Institute for Language, Logic and Information

THE AXIOMATIZATION OF RANDOMNESS

Michiel van Lambalgen

ITLI Prepublication Series
for Mathematical Logic and Foundations ML-89-06

University of Amsterdam

Instituut voor Taal, Logica en Informatie
Institute for Language, Logic and Information

Faculteit der Wiskunde en Informatica
(Department of Mathematics and Computer Science)
Plantage Muidergracht 24
1018TV Amsterdam

Faculteit der Wijsbegeerte
(Department of Philosophy)
Nieuwe Doelenstraat 15
1012CP Amsterdam

THE AXIOMATIZATION OF RANDOMNESS

Michiel van Lambalgen
Department of Mathematics and Computer Science
University of Amsterdam

Received October 1989

To appear in the Journal of Symbolic Logic
Research Supported by N.W.O.

The Axiomatization of Randomness

Michiel van Lambalgen[1]

Department of Mathematics and Computer Science
University of Amsterdam

Abstract We present a faithful axiomatization of von Mises' notion of a random sequence, using an abstract independence relation. A byproduct is a quantifier elimination theorem for Friedman's "almost all" quantifier in terms of this independence relation.

0. Introduction Characterizing randomness has always been understood as providing *explicit* definitions of what a random sequence is or should be. The literature contains examples using recursive place selections (CHURCH [1940]), recursive sequential tests (MARTIN-LÖF [1966]) and (variants of) Kolmogorov complexity (e.g. KOLMOGOROV and USPENSKY [1988]). Below, we present an axiomatic treatment of randomness. We believe that the notions that are customarily involved in characterizations of randomness, such as irregularity, complexity or independence, are best treated as *primitive,* in the sense that they need not have a unique, privileged, interpretation. These notions can sometimes be paraphrased, i.e. in some cases we can give explicit definitions of the primitive notions in terms of familiar mathematical concepts. But this does not mean that a particular paraphrase exhausts the content of randomness.

An example is the emphasis on computability in the modern definitions. Without further analysis (cf. KOLMOGOROV and USPENSKY [1988]) it is taken for granted that laws, rules or regularities as they occur in the context of randomness should be identified with rules as studied by recursion theory. By contrast, in our treatment we axiomatize one fundamental aspect of randomness (independence) and we then investigate models of these axioms. It turns out that a model for a sizable part of the axioms is definable in recursion theoretic terms (corresponding to the paraphrase of independence given in section 5 of VAN LAMBALGEN [1987a]). There exist, however, also models of an entirely different kind; and the completeness theorem given in section 2 points to a fundamental role for forcing rather than computability.

The object of the axiomatization is the notion of randomness introduced by VON MISES [1919].

According to his intuitive characterization, a sequence is random if (a) the limiting relative frequencies within the sequence exist and (b) these limiting relative frequencies are invariant under the operation of "admissible place selection". An admissible place selection is a procedure for selecting a subsequence of a given sequence x in such a way that the decision

[1] The author is indebted to Georg Kreisel, Ieke Moerdijk and Johan van Benthem for helpful comments. The author is a Huygens – fellow of the Netherlands Organization for Scientific Research (NWO).

to select a term x_n does not depend upon the outcome x_n itself. Evidently, independence is a fundamental concept here, so the main task confronting us will be the axiomatization of independence. We do not wish to imply that the resulting axiom system captures all there is to randomness; there is, e.g., also the only vaguely related idea that random sequences should satisfy "all" strong limit theorems of probability theory. But von Mises' axioms occupy a special position in that they describe the minimum properties randomness should satisfy to qualify as a foundation for probability theory (for more on this topic, see VAN LAMBALGEN [1987b]).

We now proceed to give a brief description of the contents of this paper. It consists of two parts: the first (sections 1 and 2) deals with the notion of independence and has a set theoretical emphasis, the second (sections 3,4 and 5) adds axioms for randomness and has a more probabilistic flavour.

In section 1 we state the axioms for independence and we sketch some interpretations. In section 2 we show that there is a close connection between the independence notion introduced in 1 and Friedman's "almost all" quantifier (STEINHORN [1985a,b]). We prove, using Boolean valued models and some measure theory, that any first order theory can be conservatively extended with the axioms for independence or, equivalently, the axioms for "almost all". We then proceed to show that the "almost all" quantifier can be eliminated in a first order theory which has an independence relation satisfying our axioms. This elimination allows us to prove that the theory of independence is complete with respect to a semantics given by forcing with Solovay reals.

In section 3 we present the formal analogues of von Mises' axioms for randomness and we give an informal argument to the effect that the axioms indeed say everything that can be said about admissible place selection. Section 4 contains the proof that the axioms for randomness and independence can be added consistently to second order arithmetic, thereby showing that a consistent probability theory on the basis of a notion of randomness is possible. In section 5 we investigate to what extent some familiar (explicit) definitions of randomness can be used to define models of the axioms; in particular, show that the theory of independence developed in VAN LAMBALGEN [1987a] (based on relativized prefix complexity) can be used to construct a model for (a considerable part of) the axioms.

1 Axioms for randomness and independence Von Mises proposed the following axioms for random infinite binary sequences. For ease of notation, we consider only the case of the uniform distribution on $\{0,1\}$, i.e. sequences generated by a fair coin.

Definition 1.1 A sequence $x \in 2^{\omega}$ is *random* if

(1) $\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{2}$

(2) Let $(x_{n_k})_k$ be an admissibly chosen subsequence of x, i.e. a selection of subsequence which proceeds as follows: "Aus der unendliche Folge [x wird] eine unendliche Teilfolge dadurch ausgewählt, daβ über die Indizes der auszuwählenden Elemente ohne Benützung der Merkmalunterschiede verfügt wird." Then also $\lim_{m\to\infty}\frac{1}{m}\sum_{k=1}^{m} x_{n_k} = \frac{1}{2}$.

The condition of admissibility is slightly enigmatic, so some examples may be helpful:
(a) Choose x_n if n is prime.
(b) Choose x_n if the n-9th,..., n-1st terms of x are all equal to 1.
(c) Take a second coin, supposed to be independent of the first in so far as that is possible (no strings connecting the coins, no magnetisation etc.) Then choose x_n if the outcome of the nth toss with the second coin is heads (we code "heads" as "1").
Intuitively, axiom (2) is satisfied in all three cases, although in (c) a heavy burden is put upon the word "independent". We shall call selections of type (a) and (b) *lawlike* (since they are given by some prescription) and those of type (c) *random*. Evidently combinations of the two types are possible. We shall prove below that, in a sense, *all* admissible place selections are combinations of random and lawlike selections (see theorem 3.3.5).
It will be observed that the second axiom consists in fact of two parts. The first part tries to explain admissibility in terms of some concept of independence ("ohne Benützung der Merkmalunterschiede"); the second part says that limiting relative frequencies are invariant under admissible place selection. We shall follow the same procedure: in 1.1, we first axiomatize a notion of independence; in 3.2, we then define admissible place selection in terms of this notion and state axioms on the invariance of the limiting relative frequencies.

1.1 Axioms for independence The literature contains various discussions of abstract independence relations; e.g. in algebra (VAN DER WAERDEN [1940]), combinatorial geometry (WELSH [1976]), recursion theory (METAKIDES and NERODE [1982]) and stability theory (BALDWIN [1988]). Another example is furnished by the theory of lawless sequences (TROELSTRA [1977]), where independence of lawless sequences is a fundamental concept, although it can be shown that in this case independence can be defined explicitly as extensional inequality $\neq$.
We show how to introduce an independence relation $R(x,\vec{y})$ in a countable first-order language L. Here, $\vec{y}$ denotes a vector, *of unspecified length*, of variables; hence R is a relation of indefinite arity. In analogy with linear algebra, the second parameter in $R(x,\vec{y})$ should be thought of as denoting a set; cf. axiom R3 below. (The introduction of R is of course equivalent to the introduction of infinitely many independence relations R_n, one for each arity n.) One may think of x, $\vec{y}$ as ranging over (vectors of) sequences in 2^ω, but the set up is in fact completely general. The intended interpretation of $R(x,\vec{y})$ is: "x is independent of $\vec{y}$ " or "$\vec{y}$ has no information about x". $\vec{y}$ may be empty; in that case we

write $R(x,\emptyset) =: R(x)$ and we may think of x with R(x) as random sequences (cf. also the fifth remark below).
All independence relations that arise naturally in our context satisfy the properties R1 - R5 (examples will be given below); furthermore we shall work in classical logic.

R0. Axioms and inference rules for classical predicate logic
R1. $\exists x R(x)$, $\forall \vec{y} \exists x R(x,\vec{y})$
R2. $R(x,\vec{y}\vec{z}) \rightarrow R(x,\vec{z})$
R3. $R(x,\vec{y}) \rightarrow R(x,\pi\vec{y})$ for any permutation π
R4. $\neg R(x,x\vec{y})$
R5. $R(y,\vec{z}) \wedge R(x,y\vec{z}) \rightarrow R(y,x\vec{z})$.

Before stating the last axiom, we give an example of an independence relation on infinite binary sequences that satisfies the axioms given so far.
Put $LLN(x) := \forall\varepsilon\exists n_0\forall n\geq n_0 \left|\frac{1}{n}\sum_{k=1}^{n} x_k - \frac{1}{2}\right| < \varepsilon$ and define a partial operation $/: 2^\omega\times 2^\omega \rightarrow 2^\omega$ by: $({}^x/y)_n = x_m$ if m is the index of the n^{th} 1 in y and undefined if there is no such index. (${}^x/y$ is defined if y has infinitely many 1's.) Then define $R(x,\vec{y})$ as: $LLN(x)$ & $\forall y \in \vec{y}$ (${}^x/y$ defined $\rightarrow LLN({}^x/y)$).

Definition 3.1.2 The partial operation $/: 2^\omega\times 2^\omega \rightarrow 2^\omega$ is defined by: $({}^x/y)_n = x_m$ if m is the index of the n^{th} 1 in y and undefined if there is no such index. ${}^x/y$ is defined if y has infinitely many 1's.

The last axiom, the first to establish a relation between the language L and the independence relation R, is less straightforward. It is suggested both by the connection of R to Friedman's "almost all" quantifier and by the example of Solovay forcing (see below) and leads to a completeness theorem for the whole system. To formulate this axiom, we first define a subclass of the formulas in the language $L\cup\{R\}$.

Definition 1.1.1 The class **IF** of *independence formulas* is the smallest class of formulas in the language $L\cup\{R\}$ such that
(i) if ϕ is a formula in L, then ϕ is in **IF**
(ii) **IF** is closed under $\wedge, \rightarrow, \forall$
(iii) if $\phi(x,\vec{y})$ is in **IF** (all free variables are indicated), then $\forall x(R(x,\vec{z}) \rightarrow \phi(x,\vec{y}))$ is also in **IF**, where $\vec{z}$ contains $\vec{y}$.

We are now in a position to state the last axiom, which is a kind of homogeneity principle.

R6. Suppose $\phi(x,\vec{y})$ is in **IF**, and z does not occur free in ϕ. Then

$$\forall x(R(x,z\,\vec{y}) \rightarrow \phi(x,\vec{y})) \rightarrow \forall x(R(x,\vec{y}) \rightarrow \phi(x,\vec{y})).$$

Some remarks on the meaning of these axioms follow. 1. Axiom R3 expresses that the argument $\vec{y}$ behaves as a set.

2. The axioms for independence R1 – R5 form a subsystem of the axioms for infinite dimensional combinatorial geometries. Usually these axioms are formulated in terms of a relation of *dependence*, but if we formulate them instead using the independence relation $R(x,\vec{y})$ we get a system which is almost equal to R1 – R5 (e.g. R5 corresponds to the Steinitz exchange principle), except that R2 is replaced by the stronger postulate

R2'. $R(x,\vec{z}) \rightarrow R(x,\vec{y}) \vee \exists y \in \vec{y} R(y,\vec{z})$.

The meaning of R2' becomes more transparent if we formulate it in terms of $\neg R$, or dependence; we then get

R2'. $\neg R(x,\vec{y}) \wedge \forall y \in \vec{y} \neg R(y,\vec{z}) \rightarrow \neg R(x,\vec{z})$,

in other words, the transitivity of the relation of dependence.

It might seem surprising at first sight that we do not require that the dependence relation corresponding to *our* notion of independence is transitive. The paradigm example of the dependence relation that is axiomatized by combinatorial geometry is linear dependence; so in this context, $\neg R(x,\vec{z})$ means that there exists a functional relationship between x and $\vec{z}$. However, our case is more akin to statistical dependence and allows the following situation: we have a random binary sequence y and subsequences x and z consisting of the odd and even coordinates of y respectively. Then for any reasonable concept of randomness, x and z are independent; but y depends on x as well as on z. Hence R2' fails.

3. Irreflexivity (R4) is slightly implausible from a purely probabilistic point of view; for instance if x represents a two-valued random variable which takes its values with probabilities 0 and 1, then x is (probabilistically) independent of x. We therefore think of our random sequences as sequences in which the limiting relative frequencies are distinct from 0 and 1.

4. The usual probabilistic definitions of independence are symmetric. However, in the presence of monotonicity (R2), the symmetry condition $R(x,y) \rightarrow R(y,x)$ is too strong, since we would then have $R(x,y) \rightarrow R(y)$.

5. The meaning of R6 will be clarified in the proof of the consistency theorem 2.1.3, where it is shown that R6 holds in virtue of the fact that random sequences are "indistinguishable". As an illustration of the meaning of R6 we shall show here that random sequences are not "nameable" (by closed terms of the language L).

The hypothesis that random sequences are not nameable and, even stronger, are independent of any nameable object, can be expressed by the formula

$$(*)\ R(x,\vec{y}) \to R(x,\vec{y}\,\vec{t})$$

where $\vec{t}$ is a sequence of closed terms of L.

A scheme version of (*) can easily be derived from R6: if $\phi(x,\vec{t},\vec{y})$ is an **IF** formula, $\vec{t}$ is a sequence of closed terms of L and $\vec{z}$ does not occur free in ϕ, then we have

$$\forall\vec{z}[\forall x(R(x,\vec{z}\,\vec{y}\,) \to \phi(x,\vec{t},\vec{y})) \to \forall x(R(x,\vec{y}) \to \phi(x,\vec{t},\vec{y}))]$$

hence in particular

$$\forall x(R(x,\vec{t}\,\vec{y}\,) \to \phi(x,\vec{t},\vec{y})) \to \forall x(R(x,\vec{y}) \to \phi(x,\vec{t},\vec{y})).$$

Therefore $R(x) \equiv R(x,\emptyset)$ should be interpreted as implying: x is independent of all objects denoted by closed terms of L. (This is also brought out by corollary 2.2.4 below.) Thus R6 embodies the intuition that random sequences (relative to L) cannot be constructed (with means available in L).

Henceforth, the system of axioms R0 - 6 will be called ***R***. The proof that these axioms are consistent is nontrivial, especially when we want to establish consistency of ***R*** with a given theory T (say, second order arithmetic); in other words, that the addition of ***R*** to a theory T is conservative. The consistency proof is deferred until section 2, where we study the system ***R*** by means of Friedman's theory of Borel structures and the quantifier "almost all".

Example 1.1.2 (a) (SOLOVAY [1970]) The paradigm example of an independence relation that satisfies ***R*** is provided by forcing. Let **M** be a countable transitive model for ZF plus the Axiom of Constructibility (V = L). For any sequence $\vec{y}$ of elements of 2^ω, $\mathbf{M}[\vec{y}\,]$ is well defined via relative constructibility.

We say that $x \in 2^\omega$ is *Solovay random* over $\mathbf{M}[\vec{y}]$ if x is contained in all Borel sets of full Lebesgue measure which have a Borel code in $\mathbf{M}[\vec{y}]$. Put $R(x,\vec{y})$ iff $x \in 2^\omega$ is Solovay random over $\mathbf{M}[\vec{y}]$. For this interpretation, R1 - 4 are trivially satisfied and R5 expresses the product lemma for forcing. The verification of R6 is less trivial; we need the machinery of Borel structures developed in the next section. (Cohen forcing (where "set of full measure" is replaced by "residual set") also furnishes an interpretation of R0 - 6, but is less suited to interpret the randomness axioms given in 3.2.)

(b) Let L be the language of second order arithmetic (with variables over ω and 2^ω). It is possible to establish the consistency of ***R*** by means of an interpretation into the intuitionistic theory of lawless sequences LS (for unexplained notions from intuitionism, see TROELSTRA and VAN DALEN [1988]).

If $\vec{y}$ denotes the vector $<y_1,..,y_n>$, $\#(x,\vec{y})$ is defined as: $x\neq y_1 \wedge .... \wedge x\neq y_n$. We interpret the variables x,y,.. as ranging over lawless sequences. Extend the Gödel - Gentzen negative translation * by means of the following clause for R: $(R(x,\vec{y}))^* := \#(x,\vec{y})$. Then $\boldsymbol{R} \vdash \phi$

implies LS $\vdash \phi^*$. All cases are easily checked, except axiom R6, for which we need the axioms of density and open data. In this way the consistency of $\underline{R}$ can be established, but since we want consistency of $\underline{R}$ over an arbitrary first order theory we present a different consistency proof below (2.1.4). The details of the extended Gödel - Gentzen interpretation will be given in VAN LAMBALGEN and MOERDIJK [199?].

2. Friedman's quantifier "almost all" It turns out that the axiom system $\underline{R}$ can be studied profitably in an alternative formulation, by adding the quantifier "almost all" introduced by Harvey Friedman (see STEINHORN [1985a,b]) to the language L. If Q denotes Friedman's quantifier, and ϕ is a formula of the language $L \cup \{Q\}$, then we can think of the formula $Qx\phi x$ as meaning "$\{x \mid \phi(x)\}$ has Lebesque measure 1". Alternatively, $Qx\phi x$ can be read topologically as "$\{x \mid \phi(x)\}^c$ is first category". Friedman exhibits a set $\underline{Q}$ of axioms for Q which is complete for both interpretations.

For us, the interesting point is that there is a close connection between $\underline{Q}$ and the axioms for independence listed in 1.1. Below, we shall define a translation * from $L \cup \{Q\}$ into $L \cup \{R\}$ such that ϕ is derivable in $\underline{Q}$ if and only if ϕ^* is derivable from $\underline{R}$. The advantages of this construction are threefold:

1. We obtain consistency proofs for $\underline{R}$ and $\underline{RS}$ (over a suitable base theory) by measure theoretic techniques;
2. Friedman's completeness proof for $\underline{Q}$ is transformed into a completeness result for $\underline{R}$, relating $\underline{R}$ to Solovay forcing;
3. It brings to the fore an intimate connection between measure on the one hand, and randomness and independence on the other.

2.1 The system $\underline{Q}$ The axioms for Q given below are taken from STEINHORN [1985a], except that Steinhorn uses the existential analogue of Q ("$\{x \mid \phi(x)\}$ is non - null") and does not add the nontriviality condition Q1. Again, we start with a countable first order language L.

Q0 Axioms and inference rules for classical predicate logic
Q1 $\neg Qx\ x \neq x$
Q2 $Qy\ x \neq y$
Q3 $Qx\phi(..,x,..) \rightarrow Qy\phi(..,y,..)$ provided y is free for x in ϕ
Q4 $Qx\phi \wedge \forall x(\phi \rightarrow \psi) \rightarrow Qx\psi$
Q5 $Qx\phi \wedge Qx\psi \rightarrow Qx(\phi \wedge \psi)$
Q6 $QxQy\phi \leftrightarrow QyQx\phi$

The existential analogues of axioms Q1 - 5 describe the quantifier "there are many" studied by KRIVINE and MCALOON [1973]. Q1 - 5 are clearly satisfied for the intended interpretation(s); but the consistency of the full system is a delicate matter. Without the non-triviality condition Q1, consistency is easily established: interpret $Qx\phi$ as $\forall x(\bot \rightarrow \phi)$, where $\bot$ denotes falsum. Q1, however, excludes this possibility. The intended interpretation of Q, be it measure theoretic or topological, is not very helpful in establishing consistency either. For if L is the language of second order arithmetic, L is sufficiently rich to define projective sets. But it is consistent to assume that some projective sets are non measurable and in that case we run into trouble with Q6, which is more or less a definable analogue of Fubini's theorem (but not quite, since in Fubini's theorem the measurability of $\{\langle x,y\rangle \mid \phi(x,y)\}$ is an *assumption*). For instance, assuming V = L there is a projective $\phi(x,y)$ defining the graph of a wellordering $<$ of 2^{ω} of length ω_1; but then we have $QxQy\phi(x,y)$ and $QyQx\neg\phi(x,y)$. Continuing this line of thought, one would seem to need a model in which all definable (with parameters) sets of reals are measurable; but this implies that ω_1 is inaccessible in the constructible universe, a most unwelcome assumption.

One way out of this predicament is to switch to the topological interpretation. Shelah has shown that if ZF is consistent, so is ZF + "every set of reals has the Baire property". We then get the consistency of $\underline{Q}$ by applying the Kuratowski - Ulam theorem, the topological analogue of Fubini's theorem (see theorem 15.1 in OXTOBY [1980]). While this result might set the reader's mind temporarily at rest, it is not very helpful in settling the consistency of additional assumptions with Q. For instance, we would like to have that QxLLN(x) is consistent with $\underline{Q}$; but of course this is false if Q is interpreted topologically (and LLN(x) has the standard interpretation).

There is, however, yet a third way to interpret Q, and that interpretation is the main reason for our interest in Q: read $Qx\phi(x,\vec{y})$ as "for all x *independent of* $\vec{y}$, ϕ". A little reflection shows that Q6 is now obvious (if x and y are supposed to be independent, then it does not matter which one is chosen first), whereas Q4 is more problematic. (This interpretation of Q is formalized by the elimination theorem 2.2.1, which shows that there is an embedding of $\underline{Q}$ into $\underline{R}$.) Furthermore, using a statistical concept of independence it becomes fairly easy to prove the consistency of both $\underline{Q}$ and $\underline{R}$ over arbitrary first order theories (2.1.3 and 2.1.4). The proofs use Friedman's theory of *Borel structures*.

Let **M** be a model for L, with domain a subset of 2^{ω} of positive Lebesgue measure. The satisfaction clause for Q is:

$$\mathbf{M} \models Qx\phi(x,a_1,\ldots,a_n) \text{ iff } \lambda\{x \mid \mathbf{M} \not\models \phi(x,a_1,\ldots,a_n)\} = 0, \text{ where } \lambda = (\tfrac{1}{2},\tfrac{1}{2})^{\omega} \text{ is}$$

Lebesgue measure on 2^{ω}.

The other logical constants have standard interpretation. A structure whose domain is a subset of 2^{ω} and all of whose relations and functions are Borel is called a *Borel structure* (or model). A Borel structure for $L\cup\{Q\}$ is called *totally Borel* if any relation defined using a formula from $L\cup\{Q\}$ (possibly with parameters) is Borel. Evidently the axioms for Q are valid on a totally Borel model. Henceforth we shall consider only totally Borel structures whose domain is a *full* subset of 2^ω; this is justified by the completeness theorem 2.1.2 below. Interestingly, Friedman has shown that if a theory in the language $L \cup \{Q\}$ is consistent with $\underline{Q}$, then it has a totally Borel model. It is somewhat hard to construct such models, since it is usually the case (e.g. for second order arithmetic) that the primitive relations must have a nonstandard interpretation. We therefore prefer a Boolean valued approach, where the primitive relations can be interpreted straightforwardly, while by contrast Q has a nonstandard interpretation, in terms of an independence relation.

Theorem 2.2.1 (See STEINHORN [1985a,b]) Let T be a consistent theory in L with an infinite model. Then T has a totally Borel model with domain 2^ω.

Theorem 2.1.2 (See STEINHORN [1985a]) Let T be a theory in $L \cup \{Q\}$. Then T has a totally Borel model whose domain is a full subset of 2^ω iff T is consistent in $\underline{Q}$.

Using 2.2.1 and a Boolean valued interpretation of $\forall$, $\exists$ and Q we show that any consistent theory T in L, with at least one infinite model, can be consistently extended with $\underline{Q}$, or, equivalently, with $\underline{R}$. (This technique will allow us to show the consistency of $\underline{RS}$ in section 3.1.) The proof of the consistency theorem proceeds by defining an independence relation that satisfies the Boolean valued analogue of $\underline{R}$ and showing that Q can be defined explicitly using this relation. This procedure suggests a quantifier elimination theorem for Q that will be stated and proved in 2.2.

Theorem 2.1.3 Let T be a consistent theory in L with an infinite model. Then $T \cup \underline{Q}$ is conservative over T.

Proof We construct a Boolean valued model on a universe defined as follows. (For general information on Boolean valued models, see BELL [1977].) Consider $(2^\omega)^\kappa$, where $\kappa \geq \omega_1$. We equip $(2^\omega)^\kappa$ with the product topology and the product measure λ^κ defined on the Borel σ - algebra $\mathcal{B}((2^\omega)^\kappa)$. Let I denote the σ - ideal of λ^κ nullsets, then the quotient algebra B:= $\mathcal{B}((2^\omega)^\kappa)/I$ is a complete Boolean algebra. B will be our space of truth values. The domain of the Boolean valued model consists of the (Borel) measurable functions f: $(2^\omega)^\kappa \to 2^\omega$.

Let **A** be a totally Borel model for T with universe 2^ω. If ψ is in L *and quantifier free*, we put

$$[[\,\psi(f_1,\dots,f_n)]]_{\mathcal{B}} = \{\, \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \psi(f_1(\xi),\dots,f_n(\xi))\}/I;$$

the interpretation of the Boolean connectives is obvious and $\forall$ and $\exists$ are interpreted using the infimum $\bigwedge$ and the supremum $\bigvee$, respectively.

On the space of (Borel) measurable functions $(2^\omega)^\kappa \to 2^\omega$ there are several natural independence relations that can be used to interpret Q. We choose the following option (but see remark 1 after corollary 2.1.4). Say that $\alpha < \kappa$ is in the support of f (denoted $\alpha \in$ supp(f)) if there are $\xi, \xi' \in (2^\omega)^\kappa$ which differ only at coordinate α and such that $f(\xi) \neq f(\xi')$. If f is Borel measurable, supp(f) is countable. For $\alpha < \kappa$, let $\pi_\alpha: (2^\omega)^\kappa \to 2^\omega$ denote the projection on the α^{th} coordinate. It now makes sense to define a relation (of indefinite arity) $R(g, f_1,...,f_n)$ as follows:

$R(g, f_1,...,f_n)$ if there is $\alpha < \kappa$ such that $g = \pi_\alpha$ and $\alpha \notin \text{supp}(f_1) \cup ... \cup \text{supp}(f_n)$.

We now define $[[Qx\psi(x,f_1,...,f_n)]]_B$ as

$$[[Qx\psi(x,f_1,...,f_n)]]_B = \bigwedge \{[[\psi(g,f_1,...,f_n)]]_B \mid R(g, f_1,...,f_n)\},$$

and we have to show that the Boolean truth value of the sentences T and the axioms of $\underline{Q}$ is 1_B.

1. T is valid in B. Since **A** is a model for T, it suffices to show that for *any* formula ψ in L,

$[[\psi(f_1,...,f_n)]]_B = \{ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \psi(f_1(\xi),...,f_n(\xi))\}/I$.

The proof is by induction, and the only case of interest is where ψ is $\exists x\phi$. By the maximum principle, $[[\exists x\phi(x,f)]]_B = \bigvee_g [[\phi(g,f)]]_B = [[\phi(g,f)]]_B$, for some particular measurable g: $(2^\omega)^\kappa \to 2^\omega$.

Hence, by the induction hypothesis $[[\phi(g,f)]]_b = \{ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \phi(g(\xi),f(\xi))\}/I \leq$
$\{ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \exists x\phi(x,f(\xi))\}/I$.

For the converse we need a uniformity principle, the von Neumann measurable selection theorem (see, e.g., PARTHASARATHY [1972]).

Since $\{ \langle x,\xi\rangle \in 2^\omega \times (2^\omega)^\kappa \mid \mathbf{A} \models \phi(x,f(\xi))\}$ is Borel, we can apply the selection theorem to obtain a uniformizing measurable g: $(2^\omega)^\kappa \to 2^\omega$ which satisfies:

for λ^κ a.a. $\xi \in (2^\omega)^\kappa$, $\mathbf{A} \models \phi(g(\xi),f(\xi))$ iff $\mathbf{A} \models \exists x\phi(x,f(\xi))$.

We then have $\{ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \exists x\phi(x,f(\xi))\}/I = \{ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \phi(g(\xi),f(\xi))\}/I = [[\phi(g,f)]]_B \leq [[\exists x\phi(x,f)]]_B$.

2. $\underline{Q}$ is valid in B. All axioms are trivially satisfied except (somewhat surprisingly) Q4. We have to show that

$$[[Qx\phi(x,g)]]_B \wedge \bigwedge_f ([[\phi(f,g)]]_B \to [[\psi(f,h)]]_B) \leq [[Qx\psi(x,h)]]_B .$$

Evidently, it suffices to show that if g does not occur in ψ,

$$\bigwedge \{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g) \cup \text{supp}(h)\} = \bigwedge \{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(h)\}.$$

For notational convenience we do not distinguish between an element of the Boolean algebra and a representative Borel set. Choose $\alpha < \kappa$ with $\alpha \notin$ supp(h). Let E be a Borel set such that the support of the characteristic function of E is contained in supp(h) and such that

(*) $\lambda^\kappa([[\psi(\pi_\alpha,h)]]_B \mid E) = 1$, $\lambda^\kappa([[\psi(\pi_\alpha,h)]]_B \mid E^c) < 1$.

That such an E exists can be seen as follows. We may consider $[[\psi(\pi_\alpha,h)]]_B$ to be a Borel subset of $(2^\omega)^\omega \times (2^\omega)^\omega$, where h is supported by the second coordinate and π_α lives on the first coordinate. (One has to be careful here: usually (a representative of) $[[\psi(g,h)]]_B$ is *not* measurable with respect to the σ - algebra generated by g and h. This is the analogue of the nonmeasurability of $\{<x,y> \mid \psi(x,y)\}$ in the Boolean valued case.) By differentiating $(\lambda^\omega)^2$, we obtain the conditional probability $(\lambda^\omega)^2(F \mid h)$, where F ranges over Borel subsets of $(2^\omega)^\omega$. Then define E as

$$E := \{ y \in (2^\omega)^\omega \mid (\lambda^\omega)^2([[\psi(\pi_\alpha,h)]] \mid h(y)) = 1\}.$$

E is determined up to a nullset and its support is contained in supp(h). Furthermore we may assume that $1 > \lambda^\kappa E > 0$, for otherwise we are done. Hence the conditional probabilities in (*) are welldefined.

Obviously (*) is true for *any* $\alpha \notin$ supp(h), since any two such projections can be transformed into each other by an automorphism of $<(2^\omega)^\kappa, B((2^\omega)^\kappa), \lambda^\kappa>$ that is the identity on supp(h).

It follows that for any $\alpha \notin$ supp(h), $E \le [[\psi(\pi_\alpha,h)]]_B$.

By distributivity, we have $\bigwedge\{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g)\cup\text{supp}(h)\} =$

$(\bigwedge\{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g)\cup\text{supp}(h)\} \wedge E) \vee (\bigwedge\{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g)\cup\text{supp}(h)\} \wedge E^c)$. However, the second disjunct must equal zero: let $\pi_{\alpha_1},\ldots, \pi_{\alpha_n},\ldots$ be a countable set of projections on coordinates not contained in supp(g)$\cup$supp(h), then by independence, $0_B = \bigwedge_n [[\psi(\pi_{\alpha_n},h)]]_B \wedge E^c \ge \bigwedge\{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g)\cup\text{supp}(h)\} \wedge E^c$.

Therefore $\bigwedge\{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g)\cup\text{supp}(h)\} = (\bigwedge\{[[\psi(\pi_\alpha,h)]]_B \mid \alpha \notin \text{supp}(g)\cup\text{supp}(h)\} \wedge E) \le E \le [[\psi(\pi_\alpha,h)]]_B$, for any $\alpha \notin$ supp(h). □

Since the proof of 2.1.3 in effect constructs an explicit independence relation satisfying R0 - 6, we get

Corollary 2.1.4 Let T be a consistent theory in L with an infinite model. Then T$\cup$*R̲* is conservative over T.

Remarks

1. While the above simple interpretation of R works for the purpose of showing consistency, it has an important drawback. Since mixtures (see BELL [1977]) of projections are not themselves projections, the Boolean valued model for T$\cup$*R̲* constructed above does not satisfy the maximum principle: if ϕ is a formula in L$\cup$\{R\}, then we do not necessarily have $[[\exists x\phi(x,f)]]_b = [[\phi(g,f)]]_b$, for some measurable $g: (2^\omega)^\kappa \to 2^\omega$. This means that the (Boolean valued) universe of random "sequences" is not closed under nontrivial operations. For instance, a statement like $\forall x\forall y(Rx \wedge R(x,y) \to \exists z(z = {}^x/_y \wedge Rz))$ is false (observe that

this statement is not an IF formula). Models which do satisfy the maximum principle will be studied in VAN LAMBALGEN and MOERDIJK [199?].

2. It follows from 2.1.4 that $\underline{R}$ has no consequences in, e.g., set theory. This contrasts with the approach taken by FREILING [1986], whose axioms for randomness contradict the continuum hypothesis in the form $2^{\aleph_0} = \aleph_1$ (albeit not in the aleph free formulation). This difference should be no cause for surprise, since Freiling's axioms use abstract concepts like "countable" and "nullset", whereas our axioms only involve the far more concrete notion of independence.

3. It is perhaps worth emphasizing that while the monotonicity axiom Q4 is innocuous on the "almost all" interpretation of Q, it is far from obvious on the "independence" interpretation. On the latter interpretation, Q4 turns out to be true because an independence structure has many symmetries.

2.2 Eliminating Q We define a translation * of $L \cup \{Q\}$ into $L \cup \{R\}$ as follows:

* is the identity on L formulas
* commutes with $\wedge$, $\neg$, $\forall$
$(Qx\phi(x,\vec{y}))^* := \forall x(R(x,\vec{y}) \rightarrow \phi(x,\vec{y})^*)$.

Theorem 2.2.1 * is a faithful relative interpretation of $L \cup \{Q\}$ into $L \cup \{R\}$, i.e. for all ϕ in $L \cup \{Q\}$: $\underline{Q} \vdash \phi$ iff $\underline{R} \vdash \phi^*$.

Proof. $\Rightarrow$ R1 ensures that * is correctly defined as a relative interpretation. The remainder of the argument proceeds by a routine induction on the length of proofs in $\underline{Q}$. That the axioms for Q are derivable in $\underline{R}$ can be seen as follows:

- $(\neg Qx\ x \neq x)^* = \neg\forall x(R(x) \rightarrow x \neq x) \equiv \exists xR(x)$, hence Q1 corresponds to R1 under *
- $(\forall xQy\ x \neq y)^* = \forall x\forall y(R(y,x) \rightarrow x \neq y) \equiv \forall x \neg R(x,x)$ hence Q2 corresponds to R4 under *
- Q3 holds trivially
- $(Qx\phi(x,\vec{y}) \wedge \forall x(\phi(x,\vec{y}) \rightarrow \psi(x,\vec{z})) \rightarrow Qx\psi(x,\vec{z}))^* =$
  $\forall x(R(x,\vec{y}) \rightarrow \phi(x,\vec{y})^*) \wedge \forall x(\phi(x,\vec{y})^* \rightarrow \psi(x,\vec{z})^*) \rightarrow \forall x(R(x,\vec{z}) \rightarrow \psi(x,\vec{z})^*)$;
  the antecedent implies that $\forall x(R(x,\vec{y}\vec{z}) \rightarrow \psi(x,\vec{z})^*)$, hence by R6 also $\forall x(R(x,\vec{z}) \rightarrow \psi(x,\vec{z})^*)$
- $(Qx\phi(x,\vec{y}) \wedge Qx\psi(x,\vec{z}) \rightarrow Qx(\phi(x,\vec{y}) \wedge \psi(x,\vec{z})))^* =$
  $\forall x(R(x,\vec{y}) \rightarrow \phi(x,\vec{y})^*) \wedge \forall x(R(x,\vec{z}) \rightarrow \psi(x,\vec{z})^*) \rightarrow$
  $\forall x(R(x,\vec{y}\vec{z}) \rightarrow \phi(x,\vec{y})^* \wedge \psi(x,\vec{z})^*)$, hence (Q5)* can be derived in $\underline{R}$ using R2
- $(QxQy\phi(x,y,\vec{z}) \leftrightarrow QyQx\phi(x,y,\vec{z}))^* =$
  $\forall x\forall y(R(x,\vec{z}) \wedge R(y,x\vec{z}) \rightarrow \phi(x,y,\vec{z})^*) \leftrightarrow \forall y\forall x(R(y,\vec{z}) \wedge R(x,y\vec{z}) \rightarrow \phi(x,y,\vec{z})^*)$,
  hence (Q6)* can be derived in $\underline{R}$ using R5 and R2.

The induction step is trivial, since both in the case of $\underline{Q}$ and $\underline{R}$ the inference rules are those of classical predicate logic.

$\Leftarrow$ We present two proofs. The first proof leads to a two valued interpretation for R and a completeness theorem for $\underline{R}$ with respect to a forcing semantics. The second proof exploits the construction in the proof of 2.1.3.

First proof. Suppose that $\underline{Q} \nvdash \phi$, then by the completeness theorem 2.1 there exists a totally Borel model **A** for $\underline{Q} \cup \{\neg\phi\}$. We may assume that the domain of **A** is a full subset of 2^ω. Let **M** be a model of ZF + V=L. We expand **A** to a structure <**A**,R> by defining: $R(x,\vec{y})$ iff $x \in 2^\omega$ is Solovay random over $\mathbf{M}[\vec{y}\,]$ and we show that <**A**,R> satisfies $\underline{R}$ and $\neg\phi^*$. We observed in example 1.1.2(a) above that <**A**,R> satisfies R0 - 5. To show the validity of R6 and $\neg\phi^*$, we need a

Lemma 2.2.2 For all ψ and $\vec{y}$: $\mathbf{A} \models \psi(\vec{y})$ iff $<\mathbf{A},R> \models \psi^*(\vec{y})$.

Proof of lemma The cases of the ordinary logical operations are trivial, since * commutes with these. So let $\psi(\vec{y})$ be of the form $Qx\theta(x,\vec{y})$; the induction hypothesis gives:

$$\mathbf{A} \models \theta(\vec{y}) \text{ iff } <\mathbf{A},R> \models \theta^*(\vec{y}).$$

Assume $\mathbf{A} \models Qx\theta(x,\vec{y})$; we then have $\lambda\{\, x \mid \mathbf{A} \models \theta(x,\vec{y})\} = 1$. But since $\{\, x \mid \mathbf{A} \models \theta(x,\vec{y})\}$ is a Borel set with code in $\mathbf{M}[\vec{y}]$, we have, by the definition of Solovay randomness: $\forall x(R(x,\vec{y}) \Rightarrow \mathbf{A} \models \theta(x,\vec{y}))$.

Hence by the induction hypothesis $<\mathbf{A},R> \models \forall x(R(x,\vec{y}) \to \theta^*(x,\vec{y}))$.

On the other hand, if $\mathbf{A} \not\models Qx\theta(x,\vec{y})$, then $\lambda\{\, x \mid \mathbf{A} \models \theta(x,\vec{y})\} < 1$. Since the set of Solovay random reals over $\mathbf{M}[\vec{y}]$ has measure 1, we have $\exists x(R(x,\vec{y})\ \&\ \mathbf{A} \models \neg\theta(x,\vec{y}))$, whence by the induction hypothesis $<\mathbf{A},R> \models \exists x(R(x,\vec{y}) \wedge \neg\theta^*(x,\vec{y}))$. The lemma is proved.

We are now in a position to verify the validity of R6 in the structure <**A**,R>. Suppose $<\mathbf{A},R> \models \forall x(R(x,\vec{y}z) \to \psi(x,\vec{y}))$, where ψ is an **IF** formula and z is not free in ψ. Let θ be the formula such that $\theta^* = \psi$. Then by the previous lemma for all x such that $R(x,\vec{y}z)$: $\mathbf{A} \models \theta(x,\vec{y})$. Since $\lambda\{x \mid R(x,\vec{y}z)\} = 1$, it follows that $\mathbf{A} \models Qx\theta(x,\vec{y})$, whence again by the lemma $<\mathbf{A},R> \models (Qx\theta(x,\vec{y}))^*$; and this is equivalent to $<\mathbf{A},R> \models \forall x(R(x,\vec{y}) \to \psi(x,\vec{y}))$. Hence <**A**,R> satisfies $\underline{R}$ and $\neg\phi^*$.

Second proof. As above, let **A** be a totally Borel model for $\underline{Q} \cup \{\neg\phi\}$. As in the proof of the consistency theorem 2.1.3 , we lift **A** to a Boolean valued model on $(2^\omega)^\kappa$, where $\kappa \geq \omega_1$.

Let B be the complete Boolean algebra on $(2^\omega)^\kappa$, where the latter set is equipped with product topology and product Lebesque measure λ^κ. Let I be the σ - ideal of λ^κ nullsets.

If ψ is a quantifier free formula and $f_1, ..., f_n$ a sequence of measurable functions: $(2^\omega)^\kappa \to 2^\omega$, then $[[\ \psi(f_1,...,f_n)]]_B = \{\ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \psi(f_1(\xi),...,f_n(\xi))\}/I$. The quantifiers $\forall$ and $\exists$ are interpreted using $\bigwedge$ and $\bigvee$ respectively.

In the proof of 2.1.3 , we saw that $[[\ Qx\psi(x,f_1,...,f_n)]]_B$ could be interpreted as

$$[[\ Qx\psi(x,f_1,...,f_n)]]_B = \bigwedge_g \{[[\ \psi(g,f_1,...,f_n)]]_B \mid R(g, f_1,...,f_n)\},$$

where $R(g, f_1,...,f_n)$ is defined as

$R(g, f_1,...,f_n)$ if there is $\alpha < \kappa$ such that $g = \pi_\alpha$ and $\alpha \notin \mathrm{supp}(f_1)\cup...\cup \mathrm{supp}(f_n)$, where $\pi_\alpha: (2^\omega)^\kappa \to 2^\omega$ denotes the projection on the α^{th} coordinate.

Hence if we put $[[R(g, f_1,...,f_n)]]_B = 1_B$ if $R(g, f_1,...,f_n)$ and equal to 0_B otherwise, we get

$$[[\ Qx\psi(x,f_1,...,f_n)]]_B = \bigwedge_g([[R(g, f_1,...,f_n)]]_B \to [[\ \psi(g,f_1,...,f_n)]]_B\ ).$$

Applying the von Neumann measurable selection theorem as in the proof of 2.1.3 , we obtain for *any* formula ψ in $L\cup\{Q\}$: $[[\ \psi(f_1,...,f_n)]]_B = \{\ \xi \in (2^\omega)^\kappa \mid \mathbf{A} \models \psi(f_1(\xi),...,f_n(\xi))\}/I$.

An easy induction then shows that for ψ in $L\cup\{Q\}$, $[[\ \psi(f_1,...,f_n)]]_B = [[\ \psi^*(f_1,...,f_n)]]_B$.

Hence we have obtained a B - valued model of $\underline{R}$ and $\neg\phi^*$. □

In fact, the second proof of theorem 2.2.1 proves a slightly stronger result.

Corollary 2.2.3 The addition of **R7**: $R(x,\vec{z}) \wedge R(x,\vec{y}) \to R(x,\vec{y}\vec{z})$ is conservative over $\underline{R}$ for **IF** formulas, i.e. it leads to no new identities for Q.

Proof The R defined in the second proof satisfies R7. □

Corollary 2.2.4 The addition of **R8**: $R(x,\vec{y}) \to R(x,\vec{y}\vec{t})$ (where $\vec{t}$ is a sequence of closed terms of L) is conservative over $\underline{R}$ for **IF** formulas, i.e. it leads to no new identities for Q.

Proof The R defined in the first proof satisfies R8. □

Remarks 1. Let **N** be a generic extension of **M** obtained by generically adding ω_2 Solovay random reals. Using the absoluteness properties of Borel sets, we can perform the construction of (the first proof of) theorem 2.2.1 inside **N**, in this way obtaining a countable standard interpretation of R.

2. Although we shall not pursue the matter here, we remark in passing that theorem 2.2.1 can be used to set up a natural deduction system for Q. In natural deduction systems for the predicate calculus, the rules for introduction of ∀ and elimination of ∃ have side conditions (on the non-occurrence of variables or constants) which are primitive recursive. By contrast, both the introduction and elimination rules for Q have also side conditions in terms of an independence relation R satisfying $\underline{R}$, but otherwise unspecified. In a sense, then, the rules for Q are more complicated than the rules for ∀ and ∃. This fact gives some substance to the intuition that probabilistic reasoning (for Q read: "almost surely") is more complex than ordinary logical reasoning. The subject will be explored in greater detail in VAN LAMBALGEN and MOERDIJK [199?].

2.3 Completeness The elimination theorem 2.2.1, together with the Borel completeness theorem 2.1.2 imply the following completeness theorem for R:

Theorem 2.3.1 Let T be a consistent theory in $L\cup\{R\}$ and suppose the formulas of T are either R axioms or IF formulas. Let **M** be a countable model of ZF + V = L. If T has an infinite model, than it has a (not necessarily totally) Borel model in which R is interpreted as Solovay forcing over **M**.

Proof By hypothesis, T can be identified with a theory in $L\cup\{Q\}$. This theory has a totally Borel model **A** by 2.1.2, and we may apply the construction of 2.2.1 to obtain the desired interpretation of R. Since **M** is countable, R is Borel, hence $\langle A,R\rangle$ is a Borel model for T. Lemma 2.2.2 shows that $\langle A,R\rangle$ is totally Borel with respect to IF formulas. □

3. Randomness Having obtained a theory of independence, we may now formalize von Mises' axioms for random sequences as stated in definition 1.1. We do this in two stages: first we give a simplified version which highlights the role of independence (3.1) and then we present the full version (3.2). In nsection 3.3 we consider the resulting system RS from the viewpoint of stochastic processes.

3.1 Axioms for randomness: simplified version In this section, L is a language for second order arithmetic with function variables. In this language, we can express that a sequence satisfies the law of large numbers:

Definition 3.1.1 $LLN(x) := \forall\varepsilon\exists n_0\forall n\geq n_0 \left|\frac{1}{n}\sum_{k=1}^{n} x_k - \frac{1}{2}\right| < \varepsilon.$

We could of course define a more general predicate LLN(x,p), but we take $p = \frac{1}{2}$ throughout for notational convenience. To formulate place selection, we have to introduce a defined operation in the language.

Definition 3.1.2 The partial operation $/: 2^\omega\times 2^\omega \to 2^\omega$ is defined by: $({}^x/y)_n = x_m$ if m is the index of the n^{th} 1 in y and undefined if there is no such index. ${}^x/y$ is defined if y has infinitely many 1's.

Von Mises first axiom is expressed by

RS1 $\forall x(R(x) \to LLN(x))$

We first give a weak form of the second axiom, to emphasize the role of independence.

RS2' $\forall x\forall y\forall z(R(x) \wedge R(x,y) \wedge {}^x/y = z \to LLN(z))$

Observe that by axiom R2, R(x) can be dropped. This axiom generalizes the case that LLN($^x/y$) for random x and *recursive* y; here, we require only that y has no information about x. Note that RS2' already implies (using R5)

$$\forall x \forall y (R(x) \wedge R(y,x) \rightarrow LLN(^x/y)),$$

which is the formal expression of the third example of admissible place selection (selection by means of a second coin independent of the first).

3.2 Axioms for randomness Von Mises description of admissible place selection: "Auswahl ohne Benützung der Merkmalunterschiede" is purely negative, hence difficult to formalize. The main idea of the axiomatization given here is that the negative condition "the decision to select x_n does not depend upon the value of x_n" is replaced by the positive condition "the decision to select x_n depends on (at most) x(n–1) and (perhaps) on some data which *do not contain information about* x". The italicized phrase can be made precise by means of the independence relation introduced in 1.1.

In other words, admissible place selection can be viewed as a "machine" which takes as input the random sequence x and additional (*independent*) information represented by an infinite binary sequence z; this machine produces an output y in such a way that the "computation" of y(n) uses at most the first n–1 bits of x, and an arbitrary, but finite, number of bits of z. The output y determines the subsequence $^x/y$ of x. We may think of the oracle z as produced by a stochastic mechanism such as coin tossing, but there are other possibilities, e.g. choice sequences.

The essential property of these "machines" is isolated in part (1) of the following definition:

Definition 3.2.1 A (possibly partial) continuous function F: $2^\omega \times 2^\omega \rightarrow 2^\omega$ is *admissible* if (1) for all n and all x, x' and y: x(n–1) = x'(n–1) and F(x,y), F(x',y) defined implies F(x,y)(n) = F(x',y)(n) and (2) $\lambda\{ y \mid F(x,y) \text{ defined}\} = 1$ for all x, where $\lambda = (\frac{1}{2},\frac{1}{2})^\omega$ is Lebesgue measure on 2^ω.

Part (2) of the definition might seem strange at first sight, but it is not difficult to see that an F not satisfying (2) can be extended to an F' which does satisfy (2) and such that range (F) = range (F').

The three examples of admissible selection mentioned after definition 1.1 can be represented as admissible functions. There exist, however, more general ways to construct admissible place selections.

Let ϕ: $2^{<\omega} \rightarrow \{0,1\}$ be any function, then define an admissible function F: $2^\omega \times 2^\omega \rightarrow 2^\omega$ as follows:

$$F(x,z)_n = \begin{cases} 0 & \text{if } \phi(x(n\text{-}1)) = 0 \\ z_n & \text{otherwise} \end{cases}$$

F may be used to represent the following process: to decide whether to choose x_n, we first "compute" $\phi(x(n-1))$; if the result is 1, we then toss a fair coin and choose x_n if the outcome is heads. The resulting subsequence of x is $^x/F(x,z)$ (when F(x,z) has infinitely many 1's) and the selection procedure may be viewed as a combination of lawlike and random selection. Intuitively it seems clear that the combination of an admissible function with some mechanism generating y is a general description of admissible place selection, since all information additional to x can be coded into one infinite binary sequence y. For instance, it would not make a difference if the decision to select x_n depends on tosses with several coins. There is, however, one possibility of generalizing the description, by varying the order in which the x_n are chosen, so that, for some $k > n$, x_k may be chosen before x_n. We shall not consider this possibility here. Barring this possibility, we have in fact a kind of completeness theorem: in a sense to be made precise in section 3.3, *any* admissible selection procedure can be seen as a combination of an admissible function together with a random oracle.

To formulate the final axiom, we need a slight extension of the language of second order arithmetic used in the previous subsection. In addition to the sorts for ω and 2^ω we use a subsort of 2^ω consisting of the "lawlike" or "given" elements of 2^ω. One may think of the latter sort as being of the form $2^\omega \cap M$, where M is a countable model of second order arithmetic. Corresponding to these three sorts, we have quantifiers $\forall n,m,..$ over the natural numbers, $\forall x,y,..$ over 2^ω and $\forall a$, $\forall F$ over the lawlike objects. The choice of this language is intended to bring out that we regard admissible selection procedures as "given", whereas random sequences are considered to be "incomplete" objects (cf. also the Boolean valued model constructed in 3.1).
Observe that admissible functions in the sense of 1.3.1, being continuous, can be coded into an infinite binary sequence; furthermore the condition of admissibility itself can be formulated in second order arithmetic. We may therefore state the final axiom as follows:

RS2 $\forall F(F \text{ admissible} \rightarrow \forall x \forall y \forall z \forall z'(F(x,z) = y \wedge R(x,z) \wedge {}^x/y = z' \rightarrow LLN(z')))$.
The system consisting of the axioms R0-6, RS1-2 will be denoted <u>*RS*</u>.

3.3 Admissible place selections as stochastic processes

In 3.2 we gave an informal argument to the effect that admissible place selection can be represented by an admissible function in the sense of definition 3.2.1, together with an oracle satisfying an independence condition. Here, we prove a kind of informal completeness theorem: we show that, *in so far as admissible place selection can be represented as a stochastic process*, RS2 is indeed an exhaustive description of the situation.
An admissible selection procedure, viewed as a stochastic process, can be represented as a measure on 2^ω. Because the condition of admissibility entails the dependence of the measure

on the random sequence x, the admissible place selection must in fact be represented as a parametrized measure.

Def. 3.3.1 For $w \in 2^{<\omega}$, $[w] := \{x \mid x(|w|) = w\}$, where $|w|$ is the length of w.
Def. 3.3.2 A *random measure* on 2^ω is a family of finite measures $\{\mu_x \mid x \in 2^\omega\}$, where μ_x is defined on 2^ω, such that for each w, the function $x \to \mu_x[w]$ is Borel measurable.

There seems to be only one reasonable definition for a measure to represent an admissible selection process, if the decision to select x_n is determined only by past values of x.

Def. 3.3.3 A family of probability measures $\{\nu_x \mid x \in 2^\omega\}$ is called *admissible* if for all x and x' and all n: if $x(n-1) = x'(n-1)$, then for all w, $|w| = n$: $\nu_x[w] = \nu_{x'}[w]$. (If a family of measures $\{\nu_x \mid x \in 2^\omega\}$ is admissible, we have automatically that for all w, the function $x \to \nu_x[w]$ is Borel measurable. Hence admissibility is a strengthening of the random measure condition.)

The intuitive picture behind the definition is that the properties of the process at time n are completely determined by the values of x up to time n–1. A few examples will make this clearer.

Example 3.3.4 The admissible place selections mentioned after definition 1.1 can be represented by admissible random measures:
(a) ν_x is concentrated on the sequence $z \in 2^\omega$ defined by: $z_n = 1$ iff n is prime (that is, $\nu_x[z(n)] = 1$ for all n and x);
(b) ν_x is concentrated on the sequence $z^x \in 2^\omega$ defined by: $z^x_n = 1$ iff $x_{n-9},\ldots\ldots, x_{n-1}$ are all equal to 1 (this means that $\nu_x[z^x(n)] = 1$ for all n and x);
(c) $\nu_x = \lambda$ for all x.
(d) More interesting examples are constructed as follows. If F is an admissible function in the sense of 3.2.1, then $\{\nu_x \mid x \in 2^\omega\}$ defined by $\nu_x := \lambda F_x^{-1}$ is an admissible random measure. (There is nothing special about λ here; any probability measure would do.)

The next theorem is a converse to example 3.3.4 (d).

Theorem 3.3.5 Let $\{\nu_x\}$ be an admissible random measure. Then there exists an admissible function $F: 2^\omega \times 2^\omega \to 2^\omega$ such that $\nu_x = \lambda F_x^{-1}$.

Proof. By admissibility, $\nu_x[0]$ and $\nu_x[1]$ do not depend on x. Choose disjoint open sets $A_0^{<>}$, $A_1^{<>}$ such that $\lambda A_0^{<>} = \nu_x[0]$ and $\lambda A_1^{<>} = \nu_x[1]$. (One of the $A_0^{<>}$, $A_1^{<>}$ may be empty, namely when either $\nu_x[0]$ or $\nu_x[1]$ equals zero.) Then F will be defined such that $\forall i \forall x \forall y \in A_i^{<>}$ $(F(x,y) \in [i])$. Again by admissiblity, for $|w| = 2$, $\nu_x[w]$ depends only on x_1. Choose disjoint open subsets A^i_{j0}, A^i_{j1} of $A_j^{<>}$ such that $\lambda A^{x_1}_{jk} = \nu_x[jk]$, for i,j,k in $\{0,1\}$. F will be defined such that $\forall i \forall x \in [i] \forall w(|w| = 2 \rightarrow \forall y \in A^i_w\ (F(x,y) \in [w]))$.Continuing in this way, we construct a sequence of open sets $\{A^v_w \mid v,w \in 2^{<\omega}, |v| = |w|-1\}$ with the following properties:

1. For $i \in \{0,1\}$ and $|v| = |w|-1$, A^{vi}_{w0}, A^{vi}_{w1} are disjoint open subsets of A^v_w;
2. If $m = |w|-1$, $\lambda A^{x(m)}_w = \nu_x[w]$.

Put $W := \{<x,y,w> \mid \exists n\ (y \in A^{x(n)}_w)\}$, then we can define F by:

$$F(x,y) = z \equiv \forall n\ (<x,y,z(n)> \in W).$$

Obviously F is continuous on a domain of full measure and satisfies

3. $\forall v \forall x \in [v] \forall w(|w| = |v|+1 \rightarrow \forall y \in A^v_w\ (F(x,y) \in [w]))$.

It follows from 3. that F is admissible. □

The preceding theorem shows that the examples of admissible place selection given after definition 1.1 in a sense constitute a complete description: every admissible selection can be viewed as a combination of lawlike and random selection. It also follows from the theorem that the oracle can be taken to be independent of the random sequence in a strong sense: we can think of the oracle as being generated by a fair coin. It must be emphasized, however, that this is true only on the assumption that admissible selections can be represented stochastically.

4 Consistency of <u>*RS*</u> in second order arithmetic A probability theory in the style of von Mises consists of a theory of random sequences together with a base theory which allows one to manipulate infinite sequences and real numbers. For the base theory one may take second order arithmetic.

Consider again the three sorted language L of section 3, with sorts for ω, 2^ω and the lawlike elements of 2^ω. In this language, second order arithmetic Z_2 has the following form:

- basic arithmetic, i.e. defining equations for 0,S,+ and $\cdot$
- an induction scheme: $\phi(0) \wedge \forall n(\phi(n) \rightarrow \phi(n+1)) \rightarrow \forall n \phi(n)$, where ϕ is a formula of L

- 2 comprehension schemes
  - $\exists a \forall n(a_n = 1 \leftrightarrow \phi(n))$ if ϕ has only numerical and lawlike parameters
  - $\exists x \forall n(x_n = 1 \leftrightarrow \phi(n))$ for arbitrary ϕ in L.

We now construct a Boolean valued model for Z_2 + RS. We slightly alter the construction used in the proof of 2.1.3. 2^ω is again represented by the set of Borel measurable functions $(2^\omega)^\kappa \to 2^\omega$. The sort of lawlike elements is represented by the *constant* functions $(2^\omega)^\kappa \to 2^\omega$. ω is represented by itself (alternatively we could have taken functions $(2^\omega)^\kappa \to \omega$).
The Boolean interpretations of the basic arithmetical relations are as follows.
- $[[n + m = k]] = 1_b$ if $n + m = k$ and 0_b otherwise; likewise for 0,S and $\cdot$;
- if a: $(2^\omega)^\kappa \to 2^\omega$ is constant (i.e. lawlike), we put $[[a_n = 1]] = 1_B$ if n is in the range of a; and 0_b otherwise;
- if f: $(2^\omega)^\kappa \to 2^\omega$ is measurable, then $[[f_n = 1]] = \{\xi \in (2^\omega)^\kappa \mid f(\xi)_n = 1\}/I$.

The independence relation R is interpreted as in the proof of 2.1.3; and the quantifiers and connectives have the obvious interpretation.
The Boolean validity of Z_2 now follows. The validity of basic arithmetic and the induction scheme is trivial. To establish lawlike comprehension, it suffices to note that the Boolean algebra B is homogeneous, so that for any ϕ with only numerical and lawlike parameters, $[[\phi]]$ is either 0 or 1; for then we may define a lawlike a: $(2^\omega)^\kappa \to 2^\omega$ by putting for any ξ: $a(\xi)_n = 1$ iff $[[\phi(n)]]_B = 1_B$. Finally, the validity of the second comprehension scheme is shown by defining a measurable f: $(2^\omega)^\kappa \to 2^\omega$ such that $[[f_n = 1]]_B = [[\phi(n)]]_B$.
We now turn to the validity of RS in this model. That R is valid was shown in 2.1. Consider first RS1. $[[\forall x(Rx \to LLN(x)]]_B = \bigwedge_\pi [[LLN(\pi)]]_B$, where π ranges over the projections $(2^\omega)^\kappa \to 2^\omega$. But since the quantifiers in the defining expression for LLN have standard interpretation and the projections are measure preserving, we have $[[LLN(\pi)]]_B = \{\xi \in (2^\omega)^\kappa \mid LLN(\pi\xi)\}/I = 1_B$ for any π.
The verification of RS2 is more involved. We have to show that
$[[\forall F(F \text{ admissible} \to \forall x \forall y \forall z \forall z'(F(x,z) = y \wedge R(x,z) \wedge {}^x/_y = z' \to LLN(z')))]]_B = 1_B$.
This follows if we can show for any admissible function F and any projection π and f: $(2^\omega)^\kappa \to 2^\omega$ such that $supp(\pi) \cap supp(f) = \emptyset$: $[[LLN({}^\pi/F(\pi,f))]]_B = 1_B$.

Lemma 4.1 Let f: $(2^\omega)^\kappa \to 2^\omega$ be measurable, so that f can be identified with a function $(2^\omega)^\omega \to 2^\omega$. Then $\lambda\times\lambda^\omega\{\langle x,\zeta\rangle \in 2^\omega\times(2^\omega)^\omega \mid F(x, f(\zeta))$ defined, ${}^x/F(x,f(\zeta))$ defined and $\in A\} \le \lambda A$ for all Borel sets A.

Proof: For the sake of brevity we write $\lambda\times\lambda^\omega\{\langle x,\zeta\rangle \in 2^\omega\times(2^\omega)^\omega \mid {}^x/F(x,f(\zeta)) \in A\}$ for: $\lambda\times\lambda^\omega\{\langle x,\zeta\rangle \in 2^\omega\times(2^\omega)^\omega \mid F(x, f(\zeta))$ defined, ${}^x/F(x,f(\zeta))$ defined and $\in A\}$.
We then have

$$\lambda\times\lambda^{\omega}\{<x,\zeta> \in 2^{\omega}\times(2^{\omega})^{\omega} \mid {}^{x}/F(x,f(\zeta)) \in A\} \leq \int_{(2^{\omega})^{\omega}} \lambda\{x \mid {}^{x}/F(x,f(\zeta)) \in A\} d\lambda^{\omega}(\zeta) .$$

By the properties of F, for fixed ζ the function ${}^{\bullet}/F(\bullet,f(\zeta))$: $2^{\omega} \to 2^{\omega}$ is a place selection in the sense that the choice of x_n depends only on x(n–1). Hence, by a result of Doob (see VAN LAMBALGEN [1987a], p.758), we have $\lambda\{x \mid {}^{x}/F(x,f(\zeta)) \in A\} \leq \lambda A$ for each ζ; and the result follows by integrating over $(2^{\omega})^{\omega}$. □

Since $\lambda\{x \mid LLN(x)\} = 1$, we have $\lambda^{\kappa}\{\xi \in (2^{\omega})^{\kappa} \mid F(\pi\xi, f(\xi))$ defined, ${}^{\pi\xi}/F(\pi\xi,f(\xi))$ defined and $LLN({}^{\pi\xi}/F(\pi\xi,f(\xi)))\} = \lambda\times\lambda^{\omega}\{<x,\zeta> \in 2^{\omega}\times(2^{\omega})^{\omega} \mid LLN({}^{x}/F(x,f(\zeta)))\} = 1$. Hence Z_2 + RS is consistent.

5 Recursion theoretic interpretations of R and RS With the axiom system RS, we have tried to isolate some essential features of randomness and independence. We will now investigate how these axioms are related to the most widely known explicit definitions: in terms of recursive sequential tests (MARTIN-LÖF [1966]) and the definition using the variant of Kolmogorov complexity called prefixcomplexity (cf. KOLMOGOROV and USPENSKY [1988] and CHAITIN [1987]). Since these approaches are essentially equivalent (compare theorem 5.4 below), we shall concentrate on prefixcomplexity.

The definitions of Martin-Löf and Kolmogorov differ essentially from that of von Mises in that they do not presuppose a notion of independence. Our first task will therefore be to introduce an explicit notion of independence using prefixcomplexity. It turns out that we can take the notion defined in VAN LAMBALGEN [1987a] . We first give a quick introduction to prefixcomplexity and Martin-Löf's definition of randomness. For more information see VAN LAMBALGEN [1987a].

Definition 4.1 (KOLMOGOROV and USPENSKY [1988]; CHAITIN [1987]) A set $E \subseteq 2^{<\omega}$ is called *prefixfree* if no $w \in E$ is an initial segment of some $v \in E$. A *prefixalgorithm* is a partial recursive function A: $2^{<\omega} \to 2^{<\omega}$ which has prefixfree domain. Let A: $2^{<\omega} \to 2^{<\omega}$ be a prefixalgorithm with Gödelnumber ⌜A⌝. We define a universal prefixalgorithm U by the following condition: on inputs of the form $q = 0^{\ulcorner A \urcorner}p$, U simulates the action of A on p. The *prefixcomplexity* I(w) is defined as $I(w) := \{|p| \mid U(p) = w\}$.

Definition 4.2 $x \in 2^{\omega}$ is *irregular* if $\exists m \forall n\ I(x(n)) > n - m$.

Definition 4.3 (MARTIN-LÖF [1966]) A *recursive sequential test* N is a Π^0_2 subset of 2^{ω}, $N = \bigcap_n O_n$, O_n in Σ^0_1, such that $\lambda O_n \leq 2^{-n}$ for all n. A sequence $x \in 2^{\omega}$ is *random* (denoted $x \in R$) if it is not contained in any recursive sequential test. (Obviously $\lambda R = 1$.)

Theorem 4.4 $\exists m \forall n\ I(x(n)) > n - m$ iff $x \in R$ (in other words, randomness (4.3) and irregularity (4.2) coincide).

Remark It should be noted that the analogue of 4.4 for ordinary Kolmogorov complexity K is false; as MARTIN-LÖF [1971] has shown, *no* sequence x satisfies $\exists m \forall n\, K(x(n)) > n - m$. The following alternative characterization of Martin-Löf randomness is due to Solovay.

Theorem 4.5 $x \in R$ iff for any r.e. sequence of Σ^0_1 sets O_n such that $\Sigma \lambda O_n < \infty$: $x \notin \bigcap_n O_n$.

We are now in a position to define a complexity theoretic independence relation. Let I^y be prefixcomplexity as defined above, except that we now use the oracle y, thus replacing the universal Turing machine U by U^y. Similarly, we may relativize Martin- Löf's randomness concept R (4.3) to R^y. Relativizing the proof of theorem 4.4 we obtain

Theorem 4.4' (a) $\exists m \forall n\, I^y(x(n)) > n - m$ iff $x \in R^y$; and hence (b) $\lambda\{ x \mid \exists m \forall n\, I^y(x(n)) > n - m\} = 1$.

The promised complexity theoretic characterisation of independence is given by

Definition 4.6 If x is random (or irregular, cf. 4.4), then y *has no information about* x if $\exists m \forall n\, I^y(x(n)) > n - m$. Similarly for $\vec{y}$.

The following two theorems are proved in VAN LAMBALGEN [1987a]. The first shows that almost no sequence contains information about a given random sequence x, whereas the second shows that axiom RS2' (cf. section 1.2) is valid for the notion of independence introduced above.

Theorem 4.7 If x is random, then $\lambda\{y \mid y$ has no information about $x\} = 1$.

Theorem 4.8 If x is random and y has no information about x and contains infinitely many 1's, then $^x/y$ is also random. A fortiori, LLN($^x/y$).

The main question now confronting us is: does the relation of independence defined in 4.6 satisfy ***RS***.? R1 - 4 are trivially satisfied, but R5 is not so easy. In section 2 we saw that R5 could be used to interpret the "Fubini axiom" Q6 for Friedman's quantifier "almost all'. The validity of R5 for our present notion of independence is also connected to Fubini's theorem, but in this case we have to use the effective analogue of Fubini proved in VAN LAMBALGEN [1987b].
First some notation. If A is a subset of $2^\omega \times 2^\omega$, we define $(A)^y := \{ x \mid \langle x,y\rangle \in A\}$ and $(A)_x := \{ y \mid \langle x,y\rangle \in A\}$ (the more customuray notation A^y being used for relativization).

Theorem 4.9 Let $N \subseteq 2^\omega \times 2^\omega$ be a recursive sequential test with respect to $\lambda\times\lambda$, where $N = \bigcap_n O_n$, O_n in Σ^0_1, such that $\lambda O_n \le 2^{-n}$ for all n. Then we have

(a) $y \in R$ implies $\sum_{n=0}^{\infty} \lambda(O_n)^y < \infty$.

(b) (Effective Fubini) $\{y \mid \lambda(N)^y > 0\}$ is contained in a recursive sequential test (with respect to λ).

Proof We first prove (b) from (a). It suffices to show that if $y \in R$, $\lambda(\bigcap_n O_n)^y = 0$. We have $(\bigcap_n O_n)^y \subseteq (\bigcap_n \bigcup_{k\ge n} O_k)^y = \bigcap_n \bigcup_{k\ge n}(O_k)^y$ and $\lambda \bigcup_{k\ge n}(O_k)^y \le \sum_{k=n}^{\infty} \lambda(O_k)^y \to 0$ for $k \to \infty$. It remains to prove (a). Let $O_n = \bigcup_i [w_{ni}]\times[v_{ni}]$. Define for each n a sequence of functions f_{nk}, where $k \ge 0$, as follows:

$$f_{n0}(y) = 0 \text{ for all } y$$

$$f_{nk}(y) = \sum\{\lambda[w_{ni}] \mid y\in [v_{ni}],\ i \le k\},\ k \ge 1.$$

The f_{nk} are computable stepfunctions and $f_{nk} \le f_{n(k+1)}$; moreover $f_{n(k+1)}(y) - f_{nk}(y) = \lambda[w_{n(k+1)}]$ for $y \in [v_{n(k+1)}]$ and $f_{n(k+1)}(y) - f_{nk}(y) = 0$ otherwise.

Clearly

$$\int f_{nk} d\lambda = \sum_{i=1}^{k} \int (f_{ni}-f_{n(i-1)}) d\lambda = \sum_{i=1}^{k} \lambda[w_{ni}]\cdot\lambda[v_{ni}] \le 2^{-n}.$$

Define $B_{nm} = \{ y \mid \exists k{\le}m\ f_{nk}(y) > (n+1)^{-2}\}$ and $B_n = \{ y \mid \exists k\ f_{nk}(y) > (n+1)^{-2}\}$. Obviously $B := \bigcap_n B_n$ is Π^0_2. Moreover $\lambda B_n \le (n+1)^2 \cdot 2^{-n}$: since $f_{nk} \le f_{n(k+1)}$, we must have, for all m, $\lambda B_{nm} = \int 1_{B_{nm}} d\lambda < n^2 \int f_{nm} d\lambda \le (n+1)^2 \cdot 2^{-n}$. Hence B is (contained in) a recursive sequential test.

Since the functions f_{nk} are nondecreasing in k, we can define $f_n(y) = \lim_{k\to\infty} f_{nk}(y)$. We then have $\sum_{n=0}^{\infty} \lambda(O_n)^y \le \sum_{n=0}^{\infty} \sum_{i=0}^{\infty} \{\lambda[w_{ni}] \mid y \in [v_{ni}]\} = \sum_{n=0}^{\infty} f_n(y)$.

Now if $y \in R$, then $y \notin B$, hence $\sum_{n=0}^{\infty} \lambda(O_n)^y \le \sum_{n=0}^{\infty} f_n(y) \le \sum_{n=1}^{\infty} n^{-2} < \infty$. □

Theorem 4.10 If $\exists m \forall n\ I^y(x(n)) > n - m$ and y is random, then $\exists m \forall n\ I^x(y(n)) > n - m$.

Proof Let R^y denote Martin-Löf's randomness notion relativized to y; hence (4.4') $x \in R^y$ iff $\exists m \forall n\ I^y(x(n)) > n - m$. Let R^2 denote Martin-Löf's randomness notion in $2^\omega \times 2^\omega$ (with respect to $\lambda \times \lambda$), and let $(R^2)^y$, $(R^2)_x$ be the sections of R^2 as defined above.

It suffices to show that $x \in R^y$ & $y \in R \rightarrow \langle x,y\rangle \in R^2$. For if $\langle x,y\rangle \in R^2$, then $y \in (R^2)_x$; and the following argument shows that we must have $(R^2)_x \subseteq R^x$.

The proof of theorem 4.4 establishes that $\{y \mid \forall m \exists n\ I(y(n)) \leq n - m\}$ is a recursive sequential test (and in fact the *universal* recursive sequential test). It follows by relativizing that proof to y that $\{\langle y,x\rangle \mid \forall m \exists n\ I^x(y(n)) \leq n - m\}$ is a recursive sequential test with respect to $\lambda \times \lambda$: clearly, $\{\langle y,x\rangle \mid \forall m \exists n\ I^x(y(n)) \leq n - m\}$ is Π^0_2, and for each m,

$$\lambda \times \lambda \{\langle y,x\rangle \mid \exists n\ I^x(y(n)) \leq n - m\} = \int \lambda\{y \mid \exists n\ I^x(y(n)) \leq n - m\} d\lambda \leq \int 2^{-n}\, d\lambda = 2^{-n}.$$

Hence $\langle y,x\rangle \mid \forall m \exists n\ I^x(y(n)) \leq n - m\}$ is contained in $(R^2)^c$ and therefore $(R^2)_x \subseteq R^x$.

So suppose $y \in R$ and $\langle x,y\rangle \notin R^2$. Since $(R^2)^c$ is a universal recursive sequential test (see Martin-Löf [1966]), it is of the form $\bigcap_n O_n$, O_n in Σ^0_1, such that $\lambda O_n \leq 2^{-n}$ for all n. Now if

$\langle x,y\rangle \notin R^2$, then $x \in ((R^2)^c)^y = \bigcap_n (O_n)^y$. We have assumed $y \in R$, hence by part (a) of the preceding theorem $\sum_{n=0}^{\infty} \lambda(O_n)^y$ converges. By using the relativized version of 4.5, we obtain $x \notin R^y$. □

Theorem 4.10 is perhaps interesting in its own right. As an illustration, we shall give an application to Turing degrees. Let $\leq_T$ denote Turing reducibility. We prove a weaker version of Sacks' result that $\lambda\{x \mid y \leq_T x\} = 0$ for nonrecursive y. Suppose y is a random sequence. and $y \leq_T x$. We then have $\forall m \exists n\ I^y(x(n)) \leq n - m$, for if $\exists m \forall n\ I^y(x(n)) > n - m$, in virtue of 4.10 we would also have $\exists m \forall n\ I^x(y(n)) > n - m$, which is impossible because of $y \leq_T x$. Hence it follows from 4.4' that $\lambda\{x \mid y \leq_T x\} = 0$ for *random* y.

In particular, we have $\lambda\{x \mid \emptyset' \leq_T x\} = 0$, for $\emptyset'$ contains the random sequence Ω, the halting probability of the universal Turing machine as defined in CHAITIN [1987].

We now return to the relation between independence as defined in 4.6 and <u>RS</u>. The validity of RS1 and RS2' for our notion of independence follows from theorems 4.4 and 4.8 respectively. Moreover, if the admissible function F is Σ^0_1 on its domain, we also have: $x \in R$ & y has no information about x & F(x,y) is defined and has infinitely many 1's $\rightarrow$ $^x/F(x,y) \in R$. This is so because for such F, $F(\bullet,y)$ is a place selection recursive in y.

Summarizing the discussion so far, we have

Theorem 4.11 The independence relation defined in 4.6 satisfies R1 - 5, RS1 and RS2 when the quantifier ∀F is restricted to F which are Σ_1^0 on their domain.

It can also be shown that R6 is true when restricted to the Σ_2^e formulas as defined by GAIFMAN and SNIR [1983], but we shall not go into the details.

4.12 Concluding remarks In accordance with von Mises' slogan "Erst das Kollektiv [i.e. random sequence], dann die Wahrscheinlichkeit", _RS_ does not refer to a probability measure, only to limiting relative frequency. This contrasts with the (more common) approach that takes random sequences to be sequences that satisfy all ("effective") strong limit laws of probability theory. It will be observed that _R_, the independence part of _RS_, is complete with respect to a semantics in which randomness is indeed defined as satisfaction of a large class of properties of Lebesgue measure 1 (2.3.1). Does this mean that, after all, the proper concept of randomness is "satisfaction of all effective strong limit laws of probability theory" ? It does not, because _RS_ by no means singles out a unique measure. _R_ can be intepreted using any continuous Borel measure and to interpret RS1 and RS2 we only need a product measure $\prod_n(1-p,p_n)$ such that $p_n \to \frac{1}{2}$ for $n \to \infty$. Therefore Ville's theorem, which says that Church random sequences need not satisfy the law of the iterated logarithm, has an analogue in our situation: it does not follow from _RS_ that $\forall x(R(x) \to x$ satisfies the law of the iterated logarithm), as can be shown using the arguments in section 4 of VAN LAMBALGEN [1987a].

References

BALDWIN, J.T. [1988], Stability theory, Springer Verlag.

BELL, J.L. [1977], Boolean valued models and independence proofs in set theory, Oxford University Press.

CHAITIN, G.J. [1987], Algorithmic information theory, Cambridge Tracts in Computer Science **1**, Cambridge University Press.

CHURCH, A. [1940], On the concept of a random sequence, *Bull. AMS* **46**, 130 - 135.

FREILING, C. [1986], Axioms of symmetry: throwing darts at the real number line, *J. Symb. Logic* **51**, 190 - 200.

GAIFMAN, H. and M. SNIR [1982], Probabilities over rich languages, randomness and testing, *J Symb. Log.* **47**, 495 - 548.

KOLMOGOROV, A.N. and V.A. USPENSKY [1988], Algorithms and randomness, *SIAM Theory Probab. Appl.* **32**, 389 - 412.

KRIVINE, J.L. and K. MCALOON [1973], Forcing and generalized quantifiers, *Ann. Math. Logic* **5**, 199 - 253.
VAN LAMBALGEN, M. [1987a], Von Mises' definition of random sequences reconsidered, *J. Symb. Logic* **52**, 725 - 755.
VAN LAMBALGEN, M. [1987b], Random sequences, Ph.D. thesis, Dept. of Mathematics, University of Amsterdam.
VAN LAMBALGEN, M. and I. MOERDIJK [199?], Models for randomness, in preparation.
MARTIN-LÖF, P. [1966], The definition of random sequences, *Inf. Contr.* **9**, 602 - 619.
MARTIN-LÖF, P. [1971], Complexity oscillations in infinite binary sequences, *Z. Wahrsch. verw. Geb.* **19**, 225 - 230.
METAKIDES, G. and A. NERODE [1982], Recursion theory and abstract dependence, in C.T. Chong ed. Southeast Asian Conference on Logic, North - Holland.
VON MISES, R [1919], Grundlagen der Wahrscheinlichkeitsrechnung, *Math. Z.* **5**, 52 - 99.
OXTOBY, J.C. [1980], Measure and Category, Springer Verlag.
PARTHASARATY, T [1972], Selection theorems and their applications, LNM 263, Springer Verlag.
SOLOVAY, R.M. [1970], A model of set theory in which every set of reals is Lebesgue measurable, *Ann. Math.* **90**, 1 - 55.
STEINHORN, C.I. [1985a], Borel structures and measure and category logics, in: K.J. Barwise, S. Feferman (eds.), Modeltheoretic Logics, Springer Verlag.
STEINHORN, C.I. [1985b], Borel structures for first order and extended logics, in L.A. Harrington et al (eds.), Harvey Friedman's research on the foundations of mathematics, North - Holland.
TROELSTRA, A.S. [1977], Choice sequences, Oxford University Press.
TROELSTRA, A.S. and D. VAN DALEN [1988], Constructivism in Mathematics I, II, North - Holland.
VAN DER WAERDEN, B.L. [1940], Modern Algebra, Frederick Ungar Publishing Co..
WELSH [1976], Matroid Theory, Academic Press.

The ITLI Prepublication Series

1986

86-01	The Institute of Language, Logic and Information
86-02 Peter van Emde Boas	A Semantical Model for Integration and Modularization of Rules
86-03 Johan van Benthem	Categorial Grammar and Lambda Calculus
86-04 Reinhard Muskens	A Relational Formulation of the Theory of Types
86-05 Kenneth A. Bowen, Dick de Jongh	Some Complete Logics for Branched Time, Part I Well-founded Time, Forward looking Operators
86-06 Johan van Benthem	Logical Syntax

1987

87-01 Jeroen Groenendijk, Martin Stokhof	Type shifting Rules and the Semantics of Interrogatives
87-02 Renate Bartsch	Frame Representations and Discourse Representations
87-03 Jan Willem Klop, Roel de Vrijer	Unique Normal Forms for Lambda Calculus with Surjective Pairing
87-04 Johan van Benthem	Polyadic quantifiers
87-05 Víctor Sánchez Valencia	Traditional Logicians and de Morgan's Example
87-06 Eleonore Oversteegen	Temporal Adverbials in the Two Track Theory of Time
87-07 Johan van Benthem	Categorial Grammar and Type Theory
87-08 Renate Bartsch	The Construction of Properties under Perspectives
87-09 Herman Hendriks	Type Change in Semantics: The Scope of Quantification and Coordination

1988

Logic, Semantics and Philosophy of Language:

LP-88-01 Michiel van Lambalgen	Algorithmic Information Theory
LP-88-02 Yde Venema	Expressiveness and Completeness of an Interval Tense Logic
LP-88-03	Year Report 1987
LP-88-04 Reinhard Muskens	Going partial in Montague Grammar
LP-88-05 Johan van Benthem	Logical Constants across Varying Types
LP-88-06 Johan van Benthem	Semantic Parallels in Natural Language and Computation
LP-88-07 Renate Bartsch	Tenses, Aspects, and their Scopes in Discourse
LP-88-08 Jeroen Groenendijk, Martin Stokhof	Context and Information in Dynamic Semantics
LP-88-09 Theo M.V. Janssen	A mathematical model for the CAT framework of Eurotra
LP-88-10 Anneke Kleppe	A Blissymbolics Translation Program

Mathematical Logic and Foundations:

ML-88-01 Jaap van Oosten	Lifschitz' Realizabiility
ML-88-02 M.D.G. Swaen	The Arithmetical Fragment of Martin Löf's Type Theories with weak Σ-elimination
ML-88-03 Dick de Jongh, Frank Veltman	Provability Logics for Relative Interpretability
ML-88-04 A.S. Troelstra	On the Early History of Intuitionistic Logic
ML-88-05 A.S. Troelstra	Remarks on Intuitionism and the Philosophy of Mathematics

Computation and Complexity Theory:

CT-88-01 Ming Li, Paul M.B.Vitanyi	Two Decades of Applied Kolmogorov Complexity
CT-88-02 Michiel H.M. Smid	General Lower Bounds for the Partitioning of Range Trees
CT-88-03 Michiel H.M. Smid, Mark H. Overmars Leen Torenvliet, Peter van Emde Boas	Maintaining Multiple Representations of Dynamic Data Structures
CT-88-04 Dick de Jongh, Lex Hendriks Gerard R. Renardel de Lavalette	Computations in Fragments of Intuitionistic Propositional Logic
CT-88-05 Peter van Emde Boas	Machine Models and Simulations (revised version)
CT-88-06 Michiel H.M. Smid	A Data Structure for the Union-find Problem having good Single-Operation Complexity
CT-88-07 Johan van Benthem	Time, Logic and Computation
CT-88-08 Michiel H.M. Smid, Mark H. Overmars Leen Torenvliet, Peter van Emde Boas	Multiple Representations of Dynamic Data Structures
CT-88-09 Theo M.V. Janssen	Towards a Universal Parsing Algorithm for Functional Grammar
CT-88-10 Edith Spaan, Leen Torenvliet Peter van Emde Boas	Nondeterminism, Fairness and a Fundamental Analogy
CT-88-11 Sieger van Denneheuvel, Peter van Emde Boas	Towards implementing RL

Other prepublications:

X-88-01 Marc Jumelet	On Solovay's Completeness Theorem

1989

Logic, Semantics and Philosophy of Language:

LP-89-01 Johan van Benthem	The Fine-Structure of Categorial Semantics
LP-89-02 Jeroen Groenendijk, Martin Stokhof	Dynamic Predicate Logic, towards a compositional, non-representational semantics of discourse
LP-89-03 Yde Venema	Two-dimensional Modal Logics for Relation Algebras and Temporal Logic of Intervals
LP-89-04 Johan van Benthem	Language in Action
LP-89-05 Johan van Benthem	Modal Logic as a Theory of Information
LP-89-06 Andreja Prijatelj	Intensional Lambek Calculi, Theory and Application

Mathematical Logic and Foundations:

ML-89-01 Dick de Jongh, Albert Visser	Explicit Fixed Points for Interpretability Logic
ML-89-02 Roel de Vrijer	Extending the Lambda Calculus with Surjective Pairing is conservative
ML-89-03 Dick de Jongh, Franco Montagna	Rosser Orderings and Free Variables
ML-89-04 Dick de Jongh, Marc Jumelet, Franco Montagna	On the Proof of Solovay's Theorem
ML-89-05 Rineke Verbrugge	Σ-completeness and Bounded Arithmetic
ML-89-06 Michiel van Lambalgen	The Axiomatization of Randomness

Computation and Complexity Theory:

CT-89-01 Michiel H.M. Smid	Dynamic Deferred Data Structures
CT-89-02 Peter van Emde Boas	Machine Models and Simulations
CT-89-03 Ming Li, Herman Neuféglise Leen Torenvliet, Peter van Emde Boas	On Space efficient Solutions
CT-89-04 Harry Buhrman, Leen Torenvliet	A Comparison of Reductions on Nondeterministic Space
CT-89-05 Pieter H. Hartel, Michiel H.M. Smid Leen Torenvliet, Willem G. Vree	A Parallel Functional Implementation of Range Queries
CT-89-06 H.W. Lenstra, Jr.	Finding Isomorphisms between Finite Fields

Other prepublications:

X-89-01 Marianne Kalsbeek	An Orey Sentence for Predicative Arithmetic
X-89-02 G. Wagemakers	New Foundations. a Survey of Quine's Set Theory
X-89-03 A.S. Troelstra	Index of the Heyting Nachlass
X-89-04 Jeroen Groenendijk, Martin Stokhof	Dynamic Montague Grammar, a first sketch
X-89-05 Maarten de Rijke	The Modal Theory of Inequality