$. Again f is an isomorphism into. Indeed, let $\varphi\in\mathscr{S}$ and suppose $\Delta\vdash\neg\varphi$. Then

$$\Theta\vdash(\bigwedge_{i<j<N} t_i\neq t_j)\to\neg\,\varphi .$$

Thus

$$\Theta \vdash \exists v_0 \ldots \exists v_{n-1} (\bigwedge_{i<j<N} v_i \neq v_j) \rightarrow \neg \varphi .$$

However,

$$\Theta \vdash \exists v_0 \ldots \exists v_{n-1} (\bigwedge_{i<j<N} v_i \neq v_j).$$

Thus $\Theta \vdash \neg \varphi$. Finally we observe that if φ' is obtained from $\varphi \in \Sigma$ by *eliminating quantifiers*; that is, by successively replacing all existential subformulas $\exists v \psi(v)$ by $\bigvee_{i<N} \psi(t_i)$, then $\Gamma \vdash \varphi \leftrightarrow \varphi'$. Thus we can complete the proof as before.

We conclude the proof of Theorem 6.2 with a remark concerning the proof of part (i). In this case we need only replace the refined method of the Ehrenfeucht-Mostowski Theorem (Lemma 6.6) by the somewhat cruder method of the original Henkin Completeness Proof (Henkin [1949]). The steps are quite similar though simpler than those just given. In this case, however, we clearly cannot expect the probability model $\langle T, m \rangle$ to have strict identity. This completes the proof of Theorem 6.2.

Remark: Gaifman [1964] gives a proof of Theorem 6.2 for the finitary language $\mathscr{L}^{(\omega)}$. Although this is not an immediate consequence of Theorem 6.2, we will utilize the main ideas of our proof to obtain the result in the finitary case almost immediately (see Theorem 7.3).

We now give an example to show that there are probabilities on $\mathscr{S}/\vdash$ which have a probability model but do not have a symmetric probability model. Indeed, for every $\varphi \in \mathscr{S}$ let $\mu(\varphi)=1$ iff φ holds in $\langle \omega_1, < \rangle$, the system of the countable ordinals with their natural ordering. Every $\xi < \omega_1$ is definable in $\mathscr{L}$; that is, there exists a formula φ_ξ of $\mathscr{L}$ with exactly one free variable v such that for every $\eta < \omega_1$, η satisfies φ_ξ in $\langle \omega_1, < \rangle$ iff $\eta = \xi$. Thus whenever $\xi < \omega_1$, we have $\mu(\exists v \varphi_\xi)=1$; and whenever $\eta < \xi < \omega_1$, we have $\mu(\exists v[\varphi_\eta \wedge \varphi_\xi])=0$. Now suppose μ has a symmetric probability model $\langle T, m \rangle$. Since $\bar{m}(\exists v \varphi_\xi)=1$ for every $\xi < \omega_1$, we obtain for every $\xi < \omega_1$ some $t \in T$ such that $m^*(\varphi_\xi(t))>0$. Since $\langle T, m \rangle$ is symmetric, by Lemma 5.7, we have for every $t' \in T$, $m^*(\varphi_\xi(t'))=m^*(\varphi_\xi(t))$. Thus for every $\xi < \omega_1$ and every $t \in T$, we find $m^*(\varphi_\xi(t))>0$. Now consider a fixed $t \in T$. Then for some $\varepsilon > 0$, $m^*(\varphi_\xi(t)) \geq \varepsilon$ for infinitely many $\xi < \omega_1$. However, since $\bar{m}(\exists v[\varphi_\eta \wedge \varphi_\xi])=0$ whenever $\eta < \xi < \omega_1$, we find that $m^*(\varphi_\eta(t) \wedge \varphi_\xi(t))=0$ whenever $\eta < \xi < \omega_1$, which cannot be the case.

There is a positive result concerning symmetric probability models for the finitary language $\mathscr{L}^{(\omega)}$ which is due to Gaifman [1964]. We will give a simplified version of his proof in Section 7, Theorem 7.14.

The questions of whether there is a method of deductively generating the probability consequences from a given set of probability assertions, or of deductively generating all probability laws are clouded by the fact that it is difficult to give the concept of deductively generating a workable meaning for infinitary languages. Nevertheless we can make a few remarks. Our discussion thus far certainly shows that we have not much reason to expect a positive answer to the first question. On the other hand the second problem has in a sense a positive solution which we will present now.

THEOREM 6.7. *Let* $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$ *be a probability assertion of* $\mathscr{L}$ *such that the free variables of* Φ *are* $\lambda_0, \ldots \lambda_{n-1}$; *further* $\vdash \neg (\varphi_i \wedge \varphi_j)$ *if* $i \neq j$, *and* $\vdash \bigvee_{i<n} \varphi_i$. *Let* $I = \{i<n : \vdash \neg \varphi_i\}$. *Then* $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$ *is a probability law of* $\mathscr{L}$ *iff the sentence*

$$\forall \lambda_0 \ldots \forall \lambda_{n-1} \left[\left[\bigwedge_{i \in I} \lambda_i = \mathbf{0} \wedge \bigwedge_{i<n} \lambda_i \geq \mathbf{0} \wedge \lambda_0 + \cdots + \lambda_{n-1} = \mathbf{1}\right] \rightarrow \Phi\right]$$

is a theorem of real algebra.

Proof: Suppose $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$ is a probability law of $\mathscr{L}$. Consider any sequence of real numbers $\langle x_0, \ldots, x_{n-1} \rangle$ such that $x_i = 0$ for all $i \in I$, $x_i \geq 0$ for all $i<n$, and $x_0 + \cdots + x_{n-1} = 1$. If $i \notin I$, then φ_i is consistent. By the Completeness Theorem for Sentences, there exists a two-valued probability measure μ_i on $\mathscr{S}/\vdash$ such that $\mu_i(\varphi_i) = 1$. Define $\mu(\psi) = \sum_{i \notin I} x_i \cdot \mu_i(\varphi_i \wedge \psi)$ for all $\psi \in \mathscr{S}$. It is easy to see that μ is a σ-additive probability measure on $\mathscr{S}/\vdash$ and $\mu(\varphi_i) = x_i$ for all $i<n$. By Theorem 6.2, there exists a probability model $\langle T, m \rangle$ such that $\bar{m}(\varphi_i) = x_i$ for all $i<n$. Since $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$ is a probability law, $\langle x_0, \ldots, x_{n-1} \rangle$ satisfies Φ in $\mathfrak{R}$. Thus the required sentence is a theorem of real algebra. The converse of Theorem 6.7 is trivial.

For any probability assertion $\langle \Psi, \psi_0, \ldots, \psi_{m-1} \rangle$ of $\mathscr{L}$ we can effectively find an equivalent probability assertion $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$ of $\mathscr{L}$ which satisfies the hypothesis of Theorem 6.7; they are equivalent in the sense that any probability model $\mathfrak{A}$ is a probability model of $\langle \Phi, \psi_0, \ldots, \psi_{m-1} \rangle$ iff it is a probability model of $\langle \Phi, \varphi_0, \ldots, \varphi_{n-1} \rangle$. Moreover, we can effectively generate all probability assertions of $\mathscr{L}$ which satisfy the hypothesis of Theorem 6.7. By a famous result of Tarski the theorems of real algebra are decidable. Thus for each such probability assertion of $\mathscr{L}$ we can decide whether the corresponding sentence is a theorem of real algebra or not. Theorem 6.7

therefore yields a method of generating all probability laws of $\mathscr{L}$. Whether there is a more useful way of generating the laws remains to be seen.

Theorem 6.7 also provides suggestions for the definition of an analogue of the concept of consistency in ordinary logic. The general problem of conditions under which a set of probability assertions has a probability model is completely open for the infinitary language $\mathscr{L}$, and it is apparently quite difficult. This completes our general discussion of the probability logic of the infinitary language $\mathscr{L}$, and we now turn to the finitary language $\mathscr{L}^{(\omega)}$ where the situation is somewhat more satisfactory.

7. The finitary case. The Boolean algebras $\mathscr{S}^{(\omega)}/\vdash$, $\mathscr{S}^{(\omega)}(T)/\vdash$ and $\mathscr{J}^{(\omega)}(T)/\vdash$ are subalgebras of the σ-algebras $\mathscr{S}/\vdash$, $\mathscr{S}(T)/\vdash$ and $\mathscr{J}(T)/\vdash$ respectively. Our definitions and results concerning the infinitary language $\mathscr{L}$ therefore have rather obvious applications to the finitary language $\mathscr{L}^{(\omega)}$, and in many cases they can be considerably strengthened. This is due to two important facts which we state for comment and reference.

LEMMA 7.1. *Every finitely additive probability μ on $\mathscr{S}^{(\omega)}/\vdash$ is σ-additive.*

Proof: For every $\Sigma \subseteq \mathscr{S}^{(\omega)}$ and every $\varphi \in \mathscr{S}^{(\omega)}$, $\Sigma \vdash \varphi$ iff for some finite $\Sigma' \subseteq \Sigma$ it is the case that $\Sigma' \vdash \varphi$. This implies that a set of disjoint elements of $\mathscr{S}^{(\omega)}/\vdash$ has a supremum in $\mathscr{S}^{(\omega)}/\vdash$ iff it is finite. Thus the σ-additivity of μ is trivial.

The following lemma is well-known and has an easy proof by means of elementary methods of functional analysis. For a purely algebraic proof we refer the reader to Horn and Tarski [1948].

LEMMA 7.2. *Let $\mathscr{B}$ be a Boolean algebra and let $\mathscr{A} \subseteq \mathscr{B}$ be a subalgebra. Every finitely additive probability on $\mathscr{A}$ can be extended to a finitely additive probability on $\mathscr{B}$.*

As pointed out before, Gaifman [1964] gives a proof of the next theorem. Our proof of Theorem 6.2 can be essentially simplified to yield this result by replacing the role of Lemma 6.3 by Lemmas 7.1 and 7.2. Indeed Lemma 6.3 has been designed to patch up the difficulties arising from the fact that Lemma 7.2 fails for σ-additive probabilities. This in turn corresponds to the fact that the Prime Ideal Theorem for σ-ideals fails for Boolean σ-algebras.

THEOREM 7.3. (i) *Every probability μ on $\mathscr{S}^{(\omega)}/\vdash$ has a denumerable probability model.*

(ii) *If for every* $0<n<\omega$, $\mu(\theta_n)\in\{0, 1\}$, *then* μ *has a denumerable probability model with strict identity.*

Proof: We again prove only part (ii). We first consider the case $\mu(\neg\theta_n)=1$ for all $n<\omega$. As in the proof of Theorem 6.2, we choose a set $T=\{t_i : i<\omega\}$ of new individual constants such that $t_i\neq t_j$ if $i\neq j$. The set $\mathscr{S}^{(\omega)}(T)$ is denumerable. Let $\exists v_{i_n}\varphi_n(v_{i_n})$ be an enumeration of all existential sentences of $\mathscr{S}^{(\omega)}(T)$, and define the sequence σ and the set Γ as before. Let $\Theta=\{\neg\theta_n : 0<n<\omega\}$, and $\Delta=\Gamma\cup\Theta$. Define the mapping f from $\mathscr{S}^{(\omega)}(T)/\Theta\vdash$ into $\mathscr{S}^{(\omega)}(T)/\Delta\vdash$ by $f(\varphi/\Theta\vdash)=\varphi/\Delta\vdash$. By Lemma 6.6 f is an isomorphism into. Since $\mu(\neg\theta_n)=1$ for all $0<n<\omega$, μ may be lifted to $\mathscr{S}^{(\omega)}(T)/\Theta\vdash$ and subsequently transferred to the subalgebra $\{\varphi/\Delta\vdash : \varphi\in\mathscr{S}^{(\omega)}\}$ of $\mathscr{S}^{(\omega)}(T)/\Delta\vdash$. By Lemma 7.1 and 7.2, μ may be extended to a probability ν on $\mathscr{S}^{(\omega)}(T)/\Delta\vdash$. Finally we transfer ν to $\mathscr{S}^{(\omega)}(T)/\vdash$ via the canonical homomorphism g from $\mathscr{S}^{(\omega)}(T)/\vdash$ onto $\mathscr{S}^{(\omega)}(T)/\Delta\vdash$ defined by $g(\varphi/\vdash)=\varphi/\Delta\vdash$. Since $\nu(\varphi)=1$ for every $\varphi\in\Delta$, ν satisfies the Gaifman Condition (**G**). If m is the restriction of ν to $\mathscr{S}^{(\omega)}(T)/\vdash$, then $\langle T, m\rangle$ is the desired probability model in view of the uniqueness part of Theorem 4.3. It is now also clear how to treat the remaining case and part (i) by analogy to the proof of Theorem 6.2.

Remarks: (1) As can readily be verified by an analysis of Lemma 6.6, our proof of part (ii) does not go through for a language $\mathscr{L}^{(\omega)}$ which either has infinitely many individual constants or non-denumerably many non-logical constants to begin with. This fact is substantiated by two counter-examples of Gaifman [1964]. Nevertheless, Theorem 7.3 still holds for these languages, as the proof in Gaifman [1964] shows. We will not discuss the question of adapting our method of Lemma 6.6 to this situation. (2) The cardinality statements of Theorem 7.3 obviously depend on our assumption that $\mathscr{L}^{(\omega)}$ has only denumerably many non-logical constants. If we allow non-denumerably many non-logical constants then the well-known adjustments have to be made. The same remark applies to all other theorems of this part which contain statements about the cardinality of probability systems.

Let **T** be the set of complete consistent theories in $\mathscr{L}^{(\omega)}$, that is, the set of prime ideals in $\mathscr{S}^{(\omega)}/\vdash$. As is well-known from the Representation Theorem for Boolean Algebras (see, e.g., Halmos [1963] p. 77), **T** is a compact Hausdorff space with a basis of closed-open (clopen) sets of the form $\{\Sigma\in\mathbf{T} : \varphi\in\Sigma\}$, where $\varphi\in\mathscr{S}^{(\omega)}$, and $\mathscr{S}^{(\omega)}/\vdash$ is isomorphic to the field of clopen subsets of **T** under an isomorphism which maps $\varphi/\vdash$ into $\{\Sigma\in\mathbf{T} : \varphi\in\Sigma\}$. Every model determines a complete consistent theory in $\mathscr{L}^{(\omega)}$, that is, a point in **T**. By the ordinary Completeness Theorem, every complete consistent theory of $\mathscr{L}^{(\omega)}$ has a model; thus **T** may be identified with the space of models.

Many important results in the ordinary logic of $\mathscr{L}^{(\omega)}$ may conveniently be established through topological considerations in the space **T**. This topological construction can be generalized in the strictest sense of the word; thus the space **M** of probability models of $\mathscr{L}^{(\omega)}$ can be defined as a compact Hausdorff space such that the space **T** can be homeomorphically embedded into **M**. The construction makes use of well-known definitions and methods of functional analysis. For the details we have to refer the reader to Dunford and Schwartz [1958].

Let $C(\mathbf{T})$ be the linear space of all continuous real functions on **T**. Since the characteristic functions of clopen subsets of **T** are continuous, we may regard $\mathscr{S}^{(\omega)}/\vdash$ as a subset of $C(\mathbf{T})$. Let **L** be the linear subspace of $C(\mathbf{T})$ generated by $\mathscr{S}^{(\omega)}/\vdash$ in $C(\mathbf{T})$. As is well-known, $C(\mathbf{T})$ is a Banach space under the sup-norm where for $x \in C(\mathbf{T})$ we have $\|x\| = \sup_{\xi \in \mathbf{T}} x(\xi)$. Any finitely additive probability μ on $\mathscr{S}^{(\omega)}/\vdash$ uniquely extends to a linear functional μ on $C(\mathbf{T})$ such that

(i) $\mu(x) \leq \|x\|$ for all $x \in C(\mathbf{T})$;

(ii) $\mu(\mathbf{1}) = 1$.

Conversely, any linear functional on $C(\mathbf{T})$ satisfying (i) and (ii) uniquely determines a finitely additive probability on $\mathscr{S}^{(\omega)}/\vdash$. Let $C(\mathbf{T})^*$ be the linear space of all continuous linear functionals on $C(\mathbf{T})$. As is well-known, $C(\mathbf{T})^*$ is also a Banach space with its own norm such that for every $\mu \in C(\mathbf{T})^*$, $\|\mu\| \leq 1$ iff $\mu(x) \leq \|x\|$ for all $x \in C(\mathbf{T})$.

Let $\mathbf{M} = \{\mu \in C(\mathbf{T})^* : \|\mu\| \leq 1, \mu(\mathbf{1}) = 1\}$. By our remark above, the set of all finitely additive probabilities on $\mathscr{S}^{(\omega)}/\vdash$ may be identified with **M**, and by Lemma 7.1 this set agrees with the set of probabilities on $\mathscr{S}^{(\omega)}/\vdash$. Every probability model determines a probability on $\mathscr{S}^{(\omega)}/\vdash$, and by Theorem 7.3 every probability on $\mathscr{S}^{(\omega)}/\vdash$ has a probability model, thus **M** may be identified with the space of probability models.

We now consider $C(\mathbf{T})^*$ with the so-called *weak star topology.* The basic open neighborhoods are sets of the form

$$\mathbf{N}(\mu; x_0, \ldots, x_{n-1}; \varepsilon) = \{\nu \in C(\mathbf{T})^* : |\nu(x_i) - \mu(x_i)| < \varepsilon \textit{ for all } i < n\}$$

where $n < \omega$, $\mu \in C(\mathbf{T})^*$, $x_0, \ldots, x_{n-1} \in C(\mathbf{T})$, and $\varepsilon > 0$ is a real number.

It is easy to see that **M** is a closed subset of the unit-sphere $\{\mu \in C(\mathbf{T})^* : \|\mu\| \leq 1\}$; therefore, by the Alaoglu Theorem **M** is a compact Hausdorff space with the relativized weak star topology.

Every $\Sigma \in \mathbf{T}$ uniquely determines a two-valued probability μ on $\mathscr{S}^{(\omega)}/\vdash$, and conversely. Thus there exists a natural embedding of **T** into **M**. Finally we observe that **M** is a convex set; that is, for every $\mu_1, \mu_2 \in \mathbf{M}$, and every real number $0<\alpha<1$, $\alpha\mu_1+(1-\alpha)\mu_2 \in \mathbf{M}$. For any subset **K** of a linear space, the *convex hull* of **K** is the smallest convex set containing **K**. The *closed convex hull* of a subset **K** of a linear topological space is the closure of the convex hull of **K**. We say $\mu \in \mathbf{M}$ is an *extreme point* of **M** iff for every μ_1, $\mu_2 \in \mathbf{M}$ and every $0<\alpha<1$, if $\mu=\alpha\mu_1+(1-\alpha)\mu_2$, then $\mu=\mu_1=\mu_2$.

We then have the following known basic theorem about our topological construction.

THEOREM 7.4. (i) *The natural embedding of* **T** *into* **M** *is a homeomorphism*;
(ii) **T**, *regarded as a subset of* **M**, *is the set of extreme points of* **M**;
(iii) **M** *is the closed convex hull of* **T**.

To prove Theorem 7.4 we first establish a useful lemma.

LEMMA 7.5. *Sets of the form* $\mathbf{M} \cap \mathbf{N}(\mu; x_0, \ldots, x_{n-1}; \varepsilon)$, *where* $x_0, \ldots, x_{n-1} \in \mathscr{S}^{(\omega)}/\vdash$, *constitute a basis for the weak star topology of* **M**.

Proof: Let $\mu \in C(T)^*$, $x_0, \ldots, x_{n-1} \in C(T)$, and $\varepsilon>0$. Consider $\nu \in \mathbf{M} \cap \mathbf{N}(\mu; x_0, \ldots, x_{n-1}; \varepsilon)$. Let $\delta=\varepsilon-\max_{i<n} |\nu(x_i)-\mu(x_i)|$. By the Stone-Weierstrass Theorem, there exist $y_0, \ldots, y_{n-1} \in \mathbf{L}$ such that $\|x_i-y_i\|<\frac{1}{3}\delta$ for all $i<n$. We show

$$\mathbf{M} \cap \mathbf{N}(\nu; y_0, \ldots, y_{n-1}; \tfrac{1}{3}\delta) \subseteq \mathbf{M} \cap \mathbf{N}(\mu; x_0, \ldots, x_{n-1}; \varepsilon).$$

Indeed, let $\lambda \in \mathbf{M} \cap \mathbf{N}(\nu; y_0, \ldots, y_{n-1}; \frac{1}{3}\delta)$ and $i<n$. Then since $\lambda, \nu \in \mathbf{M}$

$$\begin{aligned} |\lambda(x_i)-\mu(x_i)| &\leq |\lambda(x_i)-\lambda(y_i)| + |\lambda(y_i)-\nu(y_i)| \\ &\qquad + |\nu(y_i)-\nu(x_i)| + |\nu(x_i)-\mu(x_i)| \\ &< \|x_i-y_i\|\,\|\lambda\| + \tfrac{1}{3}\delta + \|y_i-x_i\|\,\|\nu\| + \varepsilon - \delta \\ &< \varepsilon . \end{aligned}$$

This proves that sets of the form $\mathbf{M} \cap \mathbf{N}(\mu; x_0, \ldots, x_{n-1}; \varepsilon)$, where $x_0, \ldots, x_{n-1} \in \mathbf{L}$, form a basis. Now let $x \in \mathbf{L}$. Then $x=\alpha_0 x_0+\cdots+\alpha_{r-1}x_{r-1}$, where $x_0, \ldots, x_{r-1} \in \mathscr{S}^{(\omega)}/\vdash$ and $\alpha_0, \ldots, \alpha_{r-1}$ are real numbers. Consider $\nu \in \mathbf{M} \cap \mathbf{N}(\mu; x; \varepsilon)$, and let $\delta=\varepsilon-|\nu(x)-\mu(x)|$. Then a straightforward computation yields $\mathbf{M} \cap \bigcap_{i<n} \mathbf{N}(\nu; x_i; \delta/r\cdot|\alpha_i|) \subseteq \mathbf{M} \cap \mathbf{N}(\mu; x; \varepsilon)$. This proves the lemma.

Now we proceed with the proof of Theorem 7.4. For part (i) we regard **T** as a subset of **M** and show that the topology of **T** is the topology of **M** relativised to **T**. Let $n<\omega$, $\mu\in\mathbf{T}$ and $x_0,\ldots,x_{n-1}\in\mathscr{S}^{(\omega)}/\vdash$. Then

$$\{\nu\in\mathbf{T}:|\nu(x_i)-\mu(x_i)|<\varepsilon \textit{ for all } i<n\}=\bigcap_{i<n}\{\nu\in\mathbf{T}:|\nu(x_i)-\mu(x_i)|<\varepsilon\}.$$

For every $i<n$, $\{\nu\in\mathbf{T}:|\nu(x_i)-\mu(x_i)|<\varepsilon\}$ will be either the empty set, the whole set, or the set of all prime ideals of $\mathscr{S}^{(\omega)}/\vdash$ containing x_i, depending on the choice of $\mu(x_i)$ and ε; in any case it is a clopen subset of **T**. Since a similar argument shows that *every* clopen set of **T** is the restriction of an open set of **M** to **T**, Lemma 7.5 proves assertion (i).

To prove part (ii) we let $\mu\in\mathbf{T}$ and consider $\mu_1,\mu_2\in\mathbf{M}$ and $0<\alpha<1$ so that $\mu=\alpha\mu_1+(1-\alpha)\mu_2$. Suppose $x\in\mathscr{S}^{(\omega)}/\vdash$ where $\mu(x)=1$. Then $1=\alpha\mu_1(x)+(1-\alpha)\mu_2(x)$. Since $|\mu_i(x)|\leq\|\mu_i\|\ \|x\|\leq 1$, for $i=1,2$, we have $\mu_1(x)=\mu_2(x)=1$. Thus $\mu=\mu_1=\mu_2$; that is, μ is an extreme point of **M**. Conversely let $\mu\in\mathbf{M}$ and suppose, for some $x\in\mathscr{S}^{(\omega)}/\vdash$, that $0<\mu(x)<1$. Let $\mu_1(y)=\mu(y\cap x)/\mu(x)$ and $\mu_2(y)=\mu(y\sim x)/(1-\mu(x))$ for all $y\in\mathscr{S}^{(\omega)}/\vdash$. Then $\mu_1,\mu_2\in\mathbf{M}$ and $\mu=\mu(x)\cdot\mu_1+(1-\mu(x))\cdot\mu_2$, that is μ is not an extreme point of **M**.

For (iii) note that **M** is a convex, compact subset of **C(T)***. Thus (iii) follows from (ii) by the Krein-Milman Theorem.

Our topological considerations now yield the full analogue of the Compactness Theorem of finitary first-order logic.

THEOREM 7.6. *Let Σ be a set of* ***closed*** *probability assertions of $\mathscr{L}^{(\omega)}$. Then*

(i) *Σ has a probability model iff every finite subset of Σ has a probability model;*

(ii) *Σ has a probability model with strict identity iff every finite subset of Σ has a probability model with strict identity.*

Proof: It is clear that the set of probability models of a closed probability assertion of $\mathscr{L}^{(\omega)}$ is a closed subset of **M**. Part (i) therefore follows from the compactness of **M**. To prove part (ii), let Σ be a set of closed probability assertions of $\mathscr{L}^{(\omega)}$ such that every finite subset of Σ has a probability model with strict identity. Consider the set $\{\theta_n:0<n<\omega\}$ as defined in Section 6. Observe that for all n, $\vdash[\theta_n\to\theta_{n+1}]$. For every $0<n<\omega$, let I_n be the set of all non-decreasing functions on $\{1,\ldots,n\}$ into $\{0,1\}$. Let $0<n<\omega$ be fixed. For each $i\in I_n$, define

$$\Delta_{ni}=\{\langle\,\lambda_0=\boldsymbol{i}_k,\theta_k\,\rangle:1\leq k\leq n\},$$

where $\lambda_0=\boldsymbol{i}_k$ is short for $-\lambda_0\geq\mathbf{0}$ if $i_k=0$, and $\lambda_0-\mathbf{1}\geq 0$ if $i_k=1$. Clearly Δ_{ni} is a set of closed assertions. Then for every $0<n<\omega$, there exists $i\in I_n$ such that $\Sigma\cup\Delta_{ni}$ has a probability model. Indeed, suppose that this is not the case.

Then, by part (i), there exists $0<n<\omega$ such that for all $i\in I_n$ there exists a finite subset $\Sigma_i \subseteq \Sigma$ such that $\Sigma_i \cup \Delta_{ni}$ has no probability model. Thus $(\bigcup_{i\in I_n} \Sigma_i) \cup \Delta_{ni}$ has no probability model for all $i\in I_n$. By hypothesis, $\bigcup_{i\in I_n} \Sigma_i$ has a probability model with strict identity; however, for some $i\in I_n$ this probability model has to be a model of Δ_{ni}, which is a contradiction. If for some n there exists $i\in I_n$ which is not identically zero, and $\Sigma \cup \Delta_{ni}$ has a probability model, then, by virtue of our observation above, there exists $0<N<\omega$ such that $\Sigma \cup \{\langle\lambda_0=\mathbf{0}, \theta_k\rangle : 0<k<N\} \cup \{\langle\lambda_0=\mathbf{1}, \theta_k\rangle : N\leq k\}$ has a probability model $\mathfrak{A}$. Otherwise for every n, $\Sigma \cup \{\langle\lambda_0=\mathbf{0}, \theta_k\rangle : 0<k\leq n\}$ has a probability model; and therefore, by part (i), $\Sigma \cup \{\langle\lambda_0=\mathbf{0}, \theta_n\rangle : 0<n<\omega\}$ has a probability model $\mathfrak{A}$. In both cases the probability model $\mathfrak{A}$ defines a probability μ on $\mathscr{S}^{(\omega)}/\vdash$ which satisfies the hypothesis of Theorem 7.3 (ii). Thus by Theorem 7.3 (ii), μ has a probability model $\mathfrak{B}$ with strict identity. Since $\mathfrak{A}$ is a probability model of Σ, we conclude that $\mathfrak{B}$ is also a probability model of Σ. This completes the proof of part (ii).

Not very much work has been done yet to investigate the question of satisfiability of sets of probability assertions in general. For one reason, general theorems about solvability conditions for *linear* inequalities are known, but there does not seem to be too much to say about *polynomial* inequalities. Note that at least quadratic inequalities are needed to formulate assertions about *conditional* probabilities and *independent* probabilities. Theorem 7.4, however, gives a very clear topological characterization of the space of probability models **M** in terms of the well-known space of models **T**, and should be a useful tool for further investigations.

We must next prove a few lemmas to which we have already referred repeatedly in earlier proofs.

We consider a fixed set T of new individual constants and assume throughout that μ is a probability on $\mathscr{S}^{(\omega)}(T)\vdash$ satisfying the Gaifman Condition (**G**). Also to simplify notation we assume that all sup's and inf's range over the set of finite subsets of T. The first lemma is an immediate consequence of elementary properties of sup's and inf's and will be stated without proof.

LEMMA 7.7. *If* $\forall v\varphi \in \mathscr{S}^{(\omega)}(T)$, *then* $\mu(\forall v\varphi) = \inf_F \mu(\bigwedge_{t\in F} \varphi(t))$.

If $\varphi, \psi \in \mathscr{S}^{(\omega)}(T)$, then an occurrence of φ in ψ is called *simple* if it does not occur within the scope of a quantifier or a negation sign. The next lemma is an easy consequence of the distributive laws.

LEMMA 7.8. (i) *Let ψ' be obtained from ψ by replacing a simple occurrence of $\exists v\varphi$ in ψ by $\bigvee_{t\in F} \varphi(t)$. Then $\mu(\psi)=\sup_F \mu(\psi')$.*

(ii) *Let ψ' be obtained from ψ by replacing a simple occurrence of $\forall v\varphi$ in ψ by $\bigwedge_{t\in F} \varphi(t)$. Then $\mu(\psi)=\inf_F \mu(\psi')$.*

Let $\varphi\in\mathscr{S}^{(\omega)}(T)$ be of the form $\mathsf{Q}_0 v_0 \dots \mathsf{Q}_{n-1} v_{n-1} \mathsf{M}(v_0,\dots,v_{n-1})$, where for each $i<n$, Q_i is either an existential quantifier $\exists$ or a universal quantifier $\forall$, and M is a formula. The *associated boundary prefix* is a sequence $bd_0 \dots bd_{n-1}$ such that for every $i<n$, $bd_i=\sup$ if $\mathsf{Q}_i=\exists$, and $bd_i=\inf$ if $\mathsf{Q}_i=\forall$. The *associated distribution prefix* is a sequence $\#_0 \dots \#_{n-1}$ such that for every $i<n$, $\#_i=\bigvee$ if $\mathsf{Q}_i=\exists$, and $\#_i=\bigwedge$ if $\mathsf{Q}_i=\forall$. Then we have the following basic lemma for the computation of $\mu(\varphi)$.

LEMMA 7.9.

$$\mu(\varphi) = \underset{F_0}{bd_0} \dots \underset{F_{n-1}}{bd_{n-1}} \mu\Big(\underset{t_0\in F_0}{\#_0} \dots \underset{t_{n-1}\in F_{n-1}}{\#_{n-1}} \mathsf{M}(t_0,\dots,t_{n-1})\Big).$$

Proof: By induction on n. For $n=1$ the assertion holds by condition (**G**) and Lemma 7.7. For $n>1$ we have by inductive hypothesis,

$$\mu(\varphi) = \underset{F_0}{bd_0} \dots \underset{F_{n-1}}{bd_{n-1}} \mu\Big(\underset{t_0\in F_0}{\#_0} \dots \underset{t_{n-1}\in F_{n-1}}{\#} \mathsf{Q}_n v_n \mathsf{M}(t_0,\dots,t_{n-1},v_n)\Big).$$

Let $F_0,\dots,F_{n-1}$ be fixed. Let $N=\prod_{i<n}|F_i|$, where $|F_i|$ is the number of elements in F_i. We enumerate the Cartesian product set $\mathbf{F}=\prod_{i<n} F_i$, say $\mathbf{F}=\{f_k : k<N\}$. If $\langle t_0,\dots,t_{n-1}\rangle\in\mathbf{F}$, we write $\underset{f\in\mathbf{F}}{\#} \mathsf{Q}_n v_n \mathsf{M}(f,v_n)$ as short for

$$\underset{t_0\in F_0}{\#_0} \dots \underset{t_{n-1}\in F_{n-1}}{\#_{n-1}} \mathsf{Q}_n v_n \mathsf{M}(t_0,\dots,t_{n-1},v_n),$$

where $f=\langle t_0,\dots,t_{n-1}\rangle$. For every $f\in\mathbf{F}$, $\mathsf{Q}_n v_n \mathsf{M}(f,v_n)$ occurs simply in $\underset{f\in\mathbf{F}}{\#} \mathsf{Q}_n v_n \mathsf{M}(f,v_n)$. Thus, by N consecutive applications of Lemma 7.8 and by the monotonicity of disjunction and conjunction and the elementary properties of sup's and inf's we have

$$\begin{aligned}\mu\Big(\underset{f\in\mathbf{F}}{\#} \mathsf{Q}_n v_n \mathsf{M}(f,v_n)\Big) &= \underset{G_{f_0}}{bd_n} \dots \underset{G_{f_{N-1}}}{bd_n}\, \mu\Big(\underset{f\in\mathbf{F}}{\#}\, \underset{s\in G_f}{\#_n} \mathsf{M}(f,s)\Big)\\ &= \underset{F_n}{bd_n}\, \mu\Big(\underset{f\in\mathbf{F}}{\#}\, \underset{t_n\in F_n}{\#_n} \mathsf{M}(f,t_n)\Big).\end{aligned}$$

Thus

$$\mu(\varphi) = \underset{F_0}{bd_0} \dots \underset{F_n}{bd_n} \mu\Big(\underset{t_0\in F_0}{\#_0} \dots \underset{t_n\in F_n}{\#_n} \mathsf{M}(t_0,\dots,t_n)\Big).$$

The last lemma of this series is an easy consequence of Lemma 7.9 and the distributive laws. For a convenient formulation we adopt the notation of the proof of Lemma 7.9 and write

$$\underset{F}{bd}\,\mu(\underset{f\in F}{\#}\,\mathsf{M}(f)) = \underset{F_0}{bd_0}\dots\underset{F_{n-1}}{bd_{n-1}}\,\mu(\underset{t_0\in F_0}{\#_0}\cdots\underset{t_{n-1}\in F_{n-1}}{\#_{n-1}}\,\mathsf{M}(t_0,\dots,t_{n-1})).$$

For every $k<r$, let $\varphi_i \in \mathscr{S}^{(\omega)}(T)$ be of the form $\mathsf{Q}_k\mathsf{M}_k$, where Q_k is a string of quantifiers and M_k is a formula. For every $k<r$, let bd_k and $\#_k$ be the associated boundary and distribution prefixes respectively.

LEMMA 7.10.

(i) $\mu(\bigwedge_{k<r}\varphi_k) = \underset{F_0}{bd_0}\dots\underset{F_{r-1}}{bd_{r-1}}\,\mu(\bigwedge_{k<r}\underset{t_k\in F_k}{\#_k}\,\mathsf{M}_k)$

(ii) $\mu(\bigvee_{k<r}\varphi_k) = \underset{F_0}{bd_0}\dots\underset{F_{r-1}}{bd_{r-1}}\,\mu(\bigvee_{k<r}\underset{t_k\in F_k}{\#_k}\,\mathsf{M}_k)$.

We are now in a position to prove analogues of many important results about the ordinary logic of $\mathscr{L}^{(\omega)}$. As a matter of fact, if we regard closed probability assertions of $\mathscr{L}^{(\omega)}$ as the analogues of sentences in ordinary logic, then the analogy in many respects seems to be complete. There are exceptions, however: we mentioned before our failure to define an analogue of direct products. We state next a few of the positive results and comment on their proof.

THEOREM 7.11 (DOWNWARD LÖWENHEIM-SKOLEM THEOREM). *Let $\langle T_2, m_2\rangle$ be a probability system of cardinality $\kappa_2 \geq \omega_0$, and let κ_1 be a cardinal number such that $\omega_0 \leq \kappa_1 \leq \kappa_2$. Then there exists a probability system $\langle T_1, m_1\rangle$ of cardinality κ_1 such that $\langle T_1, m_1\rangle \preccurlyeq \langle T_2, m_2\rangle$. Moreover, for any subset $T_3 \subseteq T_2$ with systemcardinality $\leq \kappa_1$, we may choose $T_3 \subseteq T_1$.*

Proof: The proof uses standard methods and utilizes the fact that, by the Gaifman Condition (G), whenever $\exists v\varphi \in \mathscr{S}^{(\omega)}(\mathrm{T})$, there exists a denumerable subset $T' \subseteq T$ such that

$$m^*(\exists v\varphi) = \sup_{F\in T'^{(\omega)}} m^*(\bigvee_{t\in F}\varphi(t)),$$

where $\langle T, m\rangle$ is the given probability system.

THEOREM 7.12 (DIRECTED UNION THEOREM). *Let $\{\langle T_i, m_i\rangle : i\in I\}$ be a $\preccurlyeq$-directed family of probability systems; that is, for all $i, j\in I$ there exists $k\in I$ such that both $\langle T_i, m_i\rangle$ and $\langle T_j, m_j\rangle \preccurlyeq \langle T_k, m_k\rangle$. Let $T = \bigcup_{i\in I} T_i$, and define, for* *every* $\varphi\in\mathscr{S}^{(\omega)}(T)$, $m(\varphi)=m_i(\varphi)$, *where* $\varphi\in\mathscr{S}^{(\omega)}(T_i)$. *Then* $\langle T, m\rangle$ *is a probability system and for every* $i\in I$, $\langle T_i, m_i\rangle\preccurlyeq\langle T, m\rangle$.

Proof: Since the family $\{\langle T_i, m_i\rangle : i\in I\}$ is $\preccurlyeq$-directed, m is a finitely additive probability on $\mathscr{S}^{(\omega)}(T)/\vdash$. By Lemma 7.1, m is σ-additive. The rest of the proof proceeds along well-known lines.

THEOREM 7.13 (UPWARD LÖWENHEIM-SKOLEM THEOREM). *Let* $\langle T_1, m_1\rangle$ *be a probability system of cardinality* $\kappa_1\geq\omega_0$, *and let* κ_2 *be a cardinal number such that* $\kappa_2\geq\kappa_1$. *Then there exists a probability system* $\langle T_2, m_2\rangle$ *of cardinality* κ_2 *such that* $\langle T_1, m_1\rangle\preccurlyeq\langle T_2, m_2\rangle$. *Moreover, if* $\langle T_1, m_1\rangle$ *has strict identity, then we may choose* $\langle T_2, m_2\rangle$ *to also have strict identity.*

Proof: Again a proof may be obtained by copying the well-known argument establishing the corresponding theorem of ordinary logic. This time we make use of Theorem 7.6 and the observation that the probability measure m_i^* on $\mathscr{S}^{(\omega)}(T_1)/\vdash$ determines a set of closed probability assertions of $\mathscr{L}^{(\omega)}(T_1)$. We also note that if $t\neq t'$, then $\langle\lambda_0=\mathbf{0},\ t=t'\rangle$ is a closed probability assertion. Thus the well-known compactness argument goes through.

We now present a result concerning symmetric probability systems which obviously has no analogue in ordinary logic. This result is due to Gaifman [1964] whose proof we have simplified by using ideas from the ultraproduct construction of probability models.

THEOREM 7.14. *Let* Σ *be a set of probability assertions of* $\mathscr{L}^{(\omega)}$. *Then* Σ *has a denumerable probability model iff* Σ *has a denumerable symmetric probability model.*

Proof: Suppose Σ has a denumerable probability model $\langle T, m\rangle$. Let $T=\{t_n : n<\gamma\}$, where $\gamma\leq\omega$ and where $t_m\neq t_n$ whenever $m\neq n$. Consider a probability λ defined on the power set of T where $\lambda(\{t_n\})>0$ for all $n<\gamma$. Let $\mathscr{P}$ be the product σ-field of subsets of T^T, and let $\bar{\lambda}$ be the product measure induced by λ. For every $\pi\in T^T$ and $\varphi\in\mathscr{S}^{(\omega)}(T)$, let φ^π be the result of replacing every individual constant t in φ by the constant $\pi(t)$. Clearly if $\vdash\varphi$ then $\vdash\varphi^\pi$. If $\varphi\in\mathscr{S}^{(\omega)}(T)$, then φ contains only finitely many individual constants. Thus for every real number a, $\{\pi\in T^T : m^*(\varphi^\pi)\leq a\}$ is a cylinder set with a finite basis. Therefore, since T is denumerable and λ is defined on the power set of T, this cylinder set belongs to $\mathscr{P}$. Hence for every $\varphi\in\mathscr{S}^{(\omega)}(T)$, the function $m^*(\varphi^\pi) : T^T\to[0, 1]$ is $\mathscr{P}$-measurable. Let Ω be the set of $\pi\in T^T$ whose range $\pi(T)=T$. Clearly $\Omega\in\mathscr{P}$, and a straightforward computation yields $\bar{\lambda}(\Omega)=1$. We define for all $\varphi \in \mathscr{S}^{(\omega)}(T)$

$$\mu_\lambda(\varphi) = \int_\Omega m^*(\varphi^\pi)\, \mathrm{d}\bar{\lambda}(\pi).$$

By the same argument as in the proof of Lemma 5.4 we show that μ_λ, regarded as a function on $\mathscr{S}^{(\omega)}(T)/\vdash$, is a finitely additive probability. By Lemma 7.1, μ_λ is σ-additive. μ_λ satisfies the condition (**G**). Indeed, let $\exists v\varphi \in \mathscr{S}^{(\omega)}(T)$. Then

$$\mu_\lambda(\exists v\varphi) = \int_\Omega m^*(\exists v\varphi^\pi)\, \mathrm{d}\bar{\lambda}(\pi).$$

For every $\pi \in \Omega$, since π is onto,

$$\begin{aligned} m^*(\exists v\varphi^\pi) &= \sup_{F \in T^{(\omega)}} m^*\big(\bigvee_{t\in F} \varphi^\pi(t)\big) \\ &= \sup_{F \in T^{(\omega)}} m^*\big(\bigvee_{t\in F} \varphi(t)\big)^\pi. \end{aligned}$$

Since T is denumerable by the Dominated Convergence Theorem,

$$\begin{aligned} \mu_\lambda(\exists v\varphi) &= \int_\Omega \sup_{F \in T^{(\omega)}} m^*\big(\bigvee_{t\in F} \varphi(t)\big)^\pi\, \mathrm{d}\bar{\lambda}(\pi) \\ &= \sup_{F \in T^{(\omega)}} \int_\Omega m^*\big(\bigvee_{t\in F} \varphi(t)\big)^\pi\, \mathrm{d}\bar{\lambda}(\pi) \\ &= \sup_{F \in T^{(\omega)}} \mu_\lambda\big(\bigvee_{t\in F} \varphi(t)\big). \end{aligned}$$

If $\varphi \in \mathscr{S}^{(\omega)}$, then $\mu_\lambda(\varphi)=m^*(\varphi)$. Thus let m_λ be the restriction of μ_λ to $\mathscr{S}^{(\omega)}(T)/\vdash$. Then $\langle T, m_\lambda\rangle$ is a probability model of Σ. It remains to be shown that $\langle T, m_\lambda\rangle$ is symmetric. Let $\tau \in \Omega$ be a finite permutation of T, and let $\varphi \in \mathscr{S}^{(\omega)}(T)$. By definition $\bar{\lambda}$ is invariant under all finite permutations of the coordinates of the product space $\mathscr{P}$. It is now easy to check by standard methods of integration theory that $\int_\Omega m^*(\varphi^{\tau\pi})\, \mathrm{d}\bar{\lambda}(\pi) = \int_\Omega m^*(\varphi^\pi)\, \mathrm{d}\bar{\lambda}(\pi)$. This completes the proof.

COROLLARY 7.15. *Every probability measure μ on $\mathscr{S}^{(\omega)}/\vdash$ has a symmetric probability model.*

Proof: By Theorems 7.3 and 7.14.

Remarks: (1) It is clear that our proof of Theorem 7.14 depends on the assumption that Σ has a countable probability model. Consequently Corollary 7.15 depends on the assumption that $\mathscr{L}^{(\omega)}$ has only countably many non-logical constants; that is, in our standard case one binary predicate **R**.

Indeed, the counter-example concerning symmetric probability models given in Section 6 can be constructed in $\mathscr{L}^{(\omega)}$ if we allow non-denumerably many unary predicates in $\mathscr{L}^{(\omega)}$.

(2) The probability system $\langle T, m_\lambda \rangle$ does not have strict identity. Indeed, suppose $t_m, t_n \in T$ and $m \neq n$. Then

$$\{\pi \in \Omega : \pi(t_m) = \pi(t_n)\} = \bigcup_{i<\gamma} (\{\pi \in \Omega : \pi(t_m) = t_i\} \cap \{\pi \in \Omega : \pi(t_n) = t_i\}).$$

Thus

$$\bar{\lambda}(\{\pi \in \Omega : \pi(t_m) = \pi(t_n)\}) = \sum_{i<\gamma} \lambda(\{t_i\})^2 > 0.$$

Consequently

$$m_\lambda(t_m = t_n) = \int_\Omega m(\pi(t_m) = \pi(t_n)) \, \mathrm{d}\bar{\lambda}(\pi) \geq \sum_{i<\gamma} \lambda(\{t_i\})^2 > 0.$$

If $\langle T, m \rangle$ has strict identity, then $\langle T, m_\lambda \rangle$ has the same cardinality as $\langle T, m \rangle$ and is "completely dispersed" in the sense that for $m \neq n$, $m_\lambda(t_m = t_n)$ is a constant strictly between 0 and 1.

(3) There are simple examples of sets of probability assertions of $\mathscr{L}^{(\omega)}$ which have a probability model with strict identity but do not have a symmetric probability model with strict identity. We give Gaifman's example: Let Σ be the set of probability assertions determined by the set of sentences $\exists v_0 \mathbf{P}(v_0)$, $\forall v_0 \forall v_1 [[\mathbf{P}(v_0) \wedge \mathbf{P}(v_1)] \rightarrow v_0 = v_1]$ and $\neg \theta_n$, for all $0 < n < \omega$. It is easy to see that Σ has no symmetric probability model with strict identity. The question of finding conditions for a set of probability assertions to have a symmetric probability model with strict identity is still open.

We conclude our discussion of the finitary language $\mathscr{L}^{(\omega)}$ with an immediate consequence of Theorem 6.7.

THEOREM 7.16. *The set of probability laws of $\mathscr{L}^{(\omega)}$ is recursively enumerable.*

Proof: In the remarks following the proof of Theorem 6.7 we can, for the finitary language $\mathscr{L}^{(\omega)}$, replace "effectively generate" everywhere by "recursively enumerate". This yields a proof of Theorem 7.16.

8. Examples. We have reason to hope that the results of probability logic may have useful applications to deductive logic, inductive logic and to probability theory. The first point was illustrated by Ryll-Nardzewski's example of a complete theory without models. The second point is rather obvious, as a matter of fact our work originally started with a study of Carnap's inductive logic. We will illustrate the third point by considering some well-known measurability problems in the theory of stochastic processes with non-denumerable index sets.

Let $\mathbb{R}$ be the two point-compactification of the set of real numbers; that is, the set of real numbers together with the points $-\infty$, ∞. Let $\mathbb{Q}$ be the set of rational numbers. Let $\mathscr{B}$ be the σ-field of Borel-sets of $\mathbb{R}$. Let T be an index set, and we will choose $T \subseteq \mathbb{R}$. Let $\mathscr{B}^T$ be the product σ-field of subsets of the Cartesian product space $\mathbb{R}^T$ induced by $\mathscr{B}$. As is well-known, a stochastic process with index set T may be identified with a probability space $\langle \mathbb{R}^T, \mathscr{B}^T, m \rangle$, where m is a probability on $\mathscr{B}^T$. (For further details see e.g., Loève [1960] p. 497 ff.) Let $\bar{\mathscr{B}}$ be the σ-field of Borel sets of $\mathbb{R}^T$; that is, the σ-field of subsets of $\mathbb{R}^T$ generated by the closed sets of the product topology on $\mathbb{R}^T$. Then $\mathscr{B}^T \subseteq \bar{\mathscr{B}}$, and $\mathscr{B}^T \neq \bar{\mathscr{B}}$ if T is non-denumerable.

During the investigation of stochastic processes one frequently would like to assign probabilities to sets $X \subseteq \mathbb{R}^T$ which are not m-measurable; that is, $X \notin \mathscr{B}^T$. It is well-known that this can always be done with finitely many sets at a time. Thus if $X_0, \ldots, X_{n-1} \subseteq \mathbb{R}^T$, then m can always be extended to a probability on the σ-field generated by $\mathscr{B}^T \cup \{X_0, \ldots, X_{n-1}\}$ (see, e.g., Halmos [1950] p. 71). The extension is not unique, however. In general, given a σ-field $\mathscr{A} \supseteq \mathscr{B}^T$, the question arises of whether there exists a probability n on $\mathscr{A}$ which extends m. Moreover, one attempts to specify convenient conditions which render such an extension unique. Nelson [1959] investigates this question for $\bar{\mathscr{B}}$ and gives a sufficient condition for the existence of a uniquely determined extension. He also shows that many interesting sets belong to $\bar{\mathscr{B}}$. Another extension result is Doob's Separability Theorem (Loève [1960] p. 507). Let $\mathscr{A}$ be the σ-field generated by $\mathscr{B}^T$ and sets of the form $\bigcap_{t \in I \cap T} \{x \in \mathbb{R}^T : x(t) \in C\}$, where $I \subseteq \mathbb{R}$ is an open interval and $C \subseteq \mathbb{R}$ is closed. Then Doob's Theorem says that every probability m on $\mathscr{B}^T$ has an extension to $\mathscr{A}$. Moreover, the extension n may be assumed to be separable; that is, there exists a denumerable subset $S \subseteq T$ such that for all open intervals I and closed sets C,

$$n\Big(\bigcap_{t \in I \cap T} \{x \in \mathbb{R}^T : x(t) \in C\}\Big) = n\Big(\bigcap_{t \in I \cap S} \{x \in \mathbb{R}^T : x(t) \in C\}\Big).$$

S is called a separating set. The separability condition makes the extension unique. Upon closer inspection the separability condition turns out to be an instance of a "Gaifman Condition." This will appear more clearly during the later development of our example. Indeed, it seems that from the earliest investigations of the extension problem conditions for the "reasonableness" of extensions of probabilities on $\mathscr{B}^T$ have been proposed which strikingly resemble particular instances of a "Gaifman Condition" (see, e.g., Doob [1947]). Thus one might be tempted to put down a Gaifman Condition on extensions of probabilities on $\mathscr{B}^T$ to a certain σ-field $\mathscr{A} \supseteq \mathscr{B}^T$ which renders such extensions unique, provided they exist, and then to investigate the problem of the existence of probabilities satisfying this condition. Our example will point in this direction.

Of course, a Gaifman Condition is most conveniently stated in terms of a language rather than in terms of certain representations of sets. For our example we use the infinitary language $\mathscr{L}$. As non-logical constants of $\mathscr{L}$ we provide a binary predicate $<$, and for every $q \in \mathbb{Q}$ a unary predicate $\mathbf{P}_q$. Moreover we augment $\mathscr{L}$ by a set of individual constants which, for convenience, we choose to be the index set $T \subseteq \mathbb{R}$.

Let $\mathbf{S}$ be the set of relational systems of the similarity type of $\mathscr{L}(T)$. We embed $\mathbb{R}^T$ pointwise into $\mathbf{S}$. For $x \in \mathbb{R}^T$ define $\mathfrak{A}_x \in \mathbf{S}$ as follows: $\mathfrak{A}_x = \langle T, P_q, <, t\rangle_{q \in Q,\, t \in T}$, where $<$ is the natural ordering of T, and $P_q = \{t \in T : x(t) \leq q\}$ for every $q \in \mathbb{Q}$. The function $\mathfrak{A}$ is certainly one-one, and we shall regard $\mathbb{R}^T$ simply as a subset of $\mathbf{S}$.

For every $\varphi \in \mathscr{S}(T)$, let $\mathbf{M}(\varphi)$ be the set of relational systems in which φ holds. It is well-known that $\{\mathbf{M}(\varphi) : \varphi \in \mathscr{S}(T)\}$ is a σ-field of subsets of $\mathbf{S}$, and, by the "weak" Completeness Theorem for $\mathscr{L}(T)$, it is isomorphic to $\mathscr{S}(T)/\vdash$ under an isomorphism which sends $\varphi/\vdash$ to $\mathbf{M}(\varphi)$. The following lemma is an immediate consequence of our definitions.

LEMMA 8.1. *For all $t, t' \in T$, $q \in \mathbb{Q}$ and $\xi \leq \omega$,*

$$\text{(i)}\quad \mathbf{M}(t = t') \cap \mathbb{R}^T = \begin{cases} \mathbb{R}^T & \text{if } t = t' \\ \emptyset & \text{if } t \neq t' \end{cases}$$

$$\text{(ii)}\quad \mathbf{M}(t < t') \cap \mathbb{R}^T = \begin{cases} \mathbb{R}^T & \text{if } t < t' \\ \emptyset & \text{if } t \nless t' \end{cases}$$

$$\text{(iii)}\quad \mathbf{M}(\mathbf{P}_q(t)) \cap \mathbb{R}^T = \{x \in \mathbb{R}^T : x(t) \leq q\}$$

$$\text{(iv)}\quad \mathbf{M}(\neg\varphi) \cap \mathbb{R}^T = \mathbb{R}^T \sim \mathbf{M}(\varphi)$$

$$\text{(v)}\quad \mathbf{M}(\bigvee_{i<\xi} \varphi_i) \cap \mathbb{R}^T = \bigcup_{i<\xi} \mathbf{M}(\varphi_i) \cap \mathbb{R}^T$$

$$\text{(vi)}\quad \mathbf{M}(\bigwedge_{i<\xi} \varphi_i) \cap \mathbb{R}^T = \bigcap_{i<\xi} \mathbf{M}(\varphi_i) \cap \mathbb{R}^T$$

$$\text{(vii)}\quad \mathbf{M}(\exists v\varphi) \cap \mathbb{R}^T = \bigcup_{t \in T} \mathbf{M}(\varphi(t)) \cap \mathbb{R}^T$$

$$\text{(viii)}\quad \mathbf{M}(\forall v\varphi) \cap \mathbb{R}^T = \bigcap_{t \in T} \mathbf{M}(\varphi(t)) \cap \mathbb{R}^T.$$

The operation of restriction to a subset is a complete homomorphism of fields of sets. By Lemma 8.1, (i)-(vi), this homomorphism maps $\{\mathbf{M}(\varphi):\varphi\in\mathscr{s}(T)\}$ onto $\mathscr{B}^T$. Accordingly, we define for all $\varphi\in\mathscr{s}(T)$,

$$\mu_m(\mathbf{M}(\varphi)) = m(\mathbf{M}(\varphi) \cap \mathbb{R}^T).$$

We obtain a probability μ_m on $\{\mathbf{M}(\varphi):\varphi\in\mathscr{s}(t)\}$ and thus also on $\mathscr{s}(T)/\vdash$. We write $\mu_m(\varphi)$ for $\mu_m(\mathbf{M}(\varphi))$ and obtain $\langle T, \mu_m\rangle$ as a probability system. We are interested in the σ-field $\mathscr{A}=\{\mathbf{M}(\varphi)\cap\mathbb{R}^T:\varphi\in\mathscr{S}(T)\}$. By our remark above $\mathscr{B}^T\subseteq\mathscr{A}$, and we will see later that $\mathscr{A}$ contains a vast assortment of interesting sets. First we must complete our series of definitions. Suppose n is a probability on $\mathscr{A}$ which extends m. Define for all $\varphi\in\mathscr{S}(T)$,

$$\nu_n(\varphi) = n(\mathbf{M}(\varphi) \cap \mathbb{R}^T).$$

We say that *n satisfies the Gaifman Condition* if ν_n satisfies $(\mathbf{G})$. We thus have as an immediate corollary of Theorem 4.3:

THEOREM 8.2. *For every probability m on $\mathscr{B}^T$ there exists at most one probability on $\mathscr{A}$ which extends m and satisfies the Gaifman Condition.*

This settles the uniqueness part of the extension problem, the existence part is of course much more difficult. Consider the probability system $\langle T, \mu_m\rangle$ induced by m. It is well-known that the equation $n(\mathbf{M}(\varphi)\cap\mathbb{R}^T)=\mu_m^*(\mathbf{M}(\varphi))$ for all $\varphi\in\mathscr{S}(T)$ defines a probability n on $\mathscr{A}$ iff whenever $\varphi\in\mathscr{S}(T)$ and $\mathbb{R}^T\subseteq\mathbf{M}(\varphi)$, then $\mu_m^*(\varphi)=1$. Indeed, just in this case n is a well-defined set function on $\mathscr{A}$ (see, e.g., Halmos [1963] p. 65). This leads to:

THEOREM 8.3. *Let m be a probability on $\mathscr{B}^T$. Then there exists a probability on $\mathscr{A}$ which extends m and satisfies the Gaifman Condition iff whenever $\varphi\in\mathscr{S}(T)$ and $\mathbb{R}^T\subseteq\mathbf{M}(\varphi)$, then $\mu_m^*(\varphi)=1$.*

The authors have been able to show that not every probability m on $\mathscr{B}^T$ has a Gaifman extension to $\mathscr{A}$. A counter-example can already be produced with the case of dependent Bronoulli trials. In this case the stochastic process is two-valued, the space $\mathbb{R}^T$ collapses to 2^T, and in our language $\mathscr{L}$ we only need one unary predicate $\mathbf{P}$ (together with $<$, of course). Further $\mathbf{M}(\mathbf{P}(t))\cap 2^T=\{x\in 2^T:x(t)=1\}$ so that $\mathbf{P}(t)$ means "success at time t". For φ we choose a finitary sentence which says "**P** *has a least upper bound*" as follows:

$$\exists v_0[\forall v_1[\mathbf{P}(v_1)\rightarrow v_1\leq v_0]\wedge\forall v_1[\forall v_2[\mathbf{P}(v_2)\rightarrow v_2\leq v_1]\rightarrow v_0\leq v_1]].$$

Then $2^T \subseteq \mathbf{M}(\varphi)$, but a measure m may be defined so that $\mu_m^*(\varphi)=0$. In fact we can take $T=[0, 1]$ and determine m so that $m(\mathbf{P}(0))=0$, $m(\mathbf{P}(1))=1$ and

$$m([\neg \mathbf{P}(t) \wedge \mathbf{P}(t')]) = \begin{cases} 0 & \text{if } t' < t; \\ t' - t & \text{if } t \leq t'. \end{cases}$$

This well-known process was suggested to us by David Blackwell.

This leaves us with two questions: (1) Which probabilities on $\mathscr{B}^T$ do have Gaifman extensions to $\mathscr{A}$? (2) Does there exist an "interesting" σ-field $\mathscr{A}'$ such that $\mathscr{B}^T \subseteq \mathscr{A}' \subseteq \mathscr{A}$ and every probability on $\mathscr{B}^T$ has a Gaifman extension to $\mathscr{A}'$? The first question has hardly been attacked yet, and the authors know only of a few meager results. The second question leads us back to Doob's Separability Theorem.

For simplicity let us assume $T=[0, 1]$, and consider $I=(t_1, t_2)$, where $0 \leq t_1 < t_2 \leq 1$. Then

$$\bigcap_{t \in I \cap T} \{x \in \mathbb{R}^T : x(t) \leq q\} = \mathbf{M}(\forall v_0 [t_1 < v_0 < t_2 \rightarrow \mathbf{P}_q(v_0)]) \cap \mathbb{R}^T,$$

and the separability condition amounts to the Gaifman Condition (**G**) for sentences of the form $\forall v_0 [t_1 < v_0 < t_2 \rightarrow \mathbf{P}_q(v_0)]$, where $q \in \mathbb{Q}$. (Note that the uniformity of the separating set S is really no strengthening of the Gaifman Condition since both the open intervals and the closed subsets of $\mathbb{R}$ have countable bases.) Doob's Separability Theorem now provides a positive answer to our second question. If we let $\mathscr{A}'$ be the σ-field generated by $\mathscr{B}^T$ and sets of the form $\mathbf{M}(\forall v_0 [t_1 < v_0 < t_2 \rightarrow \mathbf{P}_q(v_0)]) \cap \mathbb{R}^T$, where $q \in \mathbb{Q}$, and $t_1, t_2 \in T$, then every probability on $\mathscr{B}^T$ has a Gaifman extension to $\mathscr{A}'$. Indeed the authors have been able to prove the corresponding condition of Theorem 8.3, which yields a rather unorthodox proof of the Separability Theorem. Moreover, it can be shown that $\mathscr{A}'$ is in a certain sense maximal; that is, Doob's Theorem is about the best result we can expect. A detailed discussion of these results would take us too far afield here.

On the other hand it is necessary to point out that the σ-field $\mathscr{A}$ is highly unorthodox in terms of the traditional notions of probability theory. $\mathscr{A}$ is described by means of the infinitary language $\mathscr{L}$, and the nested application of Boolean operations combined with quantifiers has no clearly discernible analogue in usual classical methods of generating σ-fields. However, we can show that many interesting subsets of $\mathbb{R}^T$ which have traditionally been considered belong to $\mathscr{A}$. The striking feature of our approach is that we can show this by just writing down the ordinary definitions of these sets in the language $\mathscr{L}$. We give a few examples by first describing the set $\mathbf{M}(\varphi) \cap \mathbb{R}^T$ and then giving a suitable $\varphi \in \mathscr{L}$.

(1) *The set of non-decreasing functions*:

$$\forall v_0 \forall v_1 [v_0 \leq v_1 \rightarrow \bigwedge_{q \in Q} [\mathbf{P}_q(v_1) \rightarrow \mathbf{P}_q(v_0)]].$$

(2) *The set of functions assuming a maximum*:

$$\exists v_1 \forall v_0 \bigwedge_{q \in Q} [\mathbf{P}_q(v_1) \rightarrow \mathbf{P}_q(v_0)].$$

(3) *The set of functions assuming at most n different values*:

$$\exists v_0 \ldots \exists v_{n-1} \forall v_n \bigvee_{i<n} [\bigwedge_{q \in Q} \mathbf{P}_q(v_i) \leftrightarrow \mathbf{P}_q(v_n)].$$

In the following examples the sentence φ is more complicated to write down and we leave the details to the reader:

(4) *The set of functions assuming infinitely many different values.*

(5) *The set of continuous functions.*

(6) *The set of functions whose set of discontinuities is of first category.*

Finally we observe that we can increase the class of definable subsets of $\mathbb{R}^T$ by applying the independent union construction of Section 5; that is, by refining the ordinary model structure of the set T. For example, we might introduce two binary operation symbols $+$ and $\cdot$ which we interpret as ordinary addition and multiplication in T. We could then write down the sentence corresponding to:

(7) *The set of Rieman-integrable functions.*

By some further refinements we could also obtain the sentence for:

(8) *The set of Lebesgue-measurable functions.*

A very attractive feature of a probability on $\mathscr{A}$ satisfying the Gaifman Condition is that in many cases the condition allows us to actually compute the value of the probability for various useful sets. On the other hand, a rather disheartening aspect is the problem of its existence.

Appendix

by Peter Krauss

A measure-theoretic generalization of the Rasiowa-Sikorski Lemma

In Rasiowa and Sikorski [1950] the following theorem is proved, which now is generally known as the Rasiowa-Sikorski Lemma:

Let $\langle \mathscr{B}, \wedge, \vee, \sim \rangle$ be a Boolean algebra, let $b \in \mathscr{B}$ such that $b \neq \mathbf{1}$, and for every $m < \omega$ let $\mathscr{C}_m \subseteq \mathscr{B}$ be a subset such that $\bigwedge \mathscr{C}_m$ exists in $\mathscr{B}$. Then there exists a prime ideal $\not{p}$ in $\mathscr{B}$ such that

(i) $b \in \not{p}$;

(ii) *for every* $m < \omega$, $\bigwedge \mathscr{C}_m \in \not{p}$ *iff for some* $c \in \mathscr{C}_m$, $c \in \not{p}$.

An immediate consequence is the following relativised version:

THEOREM 1. *Let $\langle \mathscr{B}, \wedge, \vee, \sim \rangle$ be a Boolean σ-algebra, let $\mathscr{A} \subseteq \mathscr{B}$ be a σ-subalgebra, let q be a σ-prime ideal in $\mathscr{A}$ and let $[q]$ be the σ-ideal in $\mathscr{B}$ generated by q. Let $b \in \mathscr{B}$ such that $\sim b \notin [q]$ and for every $m, n < \omega$ let $b_{mn} \in \mathscr{B}$. Then there exists a prime ideal $\not{p}$ in $\mathscr{B}$ such that*

(i) $b \in \not{p}$;

(ii) *for every* $m < \omega$, $\bigwedge_{n<\omega} b_{mn} \in \not{p}$ *iff for some* $n < \omega$, $b_{mn} \in \not{p}$;

(iii) $q \subseteq \not{p}$.

Proof: Apply the Rasiowa-Sikorski Lemma to the quotient algebra $B/[q]$.

We prove the following measure theoretic generalization of Theorem 1:

THEOREM 2. *Let $\langle \mathscr{B}, \wedge, \vee, \sim \rangle$ be a Boolean σ-algebra, let $\mathscr{A} \subseteq \mathscr{B}$ be a σ-subalgebra, and let μ be a σ-additive probability measure on $\mathscr{A}$. Let v' be a finitely additive probability measure on $\mathscr{B}$ such that $v'(x) = \mu(x)$ for all $x \in \mathscr{A}$, let $x_0, \ldots, x_{n-1} \in \mathscr{B}$ and $\varepsilon > 0$. Finally, for every $m, n < \omega$ let $b_{mn} \in \mathscr{B}$. Then there exists a finitely additive probability measure v on $\mathscr{B}$ such that*

(i) $|v(x_i) - v'(x_i)| < \varepsilon$ *for all* $i < n$;

(ii) *for every* $m < \omega$, $v(\bigwedge_{n<\omega} b_{mn}) = \lim_{n \to \infty} v(\bigwedge_{i<n} b_{mi})$;

(iii) $v(x) = \mu(x)$ *for all* $x \in \mathscr{A}$.

Throughout the rest of this appendix we assume that $\langle \mathscr{B}, \wedge, \vee, \sim \rangle$ is a Boolean σ-algebra, $\mathscr{A} \subseteq \mathscr{B}$ is a σ-subalgebra and μ is a σ-additive probability measure on $\mathscr{A}$. Moreover $\mathbf{X}$ is the Stone space of $\mathscr{B}$, and we identify $\mathscr{B}$ with the set of clopen sets of $\mathbf{X}$. $C(\mathbf{X})$ is the set of continuous real functions on $\mathbf{X}$ and we regard $\mathscr{B}$ as a subset of $C(\mathbf{X})$. Finally $\mathbf{L}$ is the linear subspace of $C(\mathbf{X})$ generated by $\mathscr{B}$.

We first state two lemmas which can be found in Halmos [1950] p. 71, Exercise (2). For every subset $\mathscr{C} \subseteq \mathscr{B}$, let $[\mathscr{C}]$ be the σ-subalgebra of $\mathscr{B}$ generated by $\mathscr{C}$.

LEMMA 3. *Let $b \in \mathscr{B} \sim \mathscr{A}$. Then* $[\mathscr{A} \cup \{b\}] = \{(x \cap b) \cup (y \sim b) : x, y \in \mathscr{A}\}$.

LEMMA 4. *Let* $b \in \mathscr{B} \sim \mathscr{A}$ *and let* $a, c \in \mathscr{A}$ *such that* $a \subseteq b \subseteq c$ *and* $\mu(a) = \sup\{\mu(x) : x \subseteq b, x \in \mathscr{A}\}$, $\mu(c) = \inf\{\mu(x) : b \subseteq x, x \in \mathscr{A}\}$. *Let* $d = c \sim a$, *and let* $\alpha, \beta > 0$ *be real numbers such that* $\alpha + \beta = 1$. *Define for* $x, y \in \mathscr{A}$:

$$\nu((x \cap b) \cup (y \sim b)) = \mu((x \cap a) \cup (y \sim c)) + \alpha\mu(x \cap d) + \beta\mu(y \cap d).$$

Then ν *is a* σ*-additive probability measure on* $[\mathscr{A} \cup \{b\}]$ *such that* $\nu(x) = \mu(x)$ *for all* $x \in \mathscr{A}$.

The next lemma is also known:

LEMMA 5. *Let* $b \in \mathscr{B} \sim \mathscr{A}$ *and let* ν *be a finitely additive probability measure on* $[\mathscr{A} \cup \{b\}]$ *such that* $\nu(x) = \mu(x)$ *for all* $x \in \mathscr{A}$. *Then* ν *is* σ*-additive.*

Proof: Let $x_n \in \mathscr{A}$, $n < \omega$ be a decreasing sequence. Then

$$\nu((\bigwedge_{n<\omega} x_n) \cap b) \leq \lim_{n\to\infty} \nu(x_n \cap b).$$

Suppose

$$\nu((\bigwedge_{n<\omega} x_n) \cap b) < \lim_{n\to\infty} \nu(x_n \cap b).$$

Then

$$\begin{aligned}\mu(\bigwedge_{n<\omega} x_n) = \nu(\bigwedge_{n<\omega} x_n) &= \nu((\bigwedge_{n<\omega} x_n) \cap b) + \nu((\bigwedge_{n<\omega} x_n) \sim b) \\ &< \lim_{n\to\infty} \nu(x_n \cap b) + \lim_{n\to\infty} \nu(x_n \sim b) \\ &= \lim_{n\to\infty} \nu(x_n) = \lim_{n\to\infty} \mu(x_n),\end{aligned}$$

contradicting the hypothesis that μ is σ-additive. Lemma 3 now proves Lemma 5.

LEMMA 6. *Let* $b_n \in \mathscr{B}$, $n < \omega$ *be a decreasing sequence such that* $\bigwedge_{n<\omega} b_n = \mathbf{0}$. *Then there exists a finitely additive probability measure* ν *on* $\mathscr{B}$ *such that*

(i) $\lim_{n\to\infty} \nu(b_n) = 0$;

(ii) $\nu(x) = \mu(x)$ *for all* $x \in \mathscr{A}$.

Proof: Define by recursion: $\mathscr{A}_0 = \mathscr{A}$, $\mathscr{A}_{n+1} = [\mathscr{A}_n \cup \{b_n\}]$, for $n < \omega$. By Lemma 4 we see that without loss of generality we may assume that $b_n \notin \mathscr{A}_n$, for $n < \omega$. For every $n < \omega$ choose $a_n \in \mathscr{A}$ such that $a_n \subseteq b_n$ and $\mu(a_n) = \sup\{\mu(x) : x \subseteq b_n, x \in \mathscr{A}\}$. Since b_n is decreasing we may assume that a_n is decreasing. By recursion we wish to define a σ-additive probability measure ν_n on $\mathscr{A}_n$ such that $\nu_0 = \mu$ and for $n < \omega$, $\nu_{n+1}(x) = \nu_n(x)$ for every $x \in \mathscr{A}_n$. Suppose ν_n has been defined on $\mathscr{A}_n$. To define ν_{n+1} on $\mathscr{A}_{n+1}$, let $a'_n, c'_n \in \mathscr{A}_n$ be such that $a_n' \subseteq b_n \subseteq c_n'$ and

$$v_n(a_n') = \sup\{v_n(x): x \subseteq b_n, x \in \mathscr{A}_n\}$$
$$v_n(c_n') = \inf\{v_n(x): b_n \subseteq x, x \in \mathscr{A}_n\}.$$

Since $a_n \subseteq b_n$, $a_n \in \mathscr{A} \subseteq \mathscr{A}_n$ we may assume $a_n \subseteq a_n'$ for all $n < \omega$. Let $d_n = c_n' \sim a_n'$. In Lemma 4 let $\alpha = 0$, $\beta = 1$ and define for $x, y \in \mathscr{A}_n$

$$v_{n+1}((x \cap b_n) \cup (y \sim b_n)) = v_n((x \cap a_n') \cup (y \sim c_n')) + v_n(y \cap d_n).$$

We prove by induction on n: *For every* $n < \omega$, $x \in \mathscr{A}_n$,

$$v_{n+1}(x \cap b_n) = \sup\{\mu(z): z \subseteq x \cap b_n, z \in \mathscr{A}\}.$$

By definition, $v_{n+1}(x \cap b_n) = v_n(x \cap a_n')$. In case $n = 0$, let $x \in \mathscr{A}_0$. Then $x \cap a_0' \in \mathscr{A}$ and by the definition of a_0',

$$v_1(x \cap b_0) = \mu(x \cap a_0') = \sup\{\mu(z): z \subseteq x \cap b_0, z \in \mathscr{A}\}.$$

Now suppose for every $x \in \mathscr{A}_n$,

$$v_{n+1}(x \cap b_n) = \sup\{\mu(z): z \subseteq x \cap b_n, z \in \mathscr{A}\}.$$

We first show: $v_{n+2}(b_{n+1}) = v_{n+2}(a_{n+1})$. In fact, by definition, $v_{n+2}(b_{n+1}) = v_{n+1}(a_{n+1}')$. By the induction hypothesis, $v_{n+1}(b_n) = \mu(a_n) = v_{n+1}(a_n)$. Since $a_{n+1}' \subseteq b_{n+1} \subseteq b_n$, $v_{n+1}(a_{n+1}') = v_{n+1}(a_{n+1}' \cap a_n)$. Since $a_{n+1}' \in \mathscr{A}_{n+1}$ and $a_{n+1}' \subseteq b_n$, $a_{n+1}' = x \cap b_n$ for some $x \in \mathscr{A}_n$. Thus $v_{n+1}(a_{n+1}') = v_{n+1}((x \cap a_n) \cap b_n)$, where $x \cap a_n \in \mathscr{A}_n$. By these induction hypothesis,

$$v_{n+1}((x \cap a_n) \cap b_n) = \sup\{\mu(z): z \subseteq ((x \cap a_n) \cap b_n), z \in \mathscr{A}\}.$$

Furthermore $a_{n+1} \subseteq a_{n+1}' \subseteq b_{n+1}$ and $a_{n+1} \subseteq a_n$. Thus $a_{n+1} \subseteq a_{n+1}' \cap a_n = (x \cap a_n) \cap b_n \subseteq b_{n+1}$. Therefore, by the definition of a_{n+1}, $v_{n+1}((x \cap a_n) \cap b_n) = \mu(a_{n+1})$. This proves $v_{n+2}(b_{n+1}) = v_{n+2}(a_{n+1})$. Now let $x \in \mathscr{A}_{n+1}$. Then

$$v_{n+2}(x \cap b_{n+1}) = \sup\{\mu(z): z \subseteq x \cap b_{n+1}, z \in \mathscr{A}\}.$$

In fact, $v_{n+2}(b_{n+1}) = v_{n+2}(a_{n+1})$ and $a_{n+1} \subseteq b_{n+1}$. Therefore $v_{n+2}(x \cap b_{n+1}) = v_{n+2}(x \cap a_{n+1})$. Since $x \cap a_{n+1} \in \mathscr{A}_{n+1}$ and $x \cap a_{n+1} \subseteq b_{n+1} \subseteq b_n$, $x \cap a_{n+1} = y \cap b_n$ for some $y \in \mathscr{A}_n$. Thus $v_{n+2}(x \cap b_{n+1}) = v_{n+1}(y \cap b_n)$. By the induction hypothesis, $v_{n+1}(y \cap b_n) = \sup\{\mu(z): z \subseteq y \cap b_n, z \in \mathscr{A}\}$. Since $x \cap a_{n+1} \subseteq x \cap b_{n+1}$, $v_{n+2}(x \cap b_{n+1}) \leq \sup\{\mu(z): z \subseteq x \cap b_{n+1}, z \in \mathscr{A}\}$. Clearly $v_{n+2}(x \cap b_{n+1}) \geq \sup\{\mu(z): z \subseteq x \cap b_{n+1}, z \in \mathscr{A}\}$, which completes the inductive proof.

As a direct corollary we obtain: *For every* $n < \omega$, $v_{n+1}(b_n) = \mu(a_n)$. Now define a finitely additive probability measure v on the subalgebra $\bigcup_{n<\omega} \mathscr{A}_n$

by $\nu(x)=\nu_n(x)$ for $x\in\mathscr{A}_n$, and extend ν to a finitely additive probability measure on $\mathscr{B}$. Then $\lim_{n\to\infty}\nu(b_n)=\lim_{n\to\infty}\mu(a_n)=0$ since $\bigwedge_{n<\omega} a_n\subseteq \bigwedge_{n<\omega} b_n=0$. This proves Lemma 6.

Now we proceed to prove Theorem 2. $C(\mathbf{X})$ with the sup-norm is a Banach space. Any finitely additive probability measure ν on $\mathscr{B}$ uniquely extends to a linear functional on $C(\mathbf{X})$ such that

(i) $\nu(x)\leq\|x\|$ *for all* $x\in C(\mathbf{X})$;

(ii) $\nu(\mathbf{1})=1$.

And, conversely, any linear functional on $C(\mathbf{X})$ satisfying (i) and (ii) uniquely determines a finitely additive probability measure on $\mathscr{B}$. Let $C(\mathbf{X})^*$ be the conjugate space of $C(\mathbf{X})$ and consider the weak star topology for $C(\mathbf{X})^*$. For $\varphi\in C(\mathbf{X})^*$, $x_0,\dots,x_{n+1}\in C(\mathbf{X})$, and $\varepsilon>0$, let

$$\mathbf{N}(\varphi;x_0,\dots,x_{n-1};\varepsilon)=\{\psi\in C(\mathbf{X})^*:|\psi(x_i)-\varphi(x_i)|<\varepsilon \textit{ for } i<n\}$$
$$\mathbf{M}_{\mathscr{A}}=\{\varphi\in C(\mathbf{X})^*:\|\varphi\|\leq 1,\varphi(x)=\mu(x) \textit{ for all } x\in\mathscr{A}\}\,.$$

By the Alaoglu Theorem, $\mathbf{M}_{\mathscr{A}}$ with the weak star topology is a compact Hausdorff space.

LEMMA 7. *Let $b_n\in\mathscr{B}$, $n<\omega$ be a decreasing sequence such that $\bigwedge_{n<\omega} b_n=\mathbf{0}$. For every $r\geq 1$, let $P_r=\{\varphi\in\mathbf{M}_{\mathscr{A}}:1/r\leq\lim\varphi(b_n)\}$. Then P_r is nowhere dense in $\mathbf{M}_{\mathscr{A}}$.*

Proof: Since $P_r=\bigcap_{n<\omega}\{\varphi\in\mathbf{M}_{\mathscr{A}}:1/r\leq\varphi(b_n)\}$, P_r is closed in $\mathbf{M}_{\mathscr{A}}$. Suppose for some $\varphi\in C(\mathbf{X})^*$, $x_0,\dots,x_{n-1}\in C(\mathbf{X})$ and $\varepsilon>0$, we have

$$\varphi\in\mathbf{N}(\varphi;x_0,\dots,x_{n-1};\varepsilon)\cap M_{\mathscr{A}}\subseteq P_r.$$

By the Stone-Weierstrass Theorem there exist $v_0,\dots,v_{n-1}\in\mathbf{L}$ such that $\|x_i-v_i\|<\frac{1}{2}\varepsilon$ for all $i<n$. There exist $c_0,\dots,c_{m-1}$ such that $v_0,\dots,v_{n-1}\in\mathbf{M}$, where $\mathbf{M}$ is the linear subspace generated by $[\mathscr{A}\cup\{c_0,\dots,c_{m-1}\}]$. Let φ' be the restriction of φ to $\mathbf{M}$. By Lemma 5, φ' is σ-additive on $[\mathscr{A}\cup\{c_0,\dots,c_{m-1}\}]$. By Lemma 6, φ' has an extension ψ to $C(\mathbf{X})$ such that $\psi\in\mathbf{M}_{\mathscr{A}}$ and $\lim_{n\to\infty}\psi(b_n)=0$. Let $i<n$. Then $\varphi(v_i)=\varphi'(v_i)=\psi(v_i)$. Thus

$$\begin{aligned}|\varphi(x_i)-\psi(x_i)|&\leq|\varphi(x_i)-\varphi(v_i)|+|\psi(v_i)-\psi(x_i)|\\&\leq\|\varphi\|\,\|x_i-v_i\|+\|\psi\|\,\|x_i-v_i\|<\varepsilon\,.\end{aligned}$$

Thus $\psi\in P_r$ which is a contradiction.

Now consider the hypothesis of Theorem 2. It clearly suffices to assume that for every $m<\omega$, b_{mn} is a decreasing sequence for $n<\omega$ such that $\bigwedge_{n<\omega} b_{mn}=\mathbf{0}$. For every $m<\omega$, $r\geq 1$ let

$$P_{mr}=\{\varphi\in\mathbf{M}_{\mathscr{A}}: \frac{1}{r}\leq \lim_{n\to\infty}\varphi(b_{mn})\}.$$

By Lemma 7, $P=\bigcup_{m<\omega}\bigcup_{r\geq 1}P_r$ is of first category in $\mathbf{M}_{\mathscr{A}}$. Since the set $\mathbf{N}(v';x_0,\ldots,x_{n-1};\varepsilon)\cap\mathbf{M}_{\mathscr{A}}$ is a non-empty open set and $\mathbf{M}_{\mathscr{A}}$ is a compact Hausdorff space, Theorem 2 follows by the Baire-Category Theorem.

References

DOOB, J. L., 1947, *Probability in function space*, Bull. Am. Math. Soc., vol. 53, pp. 15–30

DUNFORD, N. and J. T. SCHWARTZ, 1958, *Linear operators* I (Interscience, New York)

EHRENFEUCHT, A. and A. MOSTOWSKI, 1961, *A compact space of models of axiomatic theories*, Bull. Acad. Polon. Sci. Ser. Sci. Math., Astron. Phys., vol. 9, pp. 369–373

FENSTAD, J. E., *A limit theorem in polyadic probabilities*, Proc. Logic Colloquium in Leicester, August 1965 (forthcoming)

GAIFMAN, H., 1964, *Concerning measures on first-order calculi*, Israel J. Math., vol. 2, pp. 1–18

HALMOS, P. R., 1950, *Measure theory* (Van Nostrand, New York)

HALMOS, P. R., 1963, *Lectures on Boolean algebras* (Van Nostrand, New York)

HENKIN, L., 1949, *The completeness of the first-order functional calculus*, J. Symbolic Logic, vol. 14, pp. 159–166

HEWITT, E. and L. SAVAGE, 1955, *Symmetric measures on Cartesian products*, Trans. Am. Math. Soc., vol. 80, pp. 470–501

HORN, A. and A. TARSKI, 1948, *Measures in Boolean algebras*, Trans. Am. Math. Soc., vol. 64, pp. 467–497

KARP, C. R., 1964, *Languages with expressions of infinite length* (North-Holland Publ. Comp., Amsterdam)

KELLEY, J. L., 1959, *Measures in Boolean algebras*, Pacific J. Math., vol. 9, pp. 1165–1177

LOÈVE, M., 1960, *Probability theory* (Van Nostrand, New York)

ŁOŚ, J., 1962, *Remarks on foundations of probability*, Proc. Intern. Congress of Mathematicians, Stockholm, 1962, pp. 225–229

ŁOŚ, J. and E. MARCZEWSKI, 1949, *Extension of measure*, Fundamenta Mathematicae, vol. 36, pp. 267–276

NELSON, E., 1959, *Regular probability measures on function space*, Ann. Math., vol. 69, pp. 630–643

RASIOWA, H. and R. SIKORSKI, 1950, *A proof of the completeness theorem of Gödel*, Fundamenta Mathematicae, vol. 37, pp. 193–200

SIKORSKI, R., 1964, *Boolean algebras* (second ed., Springer Verlag, Berlin)