Matematisk Seminar Nr.6
Universitetet i Oslo April 1964

ALGEBRAIC LOGIC AND THE FOUNDATION OF PROBABILITY

By

Jens Erik Fenstad

Revised July 1965

I. INTRODUCTION.

This paper is a preliminary report on some investigations into the foundation of probability within the framework of polyadic algebras. The purpose is not to give a complete theory, but to present some basic results which do not require an extensive knowledge of the specific system of algebraic logic here adopted. Further we wish to indicate the intuitive background for the logical notion of probability through the concept of "fair betting", and we want to argue in favor of a possible application by (very) briefly presenting the general statistical decision problem. Of course, we have no pretention of providing useful techniques for the statistician, we rather aim at supplying a conceptual framework within which a discussion of e.g. a priori probabilities would be meaningful.

II. FAIR GAMBLING.

In this section we shall give an introduction to probability via the concept of fair gambling. The theorem we state is a rather immediate extension of some results of Kemeny [11] and Lehman [16].

Let S be the set of "possible states of the world" and $\mathfrak{A}$ a σ-algebra of subsets called events. If $M_0 \in S$ is the true (but in general unknown) state of the world we say that the event $A \in \mathfrak{A}$ obtains if $M_0 \in A$. On $\mathfrak{A}$ we suppose it is defined a real-valued function $\lambda(A)$ taking values in the interval $[0,1]$. λ is our betting quotient: We assume that there are two persons betting, P_I and P_{II}. P_I may e.g. bet that A obtains, P_{II} that A does not obtain. In this case they are betting in the ratio $\lambda(A) : 1 - \lambda(A)$, i.e. if P_{II} bets the amount k_{II} that A does not obtain, P_I is willing to bet the amount

$$k_I = k_{II} \cdot \frac{\lambda(A)}{1-\lambda(A)}$$

that A obtains. (The other case, i.e. that P_I bets that A does not obtain and P_{II} bets that A obtains, is symmetric.)

A gamble G is specified by a sequence $\langle x_i \rangle$ of real numbers and a sequence $\langle A_i \rangle$ of disjoint sets from $\mathfrak{A}$. The pay-off function $G(M)$ which gives the amount of money a player would receive using λ as betting quotient, is defined by

$$G(M) = \sum x_i [I_{A_i}(M) - \lambda(A_i)],$$

where I_A is the indicator function of the event A. G is called <u>admissible</u> if $G(M)$ exists for all $M \in S$. For simplicity we shall assume that $|x_i| < K$, K some constant, for all G. If λ is a σ-additive probability measure this implies that $|G(M)| \leq K \cdot |\sum I_{A_i}(M) - \lambda(\cup A_i)| \leq K \cdot |1 - \lambda(\cup A_i)|$, i.e. $|G(M)| \leq K$.

Interpreting $G(M)$ we have that $|x_i|$ represent the combined amount the players bet on the event A_i. We assume that $G(M)$ represents the pay-off for P_I. Then if $x_i > 0$, P_I bets that A_i obtains, if $x_i < 0$, P_I bets that A_i does not obtain. $x_i = 0$ signifies that no bet is made on A_i.

In the example above we have one event A. The combined amount is

$$|x| = k_I + k_{II} = k_{II} \cdot \frac{1}{1 - \lambda(A)}$$

Thus for this gamble the pay-off function is

$$G(M) = k_{II} \cdot \frac{1}{1 - \lambda(A)} \cdot [I_A(M) - \lambda(A)],$$

as the assumption is that P_I bets that A obtains. We trivially calculate that if $M_0 \in A$, i.e. A obtains, then P_I receives the amount k_{II}. And if $M_0 \notin A$, then $G(M) = -k_I$, i.e. P_{II} receives the amount k_I from P_I.

A gamble G is called <u>fair</u> if there exists an $M \in S$ such that $G(M) \geq 0$, i.e. the gamble is fair if P_I is not certain to lose. λ is called a <u>rational</u> betting function if every admissible gamble G based on λ is fair. The main result, giving some intuitive background to the logical concept of probability, is contained in the following theorem.

THEOREM. <u>Let S be the set of possible states and $\mathfrak{A}$ a σ-algebra of events in S. Then λ is a fair betting function on $\langle S, \mathfrak{A} \rangle$ if and only if λ is a σ-additive probability measure on $\langle S, \mathfrak{A} \rangle$.</u>

PROOF. By definition $0 \leq \lambda(A) \leq 1$ for all $A \in \mathfrak{A}$. We show that $\lambda(S) = 1$: If $\lambda(S) < 1$, let $x < 0$ and consider the gamble $G_1(M) = x \cdot [I_S(M) - \lambda(S)] = x \cdot (1 - \lambda(S))$. Here $1 - \lambda(S) > 0$, thus

$G_1(M) < 0$ for all $M \in S$, which contradicts the fact that λ is rational.

To show that $\lambda(A') = 1 - \lambda(A)$, consider the gamble $G_2(M) = x([I_A(M) - \lambda(A)] + [I_{A'}(M) - \lambda(A')])$. We easily see that $G_2(M) = x\cdot(1 - \lambda(A) - \lambda(A'))$. If $1 - \lambda(A) - \lambda(A') > 0$, a choice of a negative x would give that $G_2(M) < 0$ for all $M \in S$. If $1 - \lambda(A) - \lambda(A') < 0$, a choice of a positive x would yield the same conclusion, hence $1 - \lambda(A) - \lambda(A')$ must be equal to 0, i.e. $\lambda(A') = 1 - \lambda(A)$.

In order to prove finite additivity let A and B be disjoint sets from $\mathfrak{A}$ and $C = (A \cup B)'$. Consider the gamble $G_3(M) = x([I_A(M) - \lambda(A)] + [I_B(M) - \lambda(B)] + [I_C(M) - \lambda(C)])$. One easily calculates that $G_3(M) = x\,(\lambda(A \cup B) - \lambda(A) - \lambda(B))$. Thus by suitable choice of x, one concludes that $\lambda(A \cup B) = \lambda(A) + \lambda(B)$ for disjoint sets A and B.

The finite additivity of λ implies that $\sum \lambda(A_i)$ exists for all disjoint sequences $\langle A_i \rangle$. To prove σ-additivity let $\langle A_i \rangle$ be disjoint and set $A_o = \bigcup A_i$. By assumption $A_o \in \mathfrak{A}$. Consider the disjoint sequence A_o', A_1, A_2, ... and let x be some real number, then the gamble

$$G_4(M) = x\cdot([I_{A_o'}(M) - \lambda(A_o')] + \sum [I_{A_i}(M) - \lambda(A_i)])$$

is admissible. Hence there exists $M_o \in S$ such that $G_4(M_o) \geqslant 0$, which implies, noting that $I_{A_o'} + \sum I_{A_i} \equiv 1$, that

$$x\cdot(1 - \lambda(A_o') - \sum \lambda(A_i)) \geqslant 0.$$

If x is chosen suitable, this gives $\lambda(\bigcup A_i) = \sum \lambda(A_i)$. Thus the first part of the theorem is proved.

To prove the converse we recall that the assumption $|x_i| < K$ implies that $|G(M)| \leqslant K$. Hence the integral

$$E(G) = \int_S G(M)d\,\lambda(M)$$

exists. $E(G)$ expresses the expected gain for P_I with respect to the gamble G. Letting $G_n(M) = \sum_{i=1}^{n} x_i[I_{A_i}(M) - \lambda(A_i)]$, we see that $|G_n(M)| \leqslant K$, that

$$\int_S G_n(M)d\,\lambda(M) = \sum_{i=1}^{n} x_i\left(\int_S I_{A_i}(M)d\,\lambda(M) - \lambda(A_i)\right) = 0,$$

and that $G_n(M) \to G(M)$. Thus $E(G) = 0$ for all admissible G if λ is a σ-additive probability measure on $\langle S, \mathfrak{A} \rangle$. But then for every G, there must exist $M \in S$ such that $G(M) \geq 0$, i.e. λ is a rational betting function. This completes the proof.

The above theorem goes back to Ramsey [27] and de Finetti [5] who both used it as a justification for the notion of subjective probability. Actually, they proved only one half of the theorem. Kemeny [11] and Lehman [16] first showed that if λ is a probability then λ must be a rational betting function. Their proofs concern conditional probabilities, but we obtain their results combining the above theorem with the representation theorem for polyadic probabilites proved below. They also considered only "finite" gambles.

Other methods of justifying subjective probabilities has been given by Savage [31] and Anscombe, Aumann [1]. The latter authors base their definition on utility theory (see e.g. Luce, Raiffa [20]) while Savage introduces axioms which simultaneously define utility and probability. (A variant of the Savage approach is given by Suppes [35].)

III. ALGEBRAIC LOGIC

In this section we state some definitions and results from the theory of polyadic algebras. The main reference is Halmos [9].

A polyadic algebra is the algebraic counterpart of the first order predicate logic obtained by identifying equivalent formulas. More precisely, a polyadic algebra $\langle A, J, S, \exists \rangle$ consists of a Boolean algebra A, a non-empty set I, and two maps S and $\exists$. S is a map from transformations $\tau : I \to I$ to Boolean endomorphisms on A. The image $S(\tau)$ is called a <u>substitution</u> on A. $\exists$ is a map from subsets $J \subseteq I$ to quantifiers $\exists(J)$ on A, where by a <u>quantifier</u> we understand a map $\exists(J) : A \to A$ satisfying the following three conditions: (i) $\exists(J)0 = 0$ (where $0 \in A$ denotes the zero element of the Boolean algebra), (ii) $p \leq \exists(J)p$ for all $p \in A$, and (iii) $\exists(J)(p \wedge \exists(J)q) = \exists(J)p \wedge \exists(J)q$. The reader may easily interpret the properties (i)-(iii) in terms of the existential quantifier in logic.

The interplay of the maps S and $\exists$ properly defines the notion of polyadic algebra.

$$S(\delta) = \text{identity}, \quad \text{where} \quad \delta_i = i \quad \text{for all} \quad i \in I.$$

$$S(\sigma\tau) = S(\sigma)S(\tau).$$

These axioms assert that S is a semi-group homomorphism from transformations to substitutions. Correspondingly we have for :

$$\exists(\emptyset) = \text{identity}.$$

$$\exists(J \cup K) = \exists(J)\,\exists(K).$$

For the next axiom assume that $\sigma = \tau$ on $I - J$, then

$$S(\sigma)\,\exists(J) = S(\tau)\,\exists(J).$$

And finally, if τ is injective on $\tau^{-1}J$, then

$$\exists(J)S(\tau) = S(\tau)\,\exists(\tau^{-1}J).$$

The polyadic algebra $\langle A, I, S, \exists \rangle$ is called locally finite if for all $p \in A$ there exists a finite subset $J \subseteq I$ such that $\exists(I-J)p = p$. The set J is then called a support of p. It is easily seen that the intersection of all supports of p again is a support of p which we denote by supp(p). If the set I is infinite the polyadic algebra is said to be of infinite degree. In this report all algebras are supposed to be locally finite of infinite degree.

Let $\mathcal{F}$ be a first order logic and let Γ be a set of sentences of $\mathcal{F}$. Let $\mathcal{F}_\Gamma$ denote the algebra of formulas obtained from $\mathcal{F}$ by identifying two formulas α_1 and α_2 if the equivalence $\alpha_1 \longleftrightarrow \alpha_2$ is deducible from Γ. It is well known that the propositional connectives in $\mathcal{F}$ induce a Boolean structure on $\mathcal{F}_\Gamma$. And it is fairly straight forward but rather laborious to verify that the quantifier and the substitution operator of the logic make $\mathcal{F}_\Gamma$ into locally finite polyadic algebra of infinite degree (provided that there is available an infinite set of variables in the logic $\mathcal{F}$).

The general algebraic theory of polyadic algebras is not very difficult being an immediate generalization of the Boolean counterpart. An exposition can be found in Halmos [9]. We recall that polyadic homomorphisms are Boolean homomorphisms commuting with $\exists$ and S, polyadic ideals are Boolean ideals closed under $\exists$ and S. A main result is that <u>every polyadic algebra is semi-simple</u>.

We shall also need a "computational" result. Define the relation $\sigma J_* \tau$ if $\sigma_i = \tau_i$ for all $i \in I-J$. Then for locally finite polyadic algebras of infinite degree one has

$$S(\tau)\exists(J)p = \bigvee\{S(\sigma)p \; ; \; \sigma J_* \tau\} \; .$$

Algebras of formulas are the first main examples of polyadic algebras. The second main examples are derived from the notion of interpretation or model of first order languages. Let X and I be non-empty sets and B a Boolean algebra. Define on the set of all maps $p = X^I \rightarrow B$ two operations $S(\tau)$ and $\exists(J)$ in the following way.

Let $\tau = I \rightarrow I$ and define $\tau_* x$, where $x \in X^I$, by $(\tau_* x)_i = x_{\tau i}$, then $S(\tau)$ is defined by

$$S(\tau)p(x) = p(\tau_* x) \; ,$$

for all $x \in X^I$ and $p : X^I \rightarrow B$.

Let $J \subseteq I$ and denote by xJ_*y the relation that $x_i = y_i$ for all $i \in I-J$, then $\exists(J)$ is defined by

$$\exists(J)p(x) = \bigvee\{p(y) \; ; \; xJ_*y\} .$$

A <u>B-valued functional polyadic algebra</u> A is now defined to be a Boolean subalgebra of maps from X^I to B closed under the operations $S(\tau)$ and $\exists(J)$. Usually B is taken to be the Boolean algebra $\underline{0}$ consisting of two elements $\{0,1\}$. Such a functional algebra is called a <u>model</u>. An important but easy result states that <u>every model is simple</u>. The converse statement taken in conjunction with the fact that every polyadic algebra is semi-simple, yields the following representation theorem.

THEOREM. <u>Every locally finite polyadic algebra of infinite degree is isomorphic to a subdirect product of models</u>.

A proof can be found in Halmos [9] or Fenstad [7].

We shall also need the notion of free polyadic algebras. Let X and I be non-empty sets and let j be a map from X to finite subsets of I. Then a locally finite polyadic algebra $\langle F, I, S, \exists \rangle$ is calle free on $\langle X, j \rangle$ if for every $\varphi : X \to B$, where B is a locally finite algebra with index set I and where $\mathrm{supp}(\varphi(x)) \subseteq j(x)$, there exists a polyadic homomorphism $f : F \to B$ such that $f \circ i = \varphi$, where $i = X \to F$ is some fixed injection. Free polyadic algebras exists and every locally finite algebra is the homomorphic image of some free algebra.

There is a very close connection between free polyadic algebras and algebras of formulas $\mathcal{F}_\Gamma$ where Γ is a set of logical axioms, - they are essentially the same algebras. Freedom of F expresses that no extra-logical axioms are assumed.

If we endow the class of models S of a polyadic algebra A with the topology generated by the sets A_q, q sentence of A, where $M \in A_q$ iff $q_M = 1$, it follows as a consequence of the theorem above that S is compact. This may be seen as follows: A set of sentences $\{q_n\}$ has a model iff each finite subset of $\{q_n\}$ has a model. Thus $\bigcap_n A_{q_n} = \emptyset$ iff for some finite intersection, $A_{q_1} \cap \dots \cap A_{q_n} = \emptyset$. As every closed set in S is an intersection of sets A_q, the conclusion follows. Note that if q is a sentence, then A_q is both open and closed. Thus in particular if $A_q = \bigcup A_{q_n}$, then $A_q = A_{q_1} \cup \dots \cup A_{q_n}$ for some number n. This observation will be of use below.

REMARK. In the rest of this paper we shall assume that <u>all occurring algebras A are denumarable and that the index set I can be identified with the set of natural numbers</u>.

IV. POLYADIC PROBABILITIES.

Let $\langle A, I, S, \exists \rangle$ be a denumerable polyadic algebra. A probability function $\underline{c}$ on A is a map $\underline{c} : A \to [0,1]$ which satisfies the following conditions.

(i) $\underline{c}(p) \geqslant 0$ for all $p \in A$.

(ii) $\underline{c}(1) = 1$, $\underline{c}(0) = 0$.

(iii) $\underline{c}(p \vee q) + \underline{c}(p \wedge q) = \underline{c}(p) + \underline{c}(q)$.

We do not assume that $\underline{c}$ is continuous. (See, however, a remark following the theorem given below.) In this paper we do not enter into elementary axiomatics showing in more derail how $\underline{c}$ can be related to the polyadic structure of A. In a subsequent section we shall make some preliminary remark on symmetry conditions which plausibly could be imposed on $\underline{c}$, e.g. requiring that $\underline{c}(p) = \underline{c}(S(\tau)p)$ for all transformations τ.

A probability $\underline{c}$ can be introduced on A if there is given a σ-additive probability measure λ on S (-more precisely, a σ-additive probability measure λ on $\langle S, \mathfrak{A} \rangle$, where $\mathfrak{A}$ is the σ-algebra generated by the sets A_q, q a sentence of A-) and for each $M \in S$ there is given a probability measure μ_M on the sets $p[M] \subseteq X_M^I$, where X_M is the set of individuals of the model M and $p[M] = \{x \in X_M^I ;\ p_M(x) = 1\}$, by means of the formula

$$\underline{c}(p) = \int_S \mu_M(p[M])\, d\lambda(M).$$

We omit the elementary calculations that $\underline{c}$ so defined satisfies the requirements (i) - (iii), but remark that the probability of an element $p \in A$ is obtained by first giving for each model M the probability of the set of sequences of individuals of M that satisfy p in M, and then taking a suitable average over the set S of all models using the measure λ on S. The main purpose of this section is to prove the converse of this result.

THEOREM. Let $\langle A, I, S, \exists \rangle$ <u>be a denumerable polyadic algebra and let</u> <u>$\underline{c}$ be a probability on A. Then there exists a σ-additive probability measure λ on the set of models S of A and for each model $M \in S$ a probability μ_M on the sets $p[M] = \{x \in X_M^I ;\ p_M(x) = 1\}$ such that $\underline{c}$ can be given by the formula</u>

$$\underline{c}(p) = \int_S \mu_M(p[M])\, d\lambda(M) .$$

PROOF. The proof will be given in several steps:

(1). Let q be a sentence of A, i.e. $\mathrm{supp}(q) = \emptyset$. Define the set A_q by $M \in A_q$ iff $q_M = 1$. Let $\mathfrak{A}_1$ be the collection of all sets A_q. On $\mathfrak{A}_1$ define a set function λ by

$$\lambda(A_q) = \underline{c}(q) .$$

$\lambda(A_q)$ is well-defined as $A_q = A_p$ implies that $p = q$ by the representation theorem of section III. $\mathfrak{A}_1$ is an algebra, and λ is an additive set function on this algebra. We shall prove that λ is continuous on $\mathfrak{A}_1$. Thus assume that $A_q = \bigcup A_{q_n}$. By the compactness of S this implies that $A_q = A_{q_1} \cup \dots \cup A_{q_n} = A_{q_1 \vee \dots \vee q_n}$ for some number n. Hence $q = q_1 \vee \dots \vee q_n$. Thus $\lambda(A_q) = \underline{c}(q) = \underline{c}(q_1 \vee \dots \vee q_n) = \lambda(A_{q_1} \cup \dots \cup A_{q_n}) \leqslant \lim_{n\to\infty} (A_{q_1} \cup \dots \cup A_{q_n}) \leqslant \lambda(A_q)$. This proves the continuity, thus λ may be uniquely extended to a σ-additive probability measure on the σ-algebra $\mathfrak{A}$ generated by the algebra $\mathfrak{A}_1$. (λ is a probability measure as $\lambda(S) = \lambda(A_1) = \underline{c}(1) = 1$.)

(2). Next define for each $p \in A$ a measure λ_p on $\mathfrak{A}_1$ by

$$\lambda_p(A_q) = \underline{c}(p \wedge q).$$

As above λ_p is well-defined, and it is immediate that each λ_p extends to a probability measure on $\mathfrak{A}$. Further $\underline{c}(p \wedge q) \leqslant \underline{c}(q)$, thus each measure λ_p is absolutely continuous with respect to the measure λ. Hence the Radon-Nikodym theorem applies, i.e. there exist non-negative measurable functions f_p, $p \in A$, such that

$$\lambda_p(B) = \int_B f_p(M)\, d\lambda(M) ,$$

for all sets $B \in \mathfrak{A}$. This gives

$$\underline{c}(p) = \underline{c}(p \wedge 1) = \lambda_p(S) = \int_S f_p(M)\, d\lambda(M) ,$$

for each $p \in A$. It remains to convert $f_p(M)$ into a probability measure on the sets $p[M]$.

(3). As a preliminary we shall investigate the properties of the functions f_p, $p \in A$. Each f_p can be chosen such that

(i) $0 \leq f_p \leq 1$.

(ii) $f_1 = 1$ and $f_o = 0$.

(iii) $f_{p \vee q} + f_{p \wedge q} = f_p + f_q$.

The prooof is by calculations, we indicate a few instances: Let $A_q \in \mathfrak{A}_1$, then

$$\int_{A_q} f_1(M)d\lambda(M) = \lambda_1(A_q) = \underline{c}(1 \wedge q) = \underline{c}(q) = \lambda(A_q) = \int_{A_q} d\lambda(M) ,$$

thus $f_1(M) = 1$, except for a subset of S of λ-measure 0. In the same way we obtain $f_o = 0$ for almost all $M \in S$. Next let $p_1, p_2 \in A$ and $A_q \in \mathfrak{A}_1$:

$$\int_{A_q} (f_{p_1}+f_{p_2})(M)d\lambda(M) = \int_{A_q} f_{p_1}(M)d\lambda(M) + \int_{A_q} f_{p_2}(M)d\lambda(M) =$$

$$\lambda_{p_1}(A_q)+\lambda_{p_2}(A_q) = \underline{c}(p_1 \wedge q)+\underline{c}(p_2 \wedge q) =$$

$$\underline{c}((p_1 \wedge q) \vee (p_2 \wedge q)) + \underline{c}((p_1 \wedge q) \wedge (p_2 \wedge q)) =$$

$$\underline{c}((p_1 \vee p_2) \wedge q) + \underline{c}((p_1 \wedge p_2) \wedge q) =$$

$$\int_{A_q} f_{p_1 \vee p_2}(M)d\lambda(M) + \int_{A_q} f_{p_1 \wedge p_2}(M)d\lambda(M) =$$

$$\int_{A_q} (f_{p_1 \vee p_2} + f_{p_1 \wedge p_2})(M)d\lambda(M) ,$$

thus $f_{p_1 \vee p_2} + f_{p_1 \wedge p_2} = f_{p_1} + f_{p_2}$ for almost all $M \in S$. And finally we obtain $0 \leq f_p \leq 1$ almost everywhere. Thus we have a countable set of equalities or inequalities I_n each true except for some set $B_n \subseteq S$ such that $\lambda(B_n) = 0$. As $\lambda(\bigcup B_n) \leq \sum \lambda(B_n) = 0$, the set I_n is valid except for a set B of λ-measure 0. But then by choosing some $M_o \in S - B$ we may redefine the functions f_p by setting $f_p(M) = f_p(M_o)$ if $M \in B$ thus obtaining the validity of (i) - (iii) for all $M \in S$.

(4). Before proceeding with the proof proper we shall pause to prove the following lemma which has a very plausible interpretation.

LEMMA. <u>Let</u> $p \in A$ <u>and</u> $M \in S$. <u>If</u> $p_M(x) = 1$ <u>for all</u> $x \in X_M^I$, <u>then</u> f_p <u>can be chosen such that</u> $f_p(M) = 1$.

Define $B_p = \{M \in S;\ p_M(x) = 1 \text{ for all } x \in X_M^I\}$. Let q be the universal closure of p, i.e. $q = (\exists(I)p')'$. Then $q_M = 1$ iff $M \in B_p$, hence $B_p = A_q$. We note that $q \leq p$, hence $q = p \wedge q$.

If $\lambda(A_q) = 0$, we may modify f_p on a null-set such that $f_p(M) = 1$ for $M \in B_p$. Hence assume that $\lambda(A_q) > 0$. We have

$$\lambda_p(A_q) = \int_{A_q} f_p(M)\, d\lambda(M).$$

But $\lambda_p(A_q)$ can also be evaluated in another way

$$\lambda_p(A_q) = \underline{c}(p \wedge q) = \underline{c}(q) = \lambda(A_q) = \int_{A_q} d\lambda(M).$$

Therefore

$$\int_{A_q} (1 - f_p)(M)\, d\lambda(M) = 0 ,$$

and as $1 - f_p \geq 0$ and $\lambda(A_q) > 0$, we obtain $1 - f_p = 0$ for almost all $M \in A_q$. Thus we may modify f_p on a null-set so that $f_p(M) = 1$ for all $M \in B_p$. The lemma is thus proved.

(5). We want to use the functions f_p to introduce probabilities μ_M on the models M. For each $M \in S$ consider the algebra of sets $p[M] \subseteq X_M^I$, $p \in A$, defined as $p[M] = \{x \in X_M^I ;\ p_M(x) = 1\}$. We may try to define μ_M by setting

$$\mu_M(p[M]) = f_p(M).$$

The main difficulty is to verify that the definition is legitimate, i.e. $p[M] = q[M]$ must imply that $f_p(M) = f_q(M)$.

Suppose there are elements $p, q \in A$ and a model $M_0 \in S$ such that $p[M_0] = q[M_0]$ but $f_p(M_0) \neq f_q(M_0)$. Consider the element $p \triangle q \in A$ defined by

$$p \triangle q = (p \vee q') \wedge (p' \vee q).$$

Using the formulas of section (3) we have

$$f_{p \triangle q} + f_1 = f_{p \vee q'} + f_{p' \vee q}$$

From $p[M_o] = q[M_o]$ we conclude that either $p_{M_o}(x) = q_{M_o}(x) = 1$ or $p_{M_o}(x) = q_{M_o}(x) = 0$ for $x \in X^I_{M_o}$. Thus

$$(p \triangle q)_{M_o}(x) = 1$$

for all $x \in X^I_{M_o}$. From the lemma of section (4) we conclude that $f_{p \triangle q}(M_o) = 1$. As $f_1(M_o) = 1$ we obtain $f_{p \vee q'}(M_o) + f_{p' \vee q}(M_o) = 2$, i.e. $f_{p \vee q'}(M_o) = f_{p' \vee q}(M_o) = 1$. We shall further need the equalities

$$f_p + f_{q'} = f_{p \vee q'} + f_{p \wedge q'} \quad ,$$

$$f_{p'} + f_q = f_{p' \vee q} + f_{p' \wedge q} \ .$$

As $f_{p'} = 1 - f_p$ and $f_{q'} = 1 - f_q$, we have

$$f_{p \vee q'} + f_{p \wedge q'} + f_{p' \vee q} + f_{p' \wedge q} = 2 \ ,$$

Combining this result with the values of $f_{p \vee q'}(M_o)$ and $f_{p' \vee q}(M_o)$ obtained above, we may conclude that $f_{p \wedge q'}(M_o) = f_{p' \wedge q}(M_o) = 0$. But then we have

$$f_p(M_o) + f_{q'}(M_o) = 1 + 0 = 1,$$

and as $f_{q'} = 1 - f_q$, we get $f_p(M_o) = f_q(M_o)$, contradicting our assumption above. Thus uniqueness is proved: $p[M_o] = q[M_o]$ implies that $f_p(M_o) = f_q(M_o)$.

(6). Some rather trivial calculations remains in order to finish the proof. We must show that $\mu_M(p[M])$ defined on the algebra $\mathfrak{A}_M = \{p[M] \ ; \ \ p \in A\}$ is a probability measure. We give a sample calculation:

$$\mu_M(p[M]) + \mu_M(q[M]) = f_p(M) + f_q(M) =$$
$$f_{p \vee q}(M) + f_{p \wedge q}(M) = \mu_M(p \vee q[M]) + \mu_M(p \wedge q[M]) =$$
$$\mu_M(p[M] \cup q[M]) + \mu_M(p[M] \cap q[M]),$$

using for the last equality the fact that interpretations in models are Boolean homomorphisms. Thus from section (2) we obtain

$$\underline{c}(p) = \int_S \mu_M(p[M])d\lambda(M) ,$$

and the proof is completed.

We shall make several remarks in connection with this theorem.

First, using the identity $\exists(J)p = \bigvee\{S(\sigma_n)p \ ; \ \sigma_n J_* \delta\}$, where J may be chosen finite as A is locally finite, and $\langle\sigma_n\rangle$ is an enumeration of all transformations which are the identity outside of J, it can be proved that if $\underline{c}$ satisfies the assumption

$$\underline{c}(\exists(J)p) = \lim_{n\to\infty} \underline{c}(S(\sigma_1)p \vee \ldots \vee S(\sigma_n)p) ,$$

then each μ_M can be chosen so that

$$\mu_M(\exists(J)p[M]) = \lim_{n\to\infty} \mu_M(S(\sigma_1)p[M] \cup \ldots \cup S(\sigma_n)p[M]) .$$

However, if $\underline{c}$ is continuous, we have not been able to conclude that each μ_M can be extended to a continuous probability on the σ-algebra generated by the sets $p[M]$, $p \in A$, as in this case we may need more than a countable number of modifications concerning the functions f_p

Next we may combine the above theorem and the theorem on fair gambling to conclude that a function $\underline{c}$ on the set of sentences to the interval $[0,1]$ is a probability iff it is a fair betting function when we are betting on whether or not a sentence q is true. Thus "rational betting" gives us a probaiblity measure on the "possible states of the world", i.e. the measure λ , whereas a probability $\underline{c}$ also introduces a probability μ_M within each model (-or possible state-) M. If p has non-empty support, then p can be considered as a certain predicate or property. We shall later show that under suitable restrictions $\mu_M(p[M])$ measures the relative frequency of the property p in the model M.

Finally we remark that the work above was partly inspired by the work of Carnap on inductive logic ([2], [3], [4]). More direct technical inspiration has been derived from the lecture J. Łoś gave to the International Congress of Mathematicians in Stockholm ([19]). However, we believe that our set up is more natural, working with a well-defined algebraic entity, the polyadic algebra, which in turn determines a well-defined set of models, the maximal ideal space. Further it seems that the proof indicated by Łoś is not complete. (Parts (4) and (5) of the above proof concerning the admissibility of the definition $\mu_M(p[M]) = f_p(M)$ are lacking.) Łoś also assumes that c is continuous, thus the compactness argument of section (1) seems to be an improvement. However, the idea of using the Radon-Nikodym theorem in section (2) is taken from him.

V. SOME REMARKS ON SYMMETRY, EFFECTIVE COMPUTABILITY AND CONDITIONAL PROBABILITIES.

In this section we shall briefly touch upon some further topics in the theory of polyadic probabilities.

First we shall give a very simple result bearing on the principle of insufficient reason. Let F be the free polyadic algebra generated by n elements $p_1, \ldots, p_n$ such that $\mathrm{supp}(p_i) = \emptyset$, $i = 1, \ldots, n$. (Thus F is essentially a Boolean algebra as $S(\tau)p=p$ and $\exists(J)p=p$ for all $p \in F$ and all τ and J.) We shall formalize the requirement that "the states of nature" does not depend upon the way we name them. Thus let $\sigma(p)$ denote the element obtained from p, where p is any element of F, i.e. a word in the generators p_i, by substituting $p_{\sigma i}$ for each constituent p_i in p. σ is to be regarded as a permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma 1 & \sigma 2 & \ldots & \sigma n \end{pmatrix}$$

which by the above definition operates on F to produce the element $\sigma(p)$ from p. Our substitution principle now says that $c(p) = c(\sigma(p))$ for all $p \in F$. This principle entails a version of the principle of incufficient reason.

Let p^i, i = 0,1, denote p if i = 0 and p' if i = 1. It is easily seen that the element p

$$p = \sum_{(i_1,\ldots,i_n)} p_1^{i_1} \wedge p_2^{i_2} \wedge \ldots \wedge p_n^{i_n}$$

where $(i_1,\ldots,i_n)$ is running over all n-termed sequences of 0 and 1 and $\sum$ denotes repeated unions, equals 1. It thus follows that

$$\sum_{(i_1,\ldots,i_n)} \underline{c}(p_1^{i_1} \wedge \ldots \wedge p_n^{i_n}) = 1.$$

For each pair $(i_1,\ldots,i_n)$ and $(j_1,\ldots,j_n)$ one easily constructs a σ such that $\sigma(p_1^{i_1} \wedge \ldots \wedge p_n^{i_n}) = p_1^{j_1} \wedge \ldots \wedge p_n^{j_n}$. Hence we may conclude that

$$\underline{c}(p_1^{i_1} \wedge \ldots \wedge p_n^{i_n}) = 2^{-n} ,$$

for all sequences $(i_1,\ldots,i_n)$. Thus each "atomic" or "irreducible" fact is equiprobable. By additivity one immediately calculates that $\underline{c}(p_i) = \frac{1}{2}$ for all $i = 1,\ldots,n$. This is of course a very simple result, and it remains to investigate more realistic versions (e.g. uniform distributions over infinite sets, see Jeffreys [10] and Renyi [29]). But we want to emphazise one point: Symmetry principles within polyadic probabilities are rules of language, and as such, may be defended and even considered fairly intuitive. Our algebra F was supposed free, this seems to be a precise version of the notion of ignorance. There are no factual assumption involved.

Another topic of great importance is the effective computatility of $\underline{c}$. It seems most reasonable to discuss this for languages $\mathcal{F}$. Let the associated set of axioms be Γ and the derived algebra be $\mathcal{F}_\Gamma$ Define the function $\underline{c}$ on $\mathcal{F}$ by

$$\underline{c}(\alpha) = \underline{c}([\alpha]_\Gamma),$$

where α is any formula of $\mathcal{F}$. Then under any notion of effective computability of the function $\underline{c}$, we would require that the set $\{\alpha ; \underline{c}(\alpha) = r\}$ is recursive for each r in the range of $\underline{c}$. (We, of course, are assuming a suitable gödelnumbering of $\mathcal{F}$.) The situation is analogous to the discussion of the word problem for groups. $\mathcal{F}$ is the set of words, Γ the defining relations and $\mathcal{F}_\Gamma$ the derived algebraic system.

Simple cardinality arguments yield non-computable functions $\underline{c}$. A somewhat more interesting, explicit example is as follows:

Let S_{Fin} be the set of finite models of $\mathcal{F}_\Gamma$, i.e. X_M is a finite set for $M \in S_{Fin}$. (We assume that Γ contains no axiom of infinity.) Each p_M as a function depends upon a finite number of variables, i.e there are natural numbers $i_1, \ldots, i_m$ such that $p_M(x) = p_M(y)$ if $x_{i_1} = y_{i_1}, \ldots, x_{i_m} = y_{i_m}$, where $x, y \in X_M^I$. Thus we write more simply $p_M(x_{i_1}, \ldots, x_{i_m})$ for $p_M(x)$. Note that card supp(p) = m. Let $n = \text{card } X_M$, we shall define the relative frequency function of p in M by

$$fr(p,M) = \frac{1}{n^m} \sum p_M(k_{i_1}, \ldots, k_{i_m}) ,$$

where the sum is taken over all $(k_{i_1}, \ldots, k_{i_m}) \in X_M^m$. This done define $\underline{c}$ on $\mathcal{F}_\Gamma$ by

$$\underline{c}(p) = \sum \frac{fr(p, M_n)}{2^n} ,$$

where $M_1, M_2, \ldots$ is some enumeration of S_{Fin}. It is easily seen that $\underline{c}(p) = 1$ iff $fr(p, M_n) = 1$ for all n, i.e. p is valid in every finite model. If the logic $\mathcal{F}$ and the axioms Γ are comprehensive enough, this entails by a well-known result of Trachtenbrot [36] that the set $\{\alpha \in \mathcal{F};\ \underline{c}(\alpha) = 1\}$ is not recursive. Hence the simple function $\underline{c}$ defined above is not effectively computable. Again, this is just a starting point for further investigations.

However, the point made above may have some relevance for the discussion of whether "subjective" probabilities can be unknown or not. Savage ([31], [32], [33], [34]) and de Finetti ([5]) maintain that they always can be determined by a kind of "introspection" (i.e. by reflection on how you would be willing to bet in certain imagined situations). Robbins (e.g. in [30] argue that a priori probabilities may exist but be unknown. As every "subjective" probability can be considered as a probability on some suitable polyadic algebra, we believe that our remarks above may have some relevance for this discussion, and that one ought to make a distinction between "pure" existence and effective constructibility.

A further topic which it would be of interest to treat within our framework concerns the problem of how arbitrary can a probability assignment be? Of course, the assignment depends upon the evidence offered. In our framework this means that given the language $\mathcal{F}$ and axiom sets Γ_1 and Γ_2, it may well happen that $\underline{c}_1(\alpha) = \underline{c}_1([\alpha]_{\Gamma_1})$ differs from $\underline{c}_2(\alpha) = \underline{c}_2([\alpha]_{\Gamma_2})$. But if the evidence offered is the same, must then the probability assignment be the same; i.e. is there for a given polyadic algebra a "preferred, or "objective", or "rational" probability function $\underline{c}$? Here opinions differ sharply, we shall indicate some possibilities within our framework. Every polyadic algebra A is the homomorphic image of some free algebra F, the kernel being determined by some set of sentences of F. If there exists some preferred probability on F, e.g. determined by suitable symmetry conditions (remember that F is <u>free</u>), then this probability would determine a preferred one on A by some sort of formula

$$\underline{c}_A(\phi(p)) = \underline{c}_F(p|\ker\phi) ,$$

where $\underline{c}_F$ is the probability on F, ϕ is the homomorphism $\phi: F \rightarrow A$, $\ker\phi$ is the kernel of the map (which determines the "factual assumptions" adopted in passing from F to A) and $\underline{c}_F(\cdot|\cdot)$ is the conditional probability derived from $\underline{c}_F$. (We shall give a short introduction to conditional probabilities below.) This approach would perhaps correspond to the point of views of Jeffreys [10] (see also Carnap [2], [4]).

$\underline{c}_A$ thus determined might not be effectively known. What we may know is probability functions on certain homomorphic images A' of A, i.e. by adding assumptions we may effectively determine functions $\underline{c}_{A'}$ for certain images A', - if the situation is sufficiently simple we may succeed in determining $\underline{c}_{A'}$ by "introspection". The problem is whether the various $\underline{c}_{A'}$ determine the $\underline{c}_A$. This may be so if we know sufficiently many $\underline{c}_{A'}$, otherwise there may be some indeterminacy left, giving us a "subjective" $\underline{c}_A$. This may be the situation envisaged by Savage, and, if so, the difference between the subjectivists and the objectivists is not so important after all. Their special points can all be expressed within a single framework. Of course, this is yet a speculation.

Having once applied the Radon-Nikodym theorem we may now define conditional probabilities in the following way. Let B be some set of sentences of A and denote by B^* the following λ-measurable subset of S:

$$B^* = \bigcap \{ A_q \; ; \; q \in B \} .$$

If $\lambda(B^*) > 0$ we now define $\underline{c}(p|B)$ for any $p \in A$ by

$$\underline{c}(p|B) = \frac{1}{\lambda(B^*)} \cdot \int_{B^*} \mu_M(p[M]) d\lambda(M) .$$

An elementary calcultaion then shows that if B consists of one sentence q and $\underline{c}(q) > 0$:

$$\underline{c}(p|q) = \frac{\underline{c}(p \wedge q)}{\underline{c}(q)} .$$

This definition can be used to state a formula for $\underline{c}_A(\phi(p))$ as discussed above, where $\phi : F \rightarrow A$ is a homomorphism onto. Denote by $\phi^* = \bigcap \{ A_q^F \; ; \; q$ sentence in F and $q' \in \ker \phi \}$, then

$$\underline{c}_A(\phi(p)) = \lambda_F(\phi^*)^{-1} \cdot \int_{\phi^*} \mu_M^F(p[M]) d\lambda_F(M) .$$

Renyi has in [29] introduced the notion of conditional probability algebra, several of his ideas (especially in connection with uniform prior distributions) ought to be investigated within the framework of this report.

VI. INTERPRETATIONS OF THE THEORY: A LIMIT THEOREM.

The probability calculus, i.e. the formal rules, is universally agreed upon, either in the form of a probability algebra in the sense of Kolmogorov [13] (with later refinements due to Renyi [29]), or in the form of a confirmation function - both are essentially an additive set function. The interpretation, however, of the formal rules is a highly controversial issue, and we shall not in this paper try to give any serious discussion of this topic. We shall instead present a limit theorem for polyadic probabilities, which we believe may be of relevance for any attempt of interpretation. As a preliminary a short (and very inadequate) catalogue of various foundational schools may have some interest for the non-expert.

First there is the frequency school up to now almost universally adopted after the inadequacy of the "classical" conception was clearly demonstrated. Main proponents are von Mieses [22], [23] and Reichenbach [28]. Their views as regard interpretations are essentially accepted without much questioning by e.g. Feller [6] and Neyman [24], although most probabilists and statisticians base their formal developments on the axiomatics of Kolmogorov. One version of this school interpret probability as the limit "in the long run" of observed frequency. This notion has obious intuitive appeal, seems to be rather objective, but involves great difficulties. (Consider the complexity of the axioms of von Mieses. For a philosophers critique see Kneele [14].)

The necessary, logical or objective view (there exists one and only one "correct" or "rational" c for a given language (which then incorporates the given evidence)) are held by a succession of writers such as Keynes [12], Jeffreys [10] and also Carnap [2], [3] and [4] who in addition to the logical concept, probability$_1$, also recognizes the frequency concept, probability$_2$.

The pure subjectivists consist of people such as Ramsey [27], de Finetti [5] and Savage [31], [32], [33], [34]. Savage also argues that the subjective concept has immediate applications to statistical techniques (see e.g. the text of Raiffa, Schlaifer [26]), and indeed there seems to be a tendency among people working in general decision theory to adopt a subjective concept of probability. We note that the subjectivists are severely criticized by Carnap [2], [4] who maintains that probability is a relative concept depending upon our knowledge, but it is not subjective; imposing sufficient rationality requirements one is led to a "preferred" confirmation function for a given language. (See also our remarks toward the end of section V.)

The limit theorem given in this section may have some relevance for the relationship between the logical view and the frequency conception. (For a physicist interpretation of "frequency" see Feynman [8, p.6 - 1].)

To state our results we shall need some definitions. An n-place predicate of a polyadic algebra A is a map $P : I^n \longrightarrow A$ such that

$$S(\tau)P(i_1,\ldots,i_n) = P(\tau i_1,\ldots,\tau i_n) ,$$

for all $\langle i_1,\ldots,i_n\rangle \in I^n$ and transformations $\tau : I \longrightarrow I$. An equality E for a polyadic algebra A is a binary predicate which satisfies: (i) $E(i,i) = 1$ for all $i \in I$, and (ii) $p \wedge E(i,j) \leqslant S({}^i/_j)p$ whenever $i,j \in I$ and $p \in A$. Here $S({}^i/_j)$ denotes the substitution which replaces the variable i by the variable j. For elementary properties of equality algebras we refer to Halmos [9].

An equality model M is a $\underline{0}$-valued functional algebra with equality E_o defined in the following way:

$$E_o(i,j)(x) = \begin{cases} 1 & \text{if } x_i = x_j \\ 0 & \text{if } x_i \neq x_j , \end{cases}$$

for all $x \in X_M^I$. The basic representation theorem of section III may be extended to assert that <u>every locally finite simple equality algebra of infinite degree is isomorphic to an equality model</u>. (Halmos [9 , p.228].)

Let A be an algebra with equality E. We shall assume that A contains elements $v_1, v_2, \ldots, v_n, \ldots$ which, in a sense to be made precise, denote constants. As I can be identified with the set of natural numbers, we shall denote variables by natural numbers. We now assume that for all n, $\mathrm{supp}(v_n) = \{1\}$, further that

(A) $$\exists(\{1\})[v_n \wedge \forall(\{2\})[(S({}^1/_2)v_n)' \vee E(1,2)]] = 1$$

hold in A for all n. Next we shall by t_n express that there is exactly n individuals, i.e. t_n is the following element of A

(B) $$t_n = \exists(\{1,2,\ldots,n\})[\bigwedge_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} E(i,j)' \wedge \forall(\ n+1\ )[\bigvee_{1 \leqslant i \leqslant n} E(i,n+1)]] .$$

Here $\bigwedge$ and $\bigvee$ denote repeated intersections and unions. Finally we want to express that if a model has n individuals, they are all named by some v_i, i.e. we assume that the following inequalities are satisfied

(C) $$t_n \leq \bigwedge_{\substack{1\leq i,j\leq n \\ i \neq j}} \forall(1)[v_i' \vee v_j']$$

It is not at all difficult to verify that such algebras A exist: take any equality model over a finite domain and interpret suitably the functions v_n, $n = 1, 2, \ldots$.

Let $\underline{c}$ be a probability on A, we shall impose the following requirements on $\underline{c}$. First, $\underline{c}$ shall make individuals in models equiprobable (a "sampling-type" model), and next, no axiom of infinity shall receive positive probability. The first requirement is made precise through the following set of identities

(I) $$\underline{c}(v_i \wedge q) = \underline{c}(v_j \wedge q)$$

for all sentences $q \in A$ and pair of indices $i, j \in I$. Next, if B is some set of sentences of A, $B = \{p_n\}$, we define $\underline{c}(B) = \lim_{n\to\infty} \underline{c}(p_1 \wedge \ldots \wedge p_n)$, a definition which makes $\underline{c}(B) = \lambda(B^*)$, where $B^* = \bigcap\{A_q;\ q \in B\}$. Our second requirement is then rendered by

(II) For any set B of sentences of A, if $\underline{c}(B) > 0$, there shall exist a t_n such that for all $q_1, \ldots, q_m \in B$, $q_1 \wedge \ldots \wedge q_m \wedge t_n \neq 0$.

This means, by use of the representation theorem, that if a set of sentences in A has positive probability, it is satisfied in some finite model.

One further remark, S denotes in the present context the class of <u>equality</u> models of A. But because of the representation theorem for equality algebras, the development of section IV remains valid, so that any $\underline{c}$ on A can be represented in the form

$$\underline{c}(p) = \int_S \mu_M(p[M])\, d\lambda(M) ,$$

where each $M \in S$ is an equality model.

Let S_o be the set of finite models in S, i.e. if $M \in S_o$, then $\mathrm{card}(X_M)$ is finite. We shall prove that $\lambda(S_o) = 1$. Obviously, $S_o = \bigcup A_{t_n}$, hence

$$\lambda(S_o) = \lim_{n\to\infty} (A_{t_1} \cup \dots \cup A_{t_n}) = \lim_{n\to\infty} \underline{c}(t_1 \vee \dots \vee t_n) .$$

Now $\lim \underline{c}(t_1 \vee \dots \vee t_n) = 1 - \lim \underline{c}(t_1' \wedge \dots \wedge t_n') = 1 - \underline{c}(\{t_1', t_2', \dots\})$
Suppose that $\underline{c}(\{t_1', t_2', \dots\}) > 0$, then by (II) there is a t_n consistent with the set $\{t_1', t_2', \dots\}$, which is impossible. Hence $\underline{c}(\{t_1', t_2', \dots\}) = 0$, and the validity of $\lambda(S_o) = 1$ follows. But then $\lambda(S - S_o) = 0$ and we may instead of integrating over S, restrict the domain of integration to S_o; thus for $p \in A$ we have

$$\underline{c}(p) = \int_{S_o} \mu_M(p[M]) d\lambda(M) .$$

We shall next evaluate $\mu_M(v_i[M])$ for $M \in S_o$. From the requirement (I) on $\underline{c}$ we obtain

$$\int_{A_q} \mu_M(v_i[M]) d\lambda(M) = \underline{c}(v_i \wedge q) = \underline{c}(v_j \wedge q) = \int_{A_q} \mu_M(v_j[M]) d\lambda(M) ,$$

hence, by modifying on null-sets we obtain the set of equalities

$$\mu_M(v_i[M]) = \mu_M(v_j[M])$$

for all i and j. Let $M \in S_o$, then $X_M = \{k_1, \dots, k_n\}$. Define the following equivalence relation on the set X_M^I : $x \sim y$ iff $x_1 = y_1$. Then we have

$$X_M^I = \bigcup_{i=1}^{n} [k_i] ,$$

a disjoint union where $x \in [k_i]$ iff $x_1 = k_i$. From the definition of t_n we conclude that $(t_n)_M = 1$. Hence the inequality (C) implies that $(v_1)_M, \dots, (v_n)_M$ all are different functions of M, in fact, use of (A) tells us that there is some permutation of the set $\{k_1, \dots, k_n\}$ such that

$$(v_j)_M(x) = 1 \quad \text{iff} \quad x_1 = k_{i_j} ,$$

for all $x \in X_M^I$ and $j = 1, \dots, n$. Hence we conclude that

$$v_j[M] = [k_{i_j}] ,$$

$j = 1, \dots, n$. Applying this result to the above we obtain

$$1 = \mu_M(X_M^I) = \mu_M(\bigcup_{i=1}^{n}[k_i]) = \sum_{i=1}^{n} \mu_M([k_i]) =$$

$$= \sum_{i=1}^{n} \mu_M(v_i[M]) \ .$$

From this and the equality $\mu_M(v_i[M]) = \mu_M(v_j[M])$, we conclude that

$$\mu_M(v_i[M]) = \mathrm{card}(X_M)^{-1} \ ,$$

$i = 1, \ldots, n$. Not to complicate our notation unduly we shall state a special case of our limit theorem. Let $p \in A$ and suppose that $\mathrm{supp}(p) = \{1\}$. Define the function $fr(p,M)$, $M \in S_o$, by

$$fr(p,M) = \mathrm{card}(X_M)^{-1} \cdot \sum p_M(x) \ ,$$

where we sum over one representative from each of the equivalence classes $[k_1], \ldots, [k_n]$, $n = \mathrm{card}(X_M)$. Thus $fr(p,M)$ gives the relative frequency of the "property" p in the finite model M. We propose to show that $\mu_M(p[M]) = fr(p,M)$. This follows because $p[M] = \bigcup [k_i]$, where we take the union of those $[k_i]$ such that $x \in [k_i]$ implies that $p_M(x) = 1$, hence $\mu_M(p[M]) = \mu_M(\bigcup [k_i]) = \sum \mu_M([k_i]) = fr(p,M)$. Define the following random variable X on S_o:

$$X(M) = fr(p,M) \ .$$

It is then an obvious calculation that

$$E|X| = EX = \int_{S_o} fr(p,M)d\lambda(M) = \int_{S_o} \mu_M(p[M])d\lambda(M) = \underline{c}(p).$$

Hence the conditions of the Kolmogorov strong law of large numbers is satisfied [18], and we may state the following result.

THEOREM. Let A be a polyadic algebra with equality and let the special elements v_n and t_n, $n=1,2,\ldots,$ satisfy the requirements (A), (B) and (C) above. Let further c be any probability on A satisfying (I) and (II). If S_o is the set of finite models of A and p is any element of A of support one, then

$$\frac{1}{n} \sum_{i=1}^{n} X^{(i)} \longrightarrow \underline{c}(p) \qquad \text{a.s.} ,$$

where $X^{(1)}, X^{(2)}, \ldots$ is an independent sequence of "observations of models", i.e. each $X^{(i)}$ is distributed as $X = fr(p,M)$ and independently observed.

The precise content of the above convergence assertion is as follows. Let $S^* = \prod S_o$ be a countable product where each factor equals S_o and consider on S^* the product measure λ^*, each factor having the measure λ. Then for almost all sequences $\langle X^{(i)} \rangle \in S^*$, where $X^{(i)} = fr(p,M_i)$, for some $M_i \in S_o$, $\frac{1}{n}\sum X^{(i)} \longrightarrow \underline{c}(p)$ in the usual sense. Thus the convergence assertion is true except possibly for a λ^*-null set in the product space S^*.

Interpreted the theorem says that for a property p of individuals, $\underline{c}(p)$ is our estimate of the long range relative frequency of p in models ("possible worlds"), and if λ is "adequate", then our estimate is consistent. Thus the above theorem is to be interpreted as a consistency requirement on the theory of a specific function $\underline{c}$.

This result has also some connection with the views of Reichenbach [28] on how to interpret probability assertions concerning single statements. Our space of models, S, is a precise version of the rather vague notion of "all possible worlds".

VII. THE GENERAL STATISTICAL DECISION PROBLEM.

To conclude the discussion we outline briefly the general statistical decision problem following mainly Raiffa, Schlaiffer [26] and Raiffa [25], indicating a possible connection with our approach.

The general formulation of the problem originates with Wald [37], [38]. We assume given a set A of acts, a set of possible states Θ, a family of experiments E, a sample space Z and a utility evaluation $u : E \times Z \times A \times \Theta \longrightarrow \mathbb{R}$. The situation can be described as decision making under uncertainty: We do not know which state $\theta \in \Theta$ obtains, but we may perform an experiment $e \in E$ to obtain further information $z \in Z$ on which to base our decision about which act $a \in A$ we shall perform.

The further assumption is made that there is given a probability measure $P_{\Theta,Z}(.,.\mid e)$ on $\Theta \times Z$. This probability often can be determined by a probability $P_Z(.\mid\Theta,e)$ on Z and an a priori probability P_Θ on Θ .

If we have selected an experiment $e \in E$ and observed the outcome $z \in Z$, the expected utility when performing the act $a \in A$ will be

$$u^*(e,z,a) = \int_\Theta u(e,z,a,\Theta)dP_\Theta(\Theta\mid z,e) ,$$

where $P_\Theta(.\mid z,e)$ is the a posteriori measure on Θ calculated from P_Θ and $P_Z(.\mid\Theta,e)$ by means of Bayes formula.

The decision maker seeks to maximoze his utility, hence he will choose an act a maximizing $u^*(e,z,a)$. Define

$$u^*(e,z) = \max_a u^*(e,z,a) ,$$

(to simplify we assume that such an act always exists). The expected utility of an experiment e is defined by

$$u^*(e) = \int_Z u^*(e,z)dP_Z(z\mid e) ,$$

and the maximum utility is

$$u^* = \max_e u^*(e).$$

Hence the decision maker wants to select an $e_0 \in E$ such that $u^*(e_0) = u^*$, and, after performing the experiment e_0 and observing the outcome, say z_0, wants to select an act a_0 such that $u^*(e_0,z_0,a_0) = u^*(e_0,z_0)$.

This is the general problem and its optimal solution. But any actual solution is dependent upon our knowledge of the measure $P_{\Theta,Z}(.,.\mid e)$. Usually the measure $P_Z(.\mid\Theta,e)$ is "objective", e.g. based upon extended frequency observations. (But even at this point assumptions of mathematical convenience enter which help to simplify the problem but which are not always justifiable form the point of view of a strict frequency interpretation of the probability concept.) The real difficulty is felt to be the determination of the a priori measure P_Θ on Θ .

At this point we can perhaps make a contact with our previous discussion: If it were possible to determine an "adequate" language $\mathcal{F}$ (and hence a derived algebra A) corresponding to our decision problem such that the models of this algebra in some sense determined the state space Θ , then a probability c on the algebra A would induce a probability P_Θ on Θ . And there is no reason why the determination of c should proceed upon frequency reports alone, all previous knowledge is of relevance in determining the actual c. (Compare here the duscussion toward the end of section V.)

Thus, considering the close relationship between the "states of the world", Θ , and the space of models, S , it does not seem unreasonable to suppose that the conceptual frame-work provided by polyadic probabilities may give some insight into the foundation of the general theory of decision making. In fact, some of our remarks in section IV - VI may be taken to support this point of view. However, the aim has never been to supply efficient means in actual decision making.

VIII. REFERENCES.

1. Anscombe, Aumann: A defintion of subjective probability, Annals of Math.Statist., vol.34 (1963), pp.199-205.

2. Carnap: Logical Foundation of Probability, Chicago 1951.

3. Carnap: The Continuum of Inductive Methods, Chicago 1952.

4. Carnap: The aim of inductive logic, Proceedings of the 1960 International Congress at Stanford, Stanford University Press 1962, pp.303-318.

5. de Finetti: La prévision: ses lois logiques,ses sources subjectives, Annales de l'Institute Henri Poincaré, vol.8 (1937), pp.1-68. Translated in Kyburg, Smokler: Studies in subjectivity, pp.93-158.

6. Feller: An Introduction to Probability Theory, Vol.1, J. Wiley & Sons, New York, 1962.

7. Fenstad: On representation of polyadic algebras. Kongelige Norske Videnskabers Selskabs Forh. 1964, pp.36-41.

8. Feynman: Lectures on Physics, vol.1, Addison-Wesley, 1963.

9. Halmos: Algebraic Logic, Chelsea, New York, 1962.

10. Jeffreys: Theory of Probability, Oxford University Press, 1939.

11. Kemeney: Fair bets and inductive probability, J. Symbolic Logic vol.20, 1955, pp.263-273.

12. Keynes: Treatise on Probability, MacMillan, London, 1922.

13. Kolmogoroff: Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik, vol.2, no.3, 1933.

14. Kneale: Probability and Induction, Oxford University Press, 1949.

15. Kyburg, Smokler: Studies in Subjective Probability, J. Wiley & Sons, New York, 1964.

16. Lehman: On confirmation and rational betting, J. Symbolic Logic, vol.20, 1955, pp.251-262.

18. Loève: Probability Theory, Van Nostrand, Princeton, 1960.

19. Łoś: Remarks on foundation of probability, Proceedings of 1962 International Congress of Mathematicians, Stockholm 1963, pp.225-229.

20. Luce, Raiffa: Games and Decisions, J. Wiley & Sons, New York, 1957.

21. Machol, Gray: Recent Developments in Information and Decision Processes, MacMillan, New York, 1962.

22. von Mieses: Wahrscheinlichkeit, Statistik und Wahrheit, Wien 1951.

23. von Mieses: Mathematical Theory of Probability and Statistics, Academic Press, New York, 1964.

24. Neyman: Two breakthroughs in the theory of statistical decision making, Review of the International Statistical Institute, vol.30, 1963.

25. Raiffa: Bayesian decision theory, In Machol, Gray, pp.92-101.

26. Raiffa, Schlaiffer: Applied Statistical Decision Theory, Harvard 1961.

27. Ramsey: Truth and probability, in Kyburg, Smokler, pp.61-92.

28. Reichenbach: The Theory of Probability, California University Press, 1949.

29. Renyi: On a new axiomatic theory of probability, Acta Math. Acad.Sci. Hungarica, vol.6, 1955, pp.285-335.

30. Robbins: The empirical Bayes approach to statistical decision problems, Annals of Math.Statist., vol.35, 1964, pp.1-20.

31. Savage: The Foundation of Statistics, J. Wiley & Sons, New York, 1954.

32. Savage: The Foundation of statistics reconsidered, Proceedings of Fourth Berkeley Symp., vol.1, 1961, pp.575-586. Reprinted in Kyburg, Smokler, pp.173-188.

33. Savage et al.: Statistical Inference, Methuen, London, 1962.

34. Savage: Bayesian statistics, in Machol, Gray, pp.161-194.

35. Suppes: The role of subjective probability in decision making, Proceedings of Third Berkeley Symp., vol.5, pp.61-73.

36. Trachtenbrot: Impossibility of an algorithm for the decision problem in finite classes (Russian). Dok.Ak.Nauk SSSR, n.s., vol.70, 1950, pp.569-572.

37. Wald: On the Principles of Statistical Inference, Notre Dame Math.Lectures, no.1, 1942.

38. Wald: Statistical Decision Functions, J. Wiley & Sons, New York, 1950.