MATHEMATICAL AND PHILOSOPHICAL PERSPECTIVES ON ALGORITHMIC RANDOMNESS

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

Christopher P. Porter

Peter Cholak, Co-Director

Michael Detlefsen, Co-Director

Curtis Franks, Co-Director

Graduate Program in Mathematics and Philosophy
Notre Dame, Indiana

April 2012

MATHEMATICAL AND PHILOSOPHICAL PERSPECTIVES ON ALGORITHMIC RANDOMNESS

Abstract

by

Christopher P. Porter

The mathematical portion of this dissertation is a study of the various interactions between definitions of algorithmic randomness, Turing functionals, and non-uniform probability measures on Cantor space. Chapters 1 and 2 introduce the main results and the relevant background for the remaining chapters. In Chapter 3, we study the connection between Turing functionals and a number of different definitions of randomness, culminating in a number of characterizations of these definitions of randomness in terms of a priori complexity, a notion of initial segment complexity given in terms of Turing functionals. In Chapter 4, we investigate possible generalizations of Demuth's Theorem, an important theorem in algorithmic randomness concerning the behavior of random sequences under truth-table reducibility. One technique developed in this chapter, that of inducing a probability measure by means of a special type of truth-table functional that we call a *tally functional*, proves to be very useful. We use this technique to study randomness with respect to trivial computable measures in both Chapters 5 and 6.

In the philosophical portion of this dissertation, we consider the problem of pro-

ducing a correct definition of randomness, as introduced in Chapter 7: Some have claim that one definition of randomness in particular, Martin-Löf randomness, captures the so-called intuitive conception of randomness, a claim known as the Martin-Löf-Chaitin Thesis, but some have offered alternative definitions as capturing our intuitions of randomness. Prior to evaluating the Martin-Löf-Chaitin Thesis and related randomness-theoretic theses, Chapters 8 and 9 discuss two roles of definitions of randomness, both of which motivated much early work in the development of algorithmic randomness: the *resolutory role of randomness*, which is successfully filled by a definition of randomness that allows for the solution of problems in a specific theory of probability, and the *exemplary role of randomness*, which is successfully filled by a definition of randomness that counts as random those sequences that exemplify the properties typically held by sequences chosen at random. In Chapter 10, we lay out the status of the Martin-Löf-Chaitin Thesis, discussing the evidence that has been offered in support of it, as well as the arguments that have been raised against it. In Chapter 11, we argue that the advocate of a claim like the Martin-Löf-Chaitin Thesis faces what we call the Justificatory Challenge: she must present a precise account of the so-called intuitive conception of randomness, so as to justify the claim that her preferred definition of randomness is the correct one and block the claim of correctness made on behalf of alternative definitions of randomness. Lastly, in Chapter 12, we present two further roles for definitions of randomness to play, which we call the *calibrative role of randomness* and the *limitative role of randomness*, which can be successfully filled by multiple definitions of randomness. Definitions filling the calibrative role allow us to calibrate the level of randomness necessary and sufficient for certain "almost-everywhere" results in classical mathematics to hold, while definitions filling the limitative role illuminate a phenomenon known as the *indefinite contractibility of the notion of randomness*. Moreover, we argue that in light of the fact that many definitions can successfully fill these two roles, we should accept what we call the *No-Thesis Thesis*: there is no definition of randomness that (i) yields a well-defined, definite collection of random sequences and (ii) captures everything that mathematicians have taken to be significant about the concept of randomness.

To Laura

CONTENTS

FIGURES

PREFACE

The subject of this dissertation is the notion of algorithmic randomness. The first six chapters treat the mathematics of algorithmic randomness, and the last six chapters treat the philosophy of algorithmic randomness. Chapter 1 introduces the mathematical perspective on algorithmic randomness, while Chapter 2 provides the necessary background for the remaining chapters. Chapters 3 through 6 contain the mathematical contributions of this dissertation, which involve the interactions between various notions of algorithmic randomness, Turing functionals, and probability measures on Cantor space. Chapter 7 introduces the philosophical perspective on algorithmic randomness that we take here. Chapters 8 and 9 provide an overview of the historical development of algorithmic randomness, highlighting several purposes for which definitions of randomness have been developed. Chapters 10 and 11 discuss the problem of determining whether there is some correct definition of randomness, i.e. one that captures the prevailing intuitive conception of randomness. Chapter 12 presents an alternative approach to the various definitions of randomness, according to which no single definition captures everything that mathematicians have taken to be significant about the concept of randomness. Rather, multiple definitions are needed to accomplish this task.

ACKNOWLEDGMENTS

There are many individuals who have contributed to this project over the years. First, I owe a debt of gratitude to my advisors, Peter Cholak, Michael Detlefsen, and Curtis Franks, each of whom has played a vital role bringing this project to fruition.

I am grateful to Peter for being a constant source of support and encouragement and for contributing so greatly to my mathematical education at Notre Dame. Peter has always made himself accessible to his students, spending many hours with me individually and in the infamous "Cholak seminar", where I learned how to teach, to ask questions, and to make my own ideas clear to others. Participating in his seminar was one of the most formative experiences of my graduate school career.

I am also thankful to Mic, who taught me how to operate as both a mathematician and a philosopher. From Mic I learned not only how to formulate problems and relentlessly pursue solutions to those problems, but also how to do philosophy from within mathematics, so to speak. I will always appreciate the values Mic has instilled in me as a philosopher, the ways he has pushed me to improve, and his faith in my instincts, demonstrated in his insistence that I find my own way as a scholar.

Curtis has been a tireless advocate over the last four years since he signed on to co-direct this project. Our conversations have always been rewarding; many of the ideas contained in this dissertation were fleshed out and developed in those very conversations. Moreover, Curtis helped to develop in me a profound appreciation of the intrinsic philosophical and scientific value of our mathematical theories, an approach that has influenced this project tremendously. I am deeply indebted to him.

I am also thankful to the many teachers I've had at Notre Dame, specfically, Paddy Blanchette, Tim Bays, Julia Knight, Sergei Starchenko, Noam Greenberg, and David Lippel, all of whom have contributed much to my logical education. I am also grateful to both Paddy and Tim for serving as members of my dissertation proposal committee, and to Julia Knight, who carefully read this dissertation and gave me much helpful feedback. I also owe thanks to Tim McCarthy who served as a member of my dissertation proposal committee and as one of my dissertation committee members.

I would also like to especially thank Laurent Bienvenu, who also served on the dissertation committee. In many ways, Laurent has been a catalyst for the mathematics portion of the dissertation. Results from the latter half of Chapter 3, as well as the material in Chapter 4, is the fruit of our very successful and enjoyable collaboration. He has also been extremely supportive throughout the entire project, and I look forward to working further with him in the near future.

There are many other individuals to whom I owe thanks. Members of the Cholak group, both past and present, have been helpful in so many ways: Steve Walk, Charlie McCoy, Rebecca Weber, Josh Cole, Logan Axon, Sean Walsh, Stephen Flood, and Quinn Culver. I'd like to thank Sean Walsh in particular, who, as a colleague in the joint program, set such a valuable example for me to follow. I have benefitted greatly from Sean's insightful feedback and generous encouragement over the years. I am also grateful to Mic's other students: Andy Arana, Iulian Toader, Graham Leach-Krouse, and Tony Strimple.

Many gracious conversation partners have been a source of encouragement and insight over the years: Alex Arnold, Jeremy Avigad, Chris Conidis, Walter Dean, Rod Downey, Damir Dzhafarov, Antony Eagle, Johanna Franklin, Serge Grigorieff, Brice Halimi, Denis Hirschfeldt, Kevin Kelly, Per Martin-Löf, Cristian Mihut, Joe Miller, Antonio Montalbán, André Nies, Luke Potter, David Rabouin, Jan Reimann, Ted Slaman, Ivahn Smadja, Jeff Snapper, and many others.

I'd also like to thank my family for supporting me all these years, especially my mother, Chantal Porter, and my sister, Quincyann Hillery. Lastly, I am deeply grateful to my wife Laura for being such a loving, reliable, patient, understanding and abiding friend.

CHAPTER 1

MATHEMATICAL PERSPECTIVES ON ALGORITHMIC RANDOMNESS

An algorithmically random sequence of 0s and 1s cannot be distinguished, by means of any effective procedure, from a binary sequence produced by a random process. But not all random processes yield the same probability measure on the collection of all sequences 2^ω. For instance, if our random process is given by the tosses of a fair coin, the resulting probability measure is the Lebesgue measure, but if we consider those sequence produced by the tosses of a biased coin, the resulting probability measure will *not* be the Lebesgue measure. Consequently, the collection of sequences that are algorithmically random with respect to this biased measure will be disjoint from the collection of sequences that are random with respect to the Lebegue measure.

The primary focus of the mathematical portion of this dissertation is to study algorithmic randomness with respect to various computable measures. While much of the work in the theory of algorithmic randomness over the last decade or so has been concerned with studying notions of algorithmic randomness with respect to the Lebesgue measure, randomness with respect to various computable measures is not as well-studied. In particular, one aspect of randomness with respect to computable measures that has not been fully developed is the interaction between Turing functionals, computable measures, and the sequences that are random with respect to these measures. Most of the results in this dissertation concern precisely this interaction.

An overview of the results of this portion of the dissertation is as follows. In Chapter 2, we provide a survey of the main definitions and results in the theory of algorithmic randomness. However, given the focus on computable measures as described above, this background material is presented in more generality than what is typically found in the standard presentations on algorithmic randomness (such as [Nie09] and [DH10]). In particular, we define the main notions of algorithmic randomness in terms of an arbitrary computable measure and reprove many of the standard theorems of algorithmic randomness in this more general setting.

In Chapter 3, we prove the Functional Existence Theorem, a result that guarantees, as long as several technical conditions are satisfied, the existence of a Turing functional that maps certain amounts of measure to initial segments of a sequence. Although this result is essentially a reformulation of a result found in the early work of Levin and Zvonkin [ZL70], a new proof of this result is provided here that is useful for a number of applications. In particular, the Functional Existence Theorem is used here to provide a number of characterizations of Martin-Löf randomness, Schnorr randomness, computable randomness, and Kurtz randomness in terms of various classes of Turing functionals. These latter results were obtained in collaboration with Laurent Bienvenu.

The subject of Chapter 4 is Demuth's Theorem, an important theorem concerning the behavior of randomness under transformation by truth-table functionals.

Specifically, Demuth's Theorem states that every non-computable sequence that is truth-table reducible to a Martin-Löf random sequence is Turing equivalent to a Martin-Löf random sequence. The significance of this result is this: If we apply an effectively continuous transformation to a Martin-Löf random sequence, so long as we don't completely destroy the randomness of the original sequence (by mapping it to a computable sequence), we will be able to extract an unbiased random sequence from the transformed sequence. We show that Demuth's Theorem also holds for Schnorr randomness, computable randomness and weak 2-randomness, albeit using techniques that differ from the one used to prove Demuth's Theorem for Martin-Löf randomness. Next, it is shown that several generalizations of Demuth's Theorem, given in terms of weak truth-table reducibility, fail to hold. That is, it is shown that when we extract an unbiased random sequence from a sequence obtained from applying an effectively continuous transformation to a random sequence, we cannot, in general, effectively bound the amount of oracle access needed to recover the unbiased sequence. Lastly, a characterization is provided of those random Turing degrees that contain a sequence that (i) is Martin-Löf random with respect to some computable measure, but (ii) is not Martin-Löf random with respect to any continuous computable measure. The results in this chapter were obtained in collaboration with Laurent Bienvenu and submitted for publication as an article entitled "Strong Reductions and Effective Randomness".

In Chapter 5, the technique used to prove the failure of the weak truth-table versions of Demuth's Theorem is isolated and studied. Specifically, the functionals used in these proofs, referred to as *tally functionals*, are studied in detail. In the first half of the chapter, a number of examples of tally functionals are provided. Further, two types of measures induced by these functionals, trivial measures and diminutive measures, are studied. Lastly, tally functionals used to induce trivial computable measures that witness the separation of different notions of randomness.

Lastly, in Chapter 6, the technique of defining trivial measures by means of tally functionals is used to study a correspondence between a certain class of trivial measures and the collection of finite distributive lattices.

CHAPTER 2

MATHEMATICAL BACKGROUND

2.1 Introduction

The goal of this chapter is to provide the main background necessary for the mathematics portion of the dissertation. We will proceed as follows. In Section 2.2, the requisite notation is provided, while in Section 2.3, the essentials of computability theory are reviewed, drawing mainly from [Soa87]. In Section 2.4, we discuss computable probability measures on Cantor space, 2^ω. Section 2.5 provides a survey of the main definitions of algorithmic randomness, and then in Section 2.6 several useful theorems on algorithmic randomness are discussed. While the standard references on algorithmic randomness are [Nie09] and [DH10], I will proceed in slightly more generality; whereas these standard sources are almost exclusively concerned with notions of randomness with respect to the Lebesgue measure, I will consider these notions of randomness with respect to *any* computable probability measure on 2^ω.

2.2 Basic Notation

We will primarily concern ourselves with four kinds of objects: natural numbers, finite binary strings, and infinite binary sequences. Henceforth, the collections of these objects will be respectively denoted by

ω, the set of natural numbers;

$2^{<\omega}$, the set of finite binary strings;

2^{ω}, the set of infinite binary sequences, also known as the *Cantor space*.

Members of ω will be represented by lowercase Roman letters $e, i, j, k, m, n, s, t, x, y, z$ (and also ℓ in place of l). Members of $2^{<\omega}$ will be represented by the lowercase Greek letters $\alpha, \beta, \gamma, \xi, \sigma, \tau$, and the empty string will be written as $\varnothing$. Members of 2^{ω} will be represented by uppercase Roman letters $A, B, C, D, S, T, U, V, W, X, Y, Z$, while subsets of 2^{ω} will be represented by script letters $\mathcal{O}, \mathcal{P}, \mathcal{Q}, \mathcal{R}, \mathcal{S}, \mathcal{T}, \mathcal{U}, \mathcal{V}, \mathcal{W}, \mathcal{X}, \mathcal{Y}, \mathcal{Z}$. Subsets of ω and $2^{<\omega}$ will also be represented by uppercase Roman letters, which is justified by the following two identifications:

(1) The identification of ω and $2^{<\omega}$: Each $\sigma \in 2^{<\omega}$ can be represented by a unique $n_\sigma \in \omega$, where $n_\sigma = \mathsf{bin}(1\sigma) - 1$, where $\mathsf{bin} : 2^{<\omega} \to \omega$ maps a string to the natural number it represents in binary notion, and 1σ is the concatenation of the bit 1 with σ. Thus we have,

$$\varnothing \mapsto 0$$

$$0 \mapsto 1$$

$$1 \mapsto 2$$
$$00 \mapsto 3$$
$$01 \mapsto 4$$

and so on.

(2) The identification of subsets of ω and 2^ω: Each subset A of ω can be represented by a unique $X_A \in 2^\omega$ as follows:

$$n \in A \leftrightarrow X_A(n) = 1$$
$$n \notin A \leftrightarrow X_A(n) = 0$$

where, for any $X \in 2^\omega$ and $n \in \omega$, $X(n)$ is $(n+1)$st bit of X (so that $X(0)$ is the first bit of X).

Due to this latter identification, I will sometimes also refer to elements of 2^ω as *sets*.

For $X \in 2^\omega$, $\overline{X}$ will denote the complement of X in ω; that is, $n \in \overline{X}$ if and only if $n \notin X$ for every $n \in \omega$. However, for $\mathcal{X} \subseteq 2^\omega$, $\mathcal{X}^c$ will denote the complement of $\mathcal{X}$ in 2^ω; that is, $A \in \mathcal{X}^c$ if and only if $A \notin \mathcal{X}$ for every $A \in 2^\omega$.

Given $\sigma \in 2^{<\omega}$, $|\sigma|$ will denote the length of σ, and $\sigma(n)$ will denote the $(n+1)$st bit of σ if $n \leq |\sigma| - 1$ ($\sigma(n)$ is undefined otherwise). Given $\sigma, \tau \in 2^{<\omega}$, $\sigma \preceq \tau$ means that σ is an initial segment of τ, i.e., $\sigma(n) = \tau(n)$ for every $n \leq |\sigma| - 1$, while $\sigma \prec \tau$ means that $\sigma \preceq \tau$ and $\sigma \neq \tau$. Similarly, given $\sigma \in 2^{<\omega}$ and $X \in 2^\omega$, $\sigma \prec X$ means that σ is an initial segment of X.

Both $\sigma\tau$ and $\sigma^\frown\tau$ denote the concatenation of σ and τ, i.e. the unique $\gamma \in 2^{<\omega}$ such that

$$\gamma(n) = \begin{cases} \sigma(n) & \text{if } n \leq |\sigma| - 1 \\ \tau(|\sigma| - n) & \text{if } |\sigma| \leq n \leq |\sigma| + |\tau| - 1 \end{cases}$$

and is undefined otherwise. For $\sigma \in 2^{<\omega}$ and $X \in 2^\omega$, both σX and $\sigma^\frown X$ denote the infinite sequence obtained by concatenating σ and X. Further, given $X \in 2^\omega$ and $n \in \omega$, $X\restriction n$ denotes the initial segment of X of length n. Moreover, $\sigma\restriction n$ is the length n initial segment of σ, as long as $n \leq |\sigma| - 1$.

Given a collection $\{\tau_i\}_{i\in\omega}$ such that for every $j, k \in \omega$, $\tau_j \preceq \tau_k$ or $\tau_k \preceq \tau_j$, $\bigcup\{\tau_i\}_{i\in\omega}$ is the unique $X \in 2^\omega$ such that $\tau_i \prec X$ for every $i \in \omega$ in the case that the lengths of the τ_i's are unbounded; otherwise $\bigcup\{\tau_i\}_{i\in\omega}$ is the unique τ such that $\tau_i \preceq \tau$ for every $i \in \omega$.

Sometimes, we will consider $2^{<\omega}$ and 2^ω ordered lexicographically, where given $\sigma, \tau \in 2^{<\omega}$ such that $|\sigma| = |\tau|$, $\sigma \leq_{lex} \tau$ means that $\sigma(n) = 0$ and $\tau(n) = 1$ for the least n such that $\sigma(n) \neq \tau(n)$. Similarly, for $X, Y \in 2^\omega$, $X \leq_{lex} Y$ means that $X(n) = 0$ and $Y(n) = 1$ for the least n such that $X(n) \neq Y(n)$.

Lastly, a few odds and ends. We will use $\langle \cdot, \cdot \rangle : \omega \times \omega \to \omega$ to denote some standard pairing function.[1] For $n \in \omega$, $2^n = \{\sigma \in 2^{<\omega} : |\sigma| = n\}$, the collection of strings of length n. The collection of dyadic rationals, i.e. rationals of the form $\dfrac{n}{2^m}$ for $n, m \in \omega$, will be denoted as $\mathbb{Q}_2$. Members of $\mathbb{Q}_2$ will be denoted by p, q, or ϵ. If f and g are functions from ω to ω, $f(n) \leq g(n) + O(1)$ means that there is some c such that for every n, $f(n) \leq g(n) + c$.

[1] For instance, let $\langle x, y \rangle := \frac{1}{2}(x + y)(x + y + 1) + y$.

2.3 Computability Essentials

In this section, we review the basics of computability theory, as much of this material will be useful in the chapters that follow.[2] For more details, see, for instance, [Soare], [Cooper], and [DowHir], chapter 2.

2.3.1 Computability on ω

$\phi_e : \omega \to \omega$ denotes the eth partial computable function for a fixed enumeration of the partial computable functions. $\phi_e(x)\downarrow$ means that the eth partial computable function is defined on input x (in which case we say that ϕ_e *halts on input* x), while $\phi_e(x)\uparrow$ means that the eth partial computable function is not defined on x (in which case we say that ϕ_e *diverges on input* x).

For $s \in \omega$, $\phi_{e,s}(x)\downarrow$ means ϕ_e halts on input x in no more than s steps, and thus $\phi_e(x)\downarrow$ implies that $\phi_{e,s}(x)\downarrow$ for some $s \in \omega$. Similarly, $\phi_{e,s}(x)\uparrow$ means that ϕ_e does not halt on input x in s or fewer steps, and thus $\phi_e(x)\uparrow$ implies that $\phi_{e,s}(x)\uparrow$ for every $s \in \omega$.

The universal partial computable function is the function ϕ defined by

$$\phi(\langle x, e\rangle) = \begin{cases} \phi_e(x) & \text{if } \phi_e(x)\downarrow \\ \text{undefined} & \text{if } \phi_e(x)\uparrow \end{cases}.$$

Partial computable functions that are total will henceforth be referred to as *com-*

[2] We distinguish here between computability on ω, computability on $2^{<\omega}$, and computability on 2^ω, distinctions that are, in large part, arbitrary, given the identification of ω and $2^{<\omega}$ discussed above and the fact that computability on 2^ω is merely defined in terms of computability on ω relative to an oracle. This notwithstanding, for our purposes it will be helpful to keep these notions distinct.

putable functions. We now define the collections of computable and computably enumerable sets of natural numbers.

Definition 2.1. A set $S \subseteq \omega$ is *computable* if the characteristic function of S, χ_S, is a computable function.

Definition 2.2. A set $S \subseteq \omega$ is *computably enumerable* (or c.e.) if S is the domain of a partial computable function.

The domain of the eth partial computable function, ϕ_e, will be denoted by W_e. It's not difficult to see that a set C is computable if and only if $C = W_e$ and $\overline{C} = W_i$ for some $e, i \in \omega$. Let D_e be the eth finite set, where $D_e = \{n_1, \ldots, n_k\}$ if and only if $e = 2^{n_1} + \ldots + 2^{n_k}$. One particularly important set is the halting set $\emptyset' = \{e : \phi_e(e)\downarrow\}$, which is c.e. but not computable.[3]

A useful collection of computable functions is the collection of *computable orders*, where a total function $f : \omega \to \omega$ is an *order* if f is non-decreasing and unbounded. Note that if f is a computable order, then the function f^{-1} defined by

$$f^{-1}(n) = \min\{k : f(k) \geq n\}$$

is also a computable order.

2.3.2 Computability on $2^{<\omega}$

By the identification of ω and $2^{<\omega}$ discussed above, each partial computable function ϕ_e can be viewed as a map from $2^{<\omega}$ to $2^{<\omega}$. Thus, we can extend the

[3] We adopt the convention of using the label '$\emptyset'$' rather than 'K' because we will use 'K' to refer to prefix-free Kolmogorov complexity, introduced below.

definition of computable and computably enumerable subsets of ω to computable and computably enumerable subsets of $2^{<\omega}$. Further, viewing partial computable functions as maps from $2^{<\omega}$ to $2^{<\omega}$, we can impose a useful restriction on the class of partial computable functions, namely that the domain of such a function be *prefix-free*.

Definition 2.3. A set $S \subseteq 2^{<\omega}$ is *prefix-free* if for every $\sigma, \tau \in S$, if $\sigma \preceq \tau$ or $\tau \preceq \sigma$, then $\sigma = \tau$.

We will refer to a partial computable function whose domain is prefix-free as a *prefix-free machine*, and $\{M_e\}_{e\in\omega}$ will denote the collection of prefix-free machines.

Just as there is a universal partial computable function, there is a universal prefix-free machine. For example, the function

$$U(1^e 0\sigma) = \begin{cases} M_e(\sigma) & \text{if } M_e(\sigma)\downarrow \\ \text{undefined} & \text{if } M_e(\sigma)\uparrow \end{cases}$$

for each $e \in \omega$ and $\sigma \in 2^{<\omega}$ is a universal prefix-free machine.

2.3.3 Computability on 2^ω

There are several ways to define a computable functional on 2^ω, two of which we will employ throughout this study. First, if we let $\{\phi_e^A\}_{e\in\omega}$ be the collection of partial computable functions relative to a fixed oracle $A \in 2^\omega$ and view these functions as $\{0,1\}$-valued (interpreting any non-zero output as equal to 1), then we can think of

ϕ_e as mapping A to some $B \in 2^\omega$ if

$$\phi_e^A = \chi_B,$$

where χ_B is the characteristic function of the set B. Further, given $\sigma \in 2^{<\omega}$, if we define ϕ_e^σ to be $\phi_{e,|\sigma|}^\sigma$ (the computation is run $|\sigma|$ many steps), it follows that

(i) $\phi_e^A(n)\downarrow$ implies that $\phi_e^\sigma(n)\downarrow$ for some $\sigma \prec A$;

(ii) if $\phi_e^\sigma(n)\downarrow$, then $\phi_e^\tau(n)\downarrow$ for every $\sigma \preceq \tau$; and

(iii) $\phi_e^A(n)\uparrow$ implies that $\phi_e^\sigma(n)\uparrow$ for every $\sigma \prec A$.

If $\phi_e^A(n)\downarrow$, the *use* of the computation is $x+1$, where x is the largest number such that the value $A(x)$ is queried in the course of the computation. Items (i) and (ii) above imply what is known as the Use Principle: If the use of a halting computation $\phi_e^A(n)$ is u, then for any $B \in 2^\omega$ such that $B\restriction(u+1) = A\restriction(u+1)$, $\phi_e^B(n)\downarrow = \phi_e^A(n)$.

The Turing jump of a sequence $A \in 2^\omega$ if defined to be

$$A' := \{e : \phi_e^A(e)\downarrow\}.$$

An equivalent way to define a computable functional on 2^ω, which we will adopt here, is to define a computable functional $\Phi : 2^\omega \to 2^\omega$ to be a c.e. set of pairs of strings (σ, τ) such that if $(\sigma, \tau), (\sigma', \tau') \in \Phi$ and $\sigma \preceq \sigma'$, then $\tau \preceq \tau'$. Then, for $\sigma \in 2^\omega$, if we define

$$\Phi^\sigma := \bigcup\{\tau : \exists \sigma' \preceq \sigma(\sigma', \tau) \in \Phi\},$$

then given $B \in 2^\omega$,

$$\Phi(B) := \bigcup_n \Phi^{B \upharpoonright n}.$$

Equivalently, we can define $\Phi(B) = \bigcup\{\tau : \exists n (B\upharpoonright n, \tau) \in \Phi\}$. If $\Phi(B) \in 2^\omega$, we say $\Phi(B)$ is defined, denoted $\Phi(B)\downarrow$; if there is some n such that $\Phi(B)(n)$ is undefined, we write $\Phi(B)\uparrow$. Henceforth, I will refer to such functionals as *Turing functionals*. The *domain* of a Turing functional Φ, denoted $\mathsf{dom}(\Phi)$, is thus

$$\mathsf{dom}(\Phi) = \{X \in 2^\omega : \Phi(X)\downarrow\}.$$

2.3.4 The Turing Degrees

By means of Turing functionals, we can now define Turing reducibility and the Turing degrees. For a Turing functional Φ and $A, B \in 2^\omega$, if $\Phi(B)\downarrow = A$, then A is *Turing reducible* to B, denoted $A \leq_T B$. Moreover, we will say that A and B are *Turing equivalent*, denoted $A \equiv_T B$. Lastly, the Turing degree of A, denoted $\mathsf{deg}_T(A)$, is defined to be

$$\mathsf{deg}_T(A) = \{B : A \equiv_T B\}.$$

In general, Turing degrees will be written as bold lowercase Roman letters $\mathbf{a}, \mathbf{b}$, and so on.

The Turing degrees form an upper semilattice, meaning that they are a partially ordered set with a join operation defined on every nonempty finite set of elements: For Turing degrees $\mathbf{a}$, $\mathbf{b}$ and $A \in \mathbf{a}$ and $B \in \mathbf{b}$, $\mathbf{a} \leq \mathbf{b}$ if $A \leq_T B$ for $A \in \mathbf{a}$ and

$B \in \mathbf{b}$, and $\mathbf{a} \cup \mathbf{b} = \mathsf{deg}_T(A \oplus B)$, where $A \oplus B = \{2n : n \in A\} \cup \{2n+1 : n \in B\}$. Given $A_1, \ldots, A_n \in 2^\omega$,

$$\bigoplus_{i=1}^{n} A_i = (\ldots((A_1 \oplus A_2) \oplus A_3) \oplus \ldots) \oplus A_n.$$

Clearly, $A_i \leq_T \bigoplus_{i=1}^{n} A_i$ for $i \leq n$. Further, for an infinite collection $A_1, A_2, \ldots \in 2^\omega$,

$$\bigoplus_{i \in \omega} A_i = \{\langle i, n \rangle : n \in A_i\}.$$

Again, it's not hard to see that $A_i \leq_T \bigoplus_{i \in \omega} A_i$ for every $i \in \omega$. Note that the jump operator on sets can be extended to Turing degrees, so that $A \in \mathbf{a}$, implies $\mathsf{deg}_T(A') = \mathbf{a}'$.

Two notable Turing degrees are $\mathsf{deg}_T(\emptyset) = \mathbf{0}$, the Turing degree of the computable sets, and $\mathsf{deg}_T(\emptyset') = \mathbf{0}'$ the Turing degree of the halting problem. Two notable classes of Turing degrees are the low degrees and the high degrees.

- A sequence A has *low* Turing degree if $A' \equiv_T \emptyset'$; that is, the jump of A is as low as possible.

- A sequence A has *high* Turing degree if $A' \equiv_T \emptyset''$; that is, the jump of A is as high as possible.

Lastly, the collection of Δ_2^0 Turing degrees is the collection $\{\mathbf{a} : \mathbf{a} \leq \mathbf{0}'\}$. Moreover, Schoenfield's Limit Lemma says that a set $A \in 2^\omega$ of Δ_2^0 degree has a computable approximation of finite sets $\{A_s\}_{s\in\omega}$ such that for every n,

$$\lim_{s\to\infty} A_s(n) = A(n);$$

that is, for every n, there is some $s \in \omega$ such that for every $t \geq s$, $A_t(n) = A(n)$.

2.3.5 Strong Reducibilities

In this study, two strong Turing reductions play a prominent role, truth-table reducibility and weak truth-table reducibility.

Definition 2.4. A Turing functional $\Phi : 2^\omega \to 2^\omega$ is

(i) a *weak truth-table functional* if there is some computable function f that bounds the use of Φ; and

(ii) a *truth-table functional* if Φ is total.

A well-known fact is that every truth-table functional is a weak truth-table functional. It follows for every truth-table functional Φ that there is a computable function $h : \omega \to \omega$ such that for every $A \in 2^\omega$ and every $n \in \omega$, $|\Phi^{A\restriction h(n)}| \geq n$ (if f is the computable function that bounds the use of Φ, then take $h(n) = \max\{f(k) : k \leq n\}$).

2.4 Topology and Measure on Cantor Space

2.4.1 Topological Considerations

We consider the product topology on 2^ω, given by basic open sets (also known as *cylinders*) of the form

$$[\![\sigma]\!] = \{X \in 2^\omega : \sigma \prec X\}$$

for every $\sigma \in 2^{<\omega}$. Under this topology, 2^ω is metrizable and compact. The effectively open subsets of 2^ω, also known as Σ^0_1 classes, are thus given by effective unions of basic open sets. More precisely, given a c.e. set $W \subseteq 2^{<\omega}$, $\mathcal{U} = \bigcup_{\sigma \in W} [\![\sigma]\!]$ is a Σ^0_1 class. Moreover, we can set $\mathcal{U}_s = \bigcup_{\sigma \in W_s} [\![\sigma]\!]$ for $s \in \omega$.

A Π^0_1 class is $\mathcal{P} \subseteq 2^\omega$ is an effectively closed subset of 2^ω, so that $\mathcal{P} = \mathcal{U}^c$ for some Σ^0_1 class $\mathcal{U}$. Equivalently, a Π^0_1 class is the collection of paths through a computable tree:

- $T \subseteq 2^{<\omega}$ is a *tree* if $\sigma \in T$ implies that $\tau \in T$ for every $\tau \prec \sigma$.

- If T is a tree, then $X \in 2^\omega$ is a *path* through T if $X{\upharpoonright}n \in T$ for every $n \in \omega$.

- The collection of paths through T is denoted $[T]$.

Lastly, a tree T is computable if T is computable as a subset of $2^{<\omega}$.

There are certain facts about Π^0_1 classes that will be useful for our purposes.

Theorem 2.5 (Low Basis Theorem). *For every Π^0_1 class $\mathcal{P}$, there is $X \in \mathcal{P}$ that is low, i.e., $X' \equiv_T \emptyset'$.*

$X \in 2^\omega$ has *hyperimmune-free degree* if every $f \leq_T X$ is dominated by a computable function, i.e. there is some computable function g such that $f(n) \leq g(n)$ for every $n \in \omega$.

Theorem 2.6 (Hyperimmune-Free Basis Theorem). *For every Π^0_1 class $\mathcal{P}$, there is $X \in \mathcal{P}$ that has hyperimmune-free degree.*

$X \in 2^\omega$ has *c.e. degree* if there is some c.e. set W such that $W \equiv_T X$.

Theorem 2.7 (Kreisel Basis Theorem). *For every Π^0_1 class $\mathcal{P}$, there is $X \in \mathcal{P}$ that has c.e. Turing degree.*

It is also worth noting that every isolated path through a Π^0_1 class is computable.

Starting with Σ^0_1 classes and Π^0_1 classes, we can define more complicated effectively Borel sets. First note that there are effective enumerations of all Σ^0_1 classes $\{\mathcal{U}_i\}_{i\in\omega}$ and of all Π^0_1 classes $\{\mathcal{P}_i\}_{i\in\omega}$

- A Π^0_2 class is an effective intersection of Σ^0_1 classes. Given a uniform collection of Σ^0_1 classes $\{\mathcal{U}_{f(i)}\}_{i\in\omega}$ for some computable function f,

$$\mathcal{S} = \bigcap_{i\in\omega} \mathcal{U}_{f(i)}$$

  is a Π^0_2 class.

- A Σ^0_2 class is an effective union of Π^0_1 classes. Given a uniform collection of Π^0_1 classes $\{\mathcal{P}_{g(i)}\}_{i\in\omega}$ for some computable function g,

$$\mathcal{T} = \bigcup_{i\in\omega} \mathcal{P}_{g(i)}$$

is a Σ^0_2 class.

Similarly, Π^0_{n+1} classes are effective intersections of Σ^0_n classes, and Σ^0_{n+1} classes are effective unions of Π^0_n classes.

2.4.2 Computable Measures

A Borel probability measure on 2^ω is a non-negative, countably additive function $\mu : \mathscr{P}(2^\omega) \to [0,1]$ such that $\mu(\emptyset) = 0$ and $\mu(\mathcal{S}) \in [0,1]$ for every Borel $\mathcal{S} \subseteq 2^\omega$. However, by Caratheodory's theorem from classical measure theory, a function μ defined on cylinders and satisfying for all σ the identity $\mu([\![\sigma]\!]) = \mu([\![\sigma 0]\!]) + \mu([\![\sigma 1]\!])$ can be uniquely extended to a probability measure, and hence it is sufficient to consider the restriction of probability measures to cylinders. We can therefore represent measures as functions from $2^{<\omega}$ to [0,1], where for all $\sigma \in 2^{<\omega}$, $\mu(\sigma)$ is the measure of the cylinder $[\![\sigma]\!]$. This concise representation also allows us to talk about *computable* probability measures.

Before we give the definition of a computable measure, we need to define several notions from computable analysis.

Definition 2.8. (i) A real number $r \in [0,1]$ is *left-c.e.* if there is a uniformly computable, non-decreasing sequence of rationals converging to r.

(ii) A real number $r \in [0,1]$ is *right-c.e.* if there is a uniformly computable, non-increasing sequence of rationals converging to r.

(iii) A real number $r \in [0,1]$ is *computable* if it both left-c.e. and right-c.e.

It's not hard to show that $r \in [0,1]$ is computable if and only if there is a uniformly computable sequence of rationals $\{q_n\}_{n\in\omega}$ such that

$$|r - q_n| \leq 2^{-n}.$$

Definition 2.9. A probability measure μ on 2^ω is *computable* if $\sigma \mapsto \mu(\sigma)$ is computable as a real-valued function, i.e. if there is a computable function $\widehat{\mu} : 2^{<\omega} \times \omega \to \mathbb{Q}_2$ such that

$$|\mu(\sigma) - \widehat{\mu}(\sigma, i)| \leq 2^{-i}$$

for every $\sigma \in 2^{<\omega}$ and $i \in \omega$. We further say that μ is *exactly computable* if for all σ $\mu(\sigma) \in \mathbb{Q}_2$ and $\sigma \mapsto \mu(\sigma)$ is computable as a function from $2^{<\omega}$ to $\mathbb{Q}_2$.

The collection of all measures on 2^ω will be denoted by $\mathscr{M}$, while the collection of computable measures on 2^ω will be denoted by $\mathscr{M}_c$. In what follows λ will refer exclusively to the Lebesgue measure on 2^ω, where $\lambda(\sigma) = 2^{-|\sigma|}$ for each $\sigma \in 2^{<\omega}$. The Lebesgue measure is particularly important, as it is the "default measure" for the many definitions of randomness. For our purposes, it will be useful to identify several different types of measures.

Definition 2.10. Let $\mu \in \mathscr{M}$.

(i) μ is *positive* if $\mu(\sigma) > 0$ for every $\sigma \in 2^{<\omega}$. Equivalently, μ is positive if $\mu(\mathcal{U}) > 0$ for every non-empty open set $\mathcal{U}$.

(ii) μ is *atomic* if there is some sequence $A \in 2^\omega$ such that $\mu(\{A\}) > 0$. In this case, we call A an *atom* of μ or a μ-atom. The collection of μ-atoms is denoted

Atoms_μ.

(iii) μ is *atomless* if μ has no atoms.

(iv) μ is trivial if $\mu(\mathsf{Atoms}_\mu) = 1$.

An important result is the following.

Proposition 2.11 (Kautz). *$X \in 2^\omega$ is computable if and only if $X \in \mathsf{Atoms}_\mu$ for some $\mu \in \mathscr{M}_c$.*

Proof. If X is computable, we can define $\mu \in \mathscr{M}_c$ such that

$$\mu(\sigma) = \begin{cases} 1 & \text{if } \sigma = X{\upharpoonright}n \text{ for some } n \in \omega \\ 0 & \text{otherwise} \end{cases}.$$

For the other direction, if $X \in \mathsf{Atoms}_\mu$, then it follows that $\mu(\{X\}) > q$ for some $q \in \mathbb{Q}_2$. Using the approximation $\widehat{\mu}$ of μ, define a computable tree as follows:

$$T = \{\sigma \in 2^{<\omega} : \widehat{\mu}(\sigma, |\sigma|) \geq q - 2^{-|\sigma|}\}.$$

$q \in \mathbb{Q}_2$ implies that $q = \dfrac{k}{2^m}$ for some $k, m \in \omega$. For each $n \geq m$,

$$\frac{1}{q - 2^{-n}} = \frac{1}{\frac{k}{2^m} - 2^{-n}} = \frac{2^m}{k - 2^{m-n}} \leq \frac{2^m}{k-1}.$$

It follows that T contains at most $\frac{2^m}{k-1}$ strings of length n for every $n \geq m$. Then $X \in [T]$, but as there are at most $\frac{2^m}{k-1}$ paths through T, X must be an isolated path, and is thus computable.

$\square$

2.4.3 Turing Functionals and Induced Measures

We define the class of "almost total" Turing functionals. Recall that we consider Turing functionals as maps from 2^ω to 2^ω, so that for $X \in 2^\omega$, $\Phi(X)\downarrow$ if and only if $\Phi(X) \in 2^\omega$ if and only if $X \in \mathsf{dom}(\Phi)$.

Definition 2.12. A Turing functional Φ is *almost total* if $\lambda(\mathsf{dom}(\Phi)) = 1$.

To be clear, if Φ is almost total, this means that $\Phi(X)$ is total as a map from ω to $\{0, 1\}$ for measure one many $X \in 2^\omega$. It is immediate that every truth-table functional is almost total.

We are interested in studying almost total functionals due to the fact that they can be used to induce computable measures. For a Turing functional Φ and $\mathcal{S} \in 2^\omega$,

$$\Phi^{-1}(\mathcal{S}) = \{X \in 2^\omega : \Phi(X)\downarrow \in \mathcal{S}\}.$$

Definition 2.13. Given an almost total functional $\Phi : 2^\omega \to 2^\omega$, the measure induced by Φ, denoted λ_Φ, is defined to be

$$\lambda_\Phi(\mathcal{X}) = \lambda(\Phi^{-1}(\mathcal{X}))$$

for every measurable $\mathcal{X} \subseteq 2^\omega$.

In general, we can consider the measure induced by a probability measure μ and a functional Φ, but we have to make the further restriction that Φ is almost total *with respect to* μ, i.e., $\mu(\mathsf{dom}(\Phi)) = 1$. Thus we have the following definition.

Definition 2.14. Given a measure μ on 2^ω and a μ-almost total functional $\Phi : 2^\omega \to 2^\omega$, the measure *induced by* (μ, Φ), denoted μ_Φ, is defined to be

$$\mu_\Phi(\mathcal{X}) = \mu(\Phi^{-1}(\mathcal{X}))$$

for every μ-measurable $\mathcal{X} \subseteq 2^\omega$.

Intuitively, μ_Φ can be computed if μ and Φ are given. This is formalized by the following lemma.

Lemma 2.15. *For a given measure μ on 2^ω and a functional $\Phi : 2^\omega \to 2^\omega$, the following hold.*

1. *If μ is computable and Φ is a μ-almost total reduction, then μ_Φ is computable.*

2. *If μ is exactly computable and Φ is a tt-reduction, then μ_Φ is exactly computable.*

Proof. We proceed inductively as follows: First,

$$\mu_\Phi(\emptyset) = \mu(\Phi^{-1}([\![\emptyset]\!])) = \mu(\Phi^{-1}(\mathsf{dom}(\Phi))) = 1,$$

since Φ is almost total. Now suppose that $\mu_\Phi(\sigma)$ is computable. Then $\mu_\Phi(\sigma 0)$ and $\mu_\Phi(\sigma 1)$ are both approximable from below, and since $\mu_\Phi(\sigma) = \mu_\Phi(\sigma 0) + \mu_\Phi(\sigma 1)$, it follows that both $\mu_\Phi(\sigma 0)$ and $\mu_\Phi(\sigma 1)$ are approximable from above. Thus, both are computable.

For the second part, let f be a computable function such that for every $X, Y \in 2^\omega$, if $\Phi(X) = Y$, then for every $n \in \omega$, $\Phi^{X \restriction f(n)} \succeq Y \restriction n$. Without loss of generality, we can assume that if $|\sigma| = n$ and $|\tau| < f(n)$, then $\Phi^\tau \not\succeq \sigma$. If we define

$$\mathsf{Pre}_\Phi(\sigma) := \{\tau \in 2^{<\omega} : \Phi^\tau \succeq \sigma \ \wedge\ (\forall \tau' \preceq \tau)\Phi^{\tau'} \not\succeq \sigma\},$$

(so that $[\![\mathsf{Pre}_\Phi(\sigma)]\!] = \Phi^{-1}([\![\sigma]\!])$),[4] it follows that

$$\mathsf{Pre}_\Phi(\sigma) = \{\tau \in 2^{f(|\sigma|)} : \Phi^\tau \succeq \sigma\}$$

and thus

$$\mu_\Phi(\sigma) = \mu(\Phi^{-1}([\![\sigma]\!])) = \mu\Big(\bigcup_{\tau \in \mathsf{Pre}_\Phi(\sigma)} [\![\tau]\!]\Big) = \sum_{\tau \in \mathsf{Pre}_\Phi(\sigma)} \mu(\tau),$$

which is $\mathbb{Q}_2$-valued because μ is $\mathbb{Q}_2$-valued and $\mathsf{Pre}_\Phi(\sigma)$ is finite. Moreover, since we can find, effectively in σ, the index for $\mathsf{Pre}_\Phi(\sigma)$ as a finite set, if follows that μ_Φ is a computable function from $2^{<\omega}$ to $\mathbb{Q}_2$, and thus is exactly computable. $\square$

We can also show that the induced measure μ_Φ as defined above shares certain features of the original measure μ as long as the functional Φ satisfies some additional condition:

Lemma 2.16. *Let μ be a measure on 2^ω and let Φ be a μ-almost total functional. Then the following hold.*

1. *If μ is atomless and Φ is one-to-one, then μ_Φ is atomless.*

[4] In general, $[\![\mathsf{Pre}_\Phi(\sigma)]\!] \supseteq \Phi^{-1}([\![\sigma]\!])$ holds for every Turing functional, but if Φ is total, then the reverse containment holds as well.

2. If μ is positive and Φ is onto, then μ_Φ is positive.

Proof. Suppose μ is atomless and Φ is one-to-one. Then for all $X \in 2^\omega$, $\mu_\Phi(\{X\}) = \mu(\Phi^{-1}(\{X\}))$. Since Φ is one-to-one, $\Phi^{-1}(\{X\})$ is either empty or is a singleton. In the former case, clearly $\mu(\Phi^{-1}(\{X\})) = 0$, while in the latter case, since μ is atomless, it also follows that $\mu(\Phi^{-1}(\{X\})) = 0$.

Suppose now that μ is positive and Φ is onto. Let $\mathcal{U}$ be a non-empty open set. Since Turing functionals are continuous on their domain and Φ is onto, it follows that $\Phi^{-1}(\mathcal{U})$ is non-empty and open in $\mathsf{dom}(\Phi)$. Thus, $\Phi^{-1}(\mathcal{U}) = \mathsf{dom}(\Phi) \cap \mathcal{V}$ for some non-empty $\mathcal{V} \subseteq 2^\omega$ that is open in 2^ω. But since Φ is almost total and μ is positive, we have

$$\mu_\Phi(\mathcal{U}) = \mu(\Phi^{-1}(\mathcal{U})) = \mu(\mathsf{dom}(\Phi) \cap \mathcal{V}) = \mu(\mathcal{V}) > 0.$$

$\square$

2.5 Notions of Algorithmic Randomness

In this section we review a number of definitions of algorithmic randomness. While most presentations found in the literature present these definitions with respect to the Lebesgue measure on 2^ω, here we will consider the various definitions of randomness with respect to *any* computable measure on 2^ω. Proofs will be provided for most of the results below, especially in those cases in which there is no proof available in the algorithmic randomness literature (which is the case for many results given in terms of an arbitrary computable measure, a case that usually calls for a slight modification of the standard proof found in [DH10] or [Nie09]).

2.5.1 Martin-Löf Randomness

Martin-Löf randomness is the most well-studied, and in many respects, the most well-behaved definition of algorithmic randomness. In what follows, we will consider three equivalent formulations of Martin-Löf randomness: the measure-theoretic definition, the martingale definition, and the incompressibility definition.

2.5.1.1 Measure-Theoretic Formulation

When Martin-Löf first presented his definition of randomness in 1966 (in [ML66]), he formulated it in terms of certain effective statistical tests, the idea being that a random sequence is not detected as non-random by any such test. Nowadays, these tests are referred to as *Martin-Löf tests*.[5]

Definition 2.17. Given $\mu \in \mathscr{M}_c$, a *μ-Martin-Löf test* is a uniformly computable sequence $\{\mathcal{U}_i\}_{i\in\omega}$ of effectively open classes in 2^ω such that $\mu(\mathcal{U}_i) \leq 2^{-i}$ for every $i \in \omega$. Further, a real X is *μ-Martin-Löf random* if for every μ-Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$, we have $X \notin \bigcap_{i\in\omega} \mathcal{U}_i$.

The collection of μ-Martin-Löf random reals will be written as MLR_μ. However, when $\mu = \lambda$, the Lebesgue measure, we will write MLR_λ simply as MLR.

The following result, originally proved by Martin-Löf, is very useful.

[5] For a discussion of the philosophical motivation behind Martin-Löf's definition, see Chapter 9, Section 9.5.

Proposition 2.18. *For every μ, there is a Martin-Löf test $\{\widehat{\mathcal{U}}_i\}_{i\in\omega}$ (called the* universal μ-Martin-Löf test*) such that $X \in \mathsf{MLR}_\mu$ if and only if $x \notin \bigcap_{i\in\omega} \widehat{\mathcal{U}}_i$.*

One simple result concerning Martin-Löf randomness for various computable measures, which has not been noted previously, will be useful for our later discussion.

Lemma 2.19. *For $\mu, \nu \in \mathscr{M}_c$, $\mathsf{MLR}_\mu \cup \mathsf{MLR}_\nu = \mathsf{MLR}_\rho$, where $\rho = \frac{\mu+\nu}{2}$.*

Proof. Suppose $X \notin \mathsf{MLR}_\mu \cup \mathsf{MLR}_\nu$. Then there is a μ-Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$ and a ν-Martin-Löf test $\{\mathcal{V}_i\}_{i\in\omega}$ such that $X \in \bigcap_{i\in\omega} \mathcal{U}_i$ and $X \in \bigcap_{i\in\omega} \mathcal{V}_i$. If we set $\mathcal{W}_i := \mathcal{U}_i \cap \mathcal{V}_i$, then $\mathcal{W}_i$ is Σ^0_1 (as $[\![\sigma]\!]$ is enumerated into $\mathcal{W}_i$ only after we've seen it enumerated into $\mathcal{U}_i$ and $\mathcal{V}_i$) with

$$\mu(\mathcal{W}_i) \leq \mu(\mathcal{U}_i) \leq 2^{-i}$$

since $\mathcal{W}_i \subseteq \mathcal{U}_i$, and

$$\nu(\mathcal{W}_i) \leq \nu(\mathcal{V}_i) \leq 2^{-i},$$

since $\mathcal{W}_i \subseteq \mathcal{V}_i$. It follows that

$$\rho(\mathcal{W}_i) = \frac{1}{2}(\mu(\mathcal{W}_i) + \nu(\mathcal{W}_i)) \leq 2^{-i},$$

and thus $\{\mathcal{W}_i\}_{i\in\omega}$ is a ρ-Martin-Löf test containing X.

For the other direction, suppose that $X \notin \mathsf{MLR}_\rho$. Then if $\{\mathcal{U}_i\}_{i\in\omega}$ is a ρ-Martin-Löf test such that $x \in \bigcap_{i\in\omega} \mathcal{U}_i$, it follows that

$$\frac{\mu(\mathcal{U}_i) + \nu(\mathcal{U}_i)}{2} \leq 2^{-i}$$

and hence

$$\mu(\mathcal{U}_i) \leq \mu(\mathcal{U}_i) + \nu(\mathcal{U}_i) \leq 2^{-i+1}.$$

Thus $\{\mathcal{U}_i\}_{i\geq 1}$ is a μ-Martin-Löf test containing X, and hence $X \notin \mathsf{MLR}_\mu$. Similarly, $X \notin \mathsf{MLR}_\nu$. □

An important consequence of Lemma 2.19 is that it allows us to replace a non-positive measure μ with a positive measure ν without losing any of the μ-Martin-Löf random reals.

Corollary 2.20. *If $\mu \in \mathscr{M}_c$ is not positive, there is a positive $\rho \in \mathscr{M}_c$ such that*

$$\mathsf{MLR}_\mu \subsetneq \mathsf{MLR}_\rho.$$

Proof. Let $\rho(\sigma) = \frac{\mu(\sigma)+\lambda(\sigma)}{2}$ for every $\sigma \in 2^{<\omega}$ (where λ is the Lebesgue measure). Clearly, ρ is positive, as λ is positive. □

2.5.1.2 The Martingale Formulation

An alternative formulation of Martin-Löf randomness can be given in terms of martingales, an approach developed in the 1930s by Jean Ville to formalize the notion of a betting strategy (see [Vil39]). The general idea is this. Suppose we are playing a game in which we attempt to predict the successive values of a sequence. At each round, having seen the first n bits of a sequence X, we bet a certain amount of our capital that the $(n+1)$st bit of X will be a 0 (and a certain amount that the $(n+1)$st bit will be a 1). For each such bet, there is a resulting payoff, which determines the underlying measure in terms of which we define the martingale. For instance, if we playing the game with a double-or-nothing payoff, we are play with a λ-martingale. Moreover, a supermartingale is a betting strategy that permits us to bet some of our capital on 0 and some of our capital on 1 while still setting aside some of our capital to be saved for later rounds.

We first provide the definition of martingales and supermartingales without any constraints on the effectiveness of these betting strategies (and without any constraints on the underlying probability measures), and then we define the classes of computable and computably enumerable martingales and supermartingales.

Definition 2.21. For any $\mu \in \mathscr{M}$, a *μ-martingale* is a function $d : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ such that for every $\sigma \in 2^{<\omega}$,

$$\mu(\sigma)d(\sigma) = \mu(\sigma 0)d(\sigma 0) + \mu(\sigma 1)d(\sigma 1).$$

Further, a *μ-supermartingale* is a function $d : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ such that for every $\sigma \in 2^{<\omega}$,

$$\mu(\sigma)d(\sigma) \geq \mu(\sigma 0)d(\sigma 0) + \mu(\sigma 1)d(\sigma 1). \tag{2.1}$$

Moreover, a μ-martingale (or μ-supermartingale) d *succeeds* on $x \in 2^\omega$ if

$$\limsup_{n\to\infty} d(x{\upharpoonright}n) = \infty.$$

Let $\mathcal{S}_d \subseteq 2^\omega$ denote the collection of sequences on which the martingale (or supermartingale) d succeeds. Supermartingales are often easier to work with in the context of Martin-Löf randomness (but not for alternative definitions of algorithmic randomness), and the following proposition guarantees that there is no loss if we work with supermartingales, rather than just martingales.

Proposition 2.22. *Let $\mu \in \mathscr{M}$. Then for every μ-supermartingale d, there is a μ-martingale d^* such that*

$$\mathcal{S}_d \subseteq \mathcal{S}_{d^*}.$$

Proof. Given a μ-supermartingale d, let the savings function $s : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ be

$$s(\sigma) = d(\sigma) - (\mu(\sigma 0|\sigma)d(\sigma 0) + \mu(\sigma 1|\sigma)d(\sigma 1)),$$

where $\mu(\sigma^\frown i|\sigma) = \dfrac{\mu(\sigma^\frown i)}{\mu(\sigma)}$ for $i \in \{0, 1\}$. Then we set

$$d^*(\sigma) = d(\sigma) + \sum_{\tau \prec \sigma} s(\tau)$$

and $d^*(\sigma) = d(\sigma)$ Then we claim that d^* is a μ-martingale. Suppose that

$$\mu(\sigma)d^*(\sigma) = \mu(\sigma 0)d^*(\sigma 0) + \mu(\sigma 1)d^*(\sigma 1),$$

for every σ of length at most $n-1$. Then given $\sigma \in 2^n$,

$$\begin{aligned}
\mu(\sigma 0)d^*(\sigma 0) + \mu(\sigma 1)d^*(\sigma 1) &= \mu(\sigma 0)\big(d(\sigma 0) + \sum_{\tau \preceq \sigma} s(\tau)\big) + \mu(\sigma 1)\big(d(\sigma 1) + \sum_{\tau \preceq \sigma} s(\tau)\big) \\
&= \mu(\sigma 0)d(\sigma 0) + \mu(\sigma 1)d(\sigma 1) + (\mu(\sigma 0) + \mu(\sigma 1)) \sum_{\tau \preceq \sigma} s(\tau) \\
&= \mu(\sigma 0)d(\sigma 0) + \mu(\sigma 1)d(\sigma 1) + \mu(\sigma) \sum_{\tau \preceq \sigma} s(\tau) \\
&= \mu(\sigma 0)d(\sigma 0) + \mu(\sigma 1)d(\sigma 1) + \mu(\sigma)s(\sigma) + \mu(\sigma) \sum_{\tau \prec \sigma} s(\tau) \\
&= \mu(\sigma)d(\sigma) + \mu(\sigma) \sum_{\tau \prec \sigma} s(\tau) \text{ (by the definition of } s) \\
&= \mu(\sigma)d^*(\sigma).
\end{aligned}$$

Lastly, note that $d^*(\sigma) \geq d(\sigma)$ for every $\sigma \in 2^{<\omega}$, and hence $\mathcal{S}_d \subseteq \mathcal{S}_{d^*}$. □

We will also make use of the fact that the average condition in (2.1) can be generalized:

Lemma 2.23. *Let d be a μ-supermartingale and let $\{\sigma_1, \sigma_2, \dots\} \subseteq 2^{<\omega}$ be a prefix-free set of extensions of some $\tau \in 2^{<\omega}$. Then*

$$\sum_i \mu(\sigma_i)d(\sigma_i) \leq \mu(\tau)d(\tau). \tag{2.2}$$

In particular, for any prefix-free $S \subseteq 2^{<\omega}$,

$$\sum_{\sigma \in S} \mu(\sigma)d(\sigma) \leq d(\varnothing).$$

Proof. We show this holds for every finite set $\{\sigma_1, \ldots, \sigma_k\}$, from which it will follow that (2.2) holds for every *infinite* prefix-free collection of strings by induction. First, given $\sigma \succeq \tau$, it follows from the definition of a supermartingale that

$$\mu(\sigma)d(\sigma) \leq \mu(\tau)d(\tau).$$

Now suppose the result holds for every prefix-free set of n strings. Given a prefix-free $S = \{\sigma_1, \ldots, \sigma_{n+1}\} \subseteq 2^{<\omega}$ of extensions of some τ (where τ is the longest such string), then setting

$$S_0 = \{\sigma \in S : \sigma \succeq \tau 0\} \ \wedge \ S_1 = \{\sigma \in S : \sigma \succeq \tau 1\}$$

it follows from our choice of τ that $|S_i| \leq n$ for both $i \in \{0, 1\}$. Thus, by the inductive hypothesis,

$$\sum_{\sigma \in S_0} \mu(\sigma)d(\sigma) \leq \mu(\tau 0)d(\tau 0) \ \wedge \ \sum_{\sigma \in S_1} \mu(\sigma)d(\sigma) \leq \mu(\tau 1)d(\tau 1)$$

and so

$$\sum_{\sigma \in S} \mu(\sigma)d(\sigma) \leq \mu(\tau 0)d(\tau 0) + \mu(\tau 1)d(\tau 1) \leq \mu(\tau)d(\tau).$$

$\square$

Now, we restrict the martingales and supermartingales to a nice effectively approximable collection.

Definition 2.24. A μ-martingale (or μ-supermartingale) d is *computably enumer-*

able if the collection of left-cuts given by the values $\{d(\sigma)\}_{\sigma\in 2^{<\omega}}$ is a uniformly c.e. collection of rationals.

Now according to Ville's definition, a sequence X is random if there is no martingale (in some restricted collection of martingales) that wins an unbounded amount of capital when applied to initial segments of X. The following result of Schnorr's shows that this idea, suitably formalized, yields a definition of randomness that is equivalent to Martin-Löf randomness.

Theorem 2.25. *For $\mu \in \mathscr{M}_c$, $X \in \mathsf{MLR}_\mu$ if and only if no c.e. μ-supermartingale succeeds on X.*

To prove Theorem 2.25, two lemmas are needed.

Lemma 2.26. *For any measure μ on 2^ω and a μ-martingale (or μ-supermartingale) d, if we set*

$$\mathcal{U}_i = \{X \in 2^\omega : (\exists n) d(X{\upharpoonright}n) \geq 2^i\}$$

then $\mu(\mathcal{U}_i) \leq 2^{-i}$.

Proof. Let $U_i \subseteq 2^{<\omega}$ be a prefix-free set of strings such that $\mathcal{U}_i = [\![U_i]\!]$. Then by Lemma 2.23 and the definition of $\mathcal{U}_i$,

$$\sum_{\sigma\in U_i} \mu(\sigma)2^i \leq \sum_{\sigma\in U_i} \mu(\sigma)d(\sigma) \leq d(\varnothing)$$

and thus

$$\mu(\mathcal{U}_i) = \sum_{\sigma\in U_i} \mu(\sigma) \leq 2^{-i}.$$

□

Lemma 2.27. *For $\mu \in \mathscr{M}_c$, if $(d_i)_{i\in\omega}$ is a uniform collection of c.e. μ-martingales (or μ-supermartingales) such that for each i, $d_i(\sigma) \in [0,1]$ for every $\sigma \in 2^{<\omega}$, then the function $d : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ defined, for each $\sigma \in 2^{<\omega}$, by*

$$d(\sigma) = \sum_{i\in\omega} d_i(\sigma)$$

is a c.e. μ-martingale (or μ-supermartingale).

Proof. It is routine to verify that the μ-martingale condition is satisfied. □

Proof of Theorem 2.25. (⇒) We prove the contrapositive. Suppose there is $X \in 2^\omega$ and some c.e. μ-supermartingale d such that d succeeds on X. Without loss of generality, we can suppose that $d(\varnothing) \leq 1$. Then setting

$$\mathcal{U}_i = \{X \in 2^\omega : (\exists n) d(X{\restriction}n) \geq 2^i\},$$

$\{\mathcal{U}_i\}_{i\in\omega}$ is a μ-Martin-Löf test, since $\{\mathcal{U}_i\}_{i\in\omega}$ is clearly uniformly c.e. and $\mu(\mathcal{U}_i) \leq 2^{-i}$ by Lemma 2.26. Since d succeeds on X, it follows that $X \in \bigcap_{i\in\omega} \mathcal{U}_i$, and hence $X \notin \mathsf{MLR}_\mu$.

(⇐) We also prove the contrapositive. Suppose that $X \notin \mathsf{MLR}_\mu$. Then there is a μ-Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$ such that $X \in \bigcap_{i\in\omega} \mathcal{U}_i$. Now if for fixed $i \in \omega$ we define

$$\mu(\mathcal{U}_i|\sigma) = \frac{\mu(\mathcal{U}_i \cap [\![\sigma]\!])}{\mu(\sigma)}$$

to be the conditional μ-measure of $\mathcal{U}_i$ given σ, then $\mu(\mathcal{U}_i|\cdot) : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ is a c.e. μ-martingale. To see this, observe that

$$\begin{aligned}
\mu(\sigma)\mu(\mathcal{U}_i|\sigma) &= \mu(\sigma)\frac{\mu(\mathcal{U}_i \cap [\![\sigma]\!])}{\mu(\sigma)} \\
&= \mu(\mathcal{U}_i \cap [\![\sigma]\!]) \\
&= \mu(\mathcal{U}_i \cap [\![\sigma 0]\!]) + \mu(\mathcal{U}_i \cap [\![\sigma 1]\!]) \\
&= \mu(\sigma 0)\frac{\mu(\mathcal{U}_i \cap [\![\sigma 0]\!])}{\mu(\sigma 0)} + \mu(\sigma 1)\frac{\mu(\mathcal{U}_i \cap [\![\sigma 1]\!])}{\mu(\sigma 1)} \\
&= \mu(\sigma 0)\mu(\mathcal{U}_i|\sigma 0) + \mu(\sigma 1)\mu(\mathcal{U}_i|\sigma 1).
\end{aligned}$$

Further, $\mu(\mathcal{U}_i|\sigma) \in [0,1]$ for every $\sigma \in 2^{<\omega}$. Thus by Lemma 2.27, the function $d : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ defined to be

$$d(\sigma) = \sum_{i \in \omega} \mu(\mathcal{U}_i|\sigma)$$

for each $\sigma \in 2^{<\omega}$ is a c.e. μ-martingale. Now, since $X \in \bigcap_{i\in\omega} \mathcal{U}_i$, for each i, there is some $n \in \omega$ such that $[\![X{\upharpoonright}n]\!] \subseteq U_i$, and hence for every $n' \geq n$, $\mu(\mathcal{U}_i|X{\upharpoonright}n') = 1$. Thus, for every $c \in \omega$ there is some n such that $d(X{\upharpoonright}n) \geq c$, which implies that d succeeds on X.

□

We conclude this subsection with proposition that guarantees that c.e. supermartingales can be effectively approximated.

Definition 2.28. A *supermartingale approximation* is a uniformly computable sequence $(d_i)_{i\in\omega}$ of supermartingales such that each d_i is $\mathbb{Q}_2$-valued and

$$d_{i+1}(\sigma) \geq d_i(\sigma)$$

for every $i \in \omega$ and $\sigma \in 2^{<\omega}$.

Proposition 2.29. *For every c.e. supermartingale d, there is a supermartingale approximation $(d_i)_{i\in\omega}$ such that $d(\sigma) = \sup_{i\in\omega} d_i(\sigma)$ for every $\sigma \in 2^{<\omega}$.*

Proof. See [Nie09], Fact 7.2.4. □

2.5.1.3 The Incompressibility Formulation

A third formulation of Martin-Löf randomness can be given in terms of initial segment complexity. On this approach, random sequences are those with high initial segment complexity, which means that the initial segments of such sequences cannot be compressed very much. To make this precise, we measure the amount that a string can be compressed by means of Kolmogorov complexity.

Definition 2.30. Given a partial computable function $\phi : 2^{<\omega} \to 2^{<\omega}$ and some $\sigma \in 2^{<\omega}$, the *plain Kolmogorov complexity of σ with respect to ϕ* if

$$C_\phi(\sigma) = \min\{|\tau| : \phi(\tau) = \sigma\}.$$

If we define a universal partial computable $\widehat{\phi} : 2^{<\omega} \to 2^{<\omega}$ so that $\widehat{\phi}(1^e 0\tau) = \phi_e(\tau)$

whenever $\phi_e(\tau)\downarrow$, then it follows that

$$C_{\widehat{\phi}}(\sigma) \leq C_{\phi_e}(\sigma) + e + 1$$

for $e \in \omega$ and $\sigma \in 2^{<\omega}$. Thus, we set

$$C(\sigma) := C_{\widehat{\phi}}(\sigma)$$

and call this the *plain Kolmogorov complexity of* σ.

With this definition of Kolmogorov complexity, we can now define what it means for a string to be incompressible.

Definition 2.31. For $c \in \omega$ and $\sigma \in 2^{<\omega}$, σ is c-incompressible$_C$ if

$$C(\sigma) \geq |\sigma| - c.$$

That is, c-incompressible$_C$ strings cannot be compressed more than c bits below their length. With this definition, one reasonable suggestion is to define $X \in 2^\omega$ to be incompressible if there is some c such that every $n \in \omega$, $X \restriction n$ is c-incompressible$_C$. However, Martin-Löf showed that no sequence has this property.

Proposition 2.32. *For every $X \in 2^\omega$, for every $c \in \omega$, there is some $n \in \omega$ such that*

$$C(X \restriction n) < n - c.$$

Proof. See [DH10], Theorem 3.1.4. $\square$

However, Martin-Löf was able to show the following:

Theorem 2.33. *For $X \in 2^\omega$, if there is some $c \in \omega$ such that*

$$(\exists^\infty n) C(X{\upharpoonright}n) \geq n - c$$

then $X \in \mathsf{MLR}$.

Proof. See, for instance, the proof of Theorem 5 in [ML71]. $\square$

To remedy this problem, a number of suggestions were made to restrict the collection of partial computable functions, thereby modifying the notion of complexity, so as to prevent 2.32 from occurring. One suggestion made independently by Levin ([Lev10]) and Chaitin ([Cha75]) is to restrict to prefix-free free machines, which we introduced in Subsection 2.3.2 above. Thus we have:

Definition 2.34. Given a prefix-free machine $M : 2^{<\omega} \to 2^{<\omega}$ and some $\sigma \in 2^{<\omega}$, the *prefix-free Kolmogorov complexity of σ with respect to M* if

$$K_M(\sigma) = \min\{|\tau| : M(\tau) = \sigma\}.$$

As before, if we let U be a universal prefix-free machine, so that $U(1^e 0\tau) = M_e(\tau)$ whenever $M_e(\tau)\downarrow$, then we define $K(\sigma) := K_U(\sigma)$ to be the *prefix-free Kolmogorov complexity of σ.*

The analogue of the c-incompressible$_C$ strings are the c-incompressible$_K$ strings.

Definition 2.35. For $c \in \omega$ and $\sigma \in 2^{<\omega}$, σ is c-incompressible$_K$ if

$$K(\sigma) \geq |\sigma| - c.$$

Now if we define a sequence $X \in 2^\omega$ to be incompressible if there is some $c \in \omega$ such that for every $n \in \omega$, $X{\upharpoonright}n$ is c-incompressible$_K$, then the result is a non-empty notion, as shown independently by Levin and Schnorr.

Theorem 2.36 (Levin-Schnorr)**.** *For $X \in 2^\omega$, $X \in \mathsf{MLR}$ if and only if there is some such $c \in \omega$ such that*

$$K(X{\upharpoonright}n) \geq n - c$$

for every $n \in \omega$.

This result actually holds in the more general case where we define the threshold of incompressibility in terms of a computable measure μ.

Theorem 2.37. *For $\mu \in \mathscr{M}_c$, $X \in \mathsf{MLR}_\mu$ if and only if there is some $c \in \omega$ such that*

$$K(X{\upharpoonright}n) \geq -\log \mu(X{\upharpoonright}n) - c$$

for every $n \in \omega$.

To prove one direction of the Levin-Schnorr Theorem and its more general counterpart, we need an important auxiliary result known as the Machine Existence Theorem (also referred to as the KC Theorem).

Theorem 2.38 (Machine Existence Theorem)**.** *Let $W = \{(n_i, \tau_i)\}_{i \in \omega} \subseteq \omega \times 2^{<\omega}$ be* *a c.e. set of pairs satisfying*

$$\sum_{(n_i,\tau_i)\in W} 2^{-n} \leq 1.$$

Then there is a prefix-free machine M and a prefix-free collection of strings $\{\sigma_i\}_{i\in\omega}$ such that

$$|\sigma_i| = n_i \ \wedge \ M(\sigma_i) = \tau_i$$

for every $i \in \omega$.

Proof. See [Nie09], Proposition 2.2.14. $\square$

Henceforth, the set W satisfying the conditions of the Machine Existence Theorem will be referred to as a *bounded request set*.

Let us now prove the more general version of the Levin-Schnorr Theorem.

Proof of Theorem 2.36. ($\Rightarrow$) Given $\mu \in \mathscr{M}_c$, for each $k \in \omega$, let

$$U_k = \{\sigma : K(\sigma) \leq -\log\mu(\sigma) - k\}.$$

We claim that $\mu(\llbracket U_k \rrbracket) \leq 2^{-k}$. To see this, for each σ, let τ_σ be such that $U(\tau_\sigma) = \sigma$ and $|\tau_\sigma| \leq -\log\mu(\sigma) - k$. Then

$$\mu(\llbracket U_k \rrbracket) \leq \sum_{\sigma\in U_k} \mu(\sigma) = \sum_{\sigma\in U_k} 2^{\log\mu(\sigma)} \leq \sum_{\sigma\in U_k} 2^{-|\tau_\sigma|-k} \leq 2^{-k} \sum_{\tau\in \mathsf{dom}(U)} 2^{-|\tau|} \leq 2^{-k}.$$

Setting $\mathcal{U}_k := \llbracket U_k \rrbracket$, it follows that $\{\mathcal{U}_i\}_{i\in\omega}$ is a μ-Martin-Löf test. Now given $X \in 2^\omega$ such that

$$(\forall k)(\exists n)[K(X{\upharpoonright}n) \leq -\log\mu(X{\upharpoonright}n) - k],$$

it follows that $X \in \bigcap_{i\in\omega} \mathcal{U}_i$, and hence $X \notin \mathsf{MLR}_\mu$.

($\Leftarrow$) For this direction, we use the Machine Existence Theorem. If $X \in 2^\omega$ is not μ-Martin-Löf random, then there is some μ-Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$ such that $A \in \bigcap_i \mathcal{U}_i$. For each $i \in \omega$, let $U_i \subseteq$ be a prefix-free set of strings such that $[\![U_i]\!] = \mathcal{U}_i$. Then we set

$$W = \{(\lceil -\log \mu(\sigma)\rceil - k, \sigma) : k \geq 1 \wedge \sigma \in U_{2k}\}.$$

W is a bounded request set, since it is c.e. and

$$\begin{aligned}\sum_{(n,\tau)\in W} 2^{-n} = \sum_{k\geq 1}\sum_{\sigma\in U_{2k}} 2^{-\lceil -\log \mu(\sigma)\rceil + k} &\leq \sum_{k\geq 1}\sum_{\sigma\in U_{2k}} 2^k \mu(\sigma) \\ &\leq \sum_{k\geq 1} 2^k \mu(U_{2k}) \leq \sum_{k\geq 1} 2^k 2^{-2k} \leq 1.\end{aligned}$$

Thus by the Machine Existence Theorem, there is some prefix-free machine M such that for each $(n,\sigma) \in W$, there is some τ such that $|\tau| = n$ and $M(\tau) = \sigma$. In particular, since for every k there is some n such that $A{\upharpoonright}n \in U_{2k}$, it follows that

$$K_M(A{\upharpoonright}n) \leq \lceil -\log \mu(A{\upharpoonright}n)\rceil - k \leq -\log \mu(A{\upharpoonright}n) - k + 1.$$

$\square$

2.5.1.4 Examples of Martin-Löf Random Sequences

We can provide several examples of Martin-Löf random sequences that will be useful in the sequel.

Example 2.39. Let U be a universal prefix-free Turing machine.

$$\Omega_U := \sum_{\sigma \in \mathsf{dom}(U)} 2^{-|\sigma|}.$$

In what follows, we will fix an underlying universal prefix-free machine U and write $\Omega := \Omega_U$. For a proof that $\Omega \in \mathsf{MLR}$, see [DH10, Theorem 6.1.3]. Some other facts about Ω are as follows:

(i) $\Omega \equiv_T \emptyset'$;

(ii) Ω is left-c.e., meaning that it is the limit of a computable non-decreasing sequence of rationals;

(iii) every left-c.e. Martin-Löf random sequence is equal to Ω_U for some universal prefix-free machine U.

Example 2.40. There is an incomplete Δ^0_2 Martin-Löf random sequence. If we let $\widehat{\mathcal{U}}_i$ be the ith member of the universal Martin-Löf test, then $\widehat{\mathcal{P}}_i = \widehat{\mathcal{U}}_i^c$ is a Π^0_1 class containing only Martin-Löf random sequences. Thus, by the Low Basis Theorem, there is some low $X \in \widehat{\mathcal{P}}_i$, which is thus Δ^0_2 and incomplete.

Example 2.41. There is a Martin-Löf random sequence of hyperimmune-free degree. If $\mathcal{P}$ is a Π^0_1 class consisting entirely of Martin-Löf random sequences (such as the one from the previous example), then by the Hyperimmune-Free Basis Theorem, there is some $X \in \mathcal{P}$ that has hyperimmune-free degree.

2.5.1.5 Comparing the Classes MLR_μ

Let μ be a computable measure, and let $\{p_\sigma\}_{\sigma\in 2^{<\omega}}$ be the collection of conditional probabilities given by μ, i.e.,

$$p_\sigma = \frac{\mu(\sigma 0)}{\mu(\sigma)}$$

for every $\sigma \in 2^{<\omega}$. Given that μ is computable, it follows that $\{p_\sigma\}_{\sigma\in 2^{<\omega}}$ is a uniformly computable collection of real numbers. Note further that μ is positive if and only if $p_\sigma \in (0,1)$ for every $\sigma \in 2^{<\omega}$. Further, if $p_\sigma \in \mathbb{Q}_2$ for every $\sigma \in 2^{<\omega}$, then μ is exactly computable.

Definition 2.42. Let $\mu, \nu \in \mathscr{M}_c$. Then

$$\mathscr{L}^k_{\mu/\nu} := \{X \in 2^\omega : \sup_{n\in\omega} \frac{\mu(X{\upharpoonright}n)}{\nu(X{\upharpoonright}n)} \geq k\}$$

and

$$\mathscr{L}^\infty_{\mu/\nu} := \bigcap_{k\in\omega} \mathscr{L}^k_{\mu/\nu}.$$

Proposition 2.43 (Bienvenu, Merkle)**.** *For $\mu, \nu \in \mathscr{M}_c$,*

$$\mathsf{MLR}_\mu = \mathsf{MLR}_\nu \textit{ if and only if } \mathscr{L}^\infty_{\mu/\nu} \cap \mathsf{MLR}_\mu = \mathscr{L}^\infty_{\nu/\mu} \cap \mathsf{MLR}_\nu = \emptyset.$$

Lemma 2.44. *If μ and ν are positive, computable measures such that $\{p_\sigma\}_{\sigma\in 2^{<\omega}}$ and $\{q_\sigma\}_{\sigma\in 2^{<\omega}}$ are the conditional probabilities given by μ and ν, respectively, then if for every $\sigma \in 2^{<\omega}$ we have*

1. $\dfrac{p_\sigma}{q_\sigma} \leq 2^{(2^{-|\sigma|})}$,

2. $\dfrac{q_\sigma}{p_\sigma} \leq 2^{(2^{-|\sigma|})}$,

3. $\dfrac{1-p_\sigma}{1-q_\sigma} \leq 2^{(2^{-|\sigma|})}$, *and*

4. $\dfrac{1-q_\sigma}{1-p_\sigma} \leq 2^{(2^{-|\sigma|})}$,

then $\mathsf{MLR}_\mu = \mathsf{MLR}_\nu$.

Proof. First observe that for any $X \in 2^\omega$,

$$\begin{aligned}\mu(X{\upharpoonright}n+1) &= \frac{\mu(X{\upharpoonright}1)}{\mu(\varnothing)}\frac{\mu(X{\upharpoonright}2)}{\mu(X{\upharpoonright}1)}\cdots\frac{\mu(X{\upharpoonright}n+1)}{X{\upharpoonright}n} \\ &= \left(\prod_{\{k<n:X(k+1)=0\}} p_{X\upharpoonright k}\right)\left(\prod_{\{k<n:X(k+1)=1\}} 1-p_{X\upharpoonright k}\right)\end{aligned} \tag{2.3}$$

Similarly,

$$\nu(X{\upharpoonright}n+1) = \left(\prod_{\{k<n:X(k+1)=0\}} q_{X\upharpoonright k}\right)\left(\prod_{\{k<n:X(k+1)=1\}} 1-q_{X\upharpoonright k}\right). \tag{2.4}$$

By (2.3), (2.4), and conditions 1 and 2 from the statement of the Lemma, it follows that

$$\frac{\mu(X{\upharpoonright}n+1)}{\nu(X{\upharpoonright}n+1)} = \left(\prod_{\{k<n:X(k+1)=0\}} \frac{p_{X\upharpoonright k}}{q_{X\upharpoonright k}}\right)\left(\prod_{\{k<n:X(k+1)=1\}} \frac{1-p_{X\upharpoonright k}}{1-q_{X\upharpoonright k}}\right) \leq \prod_{k\leq n} 2^{2^{-k}}.$$

It follows that

$$\sup_{k\in\omega}\frac{\mu(X{\upharpoonright}n+1)}{\nu(X{\upharpoonright}n+1)} \leq \sup_{k\in\omega}\prod_{k\leq n} 2^{2^{-k}} \leq \prod_{k\in\omega} 2^{2^{-k}} = 2^{\sum_{k\in\omega}2^{-k}} = 2.$$

and thus $\mathscr{L}^{\infty}_{\mu/\nu} = \emptyset$. Similarly, it follows from (2.3), (2.4), and Conditions 3 and 4 from the statement of the Lemma that

$$\sup_{k\in\omega} \frac{\nu(X{\upharpoonright}(n+1))}{\mu(X{\upharpoonright}(n+1))} \leq 2,$$

and hence $\mathscr{L}^{\infty}_{\nu/\mu} = \emptyset$. Consequently, we have

$$\mathscr{L}^{\infty}_{\mu/\nu} \cap \mathsf{MLR}_{\mu} = \mathscr{L}^{\infty}_{\nu/\mu} \cap \mathsf{MLR}_{\nu} = \emptyset,$$

and so by Proposition 2.43, we have $\mathsf{MLR}_{\mu} = \mathsf{MLR}_{\nu}$. $\square$

Theorem 2.45. *For every positive $\mu \in \mathscr{M}_c$, there is an exactly computable $\nu \in \mathscr{M}_c$ such that $\mathsf{MLR}_{\mu} = \mathsf{MLR}_{\nu}$.*

Proof. Let $\{p_\sigma\}_{\sigma\in 2^{<\omega}}$ be the collection of conditional probabilities given by μ, for each $\sigma \in 2^{<\omega}$, where $p_\sigma \in (0,1)$ for every $\sigma \in 2^{<\omega}$, since μ is positive. Since $\{p_\sigma\}_{\sigma\in 2^{<\omega}}$ is uniformly computable, we can approximate p_σ from below via $\{p_{\sigma,s}\}_{s\in\omega}$, where $p_{\sigma,s} \in \mathbb{Q}_2$ for every $s \in \omega$.

Now, we claim that there is a computable function $f : 2^{<\omega} \to \omega$ such that for every $\sigma \in 2^{<\omega}$,

$$\frac{p_\sigma}{p_{\sigma,f(\sigma)}} \leq 2^{(2^{-|\sigma|})} \text{ and } \frac{1-p_{\sigma,f(\sigma)}}{1-p_\sigma} \leq 2^{(2^{-|\sigma|})}.$$

Given $\sigma \in 2^{<\omega}$, to define $f(\sigma)$, look for the least s such that

$$\frac{p_\sigma}{p_{\sigma,s}} \leq 2^{(2^{-|\sigma|})}, \tag{2.5}$$

which must exist since $p_{\sigma,t} \le p_\sigma$ for all $t \in \omega$. Note that

$$\frac{p_\sigma}{p_{\sigma,s}} \le 2^{(2^{-|\sigma|})} \Rightarrow (\forall t \ge s)\frac{p_\sigma}{p_{\sigma,t}} \le 2^{(2^{-|\sigma|})},$$

for by Equation (2.5) and the fact that $\{p_{\sigma,s}\}$ is non-decreasing in s,

$$p_\sigma \le 2^{(2^{-|\sigma|})}p_{\sigma,s} \le 2^{(2^{-|\sigma|})}p_{\sigma,t}$$

for any $t \ge s$. Now, look for the least $s' \ge s$ such that

$$\frac{1-p_{\sigma,s'}}{1-p_\sigma} \le 2^{(2^{-|\sigma|})},$$

Such a t must exist since (i) $1 - p_\sigma \le 1 - p_{\sigma,t}$ for all $t \in \omega$ and (ii) $\left\{\frac{1-p_{\sigma,t}}{1-p_\sigma}\right\}_{t\in\omega}$ is non-increasing, which can be routinely verified.

Now, setting $f(\sigma) = s'$ and $q_\sigma := p_{\sigma,f(\sigma)}$ for each $\sigma \in 2^{<\omega}$, we claim that $\{p_\sigma)_{\sigma\in 2^{<\omega}}$ and $\{q_\sigma\}_{\sigma\in 2^{<\omega}}$ satisfy the four conditions of Lemma 2.44:

1. $\dfrac{p_\sigma}{q_\sigma} = \dfrac{p_\sigma}{p_{\sigma,f(\sigma)}} \le 2^{(2^{-|\sigma|})}$ by the definition of f;
2. $\dfrac{q_\sigma}{p_\sigma} = \dfrac{p_{\sigma,f(\sigma)}}{p_\sigma} \le 1 \le 2^{(2^{-|\sigma|})}$;
3. $\dfrac{1-p_\sigma}{1-q_\sigma} = \dfrac{1-p_\sigma}{1-p_{\sigma,f(\sigma)}} \le 1 \le 2^{(2^{-|\sigma|})}$; and
4. $\dfrac{1-q_\sigma}{1-p_\sigma} = \dfrac{1-p_{\sigma,f(\sigma)}}{1-p_\sigma} \le 2^{(2^{-|\sigma|})}$ by the definition of f.

Letting ν be the measure defined in terms of the conditional probabilities $\{q_\sigma\}_{\sigma\in 2^{<\omega}}$, it thus follows from Lemma 2.44 that $\mathsf{MLR}_\mu = \mathsf{MLR}_\nu$. □

In general, we cannot guarantee that for *every* $\mu \in \mathscr{M}_c$, there is some exactly computable $\nu \in \mathscr{M}_c$ such that $\mathsf{MLR}_\mu = \mathsf{MLR}_\nu$. To show this, we use the following lemma.

Lemma 2.46 (Bienvenu, Merkle)**.** *For $\mu, \nu \in \mathscr{M}_c$, $\mathsf{MLR}_\mu = \mathsf{MLR}_\nu$ implies that $\mu(\sigma) > 0$ if and only if $\nu(\sigma) > 0$ for every $\sigma \in 2^{<\omega}$.*

Proposition 2.47. *There is $\mu \in \mathscr{M}_c$ that is not positive such that for any exactly computable measure ν, $\mathsf{MLR}_\mu \neq \mathsf{MLR}_\nu$.*

Proof. For each $i \in \omega$, define $\tau_i := 0^i 1$. Then we define a computable measure μ on $\{\tau_i\}$ as follows. Let

$$S = \{\langle i, s\rangle : (\exists s)\phi_{i,s}(i)\downarrow \ \wedge \ \phi_{i,s-1}(i)\uparrow\}.$$

On the assumption that for each s there is at most one $i \in \omega$ with $\langle i, s\rangle \in S$, let

$$S^* = \{s : \exists i \langle i, s\rangle \in S\}.$$

Since by the convention that $\phi_{i,s}(i)\downarrow$ implies that $i \leq s$, S^* is a computable set and thus can be listed as

$$S^* = \{s_1 < s_2 < \dots\}$$

Then we define

$$\mu(\tau_i) = \begin{cases} 2^{-k} & \text{if } \langle i, s_k\rangle \in S \\ 0 & \text{otherwise} \end{cases}.$$

Moreover, we define $\mu(\tau_i^\frown\sigma) = 2^{-\sigma}\mu(\tau_i)$ for every $i \in \omega$ and $\sigma \in 2^{<\omega}$.

Now suppose there is some exactly computable $\nu \in \mathscr{M}_c$ such that $\mathsf{MLR}_\mu = \mathsf{MLR}_\nu$. Then by Lemma 2.46, $\mu(\tau_i) > 0$ if and only if $\nu(\tau_i) > 0$. This implies that

$$\phi_i(i)\downarrow \text{ if and only if } \nu(\tau_i) > 0,$$

but since ν is exactly computable, the relation on the right-hand side of the biconditional is a computable relation, contradicting the unsolvability of the halting problem. □

2.5.2 Weaker Definitions of Algorithmic Randomness

Now we consider definitions of algorithmic randomness that are strictly weaker than Martin-Löf randomness, in the sense that every Martin-Löf random sequence is counted as random according to these definitions, but additional sequences are counted as random as well.

2.5.2.1 Computable Randomness

Computably random sequences are those sequences on which no computable martingale succeeds.

Definition 2.48. A martingale d is *computable* if the collection of left-cuts given by the values $\{d(\sigma)\}_{\sigma \in 2^{<\omega}}$ is a uniformly computable collection of rationals.

Example 2.49. Given positive $\mu \in \mathscr{M}_c$, then for every $\nu \in \mathscr{M}_c$, the function $d_\nu : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ defined by

$$d_\nu(\sigma) = \frac{\nu(\sigma)}{\mu(\sigma)}$$

for every $\sigma \in 2^{<\omega}$ is a computable μ-martingale. Conversely, for every computable μ-martingale d, there is some $\nu \in \mathscr{M}_c$ such that $d = \dfrac{\nu}{\mu}$.

Without loss of generality, we can restrict our attention to *exactly computable* martingales, where a martingale d is exactly computable if d is computable as a function from $2^{<\omega}$ to $\mathbb{Q}_2$.

Proposition 2.50. *For every exactly computable $\mu \in \mathscr{M}_c$ and every computable μ-martingale d, there is an exactly computable μ-martingale d' such that for every X, $X \in \mathcal{S}_d$ if and only if $X \in \mathcal{S}_{d'}$.*

Proof Idea. Define d' so that $d(\sigma) \leq d'(\sigma) \leq d(\sigma) + 2$. The full proof is a slight modification of the proof of Theorem 7.3.8. of [Nie09]. $\square$

Definition 2.51. $X \in 2^\omega$ is *μ-computably random* if there is no computable μ-martingale d that succeeds on X. The collection of μ-computably random reals will be written as CR_μ (unless $\mu = \lambda$, in which case we will simply write CR).

Since every computable μ-martingale is a c.e. μ-martingale, it follows that $\mathsf{MLR}_\mu \subseteq \mathsf{CR}_\mu$. However, in general, the reverse containment does not hold.

Theorem 2.52. *There exists $X \in \mathsf{CR} \setminus \mathsf{MLR}$.*

Proof. See [Nie09], Section 7.4. $\square$

In fact, we can find computably random sequences that aren't Martin-Löf random in every high degree.

Theorem 2.53 ([NST05])**.** *For every high degree $\mathbf{a}$, there is some $X \in \mathsf{CR} \setminus \mathsf{MLR}$ such that $X \in \mathbf{a}$.*

Proof. See [Nie09], Section 7.5. □

2.5.2.2 Schnorr Randomness

Schnorr randomness is another definition of randomness that results from slightly perturbing the definition of Martin-Löf randomness.

Definition 2.54. Given a computable measure μ, a μ-Schnorr test is a uniformly computable sequence $\{\mathcal{U}_i\}_{i\in\omega}$ of effectively open classes in 2^ω such that $\mu(\mathcal{U}_i) = 2^{-i}$ for every $i \in \omega$. Further, a real x is μ-Schnorr if for every μ-Schnorr test $\{\mathcal{U}_i\}_{i\in\omega}$, we have $x \notin \bigcap_{i\in\omega} \mathcal{U}_i$. The collection of μ-Schnorr random reals will be written as SR_μ (unless $\mu = \lambda$, in which case we will simply write SR).

An alternative formulation of Schnorr randomness can be given in terms of computable martingales. To prove this, we first need a lemma.

Lemma 2.55. *For $\mu \in \mathscr{M}_c$, $X \in \mathsf{SR}_\mu$ if and only if for every uniformly computable sequence $(D_n)_{n\in\omega}$ of finite sets of strings such that $\mu([\![D_n]\!]) \leq 2^{-n}$, X is contained in $[\![D_n]\!]$ for at most finitely many n.*

Proof. The proof is a straightforward generalization of the proof of Lemma 1.5.9 in [Bie08]. □

Theorem 2.56. *For $\mu \in \mathscr{M}_c$, $X \in \mathsf{SR}_\mu$ if and only if there is no computable μ-martingale d and no computable order $g : \omega \to \omega$ such that*

$$d(X\restriction n) \geq g(n)$$

for infinitely many $n \in \omega$

Proof. See [Bie08], Theorem 1.5.10. $\square$

It's not hard to see that $\mathsf{CR}_\mu \subseteq \mathsf{SR}_\mu$, but again, the reverse containment does not hold in general.

Theorem 2.57. *There is some $X \in \mathsf{SR} \setminus \mathsf{CR}$.*

Proof. See [Nie09], Theorem 7.5.10. $\square$

As with members of $\mathsf{CR} \setminus \mathsf{MLR}$, we can find members of $\mathsf{SR} \setminus \mathsf{CR}$ in every high degree.

Theorem 2.58 ([NST05]). *For every high degree $\mathbf{a}$, there is some $X \in \mathsf{SR} \setminus \mathsf{CR}$ such that $X \in \mathbf{a}$.*

In fact, being high is also necessary for being a member of $\mathsf{SR} \setminus \mathsf{MLR}$.

Proposition 2.59 ([NST05]). *For every $\mu \in \mathscr{M}_c$ and every $X \in 2^\omega$, if $X \in \mathsf{SR}_\mu \setminus \mathsf{MLR}_\mu$, then X has high Turing degree.*

To prove this proposition, we rely on the following well-known result.

Proposition 2.60. *$X \in 2^\omega$ has high Turing degree if and only if there is some $f \leq_T X$ that dominates all computable functions. That is, for every computable function g,*

$$(\exists m)(\forall n \geq m) g(n) \leq f(n).$$

Proof of Proposition 2.59. Since $X \in \mathsf{SR}_\mu \setminus \mathsf{MLR}_\mu$, let $\{\mathcal{U}_i\}_{i\in\omega}$ be a μ-Martin-Löf test such that $X \in \bigcap_{i\in\omega} \mathcal{U}_i$. We define an X-computable function $f : \omega \to \omega$ as follows. For each $n \in \omega$, let

$$f(n) = \text{the least } s \text{ such that } (\exists k)[\![X{\upharpoonright}k]\!] \subseteq \mathcal{U}_{n,s}.$$

Clearly $f \leq_T X$. Now suppose there is some computable function g such that

$$(\exists^\infty n) f(n) < g(n).$$

Then setting $\mathcal{V}_i := \mathcal{U}_{i,g(i)}$, it follows that

(i) $\mathcal{V}_i$ is the set of extensions of a finite collection of finite strings,

(ii) $\mu(\mathcal{V}_i) \leq 2^{-i}$, and

(iii) $X \in \mathcal{V}_i$ for infinitely many i.

Then by Lemma 2.55, it follows that $X \in \mathsf{SR}_\mu$, contradicting our hypothesis. It thus follows that f dominates all computable functions.

□

Corollary 2.61. *For every $X \in 2^\omega$, if $X \in \mathsf{CR}_\mu \setminus \mathsf{MLR}_\mu$, then X has high Turing degree.*

2.5.2.3 Kurtz Randomness

Kurtz randomness is yet another definition of algorithmic randomness that has been studied. Unlike the other definitions we have encountered so far, Kurtz randomness includes a number of sequences that fail a number of basic statistical properties that hold of most sequences such as the Law of Large Numbers, which we will discuss shortly.

Definition 2.62. For $\mu \in \mathscr{M}_c$, $X \in 2^\omega$ is μ-Kurtz random, denoted $X \in \mathsf{KR}_\mu$, if for every Π^0_1 class $\mathcal{P}$ such that $\mu(\mathcal{P}) = 0$, $X \notin \mathcal{P}$.

As above, KR denotes the collection of λ-Kurtz random sequences. Like the other definitions of randomness we have considered, Kurtz randomness can be defined in terms of martingales.

Theorem 2.63. *For $\mu \in \mathscr{M}_c$, $X \in \mathsf{KR}_\mu$ if and only if there is no computable μ-martingale d and no computable order $g : \omega \to \omega$ such that*

$$d(X{\upharpoonright}n) \geq g(n)$$

for every $n \in \omega$

Proof. See, for instance, [Bie08], Theorem 1.5.12. □

The notion of martingale success in the definition of Kurtz randomness is stronger than that for Schnorr randomness, from which it follows that $\mathsf{SR}_\mu \subseteq \mathsf{KR}_\mu$. However, the converse does not hold in general. We show this indirectly.

Definition 2.64. $X \in 2^\omega$ is *weakly 1-generic*, denoted $X \in \mathsf{WG}$, if for every dense Σ^0_1 $\mathcal{S} \subseteq 2^\omega$, $X \in \mathcal{S}$.

Recall that $\mathcal{S} \subseteq 2^\omega$ is dense if $\mathcal{S}$ meets every non-empty open subset of 2^ω.

Proposition 2.65. *For every positive* $\mu \in \mathscr{M}_c$*,* $\mathsf{WG} \subseteq \mathsf{KR}_\mu$*.*

Proof. If $\mathcal{P}$ is a Π^0_1 class such that $\mu(\mathcal{P}) = 0$, then $\mathcal{U} = \mathcal{P}^c$ is a Σ^0_1 class such that $\mu(\mathcal{U}) = 1$. Since μ is positive, for any $\sigma \in 2^{<\omega}$ we have $\mu(\sigma) > 0$, and hence $\mathcal{U} \cap [\![\sigma]\!] \neq \emptyset$. This implies that $\mathcal{U}$ is dense in 2^ω. The result now immediately follows. $\square$

A sequence $X \in 2^\omega$ satisfies the Law of Large Numbers if

$$\lim_{n\to\infty} \frac{\{i < n : X(i) = 1\}}{n} = \frac{1}{2}.$$

That is, in the limit, the number of 0s and 1s in X are equal to $\frac{1}{2}$.

Proposition 2.66. $\mathsf{WG} \cap \mathsf{SR} = \emptyset$*.*

Proof Sketch. This follows from the fact that every weakly 1-generic sequence fails to satisfy the Law of Large Numbers (as the property of having sufficiently more 0s than 1s is dense) , while every Schnorr random sequence satisfies this law, as we can build a Schnorr test for this property. $\square$

2.5.3 Stronger Definitions of Randomness

We now turn to the definitions of randomness that are strictly stronger than Martin-Löf randomness, i.e. those definitions $\mathscr{D}$ with the property that every $\mathscr{D}$-random sequence is Martin-Löf random, but not vice versa.

2.5.3.1 Weak 2-Randomness

Definition 2.67. For $\mu \in \mathscr{M}_c$, $X \in 2^\omega$ is *μ-weakly 2-random*, denoted $X \in \mathsf{W2R}_\mu$, if for every Π^0_2 class $\mathcal{P}$ such that $\mu(\mathcal{P}) = 0$, $X \notin \mathcal{P}$.

It is immediate that $\mathsf{W2R}_\mu \subseteq \mathsf{MLR}_\mu$, since the collection of sequences captured by a μ-Martin-Löf defines a Π^0_2 set of μ-measure zero. However, the converse does not hold, as we now show.

Definition 2.68. $X, Y \in 2^\omega$ form a *minimal pair* in the Turing degrees if $A <_T X$ and $A <_T Y$ implies that $A \equiv_T \emptyset$.

Theorem 2.69 (Downey, Nies, Weber, Yu [DNWY06]; Hirschfeldt, Miller)**.** *For $\mu \in \mathscr{M}_c$, if X is not computable, then $X \in \mathsf{W2R}_\mu$ if and only if $X \in \mathsf{MLR}_\mu$ and X and $\emptyset'$ form a minimal pair.*

Note that this result and the fact that there are Δ^0_2 sequences in MLR immediately implies that $\mathsf{MLR} \neq \mathsf{W2R}$. Theorem 2.69 follows from the next two theorems.

Theorem 2.70 (Downey, Nies, Weber, and Yu [DNWY06])**.** *For $\mu \in \mathscr{M}_c$, if $X \in \mathsf{W2R}_\mu$ and X is not computable, then X and $\emptyset'$ form a minimal pair.*

Proof Sketch. We modify the proof given by Downey, Nies, Weber, and Yu for the case that $\mu = \lambda$. If $A \in 2^\omega$ is Δ^0_2, $Z \in \mathsf{W2R}_\mu$, and $\Phi^Z = A$ for some Turing functional A, then we argue that A is computable. Towards this end, we define

$$\mathcal{S} = \{X : \forall n \forall s \exists t > s (\Phi^X(n)[t]{\downarrow} = A_t(n)\},$$

which is Π^0_2 and contains Z. Since $Z \in \mathsf{W2R}_\mu$, it follows that $\mu(\mathcal{S}) > 0$. Now the only difference between the original proof and the situation here is that μ may be atomic, so that there is some μ-atom $Y \in \mathcal{S}$. But since μ is computable, it follows that Y is computable. Then $\Phi^Y = A$, and hence A is computable.

In the case that $\mathcal{S}$ contains no atoms, the proof proceeds exactly as in the case of the Lebesgue measure: by a "majority vote" argument, which shows that that one can compute values of A using the majority of sequences in a set of positive measure, one shows that A is computable. See, for instance, the proof of Theorem 7.2.8. in [DH10] for the details. $\square$

Theorem 2.71. *[Hirschfeldt, Miller] Let $\mu \in \mathscr{M}_c$. For any Σ^0_3 class $\mathcal{S} \subseteq 2^\omega$ such that $\mu(\mathcal{S}) = 0$, there is a noncomputable c.e. set A such that $A \leq_T X$ for every non-computable $X \in \mathsf{MLR}_\mu \cap \mathcal{S}$.*

Proof Sketch. The proof proceeds exactly in the case as the case of the Lebesgue measure, since $X \in \mathsf{MLR}_\mu \cap \mathcal{S}$ implies that $\mu(X) = 0$ and hence X is not a μ-atom. See the proof of Theorem 7.2.11 of [DH10]. $\square$

Proof of Theorem 2.69. ($\Rightarrow$) This is simply Theorem 2.70.
($\Leftarrow$) Suppose that $X \in \mathsf{MLR}_\mu \setminus \mathsf{W2R}_\mu$. Then there is a Π^0_2 μ-null set $\mathcal{S}$ such that $X \in \mathcal{S}$, and so by Theorem 2.71, there is some noncomputable c.e. set A such that $A \leq_T X$. Therefore, X and $\emptyset'$ do not form a minimal pair. $\square$

2.5.3.2 2-Randomness and n-randomness

A definition of randomness slightly stronger than weak 2-randomness is 2-randomness. In fact, for each n, we can define n-randomness.

Definition 2.72. For $\mu \in \mathscr{M}_c$ and $A \in 2^\omega$, a μ-Martin-Löf test relative to a A is a uniformly A-computable sequence $\{\mathcal{U}_i^A\}_{i\in\omega}$ of A-effectively open classes in 2^ω such that $\mu(\mathcal{U}_i^A) \leq 2^{-i}$ for every $i \in \omega$. Further, a real X is *μ-Martin-Löf random* relative to A, denoted MLR_μ^A if for every μ-Martin-Löf test $\{\mathcal{U}_i^A\}_{i\in\omega}$ relative to A, we have $X \notin \bigcap_{i\in\omega} \mathcal{U}_i^A$.

The following is immediate.

Lemma 2.73. *If $A \equiv_T B$, then $\mathsf{MLR}^A = \mathsf{MLR}^B$.*

Definition 2.74. For $\mu \in \mathscr{M}_c$, $X \in 2^\omega$ is *2-random with respect to μ*, denoted $X \in \mathsf{2MLR}_\mu$, if X is μ-Martin-Löf random relative to $\emptyset'$. Moreover, X is *n-random with respect to μ*, denoted $X \in \mathsf{nMLR}_\mu$, if X is μ-Martin-Löf random with respect to $\emptyset^{(n-1)}$.

In Subsubsection 2.5.1.3, we considered Theorem 2.33, Martin-Löf's result that

$$(\exists^\infty n)[C(X{\upharpoonright}n) \geq n - O(1)].$$

implies that $X \in \mathsf{MLR}$. A partial converse was only recently obtained by Nies, Stephan, and Terwijn, and independently by Miller.

Theorem 2.75 ([Mil04], [NST05])**.** *$X \in \mathsf{2MLR}$ if and only if*

$$(\exists^\infty n)[C(X{\upharpoonright}n) \geq n - O(1)].$$

Proof. See, for instance, [DH10], Theorem 6.11.6. □

2.5.4 An Open Case: Kolmogorov-Loveland Randomness

One important definition whose relationship to Martin-Löf randomness is still an open question is Kolmogorov-Loveland ranodmness. The basic idea behind this definition is that we allow a betting strategy to proceed non-monotonically, meaning that we need not bet along initial segments of sequences, but at different locations determined by the betting strategy, and not necessarily at strictly increasing locations. The following definitions are necessary to define Kolmogorov-Loveland randomness.

Definition 2.76. An *assignment* is a sequence (which may be finite or infinite)

$$x = (r_0, a_0), (r_1, a_1), \dots, (r_{n-1}, a_{n-1}), \dots$$

of pairs, where for each i, $r_i \in \omega$ and $a_i \in \{0, 1\}$.

We let FA denote the set of all finite assignments. Given an assignment $x = (r_0, a_0), (r_1, a_1), \dots, (r_{n-1}, a_{n-1}), \dots$, the domain of x, written $\mathrm{dom}(x)$, is the set $\{r_0, r_1, \dots\}$.

Definition 2.77. A *scan rule* is a partial function $S : FA \to \omega$ such that

$$(\forall x \in FA)(S(x) \notin \mathrm{dom}(x)).$$

Definition 2.78. Given a scan rule S and a real $X \in 2^\omega$, the *assignment given by* S, denoted σ_S^X, is defined as follows:

(i) Let $\sigma_S^X(0) = \emptyset$.

(ii) If $x_n = \sigma_S^X(n)$ and $S(x_n)$ are both defined, we let

$$\sigma_S^X(n+1) = x_n{}^\frown(S(x_n), X(S(x_n))).$$

(iii) If either $x_n = \sigma_S^X(n)$ or $S(x_n)$ are undefined, then $\sigma_S^X(n+1)$ is undefined as well.

Given a scan rule S and a real X, if

$$\sigma_S^X = (r_0, a_0), (r_1, a_1), \ldots, (r_{n-1}, a_{n-1}), \ldots$$

then we let $S(X)$ denote the string whose nth bit is the nth bit of X scanned by S, i.e. $S(X)(i) = a_i$. If σ_S^X is an infinite assignment, then $S(X) \in 2^\omega$. Otherwise, $S(X) \in 2^{<\omega}$. Further, if $\sigma_S^X(n)$ is defined, we let $S(X, n)$ be the string of length n such that $S(X, n)(i) = a_i$ for $i \in \{0, \ldots, n-1\}$. It follows that if σ_S^X is infinite, then $S(X, n) \prec S(X)$ for all n.

We next define the "stake" function. This is the function that determines how much of our capital we are to bet on our guess.

Definition 2.79. A *stake function* is a partial function $Q : FA \to [0, 2]$.

Having defined both the scan rules and the stake functions, we can now define non-monotonic betting strategies.

Definition 2.80. A *non-monotonic betting strategy* is a pair $B = (S, Q)$, where S is a scan rule and Q is a stake function, that bets on a real $X \in 2^\omega$ as follows: B

determines a payoff function $p_B^X : \omega \to [0,2]$ and a capital function $c_B^X : \omega \to \mathbb{R}^+ \cup \{0\}$, where

$$p_B^X(n+1) = \begin{cases} Q(\sigma_S^X(n)) & \text{if } S(X)(n) = 0 \\ 2 - Q(\sigma_S^X(n)) & \text{if } S(X)(n) = 1 \end{cases}$$

and

$$C_B^X(n) = C_B^X(0) \prod_{i=1}^{n} p_B^X(i),$$

where $C_B^X(0)$ is the initial capital of the betting strategy B.

Now, we define what it means for a non-monotonic betting strategy to "succeed" on a real X:

Definition 2.81. A non-monotonic betting strategy B succeeds on $X \in 2^\omega$ if

$$\limsup_{n\to\infty} C_B^X(n) = \infty.$$

In order for us to define a notion of randomness in terms of non-monotonic betting strategies, we need to consider the effectivized versions of these strategies. A partial computable non-monotonic betting $B = (S, Q)$ strategy is one in which the scan rule S and the stake function Q are partial computable functions, and the range of Q is $\mathbb{Q} \cup [0, 2]$. We now give the associated definition of randomness.

Definition 2.82. A real X is *Kolmogorov-Loveland random*, denoted $X \in \mathsf{KLR}$, if there is no partial computable non-monotonic betting strategy that succeeds on X.

Muchnik, Semonov, and Uspensky were able to show that Martin-Löf randomness is sufficient for Kolmogorov-Loveland randomness.

Theorem 2.83 ([MSU98])**.** $\mathsf{MLR} \subseteq \mathsf{KLR}$

Whether the converse holds is one of the most important open questions in the theory of algorithmic randomness.

Question 2.84. Is there some $X \in \mathsf{KLR} \setminus \mathsf{MLR}$?

For our purposes, the following result will be useful.

Proposition 2.85 ([Mer03])**.** *$X \in \mathsf{KLR}$ if and only if no* total *computable non-monotonic betting strategy succeeds on X.*

2.6 Several Useful Theorems

We conclude this chapter with several theorems concerning Martin-Löf randomness with respect to the Lebesgue measure. First, van Lambalgen's Theorem gives us conditions under which the join $X \oplus Y$ of two Martin-Löf random sequences is also Martin-Löf random.

Theorem 2.86 (van Lambalgen's Theorem)**.** *For $X, Y \in 2^\omega$,*

$$X \oplus Y \in \mathsf{MLR} \Leftrightarrow (X \in \mathsf{MLR}^Y \wedge Y \in \mathsf{MLR}).$$

Proof. See [Nie09], Theorem 3.4.6. □

Corollary 2.87. *$X \in \mathsf{2MLR}$ if and only if $X \in \mathsf{MLR}$ and $\Omega \in \mathsf{MLR}^X$.*

Proof.

$$\begin{aligned}
X \in \mathsf{2MLR} &\Leftrightarrow X \in \mathsf{MLR}^{\emptyset'} && \text{(by definition of } \mathsf{2MLR}\text{)}\\
&\Leftrightarrow X \in \mathsf{MLR}^{\Omega} && \text{(since } \emptyset' \equiv_T \Omega\text{)}\\
&\Leftrightarrow X \oplus \Omega \in \mathsf{MLR} && \text{(by van Lamgalgen's Theorem)}\\
&\Leftrightarrow \Omega \in \mathsf{MLR}^{X} \ \wedge \ X \in \mathsf{MLR}. && \text{(by van Lamgalgen's Theorem)}
\end{aligned}$$

□

If we define $\mathsf{Low}(\Omega)$ to be the set $\{X \in 2^\omega : \Omega \in \mathsf{MLR}^X\}$, then Corollary 2.87 yields

$$\mathsf{Low}(\Omega) \cap \mathsf{MLR} = \mathsf{2MLR}.$$

Lastly, we have the surprising result that every $X \in 2^\omega$ is Turing reducible to a Martin-Löf random sequence. In fact, we can computably bound the use of this reduction.

Theorem 2.88 (The Kučera-Gács Theorem, [Kuč85], [Gác86])**.** *For every $X \in 2^\omega$, there is some $A \in \mathsf{MLR}$ such that $X \leq_{wtt} A$.*

Proof. See [Nie09], Section 3.3. □

CHAPTER 3

THE FUNCTIONAL EXISTENCE THEOREM

3.1 Introduction

3.1.1 Motivation

As motivation for the Functional Existence Theorem, let us first consider the distinction between semimeasures on $2^{<\omega}$ and 2^{ω}. As the name suggests, a semimeasure can be seen as a defective measure. Whereas a probability measure $\pi : 2^{<\omega} \to [0,1]$ on ω must satisfy

$$\sum_{n\in\omega} \pi(n) = 1,$$

a *discrete semimeasure* on $2^{<\omega}$ is simply a function $m : 2^{<\omega} \to [0,1]$ on $2^{<\omega}$ satisfying

$$\sum_{n\in\omega} m(n) \leq 1.$$

Similarly, whereas a probability measure $\mu : 2^{\omega} \to [0,1]$ on 2^{ω} must satisfy

$$\mu(\varnothing) = 1 \text{ and } \mu(\sigma) = \mu(\sigma 0) + \mu(\sigma 1)$$

for every $\sigma \in 2^{<\omega}$, a *continuous semimeasure* on 2^ω is a function $\delta : 2^\omega \to [0,1]$ satisfying

$$\delta(\varnothing) \leq 1 \text{ and } \delta(\sigma) \geq \delta(\sigma 0) + \delta(\sigma 1)$$

for every $\sigma \in 2^{<\omega}$.[1]

Here we are particularly interested in computably enumerable discrete and continuous semimeasures. Paradigm examples of both discrete and continuous semimeasures arise by feeding a Turing machine or Turing functional with a randomly generated input.

Example 3.1. Given a prefix-free Turing machine $T : 2^{<\omega} \to 2^{<\omega}$, if we generate a binary string by repeated tosses of a fair coin and feed this string into our machine T until it accepts the string (unless it never accepts the string), then the function $m_T : 2^{<\omega} \to [0,1]$ defined by

$$m_T(\sigma) = \sum_{T(\tau)\downarrow=\sigma} 2^{-|\tau|}$$

is the probability that the string generated by the tosses of our coin will be accepted by T and mapped to σ. Since

$$\sum_{\sigma\in 2^{<\omega}} m_T(\sigma) = \sum_{\sigma\in 2^{<\omega}} \sum_{T(\tau)\downarrow=\sigma} 2^{-|\tau|} = \sum_{\tau\in\mathsf{dom}(T)} 2^{|-\tau|} \leq 1,$$

where inequality holds because T is prefix-free, it follows that m_T is a discrete

[1] For sake of brevity, we write $\delta(\sigma)$ instead of $\delta([\![\sigma]\!])$.

semimeasure.[2]

Example 3.2. Given a Turing functional $\Phi : 2^\omega \to 2^\omega$, we can generate a binary sequence by repeated tosses of a fair coin and feed this string into our functional, outputting longer and longer initial segments of a output sequence as we feed longer and longer initial segments of a input sequence. If we set

$$\mathsf{Pre}_\Phi(\sigma) := \{\tau \in 2^{<\omega} : \Phi^\tau \succeq \sigma \ \wedge \ (\forall \tau' \prec \tau)\Phi^{\tau'} \not\succeq \sigma\},$$

then the function $\delta_\Phi : 2^{<\omega} \to [0,1]$ defined by

$$\delta_\Phi(\sigma) = \sum_{\tau \in \mathsf{Pre}_\Phi(\sigma)} 2^{-|\tau|} = \lambda([\![\mathsf{Pre}_\Phi(\sigma)]\!])$$

is the probability that an initial segment of the sequence generated by the tosses of our coin will be accepted by Φ and mapped to an extension of σ. Now since $\mathsf{Pre}_\Phi(\sigma)$ is a prefix-free set for each $\sigma \in 2^{<\omega}$, it follows that

$$\delta_\Phi(\varnothing) = \lambda([\![\mathsf{Pre}_\Phi(\sigma)]\!]) \leq 1.$$

Moreover, since

$$[\![\mathsf{Pre}_\Phi(\sigma)]\!] \supseteq [\![\mathsf{Pre}_\Phi(\sigma 0)]\!] \cup [\![\mathsf{Pre}_\Phi(\sigma 1)]\!],$$ [3]

[2]Moreover, if we define the semimeasure m_U in terms of a *universal* prefix-free Turing machine, then

$$\sum_{\sigma \in 2^{<\omega}} m_U(\sigma) = \Omega_U,$$

Chaitin's Ω. This is why Ω is sometimes referred to as a *halting probability.*

[3]We need not have equality here, since Φ might map some τ to σ while being undefined on all

it follows that

$$\delta_\Phi(\sigma) \geq \delta_\Phi(\sigma 0) + \delta_\Phi(\sigma 1)$$

for every $\sigma \in 2^{<\omega}$. Thus

3.1.2 The Machine Existence Theorem and Discrete Semimeasures

As discussed in the previous chapter, The Machine Existence Theorem provides a recipe for building a prefix-free Turing machine M whenever we are given a bounded c.e. list of requests (n, τ); that is, the collection of requests $W \subseteq \omega \times 2^{<\omega}$ is such that

$$\sum_{(n,\tau) \in W} 2^{-n} \leq 1,$$

where each (n, τ) is a request that the machine we are building map some string of length n to the string τ. More precisely, the theorem states that given a bounded request set $W \subseteq \omega \times 2^{<\omega}$, there is a prefix-free Turing machine M such that for each $(n, \tau) \in W$, there is some $\sigma \in \mathrm{dom}(M)$ such that $|\sigma| = n$ and $M(\sigma) = \tau$. This in turn implies that

$$K_M(\tau) \leq |\sigma|$$

and hence that

$$K(\tau) \leq |\sigma| + O(1).$$

For our purposes, it's important to note that a bounded request set is essentially a discrete semimeasure. Given a bounded request set W, we can define a discrete

extensions of τ; more formally, it may be that $(\tau, \sigma) \in \Phi$ but for all $\tau' \succ \tau$, there is no $\sigma' \in 2^{<\omega}$ such that $(\tau', \sigma') \in \Phi$.

semimeasure m_W by

$$m_W(\tau) := \sum_{(n,\tau)\in W} 2^{-n}.$$

Seen in this way, the Machine Existence Theorem tells us that *every* discrete semimeasure can be seen as providing the probability that a machine with randomly generated input will accept the input and output a given string. That is, *every* c.e. discrete semimeasure is an instance of the paradigm example of a c.e. discrete semimeasure given in Example 1 above.

3.1.3 An Analogue for Continuous Semimeasures?

This fact about c.e. discrete semimeasures suggests that every c.e. *continuous* semimeasure may be an instance of the paradigm example of a c.e. continuous semimeasure given in Example 2 above. The main result in this chapter says that this is true. Toward this end, we prove an analogue of the Machine Existence Theorem for Turing functionals $\Phi : 2^\omega \to 2^\omega$, which we refer to appropriately enough as the Functional Existence Theorem. We should note that this result is implicit in early work of Levin and Zvonkin [ZL70], and more recently, Day proved a similar theorem, from which he derived many of the results in Section 3.3. See [Day10] for more details.

3.2 The Construction

3.2.1 A General Overview

Unlike the Machine Existence Theorem, our requests are pairs (q, τ), where $q \in \mathbb{Q}_2$ and $\tau \in 2^{<\omega}$. To fulfill a request (q, τ), we define the functional Φ so that it maps to τ a finite collection of strings $\{\sigma_1, \sigma_2, \ldots, \sigma_k\}$ such that $\lambda(\bigcup_{i \leq k} [\![\sigma_i]\!]) = q$. Moreover, unlike the Machine Existence Theorem, we don't require that the total weight of our requests be bounded by 1, but rather that for each n, the total weight of our requests for strings of length n be bounded by 1. That is, if $V \subseteq \mathbb{Q}_2 \times 2^{<\omega}$ is the set of requests, then for each n

$$\sum_{\tau \in 2^n} \sum_{(q,\tau) \in V} q \leq 1. \tag{3.1}$$

In the following, the set V will be called a *bounded functional request set.*

Next, since we are enumerating a functional Φ, we need to ensure that it is consistent; that is, if $\Phi^\sigma \succeq \tau$ and $\sigma' \succeq \sigma$, then $\Phi^{\sigma'} \succeq \tau$. Moreover, we must require that the weight of requests for a fixed string σ not be exceeded by the sum of the weights of its extensions $\sigma 0$ and $\sigma 1$. To put it formally, if the V-weight of σ is defined to be

$$\mathsf{wt}_V(\sigma) := \sum_{(q,\sigma) \in V} q,$$

then for every $\sigma \in 2^{<\omega}$, we must satisfy the requirement

$$\mathsf{wt}_V(\sigma) \geq \mathsf{wt}_V(\sigma 0) + \mathsf{wt}_V(\sigma 1). \tag{3.2}$$

It immediately follows from Equations 1 and 2 above that wt_v is a continuous semimeasure.

We now prove that if the above listed conditions hold, then there is a Turing functional Φ such that for every $\sigma \in \mathsf{dom}(\mathsf{wt}_V)$, the amount of measure mapped into σ by Φ, $\lambda(\llbracket \mathsf{Pre}_\Phi(\sigma) \rrbracket)$, is precisely the amount requested, $\mathsf{wt}_V(\sigma)$.

3.2.2 The Formal Details

The full statement of the Functional Existence Theorem is as follows.

Theorem 3.3. *Suppose $V \subseteq \mathbb{Q}_2 \times 2^{<\omega}$ is a bounded functional request set, i.e., V is a c.e. set of pairs (q, τ) such that for every n,*

$$\sum_{\tau \in 2^n} \sum_{(q,\tau) \in V} q \leq 1$$

and such that the V-weight function wt_V satisfies

$$\mathsf{wt}_V(\sigma) \geq \mathsf{wt}_V(\sigma 0) + \mathsf{wt}_V(\sigma 1)$$

for every $\sigma \in 2^{<\omega}$. Then there exists a Turing functional $\Phi : 2^\omega \to 2^\omega$ such that for each $\sigma \in \mathsf{dom}(\mathsf{wt}_V)$, $\lambda(\llbracket \mathsf{Pre}_\Phi(\sigma) \rrbracket)) = \mathsf{wt}_V(\sigma)$.

The following terminology will be useful in the proof of Theorem 3.3. For $\sigma \in 2^{<\omega}$, σ^- denotes the initial segment of σ of length $|\sigma| - 1$; we will sometimes refer to σ^- as the *parent* of σ. Moreover, for $\tau \in 2^{<\omega}$, $\tau 0$ and $\tau 1$ are called the *children* of τ, and $\tau 0$ and $\tau 1$ are called *siblings*.

Proof. Recall from the previous chapter we define a functional Φ to be a c.e. set of pairs of strings (σ, τ), where $(\sigma, \tau) \in \Phi$ means that $\Phi^\sigma \succeq \tau$. For $s \in \omega$, Φ_s will denote the collection of pairs (σ, τ) that have been enumerated into Φ by the end of stage s. When a pair (q, τ) is enumerated into V at stage s, we will say that *measure q is requested for τ at stage s.* Further, we *fulfill a request (q, τ) at stage s* by enumerating a finite collection of pairs $(\sigma_1, \tau), \ldots, (\sigma_k, \tau)$ into Φ_s such that $\sum_{i \leq k} \lambda(\sigma_i) = q$ and each σ_i properly extends some string ξ such that $(\xi, \tau^-) \in \Phi_{s-1}$, where τ^- is the string obtained by removing the final bit of τ.

For a fixed string $\tau \in 2^{<\omega}$ such that $\tau = \rho^\frown i$ for some $\rho \in 2^{<\omega}$ and $i \in \{0, 1\}$, the *measure mapped to τ by (the end of) stage s* is $\sum_{(\sigma,\tau) \in \Phi_s} 2^{-|\sigma|}$, while the *measure available for τ at (the beginning of) stage s* is

$$\sum_{(\sigma,\rho) \in \Phi_{s-1}} 2^{-|\sigma|} - \sum_{(\sigma',\rho^\frown(1-i)) \in \Phi_{s-1}} 2^{-|\sigma'|}.$$

That is, the measure mapped to τ by stage s is the sum of the measure of the strings σ such that (σ, τ) has been enumerated by the end of stage s, and the measure available for τ at stage s is the measure mapped to the parent of τ by the end of stage $s - 1$ minus the measure mapped to the sibling of τ by the end of stage $s - 1$.

Now, the general approach here is straightforward: for each pair (q, τ) enumerated into V, we first wait until each of the initial segments of τ are in the range of Φ; this is guaranteed to occur since by the hypothesis of our theorem, for each $\xi \prec \tau$, $\mathsf{wt}_V(\xi) \geq \mathsf{wt}_V(\tau)$. Once the condition

$$(\forall \tau' \preceq \tau)(\exists \sigma')(\sigma', \tau') \in \Phi$$

has been satisfied, say by stage t, we would like to fulfill the request (q, τ) at stage $t+1$. However, there are two problems that might prevent us from immediately mapping the requested measure's worth of strings to τ:

(P_1) The measure mapped to τ^- by stage t may be less than q.

(P_2) It may occur that enough measure has been mapped to τ^- by stage t, but the sibling of τ has taken some of it, so that the measure available for τ at stage $t+1$ is less than q.

Since the weight of τ^- exceeds the sum of the weight of τ and the weight of the sibling of τ, in principle these two problems are only temporary; however, if we naively wait until more measure is mapped to the parent τ^- before we try to fulfill the request (q, τ), it may be that in the meantime another request (q', τ) is enumerated into V, so that we will now need to map measure totalling $q + q'$ to τ.[4]

The solution is not to wait to act upon a request until enough measure is available, but to act upon whatever measure is *currently* available, perhaps only partially fulfilling a request (that will be completely fulfilled eventually).

Let V_s denote the set of pairs (q, τ) enumerated into V by the end of stage s; for convenience, we will assume that at most one pair (q, τ) is enumerated into V_s at each stage s. Pairs that are enumerated into V but which cannot be acted upon immediately are temporarily placed in a queue (one for each stage s), which will be denoted by Q_s.

[4]In the worst case, it may be that (i) the total amount of measure requested for some τ and one of its extensions, say $\tau 0$, is q (and thus no measure is requested for $\tau 1$), (ii) that one request $(q, \tau 0)$ is made, and (iii) this requests is followed by requests of the form $(2^{-i}q, \tau)$ for every $i \in \omega$. Thus, if we only wait to act on the request for measure to be mapped to $\tau 0$ only after the total measure requested to be mapped to τ is q, we will wait forever.

Let us fix a bit more terminology. We say that an enumerated request (q, τ) can be *completely fulfilled* at stage s if the amount of measure available to τ at stage s is at least q. In addition, we say that an enumerated request (q, τ) can be *partially fulfilled* at stage s if there is some measure less than q available to τ at stage s. Let us now turn to the construction.

Construction

At stage $s = 0$, we set $\Phi_0 = V_0 = Q_0 = \emptyset$.

At stage $s+1$, if no pair (q, τ) enters V, then we move on to stage $s+2$. Otherwise, we will iterate the following procedure, during which we will make a finite number of changes to the set Q_s, resulting in a sequence of sets $Q_s = Q^0_{s+1}, \dots, Q^N_{s+1}$ for some $N \in \omega$. Then we will define $Q_{s+1} := Q^N_{s+1}$.

Step 1: First, given (q, τ) that has entered V_s, there are three possibilities:

(1a) no measure is available for τ;

(1b) the measure available for τ is $r < q$; or

(1c) the measure available for τ is $r \geq q$.

If (1a) occurs, there are two further possibilities:

(1a.i) there is no pair $(r, \tau) \in Q^0_{s+1} (= Q_s)$ for any $r \in \mathbb{Q}_2$; or

(1a.ii) there is some pair $(r, \tau) \in Q^0_{s+1}$ for some $r \in \mathbb{Q}_2$.

In subcase (1a.i), we add (q, τ) to the queue by defining $Q_{s+1} = Q^0_{s+1} \cup \{(q, \tau)\}$. Then we skip Step 2 and move on the next stage. In subcase (1a.ii) we remove

(q,τ) from the queue and replace it with $(q+r,\tau)$; that is, we define $Q_{s+1} := (Q^0_{s+1} \setminus \{(q,\tau)\}) \cup \{(q+r,\tau)\}$. Having done this, we skip Step 2 and move on to the next stage.

If (1b) occurs, then we can only partially fulfill the request (q,τ).[5] Since the amount of measure available for τ is r, this means that there are strings $\sigma_1,\dots,\sigma_k$ such that for each $i \leq k$, there is a string $\xi \prec \sigma_i$ such that (ξ,τ^-) has been enumerated into Φ_t for some $t \leq s$ and $\sum_{i=1}^k 2^{-|\sigma_i|} = r$. We thus enumerate the pairs $(\sigma_1,\tau),\dots,(\sigma_k,\tau)$ into Φ_s, and update the queue by defining $Q^1_{s+1} := (Q^0_{s+1} \setminus \{(q,\tau)\}) \cup \{(q-r,\tau)\}$, and then we move on to the second step of the construction.

If (1c) occurs, we carry out the same steps as we did for (1b), except that now the request (q,τ) can be completely fulfilled, and thus we define $Q^1_{s+1} := Q^0_{s+1}$ and move on to the second step of this stage of the construction.

Step 2: Given Q^i_{s+1}, we proceed as follows: We first check to see if there is any request $(q',\tau') \in Q^i_{s+1}$ that can be partially or completely fulfilled. There are two possibilities:

(2a) No pair $(q',\tau') \in Q^i_{s+1}$ can be fulfilled partially or completely; or

(2b) There is at least one pair $(q',\tau') \in Q^i_{s+1}$ that can be fulfilled partially or completely.

[5]As we will see, in this case, there can be no $(p,\tau) \in Q_{s-1}$, since if there were, we would have already used the measure r available to τ to fulfill, at least partially, the request for measure p to be mapped to τ.

If (2a) occurs, there is no action to take, so we set $Q_{s+1} := Q^i_{s+1}$ and move on to the next stage of the construction.

If (2b) occurs, then there are two further possibilities:

(2b.i) there is exactly one pair $(q', \tau') \in Q^i_{s+1}$ that can be partially or completely fulfilled; or

(2b.ii) there are two pairs $(q', \tau'), (q'', \tau'') \in Q^i_{s+1}$ that can be partially or completely fulfilled after the action we took in Step 1.

In subcase (2b.i) it must be that τ' is an immediate successor of a string we dealt with in Step 1 or in the previous iteration of Step 2. For the pair $(q', \tau^\frown i)$, we carry out the actions from Step 1 (where this time only cases (1b) or (1c) are possible), except at the end of the action, we define Q^{i+1}_{s+1} just as we defined Q^1_{s+1}.

In subcase (2b.ii), it must be that τ' and τ'' are the immediate successors of a string we dealt with in Step 1 or in the previous iteration of Step 2 (and thus τ' and τ'' are siblings). In this case, if we have a choice between partially satisfying both requests or completely satisfying one of them, we always choose to partially satisfy both requests. To do so, we split up the available measure in some effective way and carry out the actions as in Step 1 above, and define Q^{i+1}_{s+1} as we defined Q^1_{s+1}.

As there are at most finitely many pairs in Q_s to begin with, there is some N such that after $N-1$ iterations of the procedure, we eventually arrive at case (2a) and define $Q_{s+1} := Q^N_{s+1}$, thus ending this stage of the construction.

Verification

First, note that Φ is a consistent functional: if (σ, τ) is enumerated into Φ, by our construction, σ must properly extend some string ξ such that (ξ, τ^-) has previously been enumerated into Φ. Next we verify that the every request for measure is eventually fulfilled. To this end, let us define an auxiliary function $\mathsf{wt}_{Q_s}(\tau)$, the Q_s-weight of a string $\tau \in 2^{<\omega}$, to be the unique $q \in \mathbb{Q}_2$ such that $(q, \tau) \in Q_s$, unless there is no such q, in which case we set $\mathsf{wt}_{Q_s}(\tau) = 0$. Thus it suffices to prove the following.

Lemma 3.4. *For all strings $\tau \in 2^{<\omega}$, $\lim_{s\to\infty} \mathsf{wt}_{Q_s}(\tau) = 0$.*

Proof. We prove this by induction on string length. First, all requests of the form $(q, \varnothing)$ can be completely fulfilled as soon as they are enumerated into V, and hence $\mathsf{wt}_{Q_s}(\varnothing) = 0$ for all s. Now suppose the lemma holds for all strings of length k. Given a string $\tau \in 2^{<\omega}$ of length $k+1$, suppose first that $\lim_{s\to\infty} \mathsf{wt}_{Q_s}(\tau)$ does not exist. Then there is some n such that

$$\limsup_s \mathsf{wt}_{Q_s}(\tau) - \liminf_s \mathsf{wt}_{Q_s}(\tau) > \frac{1}{n}.$$

This implies that there are infinitely many stages s such that for some $t > s$, $\mathsf{wt}_{Q_t}(\tau) - \mathsf{wt}_{Q_s} \geq \frac{1}{n}$. This further implies that $\mathsf{wt}_V(\tau) = \infty$, which is impossible. So, $\lim_{s\to\infty} \mathsf{wt}_{Q_s}(\tau)$ exists.

Now, suppose there is some $\epsilon > 0$ such that $\lim_{s\to\infty} \mathsf{wt}_{Q_s}(\tau) > \epsilon$. It follows that there is some stage t such that for every stage $s \geq t$, $\mathsf{wt}_{Q_s}(\tau) > \epsilon$. By induction, $\lim_{s\to\infty} \mathsf{wt}_{Q_s}(\tau^-) = 0$, and by assumption, $\mathsf{wt}_V(\tau^-) \geq \mathsf{wt}_V(\tau) + \mathsf{wt}_V(\tau^*)$, where τ^* is the sibling of τ. Moreover, by our construction, we always choose to partially fulfill requests of siblings rather than completely fulfill one or the other. This means that there will be some stage $t' \geq t$ at which enough measure will be available to τ and

τ^* so that we will have $\mathsf{wt}_{Q_{t'+1}}(\tau) < \epsilon$, contradicting our earlier assumption. $\square$

Since every request is eventually fulfilled, it is clear from the construction that for every $\tau \in 2^{<\omega}$, $\lambda([\![\mathsf{Pre}_\Phi(\tau)]\!]) = \mathsf{wt}_V(\tau)$.

$\square$

3.3 Applications of the Functional Existence Theorem

In this section, we will use the Functional Existence Theorem to provide a number of characterizations of several notions of randomness. The results in this section were obtained in collaboration with Laurent Bienvenu.

3.3.1 Measure-Boundedness

First, we characterize different notions of randomness in terms of the rate at which Turing functionals map measure to the initial segments of sequences, results which can be straightforwardly recast in terms of a priori complexity.

Definition 3.5. Given $A \in 2^\omega$, a Turing functional $\Phi : 2^\omega \to 2^\omega$ is *measure-bounded along* A if there is some $c \in \omega$ such that for every $n \in \omega$,

$$\lambda([\![\mathsf{Pre}_\Phi(A{\upharpoonright}n)]\!]) \leq c \cdot \lambda(A{\upharpoonright}n).$$

Theorem 3.6. *$A \in 2^\omega$ is Martin-Löf random if and only if every Turing functional Φ is measure-bounded along A.*

To prove this theorem, we need two results. First we show that for any Turing functional Φ, the function

$$L_\Phi(\sigma) := \frac{\lambda([\![\mathsf{Pre}_\Phi(\sigma)]\!])}{\lambda(\sigma)}$$

is a c.e. supermartingale.

Proposition 3.7. *For every Turing functional Φ, $L_\Phi : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ is a c.e. supermartingale.*

Proof. Clearly, every real in the range of L_Φ is a c.e. real, so we just need to show that for every $\sigma \in 2^{<\omega}$

$$2L_\Phi(\sigma) \geq L_\Phi(\sigma 0) + L_\Phi(\sigma 1).$$

As we saw in our discussion of Example 3.2,

$$[\![\mathsf{Pre}_\Phi(\sigma)]\!] \supseteq [\![\mathsf{Pre}_\Phi(\sigma 0)]\!] \cup [\![\mathsf{Pre}_\Phi(\sigma 1)]\!],$$

It follows that

$$\lambda([\![\mathsf{Pre}_\Phi(\sigma)]\!]) \geq \lambda([\![\mathsf{Pre}_\Phi(\sigma 0)]\!]) \cup \lambda([\![\mathsf{Pre}_\Phi(\sigma 1)]\!]). \tag{3.3}$$

Since $\frac{1}{2}\lambda(\sigma) = \lambda(\sigma 0) = \lambda(\sigma 1)$, it follows that $\dfrac{2}{\lambda(\sigma)} = \dfrac{1}{\lambda(\sigma 0)} = \dfrac{1}{\lambda(\sigma 1)}$. When combined with Equation (3), this yields

$$\frac{2\lambda([\![\mathsf{Pre}_\Phi(\sigma)]\!])}{\lambda(\sigma)} \geq \frac{\lambda([\![\mathsf{Pre}_\Phi(\sigma 0)]\!])}{\lambda(\sigma 0)} + \frac{\lambda([\![\mathsf{Pre}_\Phi(\sigma 1)]\!])}{\lambda(\sigma 1)}.$$

□

Next, we apply the Functional Existence Theorem to derive the following result.

Proposition 3.8. *Given a c.e. supermartingale d such that $d(\varnothing) \leq 1$, there is a Turing functional Φ such that $d = L_\Phi$*

Proof. Since d is a c.e. supermartingale, d has a supermartingale approximation. That is, there is a collection of c.e. supermartingales $\{d_i\}_{i\in\omega}$ such that each d_i is $\mathbb{Q}_2$-valued, for each $\sigma \in 2^{<\omega}$ and each s,

$$d_s(\sigma) \leq d_{s+1}(\sigma),$$

and

$$\lim_{s\to\infty} d_s(\sigma) = d(\sigma).$$

We will use the supermartingale approximation of d to enumerate a bounded functional request set as follows. First, we will assume that for all $\sigma \in 2^{<\omega}$, $d_0(\sigma) = 0$. Next, whenever we see $d_s(\sigma) \neq d_{s+1}(\sigma)$, we will enumerate the pair

$$\big(2^{-|\sigma|}(d_{s+1}(\sigma) - d_s(\sigma)), \sigma\big)$$

into our bounded functional request set V. Clearly $V \subseteq \mathbb{Q}_2 \times 2^{<\omega}$, since each d_s is $\mathbb{Q}_2$-valued. Further, we have

$$\mathsf{wt}_V(\varnothing) = \sum_{(q,\varnothing)\in V} q = \sum_{s\in\omega}(d_{s+1}(\varnothing) - d_s(\varnothing)) = d(\varnothing) \leq 1,$$

where the second equality holds because

$$\sum_{s\in\omega}\big(d_{s+1}(\sigma)-d_s(\sigma)\big)=d(\sigma)$$

for every $s\in 2^{<\omega}$. Lastly, we have

$$\mathsf{wt}_V(\sigma)=\sum_{(q,\sigma)\in V} q=\sum_{s\in\omega}2^{-|\sigma|}(d_{s+1}(\sigma)-d_s(\sigma))=2^{-|\sigma|}d(\sigma),$$

and hence

$$\mathsf{wt}_V(\sigma)=2^{-|\sigma|}d(\sigma)\geq 2^{-|\sigma 0|}d(\sigma 0)+2^{-|\sigma 1|}d(\sigma 1)=\mathsf{wt}_V(\sigma 0)+\mathsf{wt}_V(\sigma 1).$$

Thus, all the conditions for the Functional Existence Theorem are satisfied, and so there is a Turing functional Φ such that

$$\lambda(\llbracket\mathsf{Pre}_\Phi(\sigma)\rrbracket)=\mathsf{wt}_V(\sigma)=2^{-|\sigma|}d(\sigma).$$

□

With these two results, we are now ready to prove Theorem 4.

Proof of Theorem 4. First, suppose that A is Martin-Löf random. For a Turing functional Φ, by Proposition 3.7, L_Φ is a c.e. supermartingale, and thus there is some constant $c\in\omega$ such that for all n, $L_\Phi(A\upharpoonright n)\leq c$. Thus, for all n

$$\frac{\lambda(\llbracket\mathsf{Pre}_\Phi(A\upharpoonright n)\rrbracket)}{\lambda(A\upharpoonright n)}\leq c,$$

and hence for all n

$$\lambda(\llbracket \mathsf{Pre}_\Phi(A\upharpoonright n)\rrbracket) \leq c\lambda(A\upharpoonright n).$$

For the other direction, if A is not Martin-Löf random, then there is some c.e. supermartingale d that succeeds on A, i.e. for every $c \in \omega$ there is some n such that $d(A\upharpoonright n) \geq c$. Then by Proposition 3.8, there is some Turing functional Φ such that $d = L_\Phi$, and hence for every c there is some n such that $L_\Phi(A\upharpoonright n) > c$, or equivalently, for every $c \in \omega$ there is some n such that

$$\lambda(\llbracket \mathsf{Pre}_\Phi(A\upharpoonright n)\rrbracket) > c\lambda(A\upharpoonright n).$$

Hence, Φ is not measure-bounded along A.

□

Next, we provide similar characterizations for computable randomness, Schnorr randomness, and Kurtz randomness in terms of almost total Turing functionals (where a Turing functional Φ is almost total if $\lambda(\mathsf{dom}(\Phi)) = 1$). We need the following lemma for our characterizations.

Lemma 3.9. *For any almost total Turing functional Φ, L_Φ is a computable martingale.*

Proof. Since $\lambda(\mathsf{dom}(\Phi)) = 1$ implies that $\lambda(\llbracket \mathsf{Pre}_\Phi(\varnothing)\rrbracket) = 1$, L_Φ has computable initial capital. Now suppose that

$$2L_\Phi(\sigma) > L_\Phi(\sigma 0) + L_\Phi(\sigma 1)$$

for some $\sigma \in 2^{<\omega}$. This implies that

$$\lambda(\llbracket \mathsf{Pre}_\Phi(\sigma) \rrbracket) > \lambda(\llbracket \mathsf{Pre}_\Phi(\sigma 0) \rrbracket) + \lambda(\llbracket \mathsf{Pre}_\Phi(\sigma 1) \rrbracket)$$

which further implies that Φ is undefined on all the sequences in

$$\llbracket \mathsf{Pre}_\Phi(\sigma) \rrbracket \setminus \big(\llbracket \mathsf{Pre}_\Phi(\sigma 0) \rrbracket \cup \llbracket \mathsf{Pre}_\Phi(\sigma 1) \rrbracket\big).$$

This is a set of positive measure, contradicting the fact that Φ is almost total. Thus, it must be the case that

$$2L_\Phi(\sigma) = L_\Phi(\sigma 0) + L_\Phi(\sigma 1).$$

□

First, we characterize computable randomness in terms of almost total functionals.

Theorem 3.10. *$A \in 2^\omega$ is computably random if and only if every almost total Turing functional Φ is measure-bounded along A.*

Proof. First, given a computably random real A and an almost total functional Φ, since L_Φ is a computable martingale by Lemma 3.9, it follows that L_Φ cannot succeed on A. Thus there is some c such that for every n,

$$\lambda(\llbracket \mathsf{Pre}_\Phi(A{\upharpoonright}n) \rrbracket) \leq c\lambda(A{\upharpoonright}n).$$

For the other direction, suppose that A is not computably random. Then there is some computable martingale d that succeeds on A. Without loss of generality, we can assume that $d(\varnothing) = 1$. By Proposition 3.8, there is a Turing functional Φ such that $d = L_\Phi$. All we have to verify is that the Φ is almost total. To see this, note that since $d(\varnothing) = 1$ and for each n, the measure mapped to strings of length n by Φ is

$$\sum_{\sigma \in 2^n} 2^{-n} d(\sigma) = d(\varnothing) = 1,$$

where the first equality follows from the martingale condition (along the lines of Lemma 2.23). This implies that $\lambda(\Phi^{-1}(\mathsf{ran}(\Phi))) = 1$, and thus $\lambda(\mathsf{dom}(\Phi)) = 1$. $\square$

Next, we characterize Schnorr randomness and Kurtz randomness in terms of almost total functionals. We first need to generalize what it means for a functional Φ to be measure-bounded along a real $A \in 2^\omega$.

Definition 3.11. Suppose $A \in 2^\omega$, let Φ be a Turing functional, and let h be a computable order h.

(i) Φ is *effectively measure-bounded along A via h* if there is some $c \in \omega$ such that

$$\lambda([\![\mathsf{Pre}_\Phi(A{\upharpoonright}n)]\!]) \leq c \cdot h(n)\lambda(A{\upharpoonright}n)$$

for every $n \in \omega$.

(ii) Φ is *weakly effectively measure-bounded along A via h* if there is some $c \in \omega$ such that

$$\lambda([\![\mathsf{Pre}_\Phi(A{\upharpoonright}n)]\!]) \leq c \cdot h(n)\lambda(A{\upharpoonright}n)$$

for infinitely many $n \in \omega$.

Theorem 3.12. *(i) $A \in 2^\omega$ is Schnorr random if and only if for every almost total functional Φ and every computable order h, Φ is weakly effectively measure-bounded along A via h.*

(ii) $A \in 2^\omega$ is Kurtz random if and only if for every almost total functional Φ and every computable order h, Φ is effectively measure-bounded along A via h.

Proof. For the left-to-right direction of Part (i), given a Schnorr random real A, an almost total Turing functional Φ, and a computable order function h, by Theorem 2.56 applied to the martingale L_Φ, there is some c such that for every n,

$$L_\Phi(A\restriction n) \leq c \cdot h(n)$$

for every n, and hence by choosing c' sufficiently large, we have

$$\lambda([\![\mathsf{Pre}_\Phi(A\restriction n)]\!]) \leq c' \cdot h(n)\lambda(A\restriction n)$$

for every n.

For the right-to-left direction of part (i), suppose there is some computable martingale d and a computable order h such that for every $c \in \omega$, there is some n such that

$$d(A\restriction n) > c \cdot h(n)$$

for infinitely many n, where $d(\varnothing) = 1$. Then by Proposition 3.7 and Lemma 3.9, there is an almost total Turing functional Φ such that $d = L_\Phi$. Thus, for every $c \in \omega$, there is some n such that

$$L_\Phi(A{\upharpoonright}n) > c \cdot h(n),$$

so

$$\lambda(\llbracket \mathsf{Pre}_\Phi(A{\upharpoonright}n) \rrbracket) > c \cdot h(n)\lambda(A{\upharpoonright}n).$$

The proof of (ii) is very similar. For the left-to-right direction, we merely need to replace all instances of "every n" with "infinitely many n" in the proof of the left-to-right direction of Part (i).

For the right-to-left direction, suppose that there exist a computable martingale d and a computable order h such that for every $c \in \omega$, there is some $N \in \omega$ such that for every $n \geq N$,

$$d(A{\upharpoonright}n) > c \cdot h(n).$$

Again by Proposition 3.7 and Lemma 3.9, there is an almost total Turing functional Φ such that $d = L_\Phi$, and thus

$$L_\Phi(A{\upharpoonright}n) > c \cdot h(n).$$

Thus, for every $c \in \omega$, there is some $N \in \omega$ such that for every $n \geq N$,

$$\lambda(\llbracket \mathsf{Pre}_\Phi(A{\upharpoonright}n) \rrbracket) > c \cdot h(n)\lambda(A{\upharpoonright}n).$$

□

3.3.2 Characterizing Notions of Randomness via Truth-Table Functionals

As we saw in the proof of Proposition 2.50, for every computable martingale d there is an exactly computable martingale d' such that for every $\sigma \in 2^{<\omega}$, $d(\sigma) \leq d'(\sigma) \leq d(\sigma) + 2$, where a martingale d' is exactly computable if $d' : 2^{<\omega} \to \mathbb{Q}_2$ is a martingale that, given input σ outputs the exact value $d(\sigma) \in \mathbb{Q}_2$ (as opposed to outputting an index for a sequence of rationals approximating $d(\sigma)$. As a result, the characterizations of computable randomness, Schnorr randomness, and Kurtz randomness, each given in terms of computable martingales, can also be given in terms of exactly computable martingales. In this section, we use these results to characterize these three notions of randomness in terms of truth-table functionals. The key result for these characterizations is the following.

Proposition 3.13. *If d is an exactly computable martingale such that $d(\varnothing) = 1$, then there is a truth-table functional Φ such that $d = L_\Phi$.*

Proof. We apply the Functional Existence Theorem to prove this result. We let V consist of all pairs of the form $(2^{-|\sigma|}d(\sigma), \sigma)$. We claim that V is a bounded functional request set. First note that V is a c.e. set of pairs, and that

$$\mathsf{wt}_V(\sigma) = 2^{-|\sigma|}d(\sigma)$$

for every $\sigma \in 2^{<\omega}$. It thus follows that $\mathsf{wt}_V(\varnothing) = d(\varnothing) = 1$ and

$$\mathsf{wt}_V(\sigma) = 2^{-|\sigma|}d(\sigma) = 2^{-|\sigma 0|}d(\sigma 0) + 2^{-|\sigma 1|}d(\sigma 1) = \mathsf{wt}_V(\sigma 0) + \mathsf{wt}_V(\sigma 1),$$

where the second equality holds because d is a martingale. By the Functional Existence Theorem, there is a Turing functional Φ such that $d = L_\Phi$. We now show that Φ is total by induction. That is, we show that for every n,

$$\bigcup_{\sigma \in 2^n} [\![\mathsf{Pre}_\Phi(\sigma)]\!] = 2^\omega.$$

First, by the construction proof of the Functional Existence Theorem, since $(1, \varnothing) \in V$, $(\varnothing, \varnothing)$ is enumerated into Φ, and thus

$$[\![\mathsf{Pre}_\Phi(\varnothing)]\!] = [\![\varnothing]\!] = 2^\omega.$$

Now suppose that

$$\bigcup_{\sigma \in 2^n} [\![\mathsf{Pre}_\Phi(\sigma)]\!] = 2^\omega$$

for fixed $n \in \omega$. For each $\sigma \in 2^n$, we have

$$(2^{-|\sigma 0|} d(\sigma 0), \sigma 0), (2^{-|\sigma 1|} d(\sigma 1), \sigma 1) \in V,$$

where

$$2^{-|\sigma|} d(\sigma) = 2^{-|\sigma 0|} d(\sigma 0) + 2^{-|\sigma 1|} d(\sigma 1).$$

Then if $\mathsf{Pre}_\Phi(\sigma) = \{\tau_1, \ldots, \tau_\ell\}$, we can find extensions $\xi_1^0, \ldots, \xi_j^0$ and $\xi_1^1, \ldots, \xi_k^1$ of the strings $\tau_1, \ldots, \tau_\ell$ such that

$$\sum_{i \le j} \lambda(\xi_i^0) = 2^{-|\sigma 0|} d(\sigma 0) \text{ and } \sum_{i \le k} \lambda(\xi_i^1) = 2^{-|\sigma 1|} d(\sigma 1).$$

If we enumerate the pairs $(\xi_i^0, \sigma 0)$ for $i \leq j$ and $(\xi_i^1, \sigma 1)$ for $i \leq k$ into Φ, then

$$\mathsf{Pre}_\Phi(\sigma 0) = \{\xi_1^0, \ldots, \xi_j^0\} \text{ and } \mathsf{Pre}_\Phi(\sigma 1) = \{\xi_1^1, \ldots, \xi_k^1\}$$

and hence

$$[\![\mathsf{Pre}_\Phi(\sigma)]\!] = [\![\mathsf{Pre}_\Phi(\sigma 0)]\!] \cup [\![\mathsf{Pre}_\Phi(\sigma 1)]\!].$$

Continuing in this way for each $\sigma \in 2^n$, it follows that

$$\bigcup_{\tau \in 2^{n+1}} [\![\mathsf{Pre}_\Phi(\tau)]\!] = 2^\omega.$$

$\square$

Using Proposition , we can now derive the following.

Theorem 3.14. *(i) $A \in 2^\omega$ is computably random if and only if for every truth-table functional Φ, Φ is measure-bounded along A.*

(i) $A \in 2^\omega$ is Schnorr random if and only if for every truth-table functional Φ and every computable order h, Φ is effectively measure-bounded along A via h.

(ii) $A \in 2^\omega$ is Kurtz random if and only if for every truth-table functional Φ and every computable order h, Φ is weakly effectively measure-bounded along A via h.

Proof. The proof is like the proofs of Theorems 10 and 12, except now we use Proposition 13 and the fact that the martingale characterizations of computable randomness, Schnorr randomness, and Kurtz randomness all still hold if we restrict to the collection of exactly computable martingales. □

3.3.3 A Priori Complexity Characterizations of Notions of Randomness

In this section, the results from the previous section can be recast in terms of *a priori complexity*. As a first step in this direction, we need to consider optimal c.e. continuous semimeasures.

Definition 3.15. A c.e. continuous semimeasure δ is *optimal* if for every c.e. continuous semimeasure δ', there is some $c \in \omega$ such that

$$\delta(\sigma) \geq c\delta'(\sigma)$$

for every $\sigma \in 2^{<\omega}$.

With this definition, we can now define the a priori complexity of a string $\sigma \in 2^{<\omega}$.

Definition 3.16. Let δ be an optimal semimeasure. Then the *a priori complexity* of $\sigma \in 2^{<\omega}$, denoted $KM(\sigma)$, is defined to be

$$KM(\sigma) = -\log \delta(\sigma).$$

The key observation to make is that a c.e. supermartingale d can easily be transformed into a c.e. continuous semimeasure $\delta(\sigma) := 2^{-|\sigma|}d(\sigma)$. In particular, if L_Φ is a c.e. supermartingale, then the function δ_Φ, defined to be

$$\delta_\Phi(\sigma) := 2^{-|\sigma|} L_\Phi(\sigma) = \lambda(\llbracket \mathsf{Pre}_\Phi(\sigma) \rrbracket)$$

for all $\sigma \in 2^{<\omega}$, is a c.e. continuous semimeasure.

Now, if we set $KM_\Phi(\sigma) = -\log \lambda(\llbracket \mathsf{Pre}_\Phi(\sigma) \rrbracket)$, we can summarize our previous results as follows:

Theorem 3.17. *(i) $A \in 2^\omega$ is Martin-Löf random if and only if for every Turing functional Φ and for all $n \in \omega$,*

$$KM_\Phi(A{\upharpoonright}n) \geq n - O(1).$$

(ii) $A \in 2^\omega$ is computably random if and only if for every almost total functional Φ and for all $n \in \omega$,

$$KM_\Phi(A{\upharpoonright}n) \geq n - O(1).$$

(iii) $A \in 2^\omega$ is Schnorr random if and only if for every almost total functional Φ, for every computable order g, and for every $n \in \omega$,

$$KM_\Phi(A{\upharpoonright}n) \geq n - g(n) - O(1).$$

(iv) $A \in 2^\omega$ is Kurtz random if and only if for every almost total functional Φ, for

every computable order g, and for infinitely many $n \in \omega$,

$$KM_\Phi(A\restriction n) \geq n - g(n) - O(1).$$

Proof. (i) By Theorem 3.6, $A \in \mathsf{MLR}$ if and only if

$$\lambda([\![\mathsf{Pre}_\Phi(A\restriction n)]\!]) \leq c \cdot \lambda(A\restriction n)$$

for every Turing functional Φ. Hence,

$$\begin{aligned} KM_\Phi(A\restriction n) &= -\log \lambda([\![\mathsf{Pre}_\Phi(A\restriction n)]\!]) \\ &\geq -\log(c \cdot \lambda(A\restriction n)) \\ &= -\log(c) + -\log(\lambda(A\restriction n)) \\ &= n - O(1). \end{aligned}$$

(ii) By Theorem 3.10, $A \in \mathsf{CR}$ if and only if

$$\lambda([\![\mathsf{Pre}_\Phi(A\restriction n)]\!]) \leq c \cdot \lambda(A\restriction n)$$

for every almost total Turing functional Φ. Hence, by the same reasoning as above,

$$KM_\Phi(A\restriction n) \geq n - O(1).$$

(iii) By Theorem 3.12 (i), $A \in \mathsf{SR}$ if and only if

$$\lambda([\![\mathsf{Pre}_\Phi(A{\upharpoonright}n)]\!]) \leq c \cdot h(n)\lambda(A{\upharpoonright}n)$$

for every $n \in \omega$, for every almost total functional Φ, and every computable order h. Given a computable order g, let $h(n) = 2^{g(n)}$. Then

$$\begin{aligned} KM_\Phi(A{\upharpoonright}n) &= -\log \lambda([\![\mathsf{Pre}_\Phi(A{\upharpoonright}n)]\!]) \\ &\geq -\log(c \cdot h(n)\lambda(A{\upharpoonright}n)) \\ &= -\log(c) + -\log(h(n)) + -\log(\lambda(A{\upharpoonright}n)) \\ &= n - g(n) - O(1) \end{aligned}$$

for every $n \in \omega$. (iv) The proof is like the proof of (iii), except we replace "every $n \in \omega$" with "infinitely many $n \in \omega$". $\square$

Moreover, we can recast Theorem 3.14 as follows:

Theorem 3.18. *(i) $A \in 2^\omega$ is computably random if and only if for every truth-table functional Φ and for all n,*

$$KM_\Phi(A{\upharpoonright}n) \geq n - O(1).$$

(ii) $A \in 2^\omega$ is Schnorr random if and only if for every truth-table functional Φ, for every computable order h, and for almost every n,

$$KM_\Phi(A{\upharpoonright}n) \geq n - h(n).$$

(iii) $A \in 2^\omega$ *is Kurtz random if and only if for every truth-table functional* Φ*, for every computable order* h*, and for every* n*,*

$$KM_\Phi(A\restriction n) \geq n - h(n).$$

Proof. The proof is nearly identical to the proof of (ii), (iii), and (iv) of Theorem 3.17, except we use the characterization of measure-boundedness given in terms of truth-table functionals provided by Theorem 3.14. $\square$

We close with the observation that both Theorem 3.17 and Theorem 3.18 can be further generalized.

Theorem 3.19. *Let* $\mu \in \mathscr{M}_c$*.*

(i) $A \in \mathsf{MLR}_\mu$ *if and only if for every Turing functional* Φ *and for all* $n \in \omega$*,*

$$KM_\Phi(A\restriction n) \geq -\log \mu(A\restriction n) - O(1).$$

(ii) $A \in \mathsf{CR}_\mu$ *if and only if for every almost total functional* Φ *and for all* $n \in \omega$*,*

$$KM_\Phi(A\restriction n) \geq -\log \mu(A\restriction n) - O(1).$$

(iii) $A \in \mathsf{SR}_\mu$ *if and only if for every almost total functional* Φ*, for every computable order* g*, and for every* $n \in \omega$*,*

$$KM_\Phi(A\restriction n) \geq -\log \mu(A\restriction n) - g(n) - O(1).$$

(iv) $A \in \mathsf{KR}_\mu$ *if and only if for every almost total functional* Φ*, for every computable order* g*, and for infinitely many* $n \in \omega$*,*

$$KM_\Phi(A{\restriction}n) \geq -\log \mu(A{\restriction}n) - g(n) - O(1).$$

Proof Sketch. Use the characterizations of MLR_μ, CR_μ, SR_μ, and KR_μ in terms of c.e. μ-supermartingales (for MLR_μ) and computable μ-martingales (for CR_μ, SR_μ, and KR_μ). First show by the Functional Existence Theorem that for every μ-supermartingale d, there is a Turing functional Φ such that

$$d(\sigma) = \frac{\mu([\![\mathsf{Pre}_\Phi(\sigma)]\!])}{\mu(\sigma)}$$

and for every μ-martingale d', there is an almost total functional Ψ such that

$$d'(\sigma) = \frac{\mu([\![\mathsf{Pre}_\Psi(\sigma)]\!])}{\mu(\sigma)}.$$

The measure-bounded characterizations can be generalized to μ, and then one can proceed as in the proof of Theorem 3.17. □

Similarly, one can also show:

Theorem 3.20. *Let* $\mu \in \mathscr{M}_c$ *be exactly computable.*

(i) $A \in \mathsf{CR}_\mu$ *if and only if for every truth-table functional* Φ *and for all* n*,*

$$KM_\Phi(A{\restriction}n) \geq -\log \mu(A{\restriction}n) - O(1).$$

(ii) $A \in \mathsf{SR}_\mu$ *if and only if for every truth-table functional* Φ, *for every computable order* h, *and for almost every* n,

$$KM_\Phi(A\restriction n) \geq -\log \mu(A\restriction n) - h(n).$$

(iii) $A \in \mathsf{KR}_\mu$ *if and only if for every truth-table functional* Φ, *for every computable order* h, *and for every* n,

$$KM_\Phi(A\restriction n) \geq -\log \mu(A\restriction n) - h(n).$$

Proof Sketch. Since μ is exactly computable, we can proceed as in the proof of Proposition 3.3.2, and then follow the proof sketch given above for Theorem 3.19. $\square$

CHAPTER 4

DEMUTH'S THEOREM: VARIANTS AND APPLICATIONS

4.1 Introduction

The main topic of this chapter is a theorem of Oswald Demuth's concerning the behavior of random sequences under truth-table reductions. Roughly, Demuth's Theorem tells us that if we apply a tt-functional to a Martin-Löf random sequence X, if the resulting sequence Y has *any* non-trivial computational content whatsoever (i.e. if it is not a computable sequence), then from Y we can effectively recover a Martin-Löf random sequence Z. Recently, a stronger version of Demuth's Theorem has repeatedly appeared in circulated drafts, talks, and even in some published papers, which asserts that one can even require that Y is wtt-equivalent to Z. This is for example the version given in [Fra08], where the author further asks

(i) whether z can further be required to be tt-equivalent to y; and

(ii) whether Demuth's theorem also holds for Schnorr randomness.

Not only is (i) answered in the negative, but the "wtt-version" of Demuth's theorem is false, contrary to what has been reported in the literature. However, (ii) has a positive answer. In fact, we show in this chapter that Demuth's Theorem holds for computable randomness, Schnorr randomness, and weak 2-randomness.

4.2 Proving Demuth's Theorem

In this section, we give a detailed proof of Demuth's Theorem.

Theorem 4.1 (Demuth [Dem88]). *Given $X \in \mathsf{MLR}$, for every non-computable $Y \leq_{tt} X$, there is some $Z \in \mathsf{MLR}$ such that $Y \equiv_T Z$.*

Although Demuth proved his theorem using tools from constructive analysis, another way to prove the theorem is to break it down into two results, each of which has been shown independently of Demuth's work. The first result is the well-known "preservation of randomness" theorem. In Chapter 2, Section 2.4.3, we saw that total and almost total functionals induce computable measures. According to the preservation of randomness theorem, these functionals also having the property of mapping sequences that are random with respect to the initial measure to sequences that are random with respect to the induced measure.

Theorem 4.2 (Preservation of Martin-Löf randomness). *For $\mu \in \mathscr{M}_c$ and Φ an almost total functional, for every $X \in 2^\omega$, $X \in \mathsf{MLR}_\mu$ implies $\Phi(X) \in \mathsf{MLR}_{\mu_\Phi}$.*

Proof. Suppose that $\Phi(X) \notin \mathsf{MLR}_{\mu_\Phi}$; we will show that $X \notin \mathsf{MLR}_\mu$. Let $\{\mathcal{U}_i\}_{i\in\omega}$ be a μ_Φ-Martin-Löf test such that $\Phi(x) \in \bigcap_{i\in\omega} \mathcal{U}_i$. We define a μ-Martin-Löf test $\{\mathcal{V}_i\}_{i\in\omega}$ containing X as follows. First, let $\{S_i\}_{i\in\omega}$ be a prefix-free presentation of $\{\mathcal{U}_i\}_{i\in\omega}$. Then we define, for each $i \in \omega$,

$$P_i = \bigcup_{\sigma \in S_i} \mathsf{Pre}_\Phi(\sigma).$$

Note that since S_i is prefix-free, for distinct $\sigma_1, \sigma_2 \in S_i$, $\mathsf{Pre}_\Phi(\sigma_1) \cap \mathsf{Pre}_\Phi(\sigma_2) = \emptyset$, and so $\bigcup_{\sigma \in S_i} \mathsf{Pre}_\Phi(\sigma)$ is a disjoint union. Hence

$$\mu([\![P_i]\!]) = \mu\Big(\bigcup_{\sigma \in S_i} [\![\mathsf{Pre}_\Phi(\sigma)]\!]\Big) = \sum_{\sigma \in S_i} \mu([\![\mathsf{Pre}_\Phi(\sigma)]\!]) = \sum_{\sigma \in S_i} \mu(\Phi^{-1}([\![\sigma]\!])) = \mu_\Phi(\mathcal{U}_i).$$

Now if we set $\mathcal{V}_i := [\![P_i]\!]$ for each i, we have $\mu(\mathcal{V}_i) = \mu_\Phi(\mathcal{U}_i)$ for each i. In addition, since the collection $\{\mathcal{V}\}_{i\in\omega}$ is definable uniformly from $\{\mathcal{U}\}_{i\in\omega}$, it follows that $\{\mathcal{V}\}_{i\in\omega}$ is a μ-Martin-Löf test. Lastly, we must verify that $X \in \bigcap_{i\in\omega} \mathcal{V}_i$. For each i, since $\Phi(X) \in \mathcal{U}_i$, there is some $\sigma \in S_i$ and some least $n \in \omega$ such that $\Phi^{X\upharpoonright n} \succeq \sigma$. Thus $X\upharpoonright n \in \mathsf{Pre}_\Phi(\sigma)$, and so it follows that $X\upharpoonright n \in P_i$ and $X \in \mathcal{V}_i$.

□

The second result used to derive Demuth's Theorem is sometimes referred to as "Levin's theorem" or "the Levin-Kautz theorem" (although Levin proved the theorem with Zvonkin, independently of Kautz).

Theorem 4.3 (Levin/Zvonkin [ZL70], Kautz [Kau91])**.** *If $Y \in \mathsf{MLR}_\mu$ is non-computable for some $\mu \in \mathscr{M}_c$, then there is some $Z \in \mathsf{MLR}$ such that $Y \equiv_T Z$.*

In the remainder of this section, we will provide a proof of the Levin-Kautz Theorem that has not appeared explicitly in the effective randomness literature. Before we do so, we should note that the two theorems given above immediately imply Demuth's Theorem: Given a *tt*-functional Φ and $X \in \mathsf{MLR}$, since Φ is almost total, by the preservation of randomness, it follows that $\Phi(X) \in \mathsf{MLR}_{\lambda_\Phi}$, and λ_Φ is computable by Lemma 2.15. Further, if $\Phi(X)$ is not computable, then by the Levin-Kautz Theorem there is some $Z \in \mathsf{MLR}$ such that $\Phi(X) \equiv_T Z$, thus establishing the result.

Let us now turn to the proof of the Levin-Kautz Theorem. In order to prove the result, we need to prove an auxiliary result that we will refer to as the "Kautz conversion procedure". This result provides a converse to Lemmas 2.15 and 2.16, as it shows that any computable measure μ can be induced by the Lebesgue measure together with an almost total functional Φ. Moreover, we can define Φ with the additional property of being *non-decreasing*: For $X, Y \in \mathsf{dom}(\Phi)$, $X \leq_{lex} Y$ implies that $\Phi(X) \leq_{lex} \Phi(Y)$ (or, equivalently, thinking of X and Y as real numbers, $X \leq Y$ implies that $\Phi(X) \leq \Phi(Y)$.

Theorem 4.4 (The Kautz conversion procedure)**.** *Let $\mu \in \mathscr{M}_c$. Then there exists a non-decreasing, almost total functional Φ such that $\lambda_\Phi = \mu$. Moreover,*

- *if μ is atomless, then Φ is one-to-one, and*
- *if μ is positive, then Φ is onto up to a set of μ-measure 0.*

Finally, if μ is both atomless and positive, then Φ has an almost total inverse Φ^{-1} such that $\mu_{\Phi^{-1}} = \lambda$.

Proof of Theorem 4.4. The key observation in Kautz's proof is that for $\mu \in \mathscr{M}_c$, almost every $X \in 2^\omega$ has a binary representation in [0,1] given in terms of μ, which we will refer to as its "μ-representation", denoted by $\mathsf{seq}_\mu(X)$.

By removing the initial decimal point, we can simply consider $\mathsf{seq}_\mu(X)$ as a member of 2^ω. Similarly, by adding an initial decimal point, we can consider X as a member of [0,1]. Moving in this way between 2^ω and [0,1] will facilitate the presentation of this proof.

Now, using the notion of a μ-representation, we will define an almost total functional Φ so that $\Phi(X) = \mathsf{seq}_\mu(X)$. To compute the μ-representation of $X \in 2^\omega$, we make use of what we call a μ-partition of [0,1]. A *μ-partition of [0,1] at level n* is a collection of $k = 2^n$ closed intervals $I_{\sigma_0}, I_{\sigma_1}, \dots I_{\sigma_{k-1}}$ such that

1. $\sigma_0, \sigma_1, \dots, \sigma_{k-1}$ is a listing of all strings of length n in lexicographical ordering,
2. $\bigcup_{i=0}^{k-1} I_{\sigma_i} = [0,1]$,
3. $\sup I_{\sigma_i} = \inf I_{\sigma_{i+1}}$ for $0 \leq i \leq k-2$, and
4. $\mu(\sigma_i) = \lambda(I_{\sigma_i})$ for $0 \leq i \leq k-1$.

We further require that the μ-partition of level n is compatible with the μ-partition of level $n+1$ for every n, so that given a string σ of length n, we have

$$I_\sigma = I_{\sigma 0} \cup I_{\sigma 1}.$$

Given a sequence $X \in 2^\omega$, we can compute its μ-representation $\mathsf{seq}_\mu(X)$ as follows. To determine the first bit of $\mathsf{seq}_\mu(X)$, we consider the μ-partition of [0,1] at level 1, $I_0 \cup I_1$. Since μ is computable but not necessarily exactly computable, we may have to approximate I_0 and I_1 until we see that $X \in I_0$ or $X \in I_1$. This will occur, so long as X is not the right endpoint of I_0. To see this, first note that since both I_0 and I_1 are closed, there is some n such that for every $Y \in [\![X{\upharpoonright}n]\!]$, Y is contained entirely in either I_0 or I_1. Further, $I_0 = [0, \mu(0)]$ and $I_1 = [\mu(0), 1]$, and hence we can approximate I_0 and I_1 using the approximation $\widehat{\mu}$ of μ as follows:

$$I_{0,s} = [0, \widehat{\mu}(0, s+1) + 2^{-s}] \text{ and } I_{1,s} = [\widehat{\mu}(0, s+1) - 2^{-s}, 1]$$

Then we have $I_i \subseteq I_{i,s}$ for $i \in \{0,1\}$, so we can simply wait until we find $n, s \in \omega$ such that $[\![X{\upharpoonright}n]\!]$ is contained entirely in $I_{i,s}$ for some $i \in \{0,1\}$. If $X \in I_0$, the first bit of $\mathsf{seq}_\mu(X)$ is a 0, and if $X \in I_1$, the first bit of $\mathsf{seq}_\mu(X)$ is a 1. In the case where X is the right endpoint of I_0, $\Phi(X)$ is undefined.

Having determined the first n bits of $\mathsf{seq}_\mu(X)$ by finding σ such that $|\sigma| = n$ and $X \in I_\sigma$, we determine whether $X \in I_{\sigma 0}$ or $x \in I_{\sigma_1}$ (where $I_{\sigma 0}$ and $I_{\sigma 1}$ are given by the μ-partition of [0,1] at level $n+1$), and output a 0 or 1 accordingly, as in the base case described above. The only difference is the way in which we approximate the intervals $I_{\sigma 0}$ and $I_{\sigma 1}$. Suppose that $X \in I_\sigma = [\ell(\sigma), r(\sigma)]$, where

$$\ell(\sigma) = \sum_{\{\tau \in 2^n : \tau <_{lex} \sigma\}} \mu(\tau) \quad \text{and} \quad r(\sigma) = \ell(\sigma) + \mu(\sigma).$$

If X is not an endpoint of I_σ, then, as above, we can approximate $I_{\sigma 0}$ and $I_{\sigma 1}$ until we see that $X \in I_{\sigma 0}$ or $X \in I_{\sigma 1}$. That is, we wait until we find $s, n\omega$ such that $[\![X{\upharpoonright}n]\!]$ is contained entirely in $I_{\sigma i,s} = [\ell_s(\sigma i), r_s(\sigma i)]$ for some $i \in \{0,1\}$, where

$$\ell_s(\alpha) = \sum_{\{\tau \in 2^{<\omega} : |\tau| = |\alpha| \,\wedge\, \tau <_{lex} \alpha\}} \widehat{\mu}(\tau, s + |\alpha| + 2) - 2^{-(s+2)}$$

and

$$r_s(\alpha) = \ell_s(\alpha) + \widehat{\mu}(\alpha, s + |\alpha| + 2) + 2^{-(s+2)}.$$

for each $\alpha \in 2^{<\omega}$, from which it follows that

$$\begin{aligned} r_s(\alpha) - \ell_s(\alpha) &= \widehat{\mu}(\alpha, s + |\alpha| + 2) + 2^{-(s+2)} \\ &\geq \mu(\alpha) - 2^{-(s+|\alpha|+2)} + 2^{-(s+2)} \\ &= \mu(\alpha) + 2^{-(s+2)}(1 - 2^{-|\alpha|}). \end{aligned} \tag{4.1}$$

Thus, $I_\alpha \subseteq I_{\alpha,s+1} \subseteq I_{\alpha,s}$ for every $\alpha \in 2^{<\omega}$ and $s \in \omega$.

If X is not an endpoint of I_σ for any $\sigma \in 2^{<\omega}$, then $\mathsf{seq}_\mu(X)$ is the unique sequence $Y \in 2^\omega$ such that $X \in I_{Y\restriction n}$ for every n. Clearly, Φ is non-decreasing and almost total, and we have that

$$\Phi^{-1}(\llbracket\sigma\rrbracket) = \{X \in 2^\omega : \Phi(X) \succ \sigma\} = I_\sigma \setminus \mathcal{A},$$

where $\mathcal{A}$ contains all endpoints of intervals I_τ for $\tau \succeq \sigma$, and hence has measure 0. Thus,

$$\lambda_\Phi(\sigma) = \lambda(\Phi^{-1}(\llbracket\sigma\rrbracket)) = \lambda(I_\sigma \setminus \mathcal{A}) = \lambda(I_\sigma) = \mu(\sigma).$$

In the case where μ is atomless, we have that for every $Y \in 2^\omega$, $\lim_{n\to\infty} \lambda(I_{Y\restriction n}) = 0$. This implies that there is a unique X such that $\bigcap_{n\in\omega} I_{Y\restriction n} = \{X\}$. Thus, if $\Phi(X_1) = \Phi(X_2)$, we must have $X_1 = X_2$.

Note that Φ is not onto: if X is an endpoint of some I_σ, then $\mathsf{seq}_\mu(X)$ will not be in the range of Φ. However, if μ is positive, Φ is onto up to a set of measure 0, which is in fact a countable set. To see this, note that for every $\sigma \in 2^{<\omega}$, $\lambda(I_\sigma) > 0$, which means that $\Phi^{-1}(\llbracket\sigma\rrbracket)$ is non-empty. If $Y \in 2^\omega$, where $Y \neq \mathsf{seq}_\mu(Z)$ for any Z

that is an endpoint of some I_σ, then

$$\Phi^{-1}([\![Y\restriction n]\!]) \supseteq \Phi^{-1}([\![Y\restriction(n+1)]\!])$$

for every n and each is non-empty. Then there is some X such that

$$X \in \bigcap_{n\in\omega} \Phi^{-1}([\![Y\restriction n]\!]).$$

Therefore $Y = \mathsf{seq}_\mu(X)$. In the case where μ is both atomless and positive, then since Φ is one-to-one, it has an inverse Φ^{-1}. Since Φ is onto up to a countable set of measure zero (and hence μ-measure zero), it follows that Φ^{-1} is μ-almost total. Given $Y \in 2^\omega$ in the range of Φ, i.e. $\Phi(x) = y$ for some $X \in 2^\omega$, then $\Phi^{-1}(Y)$ can be computed by successively computing $\Phi^{-1}([\![Y\restriction n]\!])$ for each n and then taking the intersection. More specifically, since

$$\bigcap_{n\in\omega} \Phi^{-1}([\![Y\restriction n]\!]) = \{X\},$$

for each i, we will eventually find some n_i such that

$$Z \in \bigcap_{n\le n_i} \Phi^{-1}([\![Y\restriction n]\!]) \Rightarrow Z\restriction i = X\restriction i.$$

Thus we will have $(\Phi^{-1})^{Y\restriction n_i} \succeq X\restriction i$ for every i. $\square$

The next theorem provides a partial converse to the preservation of randomness theorem, stating that every sequence that is random with respect to some computable measure induced by a functional Φ has a random sequence in its preimage under Φ.

Theorem 4.5. *For $\mu \in \mathscr{M}_c$, let Φ be a μ-almost total functional. For every $Y \in \mathsf{MLR}_{\mu_\Phi}$ there is some $X \in \mathsf{MLR}_\mu$ such that $\Phi(X) = Y$.*

Proof. Suppose $Y \in 2^\omega$ is such that for all $X \in 2^\omega$ such that $\Phi(X) = Y$, $X \notin \mathsf{MLR}_\mu$. Let $\{\mathcal{U}_i\}_{i\in\omega}$ be the universal μ-Martin-Löf test, and consider

$$\mathcal{V}_i = \{Z \in 2^\omega : (\forall X)[X \notin \mathcal{U}_i \Rightarrow \Phi(X) \neq Z]\}.$$

We claim that $\{\mathcal{V}_i\}_{i\in\omega}$ is a μ_Φ-Martin-Löf test. First, observe that $Z \in \mathcal{V}_i$ if and only if $Z \notin \Phi(2^\omega \setminus \mathcal{U}_i)$. Since Φ is an almost total Turing functional, the image under Φ of a Π^0_1 class is also a Π^0_1 class. In particular, $\Phi(2^\omega \setminus \mathcal{U}_i)$ is a Π^0_1 class.

To see that $\mu_\Phi(\mathcal{V}_i) \leq 2^{-i}$, since $2^\omega \setminus \mathcal{U}_i \subseteq \Phi^{-1}(\Phi(2^\omega \setminus \mathcal{U}_i))$, we have

$$\mu_\Phi(\mathcal{V}_i) = 1-\mu_\Phi(\Phi(2^\omega\setminus\mathcal{U}_i)) = 1-\mu(\Phi^{-1}(\Phi(2^\omega\setminus\mathcal{U}_i))) \leq 1-\mu(2^\omega\setminus\mathcal{U}_i) \leq 1-(1-2^{-i}) = 2^{-i}.$$

Finally, since for all $X \in 2^\omega$ such that $\Phi(X) = Y$, $X \notin \mathsf{MLR}_\mu$, it follows that $X \notin \mathcal{U}_i$ implies $\Phi(X) \neq Y$ for every i, and so $Y \in \mathcal{V}_i$ for every i. Thus, $Y \notin \mathsf{MLR}_{\mu_\Phi}$. $\square$

We can now prove the Levin-Kautz Theorem.

Proof of Theorem 4.3. Given a non-computable Y and $\mu \in \mathscr{M}_c$ such that $X \in \mathsf{MLR}_\mu$, by Theorem 4.4, there is some non-decreasing almost total functional Φ such that $\mu = \lambda_\Phi$. Since $Y \in \mathsf{MLR}_{\lambda_\Phi}$, by Theorem 4.5, there is some $Z \in \mathsf{MLR}$ such that $\Phi(Z) = Y$. Moreover, suppose there is $U \in 2^\omega$ such that $\Phi(U) = Y$ and $U \neq Z$. Without loss, we can assume that $U <_{lex} Z$. Since Φ is non-decreasing, this would mean that the entire interval $[U, Z]$ is mapped by Φ to the singleton $\{Y\}$, and therefore Y would be an atom of λ_Φ and hence computable, contradicting our assumption. Therefore, Y has a unique preimage Z under Φ; in other words $\Phi^{-1}(\{Y\})$ is a $\Pi^0_1(Y)$-class containing a single element Z. Just as isolated paths in a Π^0_1 class are computable, the isolated paths in a $\Pi^0_1(Y)$ class are Y-computable, and thus Z is Y-computable. Therefore, $Z \equiv_T Y$ and $Z \in \mathsf{MLR}$. □

We conclude this section with one last result, which will be useful in Chapter 5.

Theorem 4.6. *Let Φ be almost total. Then for every $X \in \mathsf{MLR}$ such that $\Phi(X)$ is not computable,*

$$X \in \mathsf{MLR}^A \Leftrightarrow \Phi(X) \in \mathsf{MLR}^A_{\lambda_\Phi}$$

for all $A \in 2^\omega$.

Proof. ($\Rightarrow$): This follows from the proof of Theorem 4.2, the preservation of Martin-Löf randomness, where we replace the Martin-Löf tests appearing in the original proof with Martin-Löf tests relative to A.

($\Leftarrow$): This follows from the proof of Theorem 4.5, where we replace the Martin-Löf tests appearing in the original proof with Martin-Löf tests relative to A. □

4.3 Demuth's Theorem for Other Notions of Randomness

In this section, we consider Demuth's Theorem for other notions of effective randomness; namely, Schnorr randomness, computable randomness, and weak 2-randomness.

Theorem 4.7. *Let $\mu \in \mathscr{M}_c$ and Φ be an almost total functional. Then $X \in \mathsf{SR}_\mu$ implies $\Phi(X) \in \mathsf{SR}_{\mu_\Phi}$.*

Proof. In the proof of the preservation of Martin-Löf randomness (Theorem 4.2), we show that if $\Phi(X)$ is contained in a μ_Φ-Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$, then there is a μ-Martin-Löf test $\{\mathcal{V}_i\}_{i\in\omega}$ containing X. In fact, we prove more: we show that $\mu(\mathcal{V}_i) = \mu_\Phi(\mathcal{U}_i)$ for every $i \in \omega$. Thus, if $\{\mathcal{U}_i\}_{i\in\omega}$ is a μ_Φ-Schnorr test containing $\Phi(X)$, it follows that $\{\mathcal{V}\}_{i\in\omega}$ is a μ-Schnorr test containing X. □

We also have the preservation of weak 2-randomness.

Theorem 4.8. *Let $\mu \in \mathscr{M}_c$ and Φ be almost total. Then $X \in \mathsf{W2R}_\mu$ implies $\Phi(X) \in \mathsf{W2R}_{\mu_\Phi}$.*

Proof. Suppose $\Phi(X) \in \mathcal{U}$, where $\mathcal{S}$ is a Π^0_2 class such that $\mu_\Phi(\mathcal{S}) = 0$. Setting $\mathcal{S} = \bigcap_{i\in\omega}\mathcal{U}_i$, where each $\mathcal{U}_i$ is a Σ^0_1 class such that $\mu_\Phi(\mathcal{U}_i) \to 0$ as i grows without bound, each $\Phi^{-1}(\mathcal{U}_i)$ is a Σ^0_1 class such that $\mu(\Phi^{-1}(\mathcal{U}_i) \to 0$ as i approaches infinity. Thus, $\Phi^{-1}(\mathcal{S}) = \bigcap_{i\in\omega}\Phi^{-1}(\mathcal{U}_i)$ is a Π^0_2 class of μ-measure 0 containing X, and hence $X \notin \mathsf{W2R}_\mu$. □

Although we have the proved both the preservation of Martin-Löf randomness and the preservation of Schnorr randomness, perhaps surprisingly, there is no preservation of computable randomness. We prove the following.

Theorem 4.9. *There exists a tt-functional Φ and $A \in \mathsf{CR}$ such that $\Phi(A) \notin \mathsf{CR}_{\lambda_\Phi}$.*

Proof. Let $A \in \mathsf{CR} \setminus \mathsf{KLR}$. Then there is a non-monotonic betting strategy S that succeeds on A. By Proposition 2.85, we can assume that this betting strategy S is total. Then we define Φ as follows: Given $X \in 2^\omega$, let $\Phi(X)$ be the sequence of bits visited by S during the game, in order of appearance. Clearly, Φ is total, and for every $n \in \omega$, n bits of X are need to compute the first n bits of $\Phi(X)$. Thus, $\lambda_\Phi = \lambda$. Now since S succeeds on A, this means there is a computable martingale d that succeeds on $\Phi(A)$, and hence $\Phi(A) \notin \mathsf{CR} = \mathsf{CR}_{\lambda_\Phi}$.

$\square$

The proof of the Levin-Kautz Theorem that we provided in the previous section does not work for Schnorr randomness or for weak 2-randomness, since the proof relies upon the existence of a universal Martin-Löf test, and it is well-known that there is no universal Schnorr test and no universal test for weak 2-randomness. For computable randomness, the situation is even worse as we have seen that there is no randomness preservation for this notion. As a consequence, we cannot prove Demuth's theorem for these randomness notions by a direct adaptation of the proof of Demuth's theorem for Martin-Löf randomness.

There are, however, alternative approaches to proving Demuth's Theorem that allow us to overcome the difficulty, one that works both for computable randomness and Schnorr randomness and another for weak 2-randomness.

The alternative approach for Schnorr randomness and computable randomness uses Propositions 2.59: a Schnorr random (resp. computably random) sequence is either Martin-Löf random, or it is high. Armed with this dichotomy, we get Demuth's theorem for Schnorr randomness and computable randomness almost immediately from Demuth's theorem for Martin-Löf randomness. In fact, we get a slightly stronger statement that subsumes both, in the sense that it suffices to assume $X \in \mathsf{SR}$ to get

$Z \in \mathsf{CR}$ in the conclusion.

Theorem 4.10 (Demuth's Theorem for computable and Schnorr randomness). *Let $X \in \mathsf{SR}$ and let Φ be a truth-table functional. If $\Phi(X) = Y$ is not computable, then there is some $Z \in \mathsf{CR}$ such that $Y \equiv_T Z$.*

Proof. Regardless of whether $X \in \mathsf{CR} \subseteq \mathsf{SR}$ or $X \in \mathsf{SR} \setminus \mathsf{CR}$, we can still apply the preservation of randomness for Schnorr randomness (Theorem 4.7) to conclude that $\Phi(X) = Y$ is Schnorr random with respect to some computable measure μ. We now distinguish two cases.

Case 1: If Y is not high, then by Proposition 2.59 it must be μ-Martin-Löf random. Thus, we can apply the Levin-Kautz theorem (Theorem 4.3) to get a real $Z \equiv_T Y$ that is Martin-Löf random and hence computably random.

Case 2: If Y is high, then we can directly apply Theorem 2.53 to get the existence of some $Z \in \mathsf{CR}$ such that $Z \equiv_T Y$. $\square$

To show that Demuth's theorem for weak 2-randomness holds, we use the characterization of weak 2-randomness from Theorem 2.69: $X \in \mathsf{W2R}$ if and only if $X \in \mathsf{MLR}$ and X forms a minimal pair with $\emptyset'$.

Theorem 4.11 (Demuth's Theorem for weak 2-randomness). *Let $X \in \mathsf{W2R}$ and let Φ be a truth-table functional. If $\Phi(X) = Y$ is not computable, then there is some $Z \in \mathsf{W2R}$ such that $Y \equiv_T Z$.*

Proof. By the preservation of weak 2-randomness (Theorem 4.8), $\Phi(X) \in \mathsf{W2R}_{\lambda_\Phi}$. Then since $\Phi(X) = Y$ is not computable, it follows that $Y \in \mathsf{MLR}_{\lambda_\Phi}$ and Y forms a minimal pair with $\emptyset'$. Then by the Levin-Kautz theorem, Theorem 4.3, there is some $Z \equiv_T Y$ that is Martin-Löf random and forms a minimal pair with $\emptyset'$, and thus is weakly 2-random. $\square$

4.4 The failure of Demuth's theorem for *wtt*-reducibility

In the original proof of Demuth's theorem, Demuth proves a slightly stronger conclusion.

Proposition 4.12 (Demuth)**.** *Let $X \in \mathsf{MLR}$ and suppose $Y \leq_{tt} X$ is not computable. Then there is some $Z \in \mathsf{MLR}$ such that $Y \equiv_T Z$ and $Y \leq_{tt} Z$.*

Proof. The key insight needed to derive the additional conclusion that $Y \leq_{tt} Z$ is that we can take the *tt*-reduction Φ from X to Y and reorder the truth-tables used by Φ to define a non-decreasing functional $\widehat{\Phi}$ that induces the same measure as Φ. More precisely, given a function h that bounds the use of Φ, if we take the table consisting of a column of the $2^{h(n)}$ strings of length $h(n)$ listed in lexicographical order alongside a column of the images of these strings under Φ (which will be strings of length n), we can define a new table by permuting the values in the output column to list them in lexicographical order (leaving fixed the first column of input strings), thus yielding a non-decreasing map from strings on length $h(n)$ to strings of length n.

Now if we let Z be the leftmost sequence such that $\widehat{\Phi}(Z) = Y$, then using the permuted truth-tables and Y, we can effectively recover Z, and hence $Y \geq_T Z$. (Note that this will not, in general, be a *tt*-reduction, because there may be no computable bound on the amount of Y needed to determine the first n bits of the leftmost sequence that $\widehat{\Phi}$ maps to Y.) $\square$

It is therefore natural to ask whether the reverse reduction, $Y \geq_T Z$, can also be strengthened to $Y \geq_{tt} Z$ or $Y \geq_{wtt} Z$. We will prove that this is not the case in general, not only for Demuth's original theorem, but also for the versions of Demuth theorem for computable randomness and Schnorr randomness proven in the previous section.

Although not strictly necessary, our analysis will make use of the so-called complex reals, which were defined by Kjos-Hanssen, Merkle, and Stephan [KHMS11]. To prove the failure of the *wtt*-version of Demuth's theorem, we will show that (i) if a real Y *wtt*-computes some $Z \in \mathsf{MLR}$, Y must be complex and (ii) there is an $X \in \mathsf{MLR}$ and a real $Y \leq_{tt} X$ such that Y is non-computable but not complex.

Complex reals were defined by Kjos-Hanssen et al. using plain (or prefix-free) Kolmogorov complexity. We shall define them using another version of Kolmogorov complexity, called *monotone complexity*, which for our purposes is slightly easier to handle.

Definition 4.13. $X \in 2^\omega$ is called *complex* if there is some computable order $h : \omega \to \omega$ such that

$$C(X \restriction n) \geq h(n)$$

for every $n \in \omega$.

Definition 4.14. A monotone machine is a function $M : 2^{<\omega} \to 2^{<\omega} \cup 2^\omega$ such that (i) $M(\sigma_1) \preceq M(\sigma_2)$ for all $\sigma_1 \preceq \sigma_2$ and (ii) the set of pairs of strings (σ, τ) with $\tau \preceq M(\sigma)$ is c.e. Fixing a universal monotone machine M, we define the

Km-complexity of $\tau \in 2^{<\omega}$ to be

$$Km(\tau) = \min\{|\sigma| : \tau \preceq M(\sigma)\downarrow\}.$$

A real X is said to be *Km-complex* if there is a computable order g such that $Km(x\restriction n) \geq g(n)$ for every n.

Proposition 4.15. *The following are equivalent for $X \in 2^\omega$:*

(i) X is Km-complex.

(ii) X is complex

The following Lemma is needed in the proof of Proposition 4.15.

Lemma 4.16. *$A \in 2^\omega$ is complex if and only if there is some sequence $\{\sigma_n\}_{n\in\omega} \leq_{wtt} A$ such that $C(\sigma_n) \geq n$ for every $n \in \omega$.*

Proof of Proposition 4.15. $(i) \Rightarrow (ii)$: First, we appeal to a standard fact about the relationship between Km, K, and C:

$$Km(\sigma) \leq K(\sigma) \leq 2C(\sigma).$$

Thus,

$$(\forall n)Km(X\restriction n) \geq h(n)$$

implies

$$(\forall n)C(X\restriction n) \geq \frac{h(n)}{2} \geq \left\lfloor \frac{h(n)}{2} \right\rfloor.$$

$(ii) \Rightarrow (i)$: Suppose that A is complex. Then by Lemma 4.16, there is some sequence $\{\sigma_n\}_{n\in\omega} \leq_{wtt} A$ such that $C(\sigma_n) \geq n$ for every $n \in \omega$. Let Φ be the wtt-functional witnessing this reduction, and let f be a computable bound on the use, so that $\Phi^{A\restriction f(n)}(n) = \sigma_n$. Then for every $n \in \omega$,

$$Km(A\restriction f(n)) \geq C(\sigma_n) - 2\log n - O(1) \geq n - 2\log n \geq \frac{n}{2} - O(1).$$

To see that the first inequality holds, consider the following procedure to produce σ_n. Suppose $\tau \in 2^{<\omega}$ is such that $M(\tau) \succeq A\restriction f(n)$, where M is the universal monotone machine. Then, from τ and a specification of n (in no more than $2\log(n) + O(1)$ bits), we can compute $\Phi^{A\restriction f(n)}(n) = \sigma_n$.

Now given $n \in \omega$, there is some k such that $n \in [f(k), f(k+1))$ (unless $f(k) = f(k+1)$, in which case $n = f(k)$). Then since $f^{-1}(n) = \min\{k : f(k) \geq n\}$, it follows that either $k = f^{-1}(n)$ (if $f(k) = n$) or $k+1 = f^{-1}(n)$ (if $f(k) \neq n$). Moreover, since Km is monotonic, $Km(A\restriction f(k)) \leq Km(A\restriction n)$ Then since

$$Km(A\restriction f(n)) \geq \frac{n}{2} - O(1)$$

for every $n \in \omega$, it follows that

$$Km(A\restriction n) \geq Km(A\restriction f(k)) \geq \frac{k}{2} - O(1) \geq \frac{f^{-1}(n) - 1}{2} - O(1).$$

Since $\dfrac{f^{-1}(n) - 1}{2} - O(1)$ is a computable order, it follows that A is Km-complex.

□

Henceforth, in light of this result, we will not distinguish between complex and Km-complex sequences.

For the following discussion, it is important to note that a version of the Levin-Schnorr theorem holds for monotone complexity.

Theorem 4.17. *$Z \in \mathsf{MLR}$ if and only if $Km(Z{\upharpoonright}n) = n - O(1)$.*

By setting $h(n) = n$, it immediately follows that Martin-Löf random sequences are complex. Furthermore, any sequence Y that wtt-computes a Martin-Löf random sequence is itself complex. This follows from the straightforward fact that complex reals are closed upwards in the wtt-degrees.

Lemma 4.18. *Given $A, B \in 2^\omega$ such that $A \geq_{wtt} B$. If B is complex then so is A.*

Proof. Let Φ be a wtt-functional such that $\Phi(A) = B$ with h a computable bound for the use of this reduction. Suppose further that B is complex as witnessed by the computable order g. Then

$$Km(A{\upharpoonright}n) \geq Km(B{\upharpoonright}h^{-1}(n)) \geq g(h^{-1}(n)).$$

Since $g \circ h^{-1}$ is a computable order, it follows that A is complex. $\square$

The key result needed to show the failure of the wtt-version of Demuth's Theorem is this:

Theorem 4.19. *There exists $A \in \mathsf{MLR}$ and a non-computable real $B \leq_{tt} A$ which is not complex.*

Proof. To prove this theorem, we take A to be Chaitin's Ω number, where $\Omega := \sum_{U(\sigma)\downarrow} 2^{-|\sigma|}$, U being a universal prefix-free machine. It is well-known that $\Omega \in \mathsf{MLR}$ and that Ω is a left-c.e. real, which means that there is a computable sequence of rationals $(\Omega_s)_s$ that converges to Ω from below. Note that this sequence must converge very slowly, i.e. there is no computable function f such that $\Omega\restriction n = \Omega_{f(n)}\restriction n$ infinitely often, for otherwise we would be able to compress the corresponding initial segments of Ω. We use the slowness of this approximation to build the sequence B. We achieve this through the following tt-reduction. Let $\Phi : 2^\omega \to 2^\omega$ be the functional defined for all sequences by

$$\Phi(X) = 1^{t_1}01^{t_2}01^{t_3}0\dots$$

where the t_i are defined as follows: $t_0 = 0$ and

$$t_i = \min\{s : \Omega_s \geq X\restriction i\}$$

(where we think of $X\restriction i$ as a rational number) with the convention that if the set on the right-hand side is empty, then $t_i = +\infty$. Thus if some t_i is infinite, then $\Phi(X) = 1^{t_1}0\dots1^{t_k}011111\dots$ where t_{k+1} is the first t_i to be infinite.

Φ is clearly total, and hence a tt-reduction. Moreover, if $A < \Omega$, then there is some s such that $\Omega_s > A\restriction i$ for every $i \in \omega$, and hence

$$\Phi(A) = \sigma^\frown(1^k0)^\omega$$

for some $\sigma \in 2^{<\omega}$ and $k \in \omega$. If $A > \Omega$, then there is some i such that $\Omega_s < A{\upharpoonright}i$ for every $s \in \omega$, and hence

$$\Phi(A) = \sigma 1^\omega$$

for some $\sigma \in 2^{<\omega}$. The interesting case is when $A = \Omega$, for in this case, setting

$$\Phi(\Omega) = 1^{s_1}01^{s_2}01^{s_3}0\ldots,$$

we know that the function f given by $f(i) = s_i$ grows faster than any computable function, since $\Omega{\upharpoonright}n = \Omega_{f(n)}{\upharpoonright}n$ for every $n \in \omega$. If we set $\Phi(\Omega) = \Omega^*$, then we have

$$Km(\Omega^*{\upharpoonright}f(n)) \leq n + O(1)$$

and hence

$$Km(\Omega^*{\upharpoonright}n) \leq f^{-1}(n) + O(1),$$

Since f grows faster than any computable order function, f^{-1} is dominated by all computable order functions. Thus, there is no computable order function g such that

$$Km(\Omega^*{\upharpoonright}n) \geq g(n).$$

$\square$

We can now prove that the *wtt*-version of Demuth's theorem fails for Martin-Löf randomness.

Corollary 4.20. *There exist $A \in \mathsf{MLR}$ and a non-computable real $B \leq_{tt} A$ such that* *there is no* $Y \in \mathsf{MLR}$ *with* $Y \leq_{wtt} B$.

Proof. By Theorem 4.19, Ω^* is tt-reducible to a Martin-Löf random sequence but is not complex. Hence by Lemma 4.18, Ω^* cannot wtt-compute any complex sequence, and thus Ω^* cannot wtt-compute any $X \in \mathsf{MLR}$. □

We would like to show that the wtt-version of Demuth's theorem also fails for computable randomness and Schnorr randomness. It seems that the real Ω^* constructed in the proof of Theorem 4.19 is so far from complex that it should not even wtt-compute a Schnorr random real. Unfortunately, we do not know whether this is the case.

Question 4.21. Is there $Y \in \mathsf{SR}$ such that $Y \leq_{wtt} \Omega^*$?

We therefore need to slightly adapt the technique used in the proof of Theorem 4.19, while keeping the main ideas. To prove that the wtt-version of Demuth theorem fails for both computable randomness and Schnorr randomness, we will prove the following (stronger) result.

Theorem 4.22. *For almost all reals* A*, there exists a non-computable real* $B \leq_{tt} A$ *that does not wtt-compute any Schnorr random real.*

This shows in particular that there exist a Martin-Löf random (hence computably random and Schnorr random) sequence A and a non-computable $B \leq_{tt} A$ such that B does not wtt-compute any Schnorr random real. Therefore the wtt-version of Demuth's theorem fails for all three notions of randomness.

To prove Theorem 4.22, we need a few auxiliary facts.

Lemma 4.23. *Let f be an increasing function that is not dominated by any computable function. Let g be a computable order function. Then for infinitely many n,*

$$f(n) < g(f(n+1)).$$

Proof. Suppose for the sake of contradiction that $f(n+1) \leq h(f(n))$ for all $n \geq k$, where $h = g^{-1}$. By induction, it follows that

$$f(n) \leq h^{(n-k)}(f(k))$$

for all $n \geq k$. Then if we set $\psi(n) = h^{(n-k)}(f(k))$, ψ is a non-decreasing, computable function that dominates f, contradicting our hypothesis. $\square$

Recall that $A \in 2^\omega$ has hyperimmune degree if A computes a function that is not dominated by any computable function.

Theorem 4.24. *Let A be a Martin-Löf random sequence of hyperimmune degree. Then there is a sequence $B \leq_{tt} A$ such that B is not complex.*

Proof. Since A is of hyperimmune degree, let $f \leq_T A$ be a function that is not dominated by any computable function, and let Ψ be the Turing reduction from A to f. For all n, define

$$g(n) = \min\{t : \Psi^A[t](n)\downarrow\}$$

By the standard conventions on oracle computations, it follows that $g(n) \geq f(n)$ for all n (as we require that the number of steps for a halting computation always exceeds the output of the computation). It follows that g is not dominated by any computable function. Now let Θ be the reduction defined by

$$\Theta(X) = 1^{t_0}01^{t_1}01^{t_2}\dots$$

where

$$t_n = \min\{t : \Psi^X[t](n)\downarrow\}$$

(with the convention that $\Theta(X) = 1^{t_0}01^{t_1}01\dots1^{t_{i-1}}01111111\dots$ if t_i is infinite and is the smallest such t_n). The definition ensures that Θ is total and that

$$B = \Theta(A) = 1^{g(0)}01^{g(1)}01^{g(2)}\dots$$

We need to show that B is not complex. Let h be a computable order. Notice that

$$\begin{aligned} Km(1^{g(0)}01^{g(1)}\dots01^{g(n)}01^{g(n+1)}) &\leq K(1^{g(0)}01^{g(1)}\dots01^{g(n)}0) + O(1) \\ &\leq n\log g(n) + O(1) \\ &\leq g(n)\log g(n) + O(1), \end{aligned}$$

where the first inequality follows from two facts: (i) $Km(\sigma\tau) \leq K(\sigma)+Km(\tau)+O(1)$ and (ii) $Km(1^k) = O(1)$ for all k.

By Lemma 4.23, applied to the composition of h and the function ϕ such that $\phi^{-1}(n) = n\log(n)$, we have for infinitely many n, $g(n)\log g(n) + k < h(g(n+1))$ for any fixed $k \in \omega$ (as $g(n)\log g(n) + k$ is not dominated by any computable function).

Thus for infinitely many n,

$$Km(B\restriction g(n+1)) \leq Km(1^{g(0)}01^{g(1)}\ldots 01^{g(n)}01^{g(n+1)}) < h(g(n+1)).$$

Since this is the case for any order h, it follows that B is not complex. □

We are now ready to prove Theorem 4.22.

Proof of Theorem 4.22. Let A be a Martin-Löf random real of hyperimmune but non-high degree. In particular, we can take A to be any 3-random sequence, since every 2-random real has hyperimmune degree, as proven by Kurtz [Kur81], and no 3-random real is high [Nie09, Exercise 8.5.21].

Then by Theorem 4.24, A *tt*-computes a real B that is not complex. Now suppose that $B \geq_{wtt} C$ for some $C \in 2^\omega$. Then since B is not complex, by Lemma 4.18, C is not complex. In particular, C not Martin-Löf random, as a Martin-Löf random real Z is s.t. $Km(Z\restriction n) = n - O(1)$. Moreover, C is not high, as $A \geq_T B \geq_T C$ and A is not high. Therefore by Theorem 2.59, if C is not Martin-Löf random and not high, then C cannot be Schnorr random. □

4.5 Some Applications

In this section, we apply the machinery developed in this chapter to study random Turing degrees and random computably enumerable sets.

4.5.1 Random Turing Degrees

One consequence of the machinery developed in the previous section is that we can use it to provide an exact characterization of all of the Martin-Löf random Turing degrees that contain a real that is random with respect to a computable measure but not random with respect to any computable *atomless* measure (where a Turing degree is Martin-Löf random if it contains a Martin-Löf random real). Let us establish a few more definitions that will be useful in this section.

Definition 4.25. Let $\mathsf{MLR}_{\text{comp}}$ be the set of sequences A such that $A \in \mathsf{MLR}_\mu$ for some computable measure μ.

The class $\mathsf{MLR}_{\text{comp}}$ was, to the best of our knowledge, first considered in [ZL70]. It was later studied in [SF77] and [MSU98].

Definition 4.26. Let $\mathsf{NCR}_{\text{comp}}$ be the set of sequences A such that $A \notin \mathsf{MLR}_\mu$ for any computable atomless measure μ.

The motivation behind the definition of $\mathsf{NCR}_{\text{comp}}$ comes from the work of Reimann and Slaman (see, for instance, [RS07] and [RS08]), who studied the collection of sequences that are not random with respect to *any* atomless measure (computable or otherwise), referring to this class as NCR_1. Although Reimann and Slaman have established a number of facts about NCR_1, for instance, that it is countable and contains no non-Δ^1_1 reals, a number of questions about the structure of NCR_1 remain open. $\mathsf{NCR}_{\text{comp}}$, in contrast, proves to be much easier to characterize.

We will begin by showing that there is at least one $X \in \mathsf{MLR}_{\text{comp}} \cap \mathsf{NCR}_{\text{comp}}$.

Proposition 4.27. *There is $X \in 2^\omega$ that is random with respect to some computable atomic measure but not random with respect to any computable atomless measure.*

To prove this proposition, we need one further result. By the Kautz conversion procedure (Theorem 4.4), if a computable measure μ is atomless and positive and Φ is an almost total functional such that $\lambda_\Phi = \mu$, then Φ^{-1} is an almost total functional such that $\mu_{\Phi^{-1}} = \lambda$. However, if Φ is total and $\lambda_\Phi = \mu$, it is not true in general that Φ^{-1} is total with $\mu_{\Phi^{-1}} = \lambda$. Still, we can find some other *tt*-functional Θ that induces a measure ν that, while not equal to λ, is equivalent to λ, in the sense that $\mathsf{MLR}_\nu = \mathsf{MLR}$.

Proposition 4.28. *If μ is a atomless, computable measure, then there is a non-decreasing tt-functional Θ such that the induced measure μ_Θ has the property that*

$$\mathsf{MLR}_{\mu_\Theta} = \mathsf{MLR}.$$

Proof. The idea behind the proof is to define a non-decreasing *tt*-functional Θ such that μ_Θ is a generalized Bernoulli measure, where this means that for every n, there is some $p_n \in [0,1]$ such that

$$p_n = \frac{\mu_\Theta(\sigma 0)}{\mu_\Theta(\sigma)}$$

for every $\sigma \in 2^{<\omega}$ of length n. Moreover, we will define Θ so that

$$\left| p_n - \frac{1}{2} \right|^2 \leq 2^{-|\sigma|}$$

for every $n \in \omega$. Lastly, we would like to define Θ in such a way that the resulting values p_n will always be contained in some fixed interval $[\epsilon, 1-\epsilon]$ for $\epsilon \in (0, \frac{1}{2})$; such measures are called *strongly positive*. Now by the effective version of Kakutani's Theorem (see, for instance, [BM09]), given two computable, strongly positive, generalized Bernoulli measures μ_1 (with associated values $p_1, p_2, \dots$) and μ_2 (with associated values $q_1, q_2, \dots$) such that

$$\sum_{i=1}^{\infty} |p_i - q_i|^2 < \infty,$$

it follows that $\mathsf{MLR}_{\mu_1} = \mathsf{MLR}_{\mu_2}$. Thus, if we can define such Θ satisfying the given conditions, then we will have

$$\sum_{i=1}^{\infty} \left| p_i - \frac{1}{2} \right|^2 < \infty,$$

and hence $\mathsf{MLR}_{\mu_\Theta} = \mathsf{MLR}$.

To define Θ, we sketch the main idea and leave the details to the reader. To define p_1, we look for a finite, prefix-free collection of strings $\{\sigma_1, \dots, \sigma_k\}$ such that

$$I_{\sigma_1} \cup \ldots \cup I_{\sigma_k} = [0, \frac{1}{2} - \epsilon_1],$$

for some $\epsilon_1 < \frac{1}{2}$, where I_σ is as defined in the proof of Theorem 4.4 (we can find such a collection effectively because μ is atomless). Then we define Θ so that extensions of each σ_i is mapped to extensions of 0 (and reals that extend none of the σ_i's are mapped to extensions of 1). Thus $p_1 = \sum_{i \leq k} \mu(\sigma_i)$.

Now we repeat this procedure, partitioning the intervals $[0, \frac{1}{2} - \epsilon_1]$ and $[\frac{1}{2} - \epsilon_1, 1]$

each into two intervals, each of which is determined by a finite, prefix-free collection of strings, just as we partitioned the interval [0,1] above, but we must make sure that ratios of the sizes of the components of each partition is the same, i.e. the left component of each is p_2 times the length of the given interval, where p_2 is within $\frac{1}{4}$ of $\frac{1}{2}$. In so doing, we will get four collections of strings, extensions of which will be mapped to extensions of 00,01,10,11 (depending on which of the four partitions the sequences are contained). Continuing this procedure, we will eventually define Θ with the desired properties. $\square$

Proof of Proposition 4.27. The real constructed in the proof of Theorem 4.19 above, Ω^*, is random with respect to the induced measure λ_Φ (which is clearly atomic), and hence $\Omega^* \in \mathsf{MLR}_{\text{comp}}$. Suppose, for sake of contradiction, that Ω is random with respect to a computable, atomless measure μ. Then by Proposition 4.28, there is a *tt*-functional Θ such that $\mathsf{MLR}_{\mu_\Theta} = \mathsf{MLR}$. Moreover, by the preservation of Martin-Löf randomness, it follows that $\Theta(\Omega^*) \in \mathsf{MLR}_{\mu_\Theta} = \mathsf{MLR}$. But as we proved in Theorem 4.19, Ω^* can't even *wtt*-compute any $Y \in \mathsf{MLR}$, yielding the desired contradiction. Thus $\Omega^* \in \mathsf{NCR}_{\text{comp}}$. $\square$

We can use the idea of this proof to provide a full classification of the Martin-Löf random Turing degrees containing elements in $\mathsf{MLR}_{\text{comp}} \cap \mathsf{NCR}_{\text{comp}}$. In providing the classification, we will use the following.

Proposition 4.29 ([RS08], Proposition 5.7)**.** *For $A \in \mathsf{MLR}$ and $B \in 2^\omega$, if $A \equiv_{tt} B$, then $B \notin \mathsf{NCR}_{\text{comp}}$.*

Theorem 4.30. *Let* $\mathbf{a}$ *be a Martin-Löf random Turing degree. Then there is some* $A \in \mathbf{a}$ *such that* $A \in \mathsf{MLR}_{\text{comp}} \cap \mathsf{NCR}_{\text{comp}}$ *if and only if* $\mathbf{a}$ *is hyperimmune.*

Proof. For the easier direction, suppose $\mathbf{a}$ is hyperimmune-free. Then given $A \in \mathbf{a} \cap \mathsf{MLR}$, by a well-known result about sets of hyperimmune-free degree, if $B \equiv_T A$, then $B \equiv_{tt} A$. Thus for any $B \equiv_T A$, by the preservation of randomness, we have $B \in \mathsf{MLR}_{\text{comp}}$. By Proposition 4.29, $B \equiv_{tt} A$ implies that $B \notin \mathsf{NCR}_{\text{comp}}$; i.e., B is random with respect to some atomless measure. Thus, no $B \in \mathbf{a}$ is in $\mathsf{MLR}_{\text{comp}} \cap \mathsf{NCR}_{\text{comp}}$.

Now suppose that $\mathbf{a}$ is hyperimmune, and let $A \in \mathbf{a} \cap \mathsf{MLR}$. We proceed as in the proof of Theorem 4.24, with a slight modification. Let $f \in \mathbf{a}$ be a function that is not dominated by any computable function. Then there is some Turing functional Ψ such that $\Psi^A(n) = f(n)$ for every n. Then, as in the proof of Theorem 4.24, we define a functional Γ such that

$$\Gamma(C) = 1^{t_0}\, 0^{C(0)+1}\, 1^{t_1}\, 0^{C(1)+1}\, 1^{t_2}\, 0^{C(2)+1} \ldots,$$

where t_i is the least t such that $\Psi^C(i)[t]\downarrow$, unless no such t exists, in which case $t_i = +\infty$. We code the sequence C into $\Gamma(C)$ so that if the $(i+1)$st block of 0s in $\Gamma(C)$ has length 1, then $C(i) = 0$, and if the $(i+1)$st block of 0s in $\Gamma(C)$ has length 2, then $C(i) = 1$. Thus, we have $\Gamma(A) \equiv_T A$. Further, by the preservation of randomness, we have $\Gamma(A) \in \mathsf{MLR}_{\text{comp}}$. Now let $g : \omega \to \omega$ be the function such that

$$\Gamma(A) = 1^{g(0)}\, 0^{A(0)+1}\, 1^{g(1)}\, 0^{A(1)+1}\, 1^{g(2)}\, 0^{A(2)+1} \ldots$$

Given the convention that for the least t such that $\Psi^C(i)[t]\downarrow = k$, we have $k \leq t$, it follows that $f(n) \leq g(n)$, and hence $g(n)$ is not dominated by any computable function.

We verify $\Gamma(A)$ is not complex as before, with the only difference being that we now have to consider the potentially doubled 0s, yielding

$$Km(1^{g(0)}\, 0^{A(0)+1}\, 1^{g(1)}\, 0^{A(1)+1} \ldots 1^{g(n)}\, 0^{A(n)+1} 1^{g(n+1)}) \leq 2n \cdot \log g(n) \leq g(n) \log g(n).$$

All the other steps proceed as before, and thus $\Gamma(A)$ is not complex. Assuming that $\Gamma(A)$ is random with respect to some atomless measure, we can argue as in the proof of Proposition 4.27 that $\Gamma(A)$ must tt-compute a Martin-Löf random sequence. This contradicts the fact that $\Gamma(A)$ is not complex. Thus, $\Gamma(A) \in \mathsf{NCR}_{\mathrm{comp}}$. □

Every hyperimmune degree contains a weakly 1-generic real, and no weakly 1-generic real is Martin-Löf random with respect to *any* computable measure (as is shown in [MSU98], Theorem 9.10). Therefore, we have an even stronger dichotomy: Every hyperimmune-free random degree contains only reals that are random with respect to some computable atomless measure, while every hyperimmune random degree contains reals that are random only with respect to some computable atomic measure as well as reals that aren't random with respect to *any* computable measure.

4.5.2 Random Computably Enumerable Sets

In this last subsection, we will show that the preservation of randomness and related results also have consequences for the study of random computably enumerable sets. In particular, we show the existence of a computably enumerable set that is random with respect to some computable measure. This is somewhat surprising, given that computably enumerable sets are far from Martin-Löf random. For instance, every c.e. set X has low initial segment complexity: for every n, $K(X{\upharpoonright}n) \leq 2\log(n)+O(1)$. Despite this behavior, there are c.e. sets that are Martin-Löf random with respect to *some* computable measure, as we now demonstrate. This result was obtained independently by Reimann and Slaman.

Theorem 4.31. *There exists a non-computable c.e. set C and some $\mu \in \mathscr{M}_c$ such that C is random with respect to μ.*

Proof. Let $(q_n)_{n\in\omega}$ be an effective enumeration of $\mathbb{Q}_2$. Let $\Xi : 2^\omega \to 2^\omega$ be the map defined by

$$\Xi(X) = \{n \mid q_n < X\}$$

where we see the input as an infinite binary sequence and the output as a set of integers. Clearly Ξ is a computable one-to-one map, hence the measure μ it induces on 2^ω is computable and atomless, and for every random X, $\Xi(X)$ is μ-random. If X is left-c.e., then by definition of Ξ, $\Xi(X)$ is a c.e. set. Therefore, $C = \Xi(\Omega)$ is both c.e. and μ-random. $\square$

We can also show that there is a non-computable c.e. member of $\mathsf{MLR}_{\text{comp}} \cap \mathsf{NCR}_{\text{comp}}$.

Theorem 4.32. *There is a non-computable c.e. set C such that $C \in \mathsf{MLR}_\mu$ for some computable atomic measure μ but $C \notin \mathsf{MLR}_\nu$ for any computable atomless measure ν.*

Proof. To prove this result, we merely need to show that Ω^*, the real constructed in the proof of Theorem 4.19, is left-c.e. and then apply the map Ξ defined above to produce a c.e. set C that is tt-reducible to Ω (via the composition of Ξ with the functional Φ defined in the proof of Theorem 4.19). Then C cannot be random with respect to any atomless computable measure, for as we argued in the proof of Proposition 4.27, this would mean that C, and hence Ω^*, can tt-compute a Martin-Löf random sequence.

To see that Ω^* is left-c.e., notice that Φ is a non-decreasing functional and Φ is continuous at Ω. Therefore, for a rational q, we have

$$q < \Omega^* \Leftrightarrow \exists X \ [X < \Omega \wedge \Phi(X) > q]$$

The right-hand side of the equivalence is a Σ^0_1 predicate. Therefore, the left cut of Ω^* is c.e., which means that Ω^* is left-c.e.

□

A few remarks are in order. First, if a non-computable c.e. set is Martin-Löf random with respect to a computable probability measure, then it must be Turing complete. Indeed, by Demuth's theorem, such a real must be Turing equivalent to a real that is Martin-Löf random for Lebesgue measure. Kučera [Kuč85] proved that a c.e. set that can compute a Martin-Löf random real must be Turing complete.

The family of c.e. sets that are random for some computable probability measure is not downwards closed in the Turing degrees. However, this family is closed downwards in the tt-degrees, by Demuth's theorem. Given a tt-functional Φ and a c.e. set C that is random with respect to a computable measure μ, if $\Phi(C)$ is c.e., then it is either computable or Turing complete, and in both cases, it will be random with respect to the measure induced by (μ, Φ). It is natural to ask whether the family of c.e. random sets forms a *tt*-ideal. The answer is negative.

Proposition 4.33. *The c.e. random sets do not form a tt-ideal.*

Proof. Let $A \in 2^\omega$ be a left-c.e., Turing incomplete sequence and $X_1 \in 2^\omega$ a left-c.e. random sequence. Set $X_2 = X_1 + A$ and notice that X_2 is left-c.e. and random as the sum of a random left-c.e. real and a left-c.e. real [DH10, Chapter 8].

We map X_1 and X_2 to c.e. sets via Ξ: $Y_1 = \Xi(X_1)$ and $Y_2 = \Xi(X_2)$. Then both Y_1 and Y_2 are c.e. and random with respect to the measure μ induced by Ξ.

Since Ξ is a total computable map, its range is a Π^0_1 class, which we will denote by $\mathcal{D}$. Since Ξ is one-to-one, we can define the functional Ξ^{-1}, which is actually a Turing functional with domain $\mathcal{D}$. Since for all $Z \in \mathcal{D}$, the set $\{X : \Xi(X) = Z\}$ is a $\Pi^0_1(Z)$ class containing only one element, that element X can be found computably in Z. Note that a partial functional defined on a Π^0_1 class can be extended to a total functional, since we can effectively determine those σ on which the partial functional is not defined. Let Λ be a *tt*-functional which is an extension of Ξ^{-1} to the entire space 2^ω.

Now, suppose that the join $Y_1 \oplus Y_2$ is random with respect to some computable measure ν. Consider the functional Ψ defined by $\Psi(Z_1 \oplus Z_2) = |\Lambda(Z_1) - \Lambda(Z_2)|$. This is a *tt*-functional, and $\Psi(Y_1 \oplus Y_2) = A$. By the preservation of Martin-Löf randomness, this means that $A \in \mathsf{MLR}_{\nu_\Psi}$. This contradicts the fact that an incomplete left-c.e. sequence cannot be Martin-Löf random with respect to any computable measure.

□

CHAPTER 5

TALLY FUNCTIONALS, TRIVAL MEASURES, AND DIMINUTIVE MEASURES

5.1 Introduction

In this chapter, we study two classes of computable measures, diminutive computable measures and trivial computable measures. We've already encountered examples of each kind of measure in the previous chapter: the measure induced by the functional used in the proof of Theorem 4.19 is a trivial measure, while the measures induced by the functionals used in the proofs of Theorem 4.24 and Theorem 4.30 are diminutive measures. Before we look more closely at the classes of measures, we will first consider some general features of the functionals used in the proofs of Theorem 4.19, Theorem 4.24, and Theorem 4.30, which we will henceforth refer to as *tally functionals.*

5.2 Tally Functionals

The main tool used in the construction of various diminutive and trivial measures are what are called "tally functionals". Let $\Theta(X, y, z)$ be a formula in the language of second-order arithmetic with no set quantifiers and no number quantifiers (but possibly bounded number quantifiers) with free first-order variables y, z and free second-order variable X. Then we define an auxiliary function $\theta(A, n) : 2^\omega \times \omega \to \omega$ by

$$\theta(A,n) = \begin{cases} \text{the least } s \text{ such that } \Theta(A,n,s) & \text{if } s \text{ exists} \\ +\infty & \text{otherwise.} \end{cases}$$

Clearly, the function $\theta(A, \cdot) : \omega \to \omega$ is A-computable. Then the tally functional Φ_Θ determined by the formula Θ is defined to be

$$\Phi_\Theta(A) = 1^{\theta(A,0)}\ 0\ 1^{\theta(A,1)}\ 0\ 1^{\theta(A,2)}\ 0 \ldots,$$

where

$$\Phi_\Theta(A) = 1^{\theta(A,0)}\ 0\ 1^{\theta(A,1)}\ 0 \ldots\ 1^{\theta(A,k)}\ 0\ 1^\omega$$

if $\theta(A, k+1) = +\infty$ (and $k+1$ is the least number such that $\theta(A, k+1) = +\infty$).

As stated in the introduction, we've already encountered several examples of tally functionals.

Example 5.1. Let $\Theta_1(X, n, s)$ be the formula

$$X{\upharpoonright}n \leq \Omega_s,$$

where $\{\Omega_s\}_{s\in\omega}$ is a computable, non-decreasing sequence of rationals converging to

Ω. Then $\theta_1(X,n)$ is defined to be

$$\theta_1(X,n) = \begin{cases} \text{the least } s \text{ such that } X{\restriction}n \leq \Omega_s & \text{if } s \text{ exists} \\ +\infty & \text{otherwise} \end{cases}.$$

Then the tally functional Φ_{Θ_1} is precisely the functional used in the proof of Theorem 4.19.

Example 5.2. Let $\Theta_2(X,n,s)$ be the formula

$$\Psi^X(n)[s]\downarrow,$$

where Ψ is a Turing functional such that

$$\{A : \Psi(A) \text{ is total and is not dominated by any computable function}\}$$

has positive measure, which exists by the proof of Miller and Martin [MM68] that every set of hyperimmune degree computes a function not dominated by any computable function. In this case $\theta_2(X,n)$ is defined to be

$$\theta_2(X,n) = \begin{cases} \text{the least } s \text{ such that } \Psi^X(n)[s]\downarrow & \text{if } s \text{ exists} \\ +\infty & \text{otherwise} \end{cases}.$$

Then Φ_{Θ_2} is defined like the tally functional used in the proof of Theorem 4.24.

Another useful example that we have yet to discuss is the tally functional in terms of the approximation of a Δ^0_2 Martin-Löf random sequence A.

Example 5.3. Given $A \in \Delta_2^0 \cap \mathsf{MLR}$, there is a computable sequence of finite sets $\{A_s\}_{s\in\omega}$ such that

$$\lim_{s\to\infty} A_s(n) = A(n)$$

for every n. Without loss of generality, we can assume that $A_s \neq A_{s+1}$ for every s. Now let $\Theta_3(X, n, s)$ be the formula

$$X{\upharpoonright}n = A_s{\upharpoonright}n,$$

so that

$$\theta_3(X, n) = \begin{cases} \text{the least } s \text{ such that } X{\upharpoonright}n = A_s{\upharpoonright}n & \text{if } s \text{ exists} \\ +\infty & \text{otherwise} \end{cases}.$$

As the tally functional Φ_{Θ_3} defined in this way for $A \in \Delta_2^0$ will feature prominently in our later discussion, let us call Φ_{Θ_3} the *tally functional given by* A, denoted Φ_A.

5.3 Trivial and Diminutive Measures

The measures induced by various tally functionals are of central interest here.

5.3.1 On Trivial Measures

Recall that $\mu \in 2^\omega$ is trivial if $\mu(\mathsf{Atoms}_\mu) = 1$.

Proposition 5.4. *The measure μ_1 induced by the tally functional Φ_{Θ_1} is trivial.*

Proof. To see that μ_1 is trivial, note that from the proof of Theorem 4.19 it follows that if $A \neq \Omega$, then either $\Phi_{\Theta_1}(A) = \sigma^\frown(1^k0)^\omega$ for some $\sigma \in 2^{<\omega}$ and $k \in \omega$ (in the case that $A < \Omega$ as real numbers) or $\Phi_{\Theta_1}(A) = \sigma 1^\omega$ for some $\sigma \in 2^{<\omega}$ (in the case that $A > \Omega$ as real numbers). Thus Φ_{Θ_1} maps $2^\omega \setminus \{\Omega\}$ into a subset of

$$\mathcal{S} = \{Y \in 2^\omega : (\exists \sigma \in 2^{<\omega})(\exists k)[Y = \sigma^\frown(1^k0)^\omega]\} \cup \{Y \in 2^\omega : (\exists \sigma \in 2^{<\omega})[Y = \sigma 1^\omega]\}.$$

In particular, by the preservation of Martin-Löf randomness,

$$\Phi_{\Theta_1}(\mathsf{MLR} \setminus \{\Omega\}) \subseteq \mathsf{MLR}_{\mu_1} \cap \mathcal{S},$$

from which it follows that $\lambda(\Phi_{\Theta_1}^{-1}(\mathcal{S})) = 1$. Since $\mathcal{S}$ consists entirely of computable sequences, $\mathcal{S} \cap \mathsf{ran}(\Phi_{\Theta_1}) = \mathsf{Atoms}_{\mu_1}$. Therefore, $\mu_1(\mathsf{Atoms}_{\mu_1}) = 1$. □

Despite the fact that μ_1 is trivial, not every $X \in \mathsf{MLR}_{\mu_1}$ is a μ_1-atom.

Proposition 5.5. $\mathsf{MLR}_{\mu_1} \setminus \mathsf{Atoms}_{\mu_1} \neq \emptyset$.

Proof. From the proof of Theorem 4.19, we see that

$$\Phi_{\Theta_1}(\Omega) = 1^{f(1)}01^{f(2)}01^{f(3)}0\ldots,$$

where $\Omega_{f(n)}{\upharpoonright}n = \Omega{\upharpoonright}n$ for every n. Thus, $\Phi_{\Theta_1}(\Omega)$ is not computable, but by the preservation of Martin-Löf randomness, $\Phi_{\Theta_1}(\Omega) \in \mathsf{MLR}_{\mu_1}$. Therefore,

$$\Phi_{\Theta_1}(\Omega) \;\in\; \mathsf{MLR}_{\mu_1} \setminus \mathsf{Atoms}_{\mu_1}.$$

□

We postpone the description of the tally functional Φ_{Θ_2} until the next section, and instead consider Φ_{Θ_3}.

Theorem 5.6. *The measure μ_3 induced by the tally functional Φ_{Θ_3} is trivial, and yet $\mathsf{MLR}_{\mu_3} \setminus \mathsf{Atoms}_{\mu_3} \neq \emptyset$. In fact, $\mathsf{MLR}_{\mu_3} = \{\Phi_{\Theta_3}(A)\} \cup \mathsf{Atoms}_{\mu_3}$.*

To prove Theorem 5.6, we first need a lemma.

Lemma 5.7. *(i) If $X = A_s$ for some s, then there is some m such that for every $n \geq m$, $\theta_3(X, n) = \theta_3(X, m)$, i.e. the function $f(x) = \theta_3(X, x)$ is eventually constant.*

(ii) If $X \neq A$ and $X \neq A_s$ for each s, then $\theta_3(X, n) = +\infty$ for some n.

(iii) The function $g(x) = \theta_3(A, x)$ is not computable.

Proof. (i) Let s be least such that $X = A_s$. If $s = 0$, then since for all n, $X{\upharpoonright}n = A_0{\upharpoonright}n$, it follows that $\theta_3(X, n) = 0$ for all n. Now suppose that $s > 0$. There is some m such that $X{\upharpoonright}m \neq A_{s-1}{\upharpoonright}m$. It follows that $X{\upharpoonright}n = A_s{\upharpoonright}n$ for all $n \geq m$, which means that $\theta_3(X, n) = s$ for all $n \geq m$.

(ii) First we establish the following claim: There is some k such that for every s, $X{\upharpoonright}k \neq A_s{\upharpoonright}k$. Suppose not, so that for every k, there is some s such that $X{\upharpoonright}k = A_s{\upharpoonright}k$. But this implies that for every k, there exist infinitely many s such that $X{\upharpoonright}k = A_s{\upharpoonright}k$, since (a) $X \neq A_s$ for every s and (b) $X{\upharpoonright}k = A_s{\upharpoonright}k$ implies that $X{\upharpoonright}j = A_s{\upharpoonright}j$ for every $j < k$. Now for each k, there is some s_k such that $A{\upharpoonright}k = A_t{\upharpoonright}k$ for every $t \geq s_k$. It follows that there is some $t \geq s_k$ such that $X{\upharpoonright}k = A_t{\upharpoonright}k$, and hence $X{\upharpoonright}k = A{\upharpoonright}k$.

Since this holds for every k, it follows that $X = A$, contradicting our hypothesis. Let k be least such $X{\restriction}k \neq A_s{\restriction}k$ for every s. Then $\theta_3(X, k) = +\infty$.

(iii) If $g(n) = \theta_3(A, n)$ were computable, then since

$$A_{g(n)}{\restriction}n = A_{\theta(A,n)}{\restriction}n = A{\restriction}n,$$

this would imply that A is computable, contradicting our assumption that $A \in \mathsf{MLR}$. $\square$

Proof of Theorem 5.6. First we show that μ_3 is trivial. Given input X, here are three cases to consider to determine the output $\Phi_{\Theta_3}(X)$:

Case 1: $X = A_s$ for some s. In this case, by Lemma 5.7 (i), the function $f(x) = \theta(X, x)$ is eventually constant. Therefore, $\Phi_\Theta(A) = \sigma 1^k 0 1^k 0 1^k 0 \ldots$ for some $\sigma \in 2^{<\omega}$, and hence $\Phi_\Theta(X)$ is computable.

Case 2: $X \neq A$ and $X \neq A_s$ for every s. By Lemma 5.7 (ii), there is some n such that $\theta(X, n) = +\infty$. Then $\Phi_\Theta(A) = \sigma 1^\omega$ for some $\sigma \in 2^{<\omega}$, and hence, as above, $\Phi_\Theta(X)$ is computable.

Case 3: $X = A$. By Lemma 5.7 (iii), since $g(n) = \theta_3(A, n)$ is not computable, it follows that

$$\Phi_{\Theta_3}(A) = 1^{\theta_3(A,0)}\, 0\, 1^{\theta_3(A,1)}\, 0\, 1^{\theta_3(A,2)}\, 0 \ldots,$$

is not a computable sequence.

It follows that for every $B \in \mathsf{MLR}$ such that $B \neq A$, $\Phi_{\Theta_3}(B) = \sigma 1^\omega$ for some $\sigma \in 2^{<\omega}$ (since $B \neq A_s$ for every s, as each A_s is finite and hence computable).

Setting

$$\mathcal{S} = \{Y : (\exists \sigma \in 2^{<\omega})[Y = \sigma 1^{\omega}]\},$$

we have

$$\mathsf{MLR} \setminus \{A\} \subseteq \Phi_{\Theta_3}^{-1}(\mathcal{S}).$$

From this it follows that

$$1 = \lambda(\mathsf{MLR} \setminus \{A\}) \leq \lambda(\Phi_{\Theta}^{-1}(\mathcal{S})).$$

Since λ_{Φ_Θ} assigns measure one to the countable collection $\mathcal{S}$, λ_{Φ_Θ} is trivial.

To verify that $\mathsf{MLR}_{\mu_3} = \{\Phi_{\Theta_3}(A)\} \cup \mathsf{Atoms}_{\mu_3}$, first observe that by Theorem 4.5,

$$\mathsf{MLR}_{\mu_3} = \Phi_{\Theta_3}(\mathsf{MLR}).$$

As shown in Case 2 above,

$$\Phi_{\Theta_3}(\mathsf{MLR} \setminus \{A\}) \subseteq \{\sigma 1^{\omega} : \sigma \in 2^{<\omega}\}.$$

By the preservation of randomness, it follows that $\Phi_{\Theta_3}(\mathsf{MLR} \setminus \{A\}) \subseteq \mathsf{Atoms}_{\mu_3}$ (and in fact that $\Phi_{\Theta_3}(\mathsf{MLR} \setminus \{A\}) = \mathsf{Atoms}_{\mu_3}$). However, $\Phi_{\Theta_3}(A) \in \mathsf{MLR}_{\mu_3} \setminus \mathsf{Atoms}_{\mu_3}$ by the preservation of Martin-Löf randomness and the fact that $\Phi_{\Theta_3}(A)$ is not a μ-atom, as shown in Case 3 above. It follows that

$$\mathsf{MLR}_{\mu_3} = \Phi_{\Theta_3}(\mathsf{MLR}) = \{\Phi_{\Theta_3}(A)\} \cup \mathsf{Atoms}_{\mu_3}.$$

□

Not every measure induced by a tally functional is trivial.

Proposition 5.8. *The measure μ_2 induced by the tally functional Φ_{Θ_2} is not trivial.*

Proof. The functional Ψ, in terms of which we defined the tally functional Φ_{Θ_2}, has the property that

$$\mathcal{S} = \{A : \Psi(A) \text{ is total and is not dominated by any computable function}\}$$

has positive measure. In particular, one can show that $\lambda(\mathcal{S} \cap \mathsf{2MLR}) > 0$. If $A \in \mathcal{S} \cap \mathsf{2MLR}$, then

$$\Phi_{\Theta_2}(A) = 1^{g(0)}01^{g(1)}01^{g(2)}\dots$$

where g is not dominated by any computable function (since for each $n \in \omega$, $g(n)$ is the least such that $\Psi(A)(n)[g(n)]\downarrow$ and thus $\Psi(A)(n) \leq g(n)$). Then $\Phi_{\Theta_2}(A)$ is not computable. It follows that

$$\Phi_{\Theta_2}(\mathcal{S} \cap \mathsf{2MLR}) \cap \mathsf{Atoms}_{\mu_2} = \emptyset.$$

Thus, $\mu_2(\mathsf{Atoms}_{\mu_2}) < 1$.

□

5.3.2 On Diminutive Measures

We now consider the collection of "diminutive" computable measures. These are measures that are defined in terms of "computably perfect" classes.

Given $\mathcal{A} \in 2^\omega$, the collection of extendible nodes of $\mathcal{A}$ is defined to be

$$\mathsf{Ext}(\mathcal{A}) = \{\sigma \in 2^{<\omega} : (\exists X \in \mathcal{A})[\sigma \prec X]\}$$

Definition 5.9. $\mathcal{A} \subseteq 2^\omega$ is *computably perfect* if there is a computable, strictly increasing function f such that for every $n \in \omega$ and $\sigma \in \mathsf{Ext}(\mathcal{A})$ such that $|\sigma| = f(n)$, there exist $\tau_0, \tau_1 \succ \sigma$ such that $\tau_0, \tau_1 \in \mathsf{Ext}(\mathcal{A})$ and $|\tau_0| = |\tau_1| = f(n+1)$.

For $\mu \in \mathscr{M}_c$, let $\widehat{\mathcal{P}}_i$ be the complement of the ith member of the universal μ-Martin-Löf test.

Definition 5.10. A measure $\mu \in \mathscr{M}_c$ is *diminutive* if for each $i \in \omega$, $\widehat{\mathcal{P}}_i$ does not contain a computably perfect subclass.

This notion of a diminutive measure can be seen as an extension of the notion of a diminutive class, first studied by Binns in [Bin03].

Definition 5.11 ([Bin03])**.** $\mathcal{A} \subseteq 2^\omega$ is *diminutive* if $\mathcal{A}$ does not contain a computably perfect subclass.

The next lemma says that we could give an alternate definition, saying that μ is diminutive if no $\widehat{\mathcal{P}}_i$ contains a computably perfect Π^0_1 subclass.

Lemma 5.12 ([Bin08])**.** *Let $\mathcal{P} \subseteq 2^\omega$ be a Π^0_1 class. If $\mathcal{P}$ does not contain a computably perfect Π^0_1 subclass, then $\mathcal{P}$ is diminutive.*

Proof. If $\mathcal{P}$ is not diminutive, then $\mathcal{P}$ contains a computably perfect subclass $\mathcal{A}$. Let $\mathcal{P} = \bigcap_{s\in\omega} \mathcal{P}_s$ and let f be the function witnessing the fact that $\mathcal{A}$ is computably perfect. We define

$$S = \{\sigma : (\exists s)(\exists n \leq s)[\sigma \in \mathsf{Ext}(\mathcal{P}_s) \wedge |\sigma| = f(n) \wedge (\exists! \tau \in \mathsf{Ext}(\mathcal{P}_s))[\tau \succ \sigma \wedge |\tau| = f(n+1)]]\}$$

Then $[\![S]\!]$ is Σ^0_1, and hence $\mathcal{P} \setminus [\![S]\!]$ is a computably perfect Π^0_1 class containing $\mathcal{A}$. $\square$

Definition 5.13. $\mathcal{A} \subseteq 2^\omega$ is a *wtt*-cover of 2^ω if for every $X \in 2^\omega$, there is some $A \in \mathcal{A}$ such that $X \leq_{wtt} A$.

The key result about diminutive Π^0_1 classes is the following.

Theorem 5.14 (Binns, [Bin08])**.** *Let $\mathcal{P}$ be a Π^0_1 class. Then the following are equivalent:*

(1) $\mathcal{P}$ is diminutive;

(2) $\mathcal{P}$ contains no complex element;

(3) $\mathcal{P}$ does not contain a wtt-cover of 2^ω.

The equivalence of (1) and (2) is noteworthy, for it allows us to conclude that the measures we considered at the beginning of this section are diminutive. Note that this result also generalizes the Kučera-Gács Theorem, Theorem 2.88. We can thus characterize diminutive measures as those measures that yield a notion of randomness for which the Kučera-Gács Theorem fails.

Corollary 5.15. *The measures $\lambda_{\Theta_1}, \lambda_{\Theta_2}$, and λ_{Θ_3} are diminutive.*

Proof. This follows immediately from Theorem 5.14 and the fact that for $i = 1, 2, 3$, if $X \in \mathsf{MLR}_{\lambda_{\Theta_i}}$, then X is not complex. $\square$

5.3.3 Some Questions

Note that since λ_{Θ_2} is diminutive but not trivial, it follows that not every diminutive computable measure is trivial. While we have seen two examples of trivial computable measures that are also diminutive, it is not clear that every trivial computable measure is diminutive.

Question 5.16. Is every trivial computable measure diminutive?

There are a number of other related questions that might help answer Question 5.16.

Question 5.17. If $\mu \in \mathscr{M}_c$ is trivial, does it follow that no $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$ is complex?

Equivalently, one can ask:

Question 5.18. If $\mu \in \mathscr{M}_c$ is trivial, is there some $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$ such that $X \in \mathsf{MLR}_\nu$ for an atomless $\nu \in \mathscr{M}_c$?

Note that if $\mu \in \mathscr{M}_c$ is trivial and $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$, then μ-atoms must be dense along X. Otherwise, there is some n such that $[\![X{\upharpoonright}n]\!] \cap \mathsf{Atoms}_\mu = \emptyset$, and hence $\mu(X{\upharpoonright}n) = 0$. Thus, if Φ is the Turing functional that induces μ, and $Z \in \mathsf{MLR}$ is such that $\Phi(Z) = X$, then there are infinitely many n such that $\Phi^{-1}([\![X{\upharpoonright}n]\!])$ contains a Π^0_2 class that Φ maps to some $A \in \mathsf{Atoms}_\mu$ (since $\Phi^{-1}(A)$ is $\Pi^0_2(A)$ and hence Π^0_2, as A is computable). This suggests that sequences like X might only occur in some "nicely" definable Turing degrees.

Question 5.19. Which Turing degrees contain some $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$ for some trivial $\mu \in \mathscr{M}_c$?

5.4 Separating Classes of Non-Uniform Randomness

In this section, we show how one can separate notions of randomness with respect to trivial measures. As a warm-up, we consider the case of weak 3-randomness.

Proposition 5.20. *If $\mu \in \mathscr{M}_c$ is trivial and $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$, then $X \notin \mathsf{W3R}_\mu$. Hence $\mathsf{W3R}_\mu = \mathsf{Atoms}_\mu$.*

Proof. Note that MLR_μ is a Σ^0_1 class. Moreover, since

$$\mathsf{Atoms}_\mu = \{X \in 2^\omega : (\exists \epsilon \in \mathbb{Q})(\forall n)\mu(X{\restriction}n) > \epsilon\},$$

we see that Atoms_μ is a Σ^0_3 class (as the predicate $\mu(\sigma) > 0$ is Σ^0_1). Thus $\mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$ is the intersection of a Σ^0_2 class and a Π^0_3 class, and hence is Π^0_3. But since $\mu(\mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu) = 0$, it follows that it contains no weakly 3-random with respect to μ. $\square$

This result can be improved for some computable measures.

Theorem 5.21. *There is a trivial $\mu \in \mathscr{M}_c$ such that (i) $\mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu \neq \emptyset$ and (ii) $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$ implies that $X \notin \mathsf{W2R}_\mu$. Hence $\mathsf{W2R}_\mu = \mathsf{Atoms}_\mu$.*

Proof. Let $A \in \mathsf{MLR} \setminus \mathsf{W2R}$. Then there is some Turing functional Φ and a non-computable Δ^0_2 sequence B such that $\Phi(A) = B$. Let $\Theta(X, n, s)$ be the formula

$$(\forall k < n)\Phi_s(X)(k){\downarrow} = B_s(k)$$

where $\{B_s\}_{s\in\omega}$ is the Δ^0_2 approximation of B. Then θ is defined to be

$$\theta(X,n) = \begin{cases} \text{the least } s \text{ such that } (\forall k < n)\Phi_s(X)(k){\downarrow} = B_s(k) & \text{if } s \text{ exists} \\ +\infty & \text{otherwise} \end{cases}.$$

Now, it follows that

$$\Phi_\Theta(A) = 1^{h(0)}01^{h(1)}0\dots$$

where $h(n) := \theta(A,n)$. We claim that h is not dominated by any computable function. Otherwise, if f is computable such that $f(n) \geq h(n)$ for every n, then $\Phi_{f(n)}(A)(k){\downarrow} = B_{f(n)}(k)$ for every $k < n$. Since $\Phi_{f(n)}(A)(k) = \Phi(A)(k) = B(k)$, it follows that $B_{f(n)}\restriction n = B\restriction n$ for every n, contradicting the assumption that B is not computable. If μ is the measure induced by Φ_Θ, then $\Phi_\Theta(A) \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$. Moreover, we have $\Phi_\Theta(A) \notin \mathsf{W2R}$, since $\Phi_\Theta(A) \geq_T B$ and thus $\Phi_\Theta(A)$ does not form a minimal pair with $\emptyset'$.

Next, consider $X \in \mathsf{W2R}$. If $\Phi(X){\uparrow}$, then there is some least n such that $\Phi(X)(n){\uparrow}$, and so $\theta(X,n) = +\infty$. Consequently,

$$\Phi_\Theta(X) = \sigma 1^\omega \tag{5.1}$$

for some $\sigma \in 2^{<\omega}$. Otherwise, since $\Phi(X) \neq B$, we claim that

$$(\exists n)(\forall s)(\exists k < n)[\Phi_s(X)(k){\downarrow} \Rightarrow \Phi_s(X)(k) \neq B_s(k)]. \tag{5.2}$$

Suppose not. Then

$$(\forall n)(\exists s)(\forall k < n)[\Phi_s(X)(k){\downarrow} = B_s(k)]. \tag{5.3}$$

Now, let k be such that $\Phi(X)(k) \neq B(k)$, and let s be such that $\Phi_s(X)(k){\downarrow}$ and $B_t{\upharpoonright}(k+1) = B{\upharpoonright}(k+1)$ for all $t \geq s$. Then for sufficiently large n, for every $s > n$,

$$\Phi_s(X)(k){\downarrow} \neq B_s(k),$$

contradicting Equation 5.3. Then let n be the least satisfying Equation 5.2, and so we have $\theta(X, n) = +\infty$, and thus Equation 5.1 holds.

Lastly, suppose that $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$. Then by Theorem 4.5, $X \in \mathsf{MLR}_\mu$ implies that there is some $Y \in 2^\omega$ such that $\Phi_\Theta(Y) = X$. Further, since X is not computable, it follows that

$$\Phi_\Theta(Y) = 1^{\theta(Y,0)}01^{\theta(Y,1)}\dots,$$

where $\theta(Y, n) < +\infty$ for every $n \in \omega$. Note that $\Phi(Y) \neq B$ implies that $\Phi_\Theta(Y) = X$ is computable, by the same reasoning as when we considered $\Phi(X)$ for $X \in \mathsf{W2R}$, so it must be the case that $\Phi(Y) = B$. But then we have $X \geq_T B$, and hence $X \notin \mathsf{W2R}_\mu$.

$\square$

We can also separate MLR and $\mathsf{2MLR}$ with respect to a trivial measure μ. To do so, we modify the tally functional Φ_{Θ_1} from the proof of Theorem 4.19. Moreover, unlike the other measures we've considered in this chapter, we can construct μ so that uncountably many sequences are μ-Martin-Löf random. Lastly, we can even guarantee that μ is diminutive.

Theorem 5.22. *There is a trivial $\mu \in \mathscr{M}_c$ such that*

(i) $|\mathsf{MLR}_\mu| = 2^{\aleph_0}$,

(ii) $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$ *implies that* $X \notin \mathsf{2MLR}$*, and*

(iii) no $X \in \mathsf{MLR}_\mu$ *wtt-computes any* $Y \in \mathsf{MLR}$*.*

Proof. We define a new functional Ψ that on input $A \oplus B$ behaves like to the tally functional Φ_{Θ_1} considered above. However, instead of the tally being given in terms of 1s, we use the bits of B. Suppose that

$$\Phi_{\Theta_1}(A) = 1^{t_0}01^{t_1}01^{t_2}0\dots1^{t_i}0\dots.$$

Then we have

$$\Psi(A \oplus B) = b_0^{t_0}\, b_1^{t_1}\, b_2^{t_2}\dots b_i^{t_i}\dots$$

where $b_i = B(i)$ for every i. Note that Ψ is total, since Φ_{Θ_1} is total. Further, if $B \in \mathsf{2MLR}$, then $B \in \mathsf{MLR}^\Omega$ and hence $\Omega \oplus B \in \mathsf{MLR}$ by van Lambalgen's Theorem. It follows from the preservation of Martin-Löf randomness that $\Psi(\Omega \oplus B)$ is random with respect to the induced measure λ_Ψ. (i) holds, since

$$|\{\Omega \oplus B : B \in \mathsf{2MLR}\}| = |\mathsf{2MLR}| = 2^{\aleph_0}$$

and $\Psi(\Omega \oplus A) = \Psi(\Omega \oplus B)$ for $A, B \in \mathsf{2MLR}$ implies that $A = B$.

Next, if $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$, then $X = \Psi(\Omega \oplus A)$ for some $A \in \mathsf{2MLR}$. Moreover, $X \notin \mathsf{2MLR}_\mu$, since we can relativize

To show (ii), note that Theorem 4.5 can be relativized: If $X \in \mathsf{2MLR}_\mu$, where μ is induced by some almost total functional Δ, then there is some $Y \in \mathsf{2MLR}$ such that $\Delta(Y) = X$. Thus if $X \in \mathsf{MLR}_\mu \setminus \mathsf{Atoms}_\mu$, if $X \in \mathsf{2MLR}_\mu$, then there is some $Y \in \mathsf{2MLR}$ such that $\Psi(Y) = X$. If $\Psi(Y)$ is not computable, then Y must be equal to $\Omega \oplus A$ for some $A \in \mathsf{2MLR}$. No 2-random can have this form, so $\Psi(Y)$ is computable and not equal to X.

Finally, as with $\Omega^* = \Phi_{\Theta_1}(\Omega)$, $\Psi(\Omega \oplus B)$ is not complex, and thus cannot *wtt*-compute any $y \in \mathsf{MLR}$. Consequently, μ is diminutive.

$\square$

We can apply a tally functional to provide a counterexample to a claim made by Schnorr, namely, that for $\mu \in \mathscr{M}_c$, $\mathsf{MLR}_\mu = \mathsf{SR}_\mu$ if and only if μ is trivial.

Theorem 5.23. *There is a trivial $\mu \in \mathscr{M}_c$ such that*

(a) $\mathsf{MLR}_\mu = \mathsf{Atoms}_\mu$, *and*

(b) $\mathsf{SR}_\mu \setminus \mathsf{MLR}_\mu \neq \emptyset$.

Proof. To construct the desired measure μ, we define a tally functional Φ in terms of a universal μ-Martin-Löf test. More specifically, we follow the proof of Proposition 2.59. Let $\Theta(X, n, s)$ be the formula

$$(\exists k)[\![A{\upharpoonright}k]\!] \subseteq \mathcal{U}_{n,s}.$$

Further, we define

$$\theta(X,n) = \begin{cases} \text{the least } s \text{ such that } (\exists k)[\![A{\upharpoonright}k]\!] \subseteq \mathcal{U}_{n,s} & \text{if } s \text{ exists} \\ +\infty & \text{otherwise} \end{cases}.$$

Let μ be the measure induced by the tally functional Φ_Θ. There are two cases of interest to us here (the case that $X \notin \mathsf{MLR} \cup \mathsf{SR}$ has no bearing on the result here).

Case 1: $X \in \mathsf{MLR}$. In this case, there is some least n such that $X \notin \mathcal{U}_n$, and hence $\theta(X,n) = +\infty$. By the preservation of Martin-Löf randomness, and the fact that $X \in \mathsf{Atoms}_\mu$ implies that $\Phi_\Theta(Y) = X$ for some $Y \in \mathsf{MLR}$, we have $\mathsf{MLR}_\mu = \mathsf{Atoms}_\mu$.

Case 2: $X \in \mathsf{SR} \setminus \mathsf{MLR}$. Then $X \in \bigcap_{i\in\omega} \mathcal{U}_i$. Let $f \leq_T X$ be the function from the proof of Proposition 2.59, where

$$f(n) = \text{ the least } s \text{ such that } (\exists k)[\![X{\upharpoonright}k]\!] \subseteq \mathcal{U}_{n,s}.$$

This function dominates all computable functions. It follows that $\theta(X,n) = f(n)$, so that

$$\Phi(X) = 1^{f(0)}\, 0\, 1^{f(1)}\, 0\, \ldots$$

is not computable, and hence is not a μ-atom. Moreover, by the conservation of Schnorr randomness, $\Phi(X) \in \mathsf{SR}_\mu$, and by Theorem 4.5, $\Phi(X) \notin \mathsf{MLR}_\mu$.

□

CHAPTER 6

TRIVIAL MEASURES AND FINITE DISTRIBUTIVE LATTICES

6.1 Trivial Measures and Finite Distributive Lattices

As further evidence of the "non-triviality" of trivial measures, we show that a trivial measure gives rise to a certain structure, which varies as we consider different trivial measures. Specifically, if one considers the LR-degrees (or "low-for-random" degrees) associated with MLR_μ for a class of trivial measures μ, one finds that different trivial measures can give rise to non-isomorphic LR-degree structures.

Nies [Nie05] gave the following definition in the context of Martin-Löf randomness with respect to the Lebesgue measure, we say that A is LR-reducible to B, denoted $A \equiv_{LR} B$ if

$$\mathsf{MLR}^B \subseteq \mathsf{MLR}^A.$$

The intuitive idea is that B is more powerful than A as an oracle, as B de-randomizes more sequences than A does. We consider the equivalence relation given in terms of $\leq_{LR}$, so that $A \equiv_{LR} B$ if and only if $A \leq_{LR} B$ and $B \leq_{LR} A$. The collection of equivalence classes under this relation is called the collection of LR-*degrees*, denoted by $\mathscr{D}_{LR}$. Like the Turing degrees, the LR-degrees form an uncountable upper semilattice, but it is a highly complicated structure that is not well-understood.

We can extend the definition of $\leq_{LR}$ to Martin-Löf randomness with respect to any $\mu \in \mathscr{M}_c$ as follows. For $\mu \in \mathscr{M}_c$ and $A, B \in 2^\omega$, we say that A is $LR(\mu)$-reducible to B, denoted $A \leq_{LR(\mu)} B$ if

$$\mathsf{MLR}_\mu^B \subseteq \mathsf{MLR}_\mu^A.$$

We can define the $LR(\mu)$-degrees, denoted $\mathscr{D}_{LR(\mu)}$, just as we defined the LR-degrees above. Let's consider some examples.

Example 6.1. Let $\mu \in \mathscr{M}_c$ be such that $\mu(\mathsf{Atoms}_\mu) = 1$ and $\mathsf{Atoms}_\mu = \mathsf{MLR}_\mu$. Then $\mathscr{D}_{LR(\mu)}$ consists of a single equivalence class, consisting of all of 2^ω. The reason is that if $\mu(\{X\}) > 0$, then $X \in \mathsf{MLR}_\mu^A$ for every $A \in 2^\omega$.

Example 6.2. If μ is the measure induced by the tally functional Φ_A for $A \in \Delta_2^0 \cap \mathsf{MLR}$ (as in Example 5.3), then $\mathscr{D}_{LR(\mu)}$ consists of exactly two elements. If we set $A^* := \Phi_A(A)$, then by Theorem 5.6,

$$\mathsf{MLR}_\mu = \{A^*\} \cup \mathsf{Atoms}_\mu$$

where A^* is not computable. By Theorem 4.6,

$$A \in \mathsf{MLR}^B \Leftrightarrow A^* \in \mathsf{MLR}_\mu^B$$

for every $B \in 2^\omega$. Then there are exactly two $LR(\mu)$-degrees, $\mathbf{0}$ and $\mathbf{1}$:

$\mathbf{0} = \{B : A \in \mathsf{MLR}^B\}$,

$\mathbf{1} = \{B : A \notin \mathsf{MLR}^B\}$.

Example 6.3. Let $A \oplus B \in \mathsf{MLR} \cap \Delta^0_2$, and let Φ_A and Φ_B be the tally functionals for A and B, and let μ_0 and μ_1 be the measures induced by Φ_A and Φ_B, respectively. By Theorem 5.6,

$$\mathsf{MLR}_{\mu_0} = \{\Phi_A(A)\} \cup \mathsf{Atoms}_{\mu_0} \text{ and } \mathsf{MLR}_{\mu_1} = \{\Phi_B(B)\} \cup \mathsf{Atoms}_{\mu_1}$$

If we set $\nu := \dfrac{\mu_0 + \mu_1}{2}$, by Lemma 2.19, we have

$$\mathsf{MLR}_{\nu} = \mathsf{MLR}_{\mu_0} \cup \mathsf{MLR}_{\mu_1} = \{\Phi_A(A), \Phi_B(B)\} \cup \mathsf{Atoms}_{\mu_0} \cup \mathsf{Atoms}_{\mu_1}.$$

Clearly, ν is trivial. There are exactly four $LR(\nu)$-degrees; namely $\mathbf{0}, \mathbf{a}, \mathbf{b}$, and $\mathbf{1}$, where

$$\mathbf{0} = \{X : A \in \mathsf{MLR}^X \ \wedge\ B \in \mathsf{MLR}^X\},$$

$$\mathbf{a} = \{X : A \in \mathsf{MLR}^X \ \wedge\ B \notin \mathsf{MLR}^X\},$$

$$\mathbf{b} = \{X : A \notin \mathsf{MLR}^X \ \wedge\ B \in \mathsf{MLR}^X\},$$

$$\mathbf{1} = \{X : A \notin \mathsf{MLR}^X \ \wedge\ B \notin \mathsf{MLR}^X\}.$$

In particular, we have $\mathbf{0} < \mathbf{a} < \mathbf{1}$ and $\mathbf{0} < \mathbf{b} < \mathbf{1}$, but $\mathbf{a}$ and $\mathbf{b}$ are incomparable. Thus, $\mathscr{D}_{LR(\nu)}$ is isomorphic to the finite Boolean algebra on two atoms, pictured in Figure 6.1.

In the previous three examples we have defined trivial measures μ such that the associated $LR(\mu)$-degrees are isomorphic to the finite Boolean algebra of one, two,

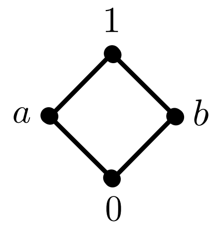

Figure 6.1. The finite Boolean algebra on two atoms

and four elements, respectively. Thus, it is natural to consider whether there is such a measure for every finite Boolean algebra.

Theorem 6.4. *For every finite Boolean algebra $\mathcal{B} = (B, \leq)$, there is a computable measure μ such that*

$$(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)}) \cong (B, \leq).$$

Proof. We proceed in four steps:

Step 1: If n is the number of atoms of $\mathcal{B}$, choose $A_1, A_2, \ldots, A_n \in \mathsf{MLR} \cap \Delta^0_2$ such that for each $J \subseteq \{1, \ldots, n\}$, if

$$X_J = \bigoplus_{j \in J} A_j$$

then

$$A_i \in \mathsf{MLR}^{X_J}$$

for every $i \notin J$.

Step 2: For each A_i, let Φ_{A_i} be the tally functional defined in terms of the Δ^0_2

approximation of A_i as in Example 5.3, and define μ_i to be the measure induced by the tally functional Φ_{A_i}. Let $A_i^* = \Phi_{A_i}(A_i)$.

Step 3: Define $\mu := \frac{1}{n}\sum_{i=1}^{n} \mu_i$. It follows from Lemma 2.19 that

$$\mathsf{MLR}_\mu = \bigcup_{i=1}^{n} \mathsf{MLR}_{\mu_i} = \{A_1^*, A_2^*, \ldots, A_n^*\} \cup \bigcup_{i=1}^{n} \mathsf{Atoms}_{\mu_i}.$$

Step 4: We verify that for $J, K \subseteq \{1, \ldots, n\}$, $deg_{LR}^{\mu}(X_J) \leq deg_{LR}^{\mu}(X_K)$ if and only if $J \subseteq K$. First, note that

$$\mathsf{MLR}_\mu^{X_J} = \{A_i^* : i \notin J\} \ \wedge \ \mathsf{MLR}_\mu^{X_K} = \{A_i^* : i \notin K\}.$$

Then

$$\begin{aligned} deg_{LR}^{\mu}(X_J) \leq deg_{LR}^{\mu}(X_K) &\Leftrightarrow \mathsf{MLR}_\mu^{X_K} \subseteq \mathsf{MLR}_\mu^{X_J} \\ &\Leftrightarrow \{A_i^* : i \notin K\} \subseteq \{A_i^* : i \notin J\} \\ &\Leftrightarrow J \subseteq K. \end{aligned}$$

Thus, $\mathscr{D}_{LR}(\mu) = \{deg_{LR}^{\mu}(X_J) : J \subseteq \{1, \ldots, n\}\}$ is isomorphic to the powerset of $\{1, \ldots, n\}$, which is isomorphic to $\mathcal{B}$. □

Theorem 6.4 can be improved further:

Theorem 6.5. *For every finite distributive lattice $\mathcal{L} = (L, \leq)$, there is a computable measure μ such that*

$$(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)}) \cong (L, \leq).$$

The following terminology will be useful in the proof of Theorem 6.5. Let $\mathcal{L}$ be a finite distributive lattice of n elements. We will consider $\mathcal{L}$ in terms of levels, where Level 1 consists of $1_{\mathcal{L}}$, Level 2 consists of the immediate predecessors of $1_{\mathcal{L}}$, Level 3 consists of the immediate predecessors of elements of Level 2, and so on. Since $\mathcal{L}$ has size n, there are only finitely many levels (in fact, at most n levels), and since it is a lattice, the lowest level consists solely of $0_{\mathcal{L}}$.

Given $a, b \in L$, the meet of a and b, denoted $a \wedge b$, is the greatest element in $\mathcal{L}$ such that $a \geq a \wedge b$ and $b \geq a \wedge b$. The element $c \in L$ is meet-reducible if there are $a, b > c$ such that $a \wedge b = c$, and it is meet-irreducible if it is not meet-reducible.

To prove Theorem 6.5, the idea is (i) construct a lattice of sets isomorphic to $\mathcal{L}$, (ii) use these sets to define a collection of tally functionals, and (iii) define a measure in terms of these tally functionals, which will give rise to an LR-structure that is isomorphic to $\mathcal{L}$. Let us first consider an example.

Let $\mathcal{L}$ be the finite distributive lattice given below in Figure 6.2.

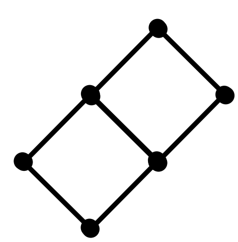

Figure 6.2. The finite distributive lattice $\mathcal{L}$

Now let $A \in \mathsf{MLR} \cap \Delta^0_2$, and let $\{A_i\}_{i\in\omega}$ be such that

$$A = \bigoplus_{i\in\omega} A_i,$$

so that each $A_i \in \mathsf{MLR} \cap \Delta^0_2$. Hereafter, the sequences $A_1, A_2, \ldots$ will be referred to as *basic sequences*. An important feature of these basic sequences is that each A_i is Martin-Löf random relative to a finite join of any basic sequences that differ from A_i.

We proceed by associating to each element at each level of $\mathcal{L}$ a set consisting of some of the A_i's or joins of the A_i's, yielding a finite distributive lattice of sets that is isomorphic to $\mathcal{L}$, as in Figure 6.3.

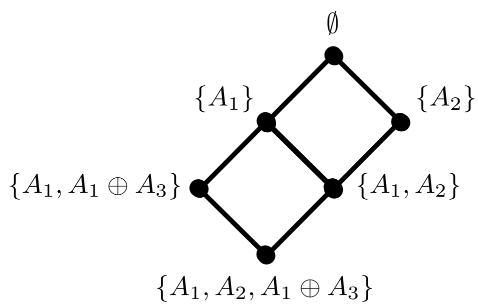

Figure 6.3. A finite distributive lattice of sets isomorphic to $\mathcal{L}$

Level 1: We associate to the top element $1_{\mathcal{L}}$ the empty set.

Level 2: There are two elements in Level 2, and so we associate to one the set $\{A_1\}$ and to the other $\{A_2\}$.

Level 3: There are two elements in Level 3, one of which is meet-reducible and the other meet-irreducible. To the meet-reducible element, we associate the set $\{A_1, A_2\}$, and to the meet-irreducible element (which is below the element associated to the set $\{A_1\}$, we associate the set $\{A_1, A_1 \oplus A_3\}$, where A_3 is the first basic sequence in $\{A_i\}_{i\in\omega}$ (in the order given by the indices) that has not appeared in the construction thus far. Note that any sequence that derandomizes A_1 also derandomizes $A_1 \oplus A_3$, but not every element that derandomizes $A_1 \oplus A_3$ also derandomizes A_1 (such as A_3 itself).

Level 4: The only element at Level 4 is $0_{\mathcal{L}}$, the meet of the two Level 3 elements, and thus we associate to this element the set $\{A_1, A_2, A_1 \oplus A_3\}$.

Next, for each element in the set associated with $0_{\mathcal{L}}$, namely A_1, A_2, and $A_1 \oplus A_3$, let μ_{A_1}, μ_{A_2}, and $\mu_{A_1\oplus A_3}$ be the measures induced by the tally functionals Φ_{A_1}, Φ_{A_2}, and $\Phi_{A_1\oplus A_3}$ defined in terms of the Δ^0_2 approximations of A_1, A_2, and $A_1 \oplus A_3$. We define

$$\mu := \frac{1}{3}(\mu_{A_1} + \mu_{A_2} + \mu_{A_1\oplus A_3}),$$

Let

$$\Phi_{A_1}(A_1) = A_1^*, \Phi_{A_2}(A_2) = A_2^*, \text{ and } \Phi_{A_1\oplus A_3}(A_1 \oplus A_3) = (A_1 \oplus A_3)^*.$$

Then for any $X \in 2^\omega$,

$$\mathsf{MLR}^X_\mu = \mathsf{Atoms}_\mu \cup \mathcal{S},$$

where $\mathcal{S}$ is equal to one of the following:

$$\emptyset, \{A_1^*\}, \{A_2^*\}, \{A_1^*, A_2^*\}, \{A_1^*, (A_1 \oplus A_3)^*\}, \text{ or } \{A_1^*, A_2^*, (A_1 \oplus A_3)^*\}.$$

Thus we have a one-to-one correspondence (that preserves $\subseteq$) between the above sets and those sets associated to the elements of $\mathcal{L}$, and thus $(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)})$ is isomorphic to $\mathcal{L}$. Now we proceed in full generality.

Proof of Theorem 6.5. Let $\mathcal{L}$ be a finite distributive lattice. We proceed as in the example. We first associate basic sequences and joins of basic sequences to elements of the various levels of $\mathcal{L}$.

Level 1: We associate to the top element $1_{\mathcal{L}}$ the empty set.

Level 2: To each of the $j \leq K$ elements in Level 2, we associate a singleton consisting of a basic sequence $A_1, A_2, \ldots, A_k$.

Level $\mathbf{n+1}$: The set we associate to a Level $n+1$ element depends on whether it is meet-reducible or meet-irreducible.

- The meet-reducible case: Let $a = b \wedge c$, where b and c are Level n elements. If $\mathcal{S}_b$ is the set of sequences associated to b and $\mathcal{S}_c$ is the set of sequences associated to c, then we associate the set $\mathcal{S}_b \cup \mathcal{S}_c$ to a. (Note: We will have to verify that this is well-defined, for it may be the case that there are Level n elements b' and c' that differ from b and c but also satisfy $b' \wedge c' = a$.)

- The meet-irreducible case: If a is meet-irreducible, then there is only one Level n element such that $a \leq b$. If $\mathcal{S}_b$ is the set associated to b, then we proceed as follows. First, let $\{A_{i_1}, A_{i_2}, \ldots, A_{i_k}\}$ be the collection of basic sequences appearing in S_b. That is, these are either elements of $\mathcal{S}_b$ or are contained in joins in $\mathcal{S}_b$ (so that, for instance, the basic sequences appearing in $\{A_1, A_2 \oplus A_3\}$ are A_1, A_2, and A_3). Let $N \in \omega$ be the least such that the basic sequence A_N has not appeared in any set associated to an element of $\mathcal{L}$. Then to a we associate the set

$$\left\{\bigoplus_{j=1}^{k} A_{i_j} \oplus A_N\right\} \cup \mathcal{S}_b.$$

To verify that $\mathcal{S}_{\mathcal{L}} = (\{\mathcal{S}_a : a \in L\}, \leq)$ is a finite distributive lattice isomorphic to $\mathcal{L}$ (where $\mathcal{S}_a \leq \mathcal{S}_b$ if and only if $\mathcal{S}_a \supseteq \mathcal{S}_b$), we first show that meets in $\mathcal{S}_{\mathcal{L}}$ are well-defined. First, suppose that $a, b, c, d \in L$ are distinct elements such $a = b \wedge c = c \wedge d$, as in Figure 6.4.

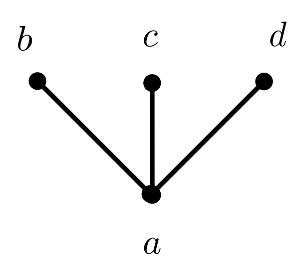

Figure 6.4. The case in which $a = b \wedge c = c \wedge d$

We claim that $b \vee c \neq c \vee d$. For otherwise, the lattice M_3 (pictured in Figure 6.5 below) would be embeddable into $\mathcal{L}$, contradicting the fact that $\mathcal{L}$ is distributive.

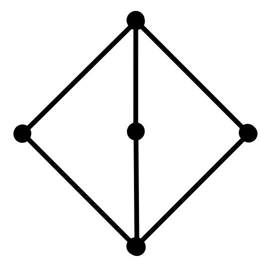

Figure 6.5. The lattice M_3

Thus, we must have:

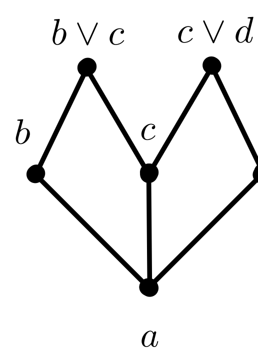

Figure 6.6. $b \vee c \neq c \vee d$

We must have $b \vee d \neq b \vee c$ and $b \vee d \neq c \vee d$, for otherwise M_3 is embeddable into $\mathcal{L}$ (for instance, Figure 6.7 shows the case that $b \vee d = c \vee d$).

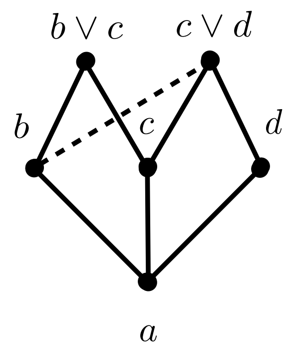

Figure 6.7. The case that $b \vee d = c \vee d$

If $e = b \vee c, f = c \vee d$, and $g = b \vee d$, then $b = e \wedge f, c = e \wedge g$, and $d = f \wedge g$, and thus none of b, c, or d is meet-irreducible.

Now, suppose that

$$\mathcal{S}_b \cup \mathcal{S}_c \neq \mathcal{S}_c \cup \mathcal{S}_d \tag{6.1}$$

Since $b = e \wedge f, c = e \wedge g$, and $d = f \wedge g$, it follows that

$$\mathcal{S}_b = \mathcal{S}_e \cup \mathcal{S}_f \ \wedge \ \mathcal{S}_c = \mathcal{S}_e \cup \mathcal{S}_g \ \wedge \ \mathcal{S}_d = \mathcal{S}_f \cup \mathcal{S}_g.$$

By (6.1), we have

$$\mathcal{S}_e \cup \mathcal{S}_f \cup \mathcal{S}_g = \mathcal{S}_b \cup \mathcal{S}_c \neq \mathcal{S}_c \cup \mathcal{S}_d = \mathcal{S}_e \cup \mathcal{S}_f \cup \mathcal{S}_g,$$

which is impossible. In the case where there are distinct Level n elements b, c, b', c' such that $a = b \wedge c = b' \wedge c'$, it follows that $a = b \wedge c'$, and so we apply the above argument to b, c, c', and then to b, c, b' to conclude that

$$\mathcal{S}_b \cup \mathcal{S}_c = \mathcal{S}_{b'} \cup \mathcal{S}_{c'}.$$

Thus, meets are well-defined.

Next we show that the isomorphism between $\mathcal{S}_{\mathcal{L}} = (\{\mathcal{S}_a : a \in L\}, \leq)$ and $\mathcal{L}$ holds level by level. In particular, we show that meets and joins are preserved level by level. First, it is clear that the top two levels of $\mathcal{S}_{\mathcal{L}}$ and $\mathcal{L}$ are isomorphic. Now suppose that $\mathcal{S}_{\mathcal{L}}$ and $\mathcal{L}$ are isomorphic from Level 1 to Level n. Having associated to Level n elements a and b the sets $\mathcal{S}_a$ and $\mathcal{S}_b$, we associate the set $\mathcal{S}_a \cup \mathcal{S}_b$ to $a \wedge b$.

Suppose Level $n+1$ elements a and b are associated with $\mathcal{S}_a$ and $\mathcal{S}_b$. To show that $\mathcal{S}_{a \vee b}$, the set associated to $a \vee b$, is $\mathcal{S}_a \cap \mathcal{S}_b$, we consider three cases.

Case 1: First, if both a and b are meet-irreducible, then either (i) there is some Level n element c such that $a, b \leq c$, or (ii) there are Level n elements c and d such that $c \neq d$, $a \leq c$, and $b \leq d$.

Subcase 1.i: By the procedure given above,

$$\mathcal{S}_a = \mathcal{S}_c \cup \{B \oplus A_\ell\}$$

and

$$\mathcal{S}_b = \mathcal{S}_c \cup \{B \oplus A_{\ell'}\},$$

where B is the join ($\bigoplus$) of the basic sequences appearing in $\mathcal{S}_c$ and $A_\ell, A_{\ell'}$ are basic sequences not contained in any set associated to elements of Levels $k \leq n$. Thus $\mathcal{S}_{a \vee b} = \mathcal{S}_c = \mathcal{S}_a \cap \mathcal{S}_b$.

Subcase 1.ii: In this subcase,

$$\mathcal{S}_a = \mathcal{S}_c \cup \{B \oplus A_\ell\},$$

and

$$\mathcal{S}_b = \mathcal{S}_d \cup \{D \oplus A_{\ell'}\},$$

where B and D are the joins of the basic sequences appearing in $\mathcal{S}_c$ and $\mathcal{S}_d$, respectively, and $A_\ell, A_{\ell'}$ are basic sequences not contained in any set associated to elements of Levels $k \leq n$. By induction, there is some $e \in L$ such that $e = c \vee d$ and $\mathcal{S}_e = \mathcal{S}_c \cap \mathcal{S}_d$. Then we have $e = a \vee b$ and

$$\mathcal{S}_a \cap \mathcal{S}_b = (\mathcal{S}_c \cup \{B \oplus A_\ell\}) \cap (\mathcal{S}_c \cup \{D \oplus A_{\ell'}\}) = \mathcal{S}_c \cap \mathcal{S}_d = \mathcal{S}_e = S_{a \vee b}.$$

Case 2: If a is meet-irreducible but b is meet-reducible, then again there are two subcases to consider: Either (i) there is some Level n element c such that $a, b \leq c$, or (ii) there are distinct Level n elements c, d, e such that $a \leq c$ and $b = d \wedge e$.

Subcase 2.i: We have

$$\mathcal{S}_a = \mathcal{S}_c \cup \{B \oplus A_\ell\},$$

where B is the join of the basic sequences appearing in $\mathcal{S}_c$ and A_ℓ is a basic sequence not contained in any set associated to any element of Levels $k \leq n$, and

$$\mathcal{S}_b = \mathcal{S}_c \cup \mathcal{S}_d$$

for some Level n element $d \neq c$. Again it follows that

$$\mathcal{S}_a \cap \mathcal{S}_b = (\mathcal{S}_c \cup \{B \oplus A_\ell\}) \cap (\mathcal{S}_c \cup \mathcal{S}_d) = \mathcal{S}_c = \mathcal{S}_{a \vee b}.$$

Subcase 2.ii: In this subcase, $c \vee (d \wedge e) = a \vee b$. As above,

$$\mathcal{S}_a = \mathcal{S}_c \cup \{B \oplus A_\ell\}$$

and

$$\mathcal{S}_b = \mathcal{S}_d \cup \mathcal{S}_e.$$

By the inductive hypothesis, we have $\mathcal{S}_c \cap (\mathcal{S}_d \cup \mathcal{S}_e) = \mathcal{S}_{c \vee (d \wedge e)}$, and thus

$$\mathcal{S}_a \cap \mathcal{S}_b = (\mathcal{S}_c \cup \{B \oplus A_\ell\}) \cap (\mathcal{S}_d \cup \mathcal{S}_e) = \mathcal{S}_c \cap (\mathcal{S}_d \cup \mathcal{S}_e) = \mathcal{S}_{c \vee (d \wedge e)} = \mathcal{S}_{a \vee b}.$$

Case 3: Lastly, in the case that a and b are both meet-reducible, either (i) there are distinct Level n elements c, d, e such $a = c \wedge d$, and $b = d \wedge e$ or (ii) there are distinct

Level n elements c, d, e, f such that $a = c \wedge d$ and $b = e \wedge f$.

Subcase 3.i: Since $a, b \leq d$, it follows that $\mathcal{S}_a = \mathcal{S}_c \cup \mathcal{S}_d$, $\mathcal{S}_b = \mathcal{S}_d \cup \mathcal{S}_e$, $\mathcal{S}_c \cap \mathcal{S}_e = \emptyset$, and thus

$$\mathcal{S}_a \cap \mathcal{S}_b = (\mathcal{S}_c \cup \mathcal{S}_d) \cap (\mathcal{S}_d \cup \mathcal{S}_e) = \mathcal{S}_d = \mathcal{S}_{a \vee b}.$$

Subcase 3.ii: Note that $a \vee b = (c \wedge d) \vee (e \wedge f)$. Since $\mathcal{S}_a = \mathcal{S}_c \cup \mathcal{S}_d$ and $\mathcal{S}_b = \mathcal{S}_e \cup \mathcal{S}_f$, by the inductive hypothesis, it follows that

$$\mathcal{S}_a \cap \mathcal{S}_b = (\mathcal{S}_c \cup \mathcal{S}_d) \cap (\mathcal{S}_e \cup \mathcal{S}_f) = \mathcal{S}_{(c \vee d) \wedge (e \vee f)} = \mathcal{S}_{a \vee b}.$$

Having verified that $\mathcal{S}_{\mathcal{L}}$ is a finite distributive lattice, we now turn to defining the measure μ. Let

$$\{B_1, \ldots, B_k\}$$

be the set in $\mathcal{S}_{\mathcal{L}}$ associated to $0_{\mathcal{L}}$. By our construction, each B_i is either a basic sequence or the join of some basic sequences. Further, since the basic sequences are all in $\mathsf{MLR} \cap \Delta^0_2$, and further, each is Martin-Löf random relative to the finite join of any number of basic sequences that differ from it, it follows from van Lambalgen's Theorem (Theorem 2.86) that each $B_i \in \mathsf{MLR} \cap \Delta^0_2$.

Let Φ_i be the tally functional defined in terms of the Δ^0_2 approximation of B_i, and let $B_i^* = \Phi_i(B_i)$, so that $\mathsf{MLR}_{\mu_i} = \{B_i^*\} \cup \mathsf{Atoms}_{\mu_i}$. Setting

$$\mu := \frac{1}{k} \sum_{i=1}^{k} \mu_i,$$

we claim that $(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)}) \cong (\mathcal{S}_{\mathcal{L}}, \leq)$. First we show that for each $\mathcal{S}^*_a$, there is some $X \in 2^\omega$ such that

$$\mathcal{S}_a^* \cup \mathsf{Atoms}_\mu = \mathsf{MLR}_\mu^X.$$

If we let $\mathsf{RScope}(B) = \{X \in 2^\omega : B \in \mathsf{MLR}^X\}$ be the *randomness scope* of B, note that by Theorem 4.6,

$$B_i \in \mathsf{MLR}^X \Leftrightarrow B_i^* \in \mathsf{MLR}_\mu^X,$$

and hence $X \in \mathsf{RScope}(B_i)$ if and only if $B_i^* \in \mathsf{MLR}_\mu^X$. Observe that $\lambda(\mathsf{RScope}(B)) = 1$ for every $B \in \mathsf{MLR}$, since by van Lambalgen's Theorem, $\mathsf{MLR}^B \subseteq \mathsf{RScope}(B)$ and $\lambda(\mathsf{MLR}^B) = 1$.

If

$$\mathcal{S}_a^* = \{B_{i_1}^*, \dots, B_{i_j}^*\},$$

then

$$\mathcal{X} = \bigcap_{j=1}^{k} \mathsf{RScope}(B_{i_j}) \neq \emptyset,$$

since the finite intersection of measure one sets has measure one. Thus for any $X \in \mathcal{X}$, we have $\mathsf{MLR}_\mu^X = \mathcal{S}_a^* \cup \mathsf{Atoms}_\mu$.

We claim that for each $X \in 2^\omega$, there is some $a \in L$ such that $\mathsf{MLR}_\mu^X = \mathcal{S}_a^* \cup \mathsf{Atoms}_\mu$. Let $\{A_1, \dots, A_k\}$ be the collection of basic sequences appearing in the elements of $\mathcal{S}_\mathcal{L}$. Further, for $j \leq k$, let $\{A_1, \dots, A_j\}$ be the basic sequences that make up the singletons assigned to Level 2 elements of $\mathcal{L}$ (which we'll call the *Level Two basic sequences*), and let $\{A_{j+1}, \dots, A_k\}$ be the basic sequences that added when we assign sets to meet-irreducible elements of $\mathcal{L}$ (which we'll call the *meet-irreducible basic sequences*).

By our construction, every non-empty $\mathcal{S}_a$ contains some Level Two basic sequence. Thus, given $X \in 2^\omega$, $deg^\mu_{LR}(X)$ is determined in part by which Level Two basic sequences it derandomizes. In particular, $\{A_1, \ldots, A_j\} \cap \mathsf{MLR}^X = \emptyset$ implies that $\mathsf{MLR}^X_\mu = \mathsf{Atoms}_\mu$. For each $X \in 2^\omega$, there is some $J \subseteq \{1, \ldots, j\}$ such that

$$X \in \bigcap_{i \in J} \mathsf{RScope}(A_i) \ \wedge \ X \notin \bigcap_{i \in \{1,\ldots,j\} \setminus J} \mathsf{RScope}(A_i). \tag{6.2}$$

Moreover, for each $X \in 2^\omega$ there is some $K \subseteq \{j+1, \ldots, k\}$ such that

$$X \in \bigcap_{i \in K} \mathsf{RScope}(A_i) \ \wedge \ X \notin \bigcap_{i \in \{j+1,\ldots,k\} \setminus K} \mathsf{RScope}(A_i). \tag{6.3}$$

Thus, for each $X \in 2^\omega$ there is thus a unique $\mathcal{S}_a = \{B_{i_1}, \ldots, B_{i_j}\}$ such that $\mathcal{S}_a \subseteq \mathsf{MLR}^X$, determined by the basic sequences for which (6.2) and (6.3) hold. Consequently, $\mathsf{MLR}^X_\mu = \{B^*_{i_1}, \ldots, B^*_{i_j}\} \cup \mathsf{Atoms}_\mu$.

Every $\mathcal{S}^*_a$ is the collection of non-atoms in MLR^X_μ for some $X \in 2^\omega$, and for every $X \in 2^\omega$, there is some $\mathcal{S}^*_a$ such that $\mathsf{MLR}^X_\mu = \mathcal{S}^*_a \cup \mathsf{Atoms}_\mu$. Since $X \leq_{LR(\mu)} Y$ if and only if $\mathsf{MLR}^Y_\mu \subseteq \mathsf{MLR}^X_\mu$, it follows that

$$(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)}) \cong (\mathcal{S}_L, \leq).$$

□

6.2 Open Questions

We conclude with several questions.

Question 6.6. If $\mathcal{L} = (L, \leq)$ is an infinite, computable, distributive lattice, is there a trivial measure $\mu \in \mathscr{M}_c$ such that

$$(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)}) \cong (L, \leq)?$$

Question 6.7. Is there an example of a finite non-distributive lattice $\mathcal{L} = (L, \leq)$ and a trivial measure $\mu \in \mathscr{M}_c$ such that

$$(\mathscr{D}_{LR(\mu)}, \leq_{LR(\mu)}) \cong (L, \leq)?$$

CHAPTER 7

PHILOSOPHICAL PERSPECTIVES ON ALGORITHMIC RANDOMNESS

7.1 Motivating the Problem

In the standard presentations of the subject of algorithmic randomness,[1] one regularly finds as motivation for the main definitions of randomness the following question:

What does it mean for a sequence of 0s and 1s to be random?

There are many ways one can attempt to answer this question, but the various answers provided by the general theory of algorithmic randomness are of a specific kind: An infinite sequence is random if it is the sort of sequence that is typically produced by a random process. Of course, this is imprecise and generally not very helpful. Which sequences are those that are typically produced by a random process? Any random process? And what does it mean to be *that* "sort of sequence"? Slight progress is made with the restriction to those random processes that are unbiased: each individual outcome of such a process is equiprobable. In this case, the paradigm

[1] See, for instance, *Computability and Randomness* [Nie09], by Andre Nies, published in 2009, and *Algorithmic Randomness and Complexity* [DH10] by Rod Downey and Denis Hirschfeldt, published in 2010.

example of a random process is the repeated tosses of a fair coin. But still, what is the typical output of such a process? Instead of taking this approach, one might hold that we should consider as random those sequences that are indistinguishable from a sequence that is generated by tosses of a fair coin. But then we can ask: Indistinguishable from whose point of view? And which is the sequence generated by the tosses of a fair coin to which we are comparing the putatively random sequences?

These are difficult questions, but what the general theory of algorithmic randomness provides for us is a number of ways to make precise notions such as 'the typical sequence produced by the tosses of a fair coin' or 'being indistinguishable from such as sequence'. There are, in fact, many ways of sharpening this notion of random sequence, each of which yields a definition of randomness. What's more, while a number of these ways of sharpening yield extensionally equivalent definitions of randomness, a number of these ways are incompatible with one another; some of the resulting definitions count certain sequences as random that are not counted as random by other definitions.

This fact notwithstanding, one particular definition, Martin-Löf randomness, has been singled out by some as *capturing the intuitive conception of randomness*, where this means at a minimum that it captures our commonly held intuitions about randomness.[2] Some have even gone so far as to suggest that there is a randomness-theoretic thesis about Martin-Löf randomness akin to the Church-Turing Thesis (henceforth, the CTT), the thesis that the collection of effectively calculable number-

[2] I provide this particular gloss on 'capturing the intuitive conception of randomness' to avoid giving the impression that there is some single conception that is *the* intuitive conception of randomness.

theoretic functions is coextensive with the collection of computable number-theoretic function. Following Jean-Paul Delahaye, I will refer to this thesis as the "Martin-Löf-Chaitin Thesis" (hereafter, the MLCT).[3]

According to the MLCT, a sequence is intuitively random if and only if it is Martin-Löf random. That is, the MLCT asserts the extensional adequacy of Martin-Löf randomness.

Clearly, establishing the MLCT certainly would be a significant philosophical development, just as the formulation of a definition of computability that captures our commonly held intuitions about computability is widely regarded to be philosophically significant. As Gödel noted, with the definition of computability "one has for the first time succeeded in giving an absolute definition of an interesting epistemological notion, i.e., one not depending on the formalism chosen" ([Göd46], p. 150). Do we now have yet another "absolute definition of an interesting epistemological notion" on our hands?

Clearly, this depends on what we take an "absolute definition" to be, as well as what notion we identify as the relevant "interesting epistemological notion". If we understand absoluteness in the sense that it has in Gödel's remark, the formally defined class of computable number-theoretic functions is absolute due to the fact that every formal definition of the class of intuitively computable number-theoretic functions has the same extension. But this is far from the case with Martin-Löf randomness: there are a number of alternative definitions of the notion of intuitively random sequence that do not yield the same extension of random sequences.

[3] See [Del11].

Of course, this lack of absoluteness in Gödel's sense does not necessarily imply that Martin-Löf fails to capture an interesting epistemological notion: Some have simply rejected all other alternatives. For instance, in a recent survey on the various definitions of randomness by A. Dasgupta, we find the following claim:

> [W]e believe that while the Martin-Löf-Chaitin thesis is not (yet) as strong as the Church-Turing thesis, the [... problem of ...] defining randomness for sequences [...] that captures our mathematical intuition of these objects [has] essentially been solved quite satisfactorily. ([Das11], p. 708)

What's more, this claim comes after a survey of many non-equivalent definitions of randomness. For Dasgupta and others, then, the interesting epistemological notion captured by Martin-Löf randomness is "our mathematical intuition" of random sequences or our commonly held intuitions of randomness, even in spite of the presence of alternative definitions of randomness.

But there is cause for concern. Those who have argued in support of the MLCT have failed to address a number of fundamental questions that must be answered before we can count as solved the problem of finding a definition of randomness that captures our commonly held intuitions about randomness: What exactly are these commonly held intuitions? Is it even proper to speak of "our mathematical intuition" of randomness in the singular, as Dasgupta does, or to speak of *the* intuitive conception of randomness, as others have done? Supposing that there is one conception that falls under the descriptions "the intuitive conception of randomness" or "our mathematical intuition" of random sequence, what does it mean for a definition of randomness to capture this intuitive conception or this mathematical intuition?

On what grounds should we judge that a given definition of randomness adequately captures our mathematical intuitions of randomness? On what grounds should we judge that a given definition *fails* to adequately capture these intuitions?

Moreover, the need for answers to these questions is heightened by the fact that the alternative definitions of randomness cannot be so easily dismissed. In fact, some of these alternative definitions are the subject of extensional adequacy theses that are very similar in form to the MLCT. For instance, Schnorr has claimed that his definition of randomness, which is strictly weaker than Martin-Löf randomness,[4] "captures the true concept of randomness", while others have claimed that weak 2-randomness, which is strictly stronger than Martin-Löf randomness, captures the so-called intuitive conception.[5]

The fact that the above questions have not been addressed, as well as the presence of alternative theses, should give us some pause before we accept the MLCT. But there are several more pressing questions that are also left unanswered: Why would we require of a formal definition of randomness that it capture all of the commonly held intuitions about randomness to begin with? Are there some purposes for a formal definition of randomness to fulfill or roles for a definition of randomness to play that can only be fulfilled or played by a definition that captures these intuitions? And what would be lost if it proved to be the case that *no* definition of randomness could fully capture these intuitions?

[4]One definition $\mathscr{D}_1$ of randomness is *weaker* than another $\mathscr{D}_2$ if every $\mathscr{D}_2$-random sequence is $\mathscr{D}_1$-random, but not conversely. In this case, we also say that $\mathscr{D}_2$ is *stronger* than $\mathscr{D}_1$.

[5]See, for instance, [OW08].

The main problems that I will consider in this portion of the dissertation are thus to determine (i) for which purposes we might seek a definition of randomness that captures our commonly held intuitions of randomness, (ii) how one might go about justifying the claim that some definition $\mathscr{D}$ captures these intuitions, and (iii) whether there is an alternative approach to the definitions of randomness that does not require of them that they capture all of our intuitions of randomness but still acknowledges their use as legitimate formalizations of these intuitions.

7.2 A Conceptual Analysis of Randomness?

One natural starting point for addressing these questions is the suggestion that a formal definition of randomness can serve as the basis of a *conceptual analysis* of the notion of randomness; that is, such a definition can provide necessary and sufficient conditions for the correct application of the concept of randomness. On this approach, the MLCT can be read as stating that Martin-Löf randomness provides a conceptual analysis of the notion of randomness, fulfilling what I'll henceforth call the *conceptual-analytic role of randomness*.

Should we thus accept the claim that Martin-Löf randomness provides a conceptual analysis of the notion of randomness? Or should we accept the claim that some other currently available definition of randomness provides a conceptual analysis of the notion of randomness? Or might we hold that the formal definitions of randomness that we find in the general theory of algorithmic randomness do not successfully play the conceptual-analytic role, but some other role(s)?

It is this latter approach that I develop and argue for here. But my goal is not merely to highlight some roles other than the conceptual-analytic role and argue that these other roles are successfully filled by definitions such as Martin-Löf randomness and other definitions of algorithmic randomness. Rather, my goal is three-fold:

(1) to argue that for any definition of randomness $\mathscr{D}$ that has a well-defined, definite extension, the advocate of the claim that $\mathscr{D}$-randomness captures our commonly held intuitions of randomness faces a serious challenge to justify this claim, which I call the *Justificatory Challenge*;

(2) to present an alternative approach to the various definitions of randomness that does not face the Justificatory Challenge, as it is based on two roles of randomness, which I call the *calibrative role of randomness* and the *limitative role of randomness*, each of which is successfully filled by multiple definitions of randomness; and

(3) to argue that this alternative approach further shows that what mathematicians have considered to be significant about the concept of randomness cannot be captured by one single definition, but that multiple definitions are necessary to capture these truths.

Moreover, I claim that (1)-(3) provide good reason to hold what I call the *No-Thesis Thesis*:

> **The No-Thesis Thesis:** No definition of randomness that has a definite, well-defined extension can capture the prevailing intuitive conception of randomness.

The details of my argument will be sketched briefly in the following chapter outline.

7.3 Outline of the Chapters

Prior to carrying out the steps (1)-(3) described above, I carry out several tasks. First, I consider the contributions of two figures central to the development of algorithmic randomness, Richard von Mises (in Chapter 8) and Jean Ville (in Chapter 9). Although both von Mises' and Ville's contributions predate the use of effective methods in the task of defining randomness, both made serious attempts at defining randomness which led to important breakthroughs that would later inform the work of such figures as Martin-Löf and Schnorr, whose work on algorithmic randomness in the late 1960s and early 1970s laid the foundation for much of the research in algorithmic randomness that continues today.

Not only did von Mises and Ville make important technical contributions, but each identified a role for definitions that informed the definitional task in which they were engaged. On von Mises' approach, a definition of randomness was necessary to serve as a foundation of his theory of probability, and more specifically, randomness play a central role in the task of solving problems in the probability calculus, a role I refer to as the *resolutory role of randomness*. Ville, on the other hand, was not interested in using randomness to solve problems, but rather, Ville sought a definition $\mathscr{D}$ of randomness with the property that every $\mathscr{D}$-random sequence is a paradigmatic instance of a sequence chosen at random. Such a definition would fill what I call the *exemplary role of randomness*, as the $\mathscr{D}$-random sequences would be exemplars of randomly chosen sequences.

But this role is a puzzling one, and it proves to be quite problematic. For not only is it not clear which properties a sequence should satisfy in order for it to exemplify the randomness of a sequence chosen at random, it is not clear what we gain from a definition that successfully fills the exemplary role.

After considering the accounts of both von Mises and Ville in Chapters 8 and 9, in Chapter 10 I consider the status of the MLCT. In particular, I consider the main evidence offered in support of the MLCT, as well as the main arguments provided against it. Not only does this chapter help set the stage for my arguments in Chapters 11 and 12, but it also provides a survey of the many arguments both for and against the MLCT, which is not currently available in the philosophical literature on algorithmic randomness.

In Chapter 11, I present the Justificatory Challenge to the advocate of the claim of extensional adequacy of some currently available definition of randomness $\mathscr{D}$. As I argue, to establish the claim that $\mathscr{D}$ captures the intuitive conception of randomness, the advocate of this claim (henceforth the $\mathscr{D}$-advocate) must provide a sharpening of the prevailing intuitive conception of randomness that is precise enough to block the claims of extensional adequacy made concerning alternative definitions of randomness without undermining the claim of the extensional adequacy of $\mathscr{D}$; this is precisely the Justificatory Challenge. Further, I argue that there is no reason to hold that the $\mathscr{D}$-advocate can meet this challenge.

Lastly, in Chapter 12, I present the calibrative and limitative roles of randomness. The salient feature of these roles is that can be successfully filled by a definition of randomness that is not extensionally adequate; in fact, each of these roles are filled not by an individual definition of randomness, but by an entire family of definitions of randomness.

The main features of these two roles are as follows. First, a definition $\mathscr{D}$ fills the calibrative role if and only if there is some notion of "almost-everywhere" typicality $\mathcal{T}$ occurring in classical mathematics such that the $\mathscr{D}$-randomness of a sequence is necessary and sufficient for that sequence to be $\mathcal{T}$-typical. For instance, given a theorem of classical analysis of the form "for almost every real number x, $\Phi(x)$", we can find a corresponding formula Φ^* (given by restricting Φ in some way) that still holds of almost every real number x, and which satisfies

$$(\forall x)[x \text{ is } \mathscr{D}\text{-random if and only if } \Phi^*(x)]$$

for some definition of randomness $\mathscr{D}$. What these results suggest is that many definitions in the family of definitions of randomness can be used to calibrate the degree of randomness necessary and sufficient to instantiate certain almost-everywhere behavior that occurs in classical mathematics.

Second, a definitions of randomness fills the limitative role of randomness by illuminating an interesting phenomenon that I refer to as the *indefinite contractibility of the notion of absolute randomness.* Broadly speaking, that the notion of absolute randomness is indefinitely contractible means that for every extension $\mathcal{E}$ of sequences that purportedly contains all absolutely random sequences, there is some $X \in \mathcal{E}$ that is not absolutely random, where, following a suggestion of Myhill's, a sequence is absolutely random if and only if it satisfies no property that is (i) satisfied by only measure zero many sequences and (ii) is definable without parameters in some formal system. By means of the various families of definitions of randomness, we can systematically study this phenomenon of contractibility.

That multiple definitions of randomness successfully fill both of these roles strongly suggests that no single definition of randomness captures everything that mathematicians have taken to be significant about the concept of randomness. There are settings in which one definition of randomness is adequate and all others are inadequate (as illustrated by the calibrative role), but there is no definition that is adequate for all such settings (as illustrated by both the calibrative and the limitative roles).

This, I claim, gives us good reason to accept the No-Thesis Thesis. For given a definition of randomness $\mathscr{D}$ that purportedly captures the prevailing intuitive conception of randomness, it will fail to yield the precise amount of randomness necessary and sufficient for certain purposes. Moreover, if the definition satisfies certain general conditions (which is a rather mild assumption), it will be contractible in the sense mentioned above: there will inevitably be $\mathscr{D}$-random sequences counted as non-random by a stronger definition of randomness.

If one nonetheless insists that we continue the search for an extensionally adequate definition of randomness in spite of the evidence that suggests that there is no such definition to be found, so be it. However, freed from the constraints of the search for the one correct definition of randomness, we can take these definitions at face value: the definitions that we already have are *good enough* for many purposes, and not just instrumental ones. Not only do these definitions illuminate a number of mathematical uses of the notion of randomness, but they also illustrate the limitations of the formal machinery in terms of which such definitions are given. This, I claim, is significant progress in understanding the concept of randomness.

CHAPTER 8

THE RESOLUTORY ROLE OF RANDOMNESS

8.1 Introduction

The first explicit definition of randomness for infinite sequences was given by Richard von Mises in 1919, who sought to develop a theory of probability in which probability is defined in terms of a notion of randomness, which is itself taken as primitive. The role that randomness plays in this account, which I call the *resolutory role of randomness* is an idiosyncratic one, insofar as it only appears in the context of a frequentist theory of probability such as von Mises'.[1] For on von Mises' frequentist approach, the probability of an event is the limiting relative frequency of that event in an infinite sequence of events, provided that the sequence is sufficiently random. But why not define probability in terms of *any* sequence, random or not? Here is where the resolutory role of randomness comes into play: without some requirement of randomness, there are certain problems in the probability calculus that cannot be solved on von Mises' approach.[2] In other words, the role that randomness plays in

[1]Thus, for instance, this role has no connection to Venn's frequency theory, as Venn's definition does not involve random sequences, nor does it have a place in Kolmogorov's finite frequency theory.

[2]The phrase "the probability calculus" or "the calculus of probability" is used by von Mises to refer to an exact theory of probability (see, for instance, [Mis51], p. 166, in which von Mises states this explicitly).

von Mises' theory of probability is to guarantee that every problem of the probability calculus can be solved within his theory, an ideal of completeness that I call the *resolutory ideal of completeness*.

In this chapter, the bulk of my effort will be directed towards locating the resolutory role of randomness in von Mises' theory of probability. Having sufficiently outlined the resolutory role, I will then lay out two views as to how a definition or family of definitions of randomness might adequately fill the resolutory role, thereby attaining the resolutory ideal of completeness. In particular, whereas von Mises' held that the resolutory ideal could not be attained by any *single* definition of randomness, fixed once and for all but rather by a family of definitions, Alonzo Church suggested that we should not seek to attain the resolutory ideal in full generality (i.e., we should not seek a definition of randomness that would enable to solve, in principle, *all* problems of the probability calculus), but instead he held that a restricted version of von Mises' definition *could* attain a correspondingly restricted version resolutory ideal (thereby enabling us to solve all problems in a proper subclass of the class of problems of the probability calculus).

The outline of the chapter is as follows. First, in Section 8.2, I outline the basic features of von Mises' definition of probability as based on the notion of "collective", which von Mises' took to be a formalization of the notion of random sequence. In Section 8.3, I consider two important objections to von Mises' definition of probability and randomness raised by his contemporaries that eventually led von Mises to modify his account, objections that (i) the theory of collectives is either trivial or inconsistent and (ii) that collectives are undefinable. In Section 8.4, I consider the most well-developed response to these objections, given by Abraham Wald, who proved the relative consistency of collectives, a result that in turn led to the above-mentioned modification of von Mises' account. Next, in Section 8.5, I discuss the way in which one solves problems of the probability calculus on von Mises' account. To this end, I will lay out some of the formal machinery of von Mises' account, namely four fundamental operations that allow for the solution of problems of the probability calculus, and explain how these operations allow for the solution of a problem of the probability calculus. Moreover, I will highlight the role that randomness plays in obtaining such a solution. Lastly, in Sections 8.6 and 8.7, I highlight the resolutory ideal of completeness and outline the two accounts of the resolutory ideal as discussed above, that of von Mises and that of Church.[3]

8.2 Von Mises' Account of Probability

8.2.1 Motivating von Mises' Definition

In 1919, von Mises presented what he considered to be a "scientifically adequate" definition of probability in his paper "Grundlagen der Wahrscheinlichkeitsrechnung" [vM19]. For von Mises, this meant that his definition of probability would be formed

[3] The main source for the following discussion of von Mises' work is the second English edition of von Mises' book *Probability, Statistics, and Truth*, which is the translation of the third German edition, published in 1951. Although von Mises spelled out his ideas in a number of places (see, for instance, [vM19], [vM31], [vM41], [vM46] and [vM64]), *Probability, Statistics, and Truth* is perhaps the best source for the philosophical motivations of von Mises' theory of probability for two reasons. First, the book consists of a series of lectures given to a general audience, and as such it is more conversant with philosophical issues than his more technical work. Second, the book likely contains von Mises' views in their most fully matured form, as third German edition contains a number of changes to the second German edition, including his responses to criticisms of his ideas as they appeared in earlier work, and was published just two years before his death in 1953.

following "a method of forming and defining concepts" that "has been developed by the exact sciences" and "which shows us the way clearly and with certainty" ([vM81], p. 3). In following this method, one stipulates the scope of a concept such as probability, and then the merits of that stipulation are judged by determining the extent to which "it is agreement with what we generally regard as the purpose of science" ([vM81], pp. 6-7). Precisely what von Mises takes this to mean is unclear,[4] but for our purposes, it is sufficient to observe the stipulative character of von Mises' approach, for in applying this method to provide a definition of probability, von Mises stipulates that a scientifically adequate definition of probability is only applicable in those cases in which we "have a practically unlimited sequence of uniform observations" ([vM81], p. 11). It follows that single-case probabilities, assigned to events that do not occur in a sequence of trials (or at least are not considered as part of a sequence of events), fall outside of the scope of von Mises' account.

Von Mises enforces this restriction of probability to those events that occur in

[4]The most detailed description of this method of concept formation as provided by von Mises is the following:

> [I]n the first place, the content of a concept is not derived from the meaning popularly given to a word, and it is therefore independent of current usage. Instead, the concept is first established and its boundaries are purposely circumscribed, and a word, as a suitable kind of label, is affixed later. In the second place, the value of a concept is not gauged by its correspondence with some usual group of notions, but only by its usefulness for further scientific development, and so, indirectly, for everyday affairs ([vM81], p. 4).

This description leaves much to be desired. For instance, "usefulness for further scientific development" is not a very clear criterion of value for a concept, nor is it clear that this notion of usefulness, whatever it might amount to, should be the only criterion according to which one judges the value of a concept. Further, one might worry that von Mises drives too large a wedge between the everyday uses of a concept and the scientific uses of that concept: Why can't there be a substantial overlap between the two? I will set these issues aside, for to pursue them would take us too far from the task at hand.

potentially unlimited sequence of trials by defining probability in terms of what he calls *collectives.* Roughly, a collective is (i) an infinite sequence $X = x_0x_1x_2\ldots$ consisting of what von Mises refers to as *attributes* (where each of the attributes x_i belongs to a fixed collection A) that (ii) satisfies certain conditions of randomness (which will be specified shortly). Thus, given an attribute $a \in A$, we can only determine the probability of a insofar as it is a member of some collective.[5]

Now, given a collective X made up of the attributes from A (which I'll write as $X \in A^\omega$), we determine the probability of an attribute a in X by computing the relative frequency of a in X, where the relative frequency of an attribute in a collective is defined as follows. Let $X = x_0x_1x_2\ldots$ be a collective in which the attributes $a_0, a_1, \ldots, a_n, \ldots$ occur (where the collection of attributes can be either finite or infinite); that is, let X be an infinite sequence of elements from the set $A = \{a_0, a_1, \ldots, a_n, \ldots\}$. Then the relative frequency of the attribute a_j in the initial segment of X of length n is the value

$$\frac{\#\{i < n : x_i = a_j\}}{n}.$$

To compute the probability of the attribute a_j in the collective X, we compute the limit of the relative frequency of a_j in the initial segments of X. That is, the

[5]It's not altogether clear what von Mises takes an attribute to be: in some cases, he speaks of attributes as observations or events, but in other cases, attributes seem to function as labels that we affix to given events. The most reasonable approach is to consider events as attribute-instances; strictly speaking, we don't consider the probability of a given attribute, but rather the probability that an event will *instantiate* that attribute. I am indebted to Antony Eagle for this point.

probability of the attribute a_j in the collective X is

$$p_j = \lim_{n\to\infty} \frac{\#\{i < n : x_i = a_j\}}{n}, \tag{8.1}$$

which we'll refer to as *the limiting relative frequency of* a_j *in* X. In this way, we get a probability distribution on the collection $A = \{a_0, a_1, \ldots, a_n, \ldots\}$.

For von Mises, then, probability is limiting relative frequency. This much of von Mises' account was not novel, for as he notes, the frequency theory of probability had already appeared in previous works, most notably in John Venn's 1866 *The Logic of Chance* [Ven66]. What *is* novel about von Mises' frequentist account of probability is the role that randomness plays in his account. Specifically, von Mises required that collectives satisfy a condition of randomness, which is enforced by the second of his two axioms of collectives. Let us thus turn to von Mises' two axioms of collectives.

8.2.2 Von Mises' Axioms of Collectives

As motivation for von Mises' first axiom of collectives, note that there are some infinite sequences of events consisting of attributes for which the limit in (1) above does not exist.[6] It thus follows that the probabilities of the attributes in those sequences are not defined. For an infinite sequence to count as a collective, it must

[6]For example, let $A = \{0, 1\}$ and let X be the sequence formed inductively as follows: At stage 1, let the first value of X be 0. If σ is the initial segment of X formed at stage k, then at stage $k+1$ there are two cases to consider, depending on whether k is odd or even. If $k+1 = 2n$ for some n, then we add $2|\sigma|$ 1's to the end of σ, and if $k+1 = 2n+1$ for some n, then we add $2|\sigma|$ 0's to the end of σ. Thus, we have

$$X = 011\overbrace{000000}^{6}\underbrace{111111111111111111}_{18}\overbrace{000000000000000000000000000000000000000000000000000000}^{54}\ldots$$

satisfy von Mises' first axiom of collectives:

(VM_1) If $X = x_0x_1x_2\ldots \in A^\omega$ is a collective, then for each $a_j \in A$,

$$\lim_{n\to\infty} \frac{\#\{i < n : x_i = a_j\}}{n}$$

exists.

According to von Mises, sequences satisfying (VM_1) still may not be appropriate for defining probabilities; such sequences must also satisfy a requirement of randomness. That is, the "limiting values must remain the same in all partial sequences which may be selected from the original one in an arbitrary way" ([vM81], p. 25).[7] The reason for this requirement of randomness is that it guarantees that certain calculations in the probability calculus, namely those that involve the product rule for probabilities, can be carried out. In his discussion on probability in his book *Positivism*, von Mises explicitly makes this point, writing,

---

Then for infinitely many n, we have

$$\frac{\#\{i < n : x_i = 0\}}{n} < \frac{1}{3}$$

and for infinitely many n,

$$\frac{\#\{i < n : x_i = 0\}}{n} > \frac{2}{3},$$

and so

$$\lim_{n\to\infty} \frac{\#\{i < n : x_i = 0\}}{n}$$

does not exist (and similarly for the attribute 1).

[7]Von Mises' use of the word 'arbitrary' here is misleading, since it gives the appearance that he is requiring that the limiting values remain the same in subsequences selected from the original sequence in *any way whatsoever*. But this can't be what von Mises has in mind, for as will be discussed shortly, this requirement would result in a highly defective definition of randomness. Let us set aside this point for now, and fill out the rest of the details of von Mises' condition of randomness.

> In order to derive the multiplicative law, which expresses a well-known empirical fact, in a sufficiently general form, one has to subject the collective to another axiom besides the one that requires the existence of a limiting value of the relative frequency. The second axiom demands that the succesion [sic] of the various labels or trial results within a collective is in a specific sense "random" ([Mis51], p. 170).

In order to illustrate this point, we need to consider how von Mises defines the selection of partial sequences from a given sequence, and how this factors into his definition of randomness.

As noted above, (VM_2) guarantees that the partial sequences that we select from a given sequence (which are more commonly referred to as *subsequences*) have the same limiting values as those of the original sequence. These subsequences are to be selected by what von Mises calls *place selections*. Formally, a place selection is a function $S : A^{<\omega} \to \{0, 1\}$ (where $A^{<\omega}$ is the collection of finite strings of members from A) that takes as input a finite initial segment of a sequence X and outputs a 0 or 1, indicating whether or not we should include the next value of X in the selected subsequence, which I'll write as X_S.[8] For instance, we can define a place selection that selects every odd-indexed place of the sequence X, or one that selects every place that is preceded by two consecutive a's for some fixed $a \in A$.

[8] Here are the details: For a given sequence $X \in A^\omega$, the partial sequence or subsequence of $X = x_0x_1x_2\dots$ selected by a place selection S is extracted as follows. First, we determine if x_0 is to be included in the subsequence by determining the value $S(\varnothing)$, where $\varnothing$ is the empty string. If $S(\varnothing) = 0$, we do not include x_0 in our subsequence; if $S(\varnothing) = 1$, we do. Similarly, to determine if x_n is to be included in our subsequence, where $n > 0$, having already made this determination for $x_0, x_1, \dots, x_{n-1}$, we calculate $S(x_0x_1\dots x_{n-1})$. If $S(x_0x_1\dots x_{n-1}) = 0$, then as before, we do not select x_n, but if $S(x_0x_1\dots x_{n-1}) = 1$, we do. In this way, we select a subsequence $X_S = x_{\ell_0}x_{\ell_1}x_{\ell_2}\dots$, where x_{ℓ_n} is the $(n+1)$st value of X such that $S(x_0x_1\dots x_{\ell_n-1}) = 1$. Note that X_S need not be an infinite sequence, as S might only select a finite number of values from X.

For any place selection S and sequence $X = x_0x_1x_2\ldots \in A^\omega$, once we apply S to X to extract a subsequence $X_S = x_{\ell_0}x_{\ell_1}x_{\ell_2}\ldots$, we can check to see if the limiting relative frequencies of the attributes of X_S are equal to those in X. Suppose that $X \in A^\omega$ satisfies (VM_1), so that for each $a_j \in A$,

$$\lim_{n\to\infty} \frac{\#\{i < n : x_i = a_j\}}{n} = p_j.$$

Then the limiting relative frequencies of the selected subsequence $X_S = x_{\ell_0}x_{\ell_1}x_{\ell_2}\ldots$ are the same as those of X if for each $a_j \in A$,

$$\lim_{n\to\infty} \frac{\#\{i < n : x_{\ell_i} = a_j\}}{n} = p_j.$$

In this case, we will write $\mathsf{relfreq}(X) = \mathsf{relfreq}(X_S)$ and we will say that the limiting relative frequencies of X are *invariant under the place selection S*.

There is one last step to take before we formulate the condition of randomness referred to above: we need to isolate a specific collection of place selections, referred to by von Mises as the *admissible* place selections. According to von Mises, a place selection is admissible only if the choice to select an attribute in a sequence does not depend on the value of the attribute, but only on the value of the previous attributes in the sequence, or the index of the given attribute in the sequence, or both, a condition I will call the *admissibility condition.*

Even though von Mises never provides an exact account of admissibility, he does offer several hints as to what he has in mind. First, von Mises gives a number of examples of admissible place selections: those places whose indices are (i) odd numbers, (ii) square numbers, (iii) prime numbers, (iv) divisible by 3, (v) equal to p^2+2, for any prime number p, and (vi) are preceded by the attribute 0 three indices earlier. Elsewhere, von Mises gives a more general characterization, stating that subsequences selected by admissible place selections are "selected by means of a pre-established arithmetical rule, independent of their attributes" ([vM81], p. 50). Unfortunately, von Mises doesn't specify what is to count as a rule, a problem his contemporaries seized upon in objecting to his definition, as discussed below in Section 8.3.

For the sake of our discussion, let us assume that the collection of admissible place selections is some well-defined collection of place selections.[9] Then von Mises' second axiom of collectives, the so-called condition of randomness, can be formulated as follows:

(VM_2) If $X \in A^\omega$ is a collective, then for every admissible place selection S,

$$\mathsf{relfreq}(X) = \mathsf{relfreq}(X_S)$$

whenever X_S is an infinite sequence.[10]

Von Mises also refers to (VM_2) as the "principle of the impossibility of a gambling system", the idea being that in the games of chance, no one has found a gambling system that will produce a subsequence on which a gambler using the system has a

[9] Here I make no assumption about whether these selections are computable or even arithmetical. Later in Section 8.7, we will consider the collection of computable selection rules, but for now, I merely want to assume that we've fixed some collection that does not contain all possible place selections.

[10] This last qualification is necessary since, as noted in footnote 8, for some place selections S, X_S is finite.

better chance at winning than not. In fact, von Mises appeals to this lack of gambling systems as evidence that collectives exist.[11] Any sequence satisfying (VM_1) and (VM_2) is thus a collective; further, we can define the probability of the occurrence of an attribute *in a given collective* to be the limiting relative frequency of the occurrence of the attribute in that collective.

We're now in a position to appreciate von Mises' rationale for (VM_2). The salient point is that from the assumption of the product rule for probabilities, we can derive the invariance of a collective under certain place selections. This is illustrated by the following example, which is found in Michiel van Lambalgen's dissertation, *Random Sequences*.[12] Let $X \in 2^\omega$ be a sequence generated by the tosses of a fair coin. If

(i) the limiting relative frequency of tails (represented by 1) in X is $\frac{1}{2}$ and

[11] For instance, von Mises writes,

> 'How do we know that collectives satisfying this new and more rigid requirement really exist?' Here again we may point to experimental results, and these are numerous enough. Everybody who has been to Monte Carlo, or who has read descriptions of a gambling bank, knows how many 'absolutely safe' gambling systems, sometimes of an enormously complicated character, have been invented and tried out by gamblers; and new systems are still being suggested every day. The authors of such systems have all, sooner or later, had the sad experience of finding out that no system is able to improve their chances of winning in the long run, i.e., to affect the relative frequencies with which different colours or numbers appear in a sequence selected from the total sequence of the game. This experience forms the experimental basis of our definition of probability ([vM81], p. 25).

[12] Van Lambalgen uses this example in an attempt to show that "anyone who interprets probability as relative frequency and accepts the Kolmogorov axioms [of the probability calculus] plus the product rule for (physically) independent events, also has to believe in [collectives]" ([Lam87], p. 36). Moreover, van Lambalgen takes himself to be showing with this example that collectives are necessary to explain the applicability of probability, a concern that van Lambalgen claims is shared by von Mises. While Kolmogorov was explicitly concerned with providing an explanation the applicability of probability, it's not clear that this was a problem that von Mises took himself to be addressing. See [Por12] for a detailed discussion of this matter.

(ii) if the product rule holds for two consecutive tosses,

then X is invariant under the following three place selections:

(S_1) if n is odd, choose the nth value of the sequence;

(S_2) if n is even, choose the nth value of the sequence; and

(S_3) if n is even and the $(n-1)$st value of the sequence is 1, choose the nth value of the sequence.[13]

Summing up the significance of this result, van Lambalgen writes,

> [I]nterpreting probability as limiting relative frequency and applying the deductions of probability theory to a sequence X entails assuming that X is a [collective], or at least that is has the [collective]-like properties required for the particular deduction at hand (and one is tempted to argue: since we could have chosen to perform a different calculation, e.g. that of the probability of n times heads on n consecutive tosses, X must in fact be a [collective], invariant under all admissible place selections ([Lam87], pp. 37).

There is nothing special about the three place selections specified above. If we require the product rule to hold for the product of the probabilities of, say, five tosses of the coin, then we will be able to conclude that our collective will be invariant under even more place selections. Thus the motivation for (VM_2) should be clear: if we want to have collectives that satisfy the product rule for probabilities, there *must* be some invariance under a rather large class of place selection rules, each of which is definable by a simple arithmetical formula, like S_1, S_2, and S_3 above.

[13] For the proof of this fact, see [Lam87], pp. 36-37.

Let us take stock here. First, (VM_1) guarantees that the limiting relative frequencies of the attributes in a collective exist. As discussed above, this is a necessary condition for a sequence to define a probability distribution on its attributes. Second, (VM_2) guarantees that the occurrence of attributes in a collective are sufficiently independent from one another, a property that is necessary for the probabilities defined by the collective to satisfy the product rule for probabilities.

Despite certain attractive features of von Mises' definition, it was considered to be problematic by many of his contemporaries. Specifically, it was objected that without a more precise statement of the admissibility condition, von Mises' definition would either be trivial or inconsistent. Further, it was also objected that von Mises' account was problematic due to the undefinability of collectives. Let us now consider these two objections in detail.[14]

8.3 Objections and Replies

Before I lay out the two objections raised against von Mises' account and discuss several responses to these objections, I should explain why these objections merit our attention. First, in light of the first objection we consider below, von Mises eventually modified his account, and moreover, this specific modification had consequences for von Mises' approach to the resolutory ideal of completeness, which, as discussed in the introduction, is attained by a theory of probability in which all problems of the probability calculus are solvable. Second, these objections merit our attention because one notable response to them, provided by the American mathematician A.H.

[14]There is, in fact, a third objection to von Mises account based on a theorem proved by Jean Ville. We will consider this theorem and associated objection in the next chapter.

Copeland, was explicitly rejected by von Mises, who held that Copeland's response would result in a definition of probability that could *not* attain the resolutory ideal.

With these reasons in mind, we now turn to the first of the two main objections to von Mises' account.

8.3.1 The Admissibility Objection

The first objection considered here, which I call the *admissibility objection*, is well-known among those familiar with von Mises' definition. The key idea behind the admissibility objection is this: any definition of randomness that requires the invariance of limiting relative frequencies in subsequences selected by *every* place selection is a defective one, in that it is either trivial or inconsistent. More specifically, the resulting definition is defective because the only sequences random according to this definition are those sequences that contain only one attribute that occurs infinitely often, so that all other attributes occur only finitely often (I will henceforth refer to collectives satisfying this property as *trivial* collectives).[15] Moreover, if we take as part of the theory of collectives the claim that non-trivial collectives exist, the resulting theory would thus be inconsistent.[16]

This observation was made by a number of von Mises' contemporaries, most notably Erich Kamke, who included criticisms of von Mises' axioms in his 1933 article

[15]Not even those sequences in which one attribute occurs with probability one and in which the other attributes, while occurring infinitely often, occur with probability zero (as they occur less and less frequently as we proceed along the sequence) would satisfy (VM_2) if we were to count all place selections as admissible.

[16]Most of von Mises' objectors glossed over this point, immediately concluding that von Mises' theory was inconsistent. Of course, the conclusion that the theory is trivial, although not quite as damning, still implies that von Mises' theory is, for all practical purposes, worthless.

"Über neuere Begründungen der Wahrscheinlichkeitsrechnung" [Kam33] and also in his monograph *Einführung in die Wahrscheinlichkeitstheorie* [Kam32]. Kamke's version of the argument is roughly as follows: Suppose we are given a collective X made up of the attributes in A, and suppose further that there are attributes $a, b \in A$ that occur infinitely often in X, so that X is non-trivial. Then if p_a and p_b are the limiting relative frequencies of a and b, respectively, then it follows that either $p_a \neq 1$ or $p_b \neq 1$ (or both). For the sake of argument, let us assume that $p_a \neq 1$. If we let $n_1, n_2, \ldots$ be a sequence of natural numbers such that $X(n) = a$ if and only if $n = n_i$ for some $i \in \omega$ (that is, the sequence $(n_i)_{i\in\omega}$ is the set of positions at which X has the value a), then we can define a place selection S such that S only selects elements for the subsequence at the indices $(n_i)_{i\in\omega}$. Thus when given the collective X, S extracts the subsequence $X_S = aaaaaaaaaa\ldots$, and thus (VM_2) is not satisfied, as the limiting relative frequency of a in the selected subsequence X_S is 1 and not p_a, the limiting relative frequency of a in the original sequence X.

Kamke's argument has been repeated numerous times over the years. Most notably, during a session on the foundations of probability at a conference on probability theory at the Université de Genève in 1937, Maurice Fréchet presented a variant of Kamke's argument (as well as a number of other criticisms of von Mises' definition of probability). According to Fréchet, either the collection of admissible place selections contains *all* place selections, in which case von Mises definition is "without concrete, precise significance" ("sans signification concrète precise"), or the collection of admissible place selections is a well-defined ("bien défini") subcollection of the collection of all place selections, but one that includes, for every collective X, at least one place selection that, like the above-defined place selection S, extracts from X a sequence consisting of a single attribute ([Fré38], p. 29).

8.3.2 Von Mises' Response to the Admissibility Objection

In response to Fréchet's version of Kamke's argument, von Mises argued that the collection of admissible place selections does not include every place selection, and more to the point, it doesn't contain those problematic place selections that extract from a given collective a sequence consisting of only one attribute.[17] The reason these problematic place selections are to be excluded, according to von Mises, is that the indices chosen by these place selections are *not* independent of the attribute at those indexed places. Thus, the place selection S constructed above in the discussion of Kamke's argument is to be excluded, since S chooses the nth place of the collective X whenever the attribute in the nth place of X is a.

There is, however, a problem with this response that doesn't appear to have been noted elsewhere: while it is true that the place selection rule S, which can be described by the rule "choose the nth value of X if and only if the nth value of X is n", is clearly in violation of the admissibility condition, this doesn't eliminate the possibility that there is some *other* rule that extracts the same subsequence from X that S does, but which doesn't violate the admissibility condition.

Although von Mises never explicitly rejects his initial response to the admissibility objection in print, there is reason to think that he ultimately found it to be

[17]As noted by van Lambalgen, von Mises didn't attend the Geneva conference, but he wrote a response to the arguments presented at the conference that was included in the proceedings of the conference. See ([vM38], pp. 61-62, 64-66) for more details.

unsatisfactory, namely that he later modified his theory by replacing the condition of admissibility with an alternative condition, which I will discuss in Section 8.4.

Let us now turn to the second main objection raised against von Mises' account.

8.3.3 The Undefinability Objection

The second objection, which I call the *undefinability objection*, concerns the undefinability of individual collectives. Von Mises motivates this objection in the following passage:

> A sequence of zeros and ones which satisfies the principle of randomness cannot be described by a formula or by a rule such as: 'Each element whose place number is divisible by 3 has the attribute 1; all the others the attribute 0'; or 'All elements with place numbers equal to squares of prime numbers plus 2 have the attribute 1, all the others the attribute 0'; and so on. If a collective could be described by such a formula, then, using the same formula for a place selection, we could select a sequence consisting of 1's (or 0's) only. The relative frequency of the attribute 1 in this selected sequence would have the limiting value 1, i.e., a value different from that of the same attribute in the initial complete sequence ([vM81], p. 88).

The problem von Mises raises here is this: if a given collective C is definable by some formula or rule, then by means of that formula or rule, we can select a subsequence C^* from C so that the limiting relative frequencies of C^* are not equal to those of C, thus violating (VM_2). Thus we have what one might call a *definability tradeoff*: if we desire that at least one collective is definable by some formula or rule, we need to ensure that the formula or rule in question cannot be used to define an admissible place selection. Formally, for a fixed collective $X \in 2^\omega$ (i.e. $\{0,1\}^\omega$)[18], suppose there are formulas Φ_0, Φ_1 in a fixed language $\mathcal{L}$ such that

$$\{Y : (\forall n)[Y(n) = 0 \iff \Phi_0(n)] \wedge (\forall n)[Y(n) = 1 \iff \Phi_1(n)]\} = \{X\}.$$

Then the worry is that one of the following place selections is counted as admissible:

S_0 : Given a sequence Y, select the nth value of Y if and only if $\Phi_0(n)$ holds.

S_1 : Given a sequence Y, select the nth value of Y if and only if $\Phi_1(n)$ holds.

The subsequence selected by S_0 from the original collective X consists of only 0's, while the subsequence selected from X by S_1 consists of only 1's, and thus (VM_2) is violated (unless, of course, X is trivial). More generally, if Ψ is a formula in a fixed language $\mathcal{L}$ such that $\{Y : \Psi(Y)\} = \{X\}$, the same problem will arise if the sets $\{n : \Phi_0(n)\}$ and $\{n : \Phi_1(n)\}$ can be defined in terms of Ψ, for then the place selections S_0 and S_1 will also be definable in terms of Ψ. Thus, if we require that at least one collective be definable, we must take care to ensure that any definition we provide cannot be used to define an admissible place selection in the above manner.

But why does it matter whether collectives are definable? Von Mises answers the question as follows:

> It is to this consideration, namely, to the impossibility of explicitly describing the succession of attributes in a collective by means of a formula that critics of the randomness principle attach their arguments. Reduced to its simplest form, the objection which we shall have to discuss first

[18]We restrict the number of attributes here to two, but the characterization we provide holds for any finite collection of attributes.

asserts that sequences which conform to the condition of randomness do not exist. Here, 'nonexistent' is equivalent to 'incapable of representation by a formula or rule' ([vM81], pp. 88-89).

Moreover, he adds,

> The existence or nonexistence of limiting values of the frequencies of numbers composing a sequence, say 1's and 0's, can be proved only if this sequence conforms to a rule or formula. Since, however, in a sequence fulfilling the condition of randomness the succession of attributes never conforms to a rule, it is meaningless to speak of limiting values in sequences of this kind ([vM81], pp. 89).

Thus, the undefinability objection is this: without representing a collective X by a formula or a rule, we cannot prove that the limiting relative frequencies of the attributes in X exist. That is, we cannot prove that X is a collective to begin with. To be clear, this objection is not that there are *no* collectives, but rather that we can never *prove* that a given sequence is a collective. In particular, the objection goes, if a collective X isn't representable by means of some rule or formula, then there is no way to calculate the limiting values of the relative frequencies of the attributes in X. And without providing some way to calculate these probabilities, von Mises' account cannot get off the ground.[19]

[19] I should emphasize the distinctness of the admissibility objection and the undefinability objection. Specifically, the admissibility objection is directed at the problem of the imprecise definition of admissible place selections in (VM_2), whereas the undefinability is directed at the problem of determining whether a sequence satisfies (VM_1) and (VM_2) to begin with. However, one might hold that collectives fail to be definable *in virtue of* the imprecise notion of admissibility, in which case it would be reasonable to hold that the imprecision of admissibility in von Mises' account results in its being vulnerable to both objections.

8.3.4 Von Mises' Response to the Undefinability Objection

For von Mises, this worry about representing collectives for the purposes of calculating probabilities, as well as general worries concerning the definability tradeoff, aren't really problematic at all. In von Mises' view, "In a problem of probability calculus, the data as well as the results are probabilities", a view summarized by the pithy phrase "the beginning and the end of each problem must be probabilities" ([vM81], p. 32). More specifically, von Mises held that "the exclusive purpose of [his] theory is to determine, from the given probabilities in a number of initial collectives, the probabilities in a new collective derived from the initial ones" ([vM81], p. 32).

To illustrate this point, von Mises takes the task of deriving probabilities from initial probabilities to be analogous to the solving of problems in geometry, wherein, starting with the data of certain known quantities (for instance, the lengths of the sides of a right triangle), one determines the unknown quantities (say, the angles between the sides of the triangle). Moreover, he adds, "The source from which these values are known is irrelevant, in the same way in which the source of knowledge of the geometrical data is irrelevant for the solution of the geometrical problem in which these data are used" ([vM81], p. 32). Elaborating on this point, von Mises writes,

> A mathematician teased with the question, 'Can you calculate the probability that I shall miss the next train?', must decline to answer it in the same way as he would decline to answer the question, 'Can you calculate the distance between these two mountain peaks?'—namely, by saying that a distance can only be calculated if other appropriate distances and angles are known, and that a probability can only be determined from the knowledge of other probabilities on which it depends ([vM81], p. 32).

Thus, on von Mises' account of probability, which one might refer to as a *transformational account of probability*, what is important is how initial collectives are transformed into new collectives, not where the initial collectives come from.[20] Thus, on von Mises' account, worries about the definability of collectives simply miss the mark.[21]

8.3.5 Copeland's Response to the Two Objections

The responses to the two objections as discussed above were not the only responses available to von Mises. For instance, von Mises could have offered another response to these objections, based on the work of the American mathematician A.H. Copeland. As von Mises observes,

> One way to avoid all these difficulties would seem to consist in effectively

[20]Von Mises actually anticipates this objection in his 1919 *Grundlagen*. There he wrote the following, as quoted in *Probability, Statistics, and Truth*,

> [T]he existence of a collective cannot be proved by means of the actual analytical construction of a collective in a way similar, for example, to the proof of existence of continuous but nowhere differentiable functions, a proof which consists in actually writing down such a function. In the case of the collective, we must be satisfied with its abstract "logical" existence. The proof of this "existence" is that it is possible to operate with the concept of a collective without contradictions arising ([vM81], p. 88).

Although von Mises later claims he would "perhaps express this thought in different words", he does assert that "the essential point remains", by which he means that the formula defining a collective cannot be used to define a place selection rule, which is precisely the definability tradeoff discussed above. As for the notion of 'abstract logical existence', I don't know what to make of it, although the fact that von Mises would appeal to it does appear to put pressure on those who seem to overemphasize constructivist tendencies in von Mises' thought (most notably, van Lambalgen; see e.g. [Lam87], pp. 9, 29).

[21]I should add that von Mises was also not concerned with proving that sequences we encounter in nature really are collectives. For as discussed briefly in Subsection 8.2.2, experience with games of chance shows that collectives exist.

> restricting the postulate of randomness. Instead of requiring that the limiting value of the relative frequency remain unchanged for *every* place selection, one may consider only a predetermined group of place selections ([vM81], p. 89).

In particular, Copeland suggested that the collection of all place selections be restricted to the collection of place selections $S_{a,b}$ that select every place whose index is of the form $an+b$ for some $n \in \omega$, a collection I will write as $\mathscr{S} = \{S_{a,b}\}_{a,b\in\omega}$. Further, Copeland defined the collection of *Bernoulli sequences* to be the collection of sequences that satisfy (VM_1) and a modified version of (VM_2), where the admissible place selections are replaced with the place selections in $\mathscr{S}$.

How would such a restriction allow for responses to the admissibility objection and the undefinability objection? The answer is that Copeland's restriction results in a definition of collectives that is neither trivial nor inconsistent, as he (and von Mises, independently) showed that an individual Bernoulli sequence can be explicitly constructed. Thus, the admissibility objection is deflected, as the place selections that feature in the counterexamples offered by Kamke and Fréchet, which lead to violations of (VM_2), are not included in $\mathscr{S}$. Moreover, the undefinability objection is also deflected, as at least one collective satisfying the condition of randomness (relative to the selection rules in $\mathscr{S}$) is definable.

Despite these apparent virtues,[22] von Mises did not accept Copeland's restriction to a predetermined collection of admissible place selections, nor *any* restriction to a

[22] One further virtue: Copeland also showed that Bernoulli sequences are *free from aftereffect*, meaning that any subsequence selected by any rule of the following form has the same limiting values as the original sequence: for $k \in \omega$ and $i \in \{0, 1\}$, given a sequence X, $S_{k,i}$ selects the $(n+k)$th value of X if and only if the nth value of X is i. This property was independently introduced by Karl Popper, and played an important role in Reichenbach's definition of randomness. See [Cop28], [Rei32], or [Pop35] for details.

predetermined collection, for reasons that will be offered in Section 8.6. However, Abraham Wald offered a solution to both objections that von Mises *did* find to be acceptable: Not only did Wald prove the consistency of the collectives, but he also identified general conditions under which collectives can be explicitly constructed.

8.4 Wald's Theorems and von Mises' Modified Account

Wald's response to the admissibility objection and the undefinability objection can be found in his 1937 paper "Die Widerspruchsfreiheit des Kollektivbegriffes der Warhrscheinlichkeitsrechnung" [Wal37], presented at Karl Menger's colloquium in Vienna in 1937.[23] Before we consider the details of Wald's response, let us fix some notation and terminology that will be useful for the ensuing discussion. First, for the sake of simplicity, let us restrict to a finite set of attributes $A = \{a_0, a_1, \ldots, a_k\}$. Second, if A is a collection of attributes and $\mathscr{S}$ is a collection of place selections, then $\mathcal{C}(\mathscr{S}, A)$ is the collection of all sequences in A^ω that are invariant under the place selections in $\mathscr{S}$. Third, in the case that $A = 2 = \{0, 1\}$, I will write $\mathcal{C}(\mathscr{S}, A)$ as $\mathcal{C}(\mathscr{S})$. Lastly, given $X \in \mathcal{C}(\mathscr{S}, A)$, when I refer to the *probability distribution on A induced by X*, I simply mean the probability distribution on A given by the limiting values of the relative frequencies of the attributes in X.

[23] Wald also presented a shorter version of his paper, entitled "Die Widerspruchsfreiheit des Kollektivbegriffes", at the Geneva conference on the foundations of probability in 1937, which was later published with the rest of the conference proceedings in *Actualités Scientifiques et Industrielles* [Wal38]. This is the version of the paper that we will draw from in the ensuing discussion.

8.4.1 Wald's Two Problems and their Solutions

In response to the admissibility objection and the undefinability objection, Wald formulated the following two problems:

Problem 1. Given a collection of attributes A (finite or infinite) and a probability distribution μ on A, what conditions must the collection $\mathscr{S}$ of place selections satisfy so that a collective in $\mathcal{C}(\mathscr{S}, A)$ exists and induces a probability distribution on A that is identical to μ?[24]

Problem 2. What conditions must be satisfied by a countable collection $\mathscr{S}$ of place selections and a probability distribution μ on A so that a collective in $\mathcal{C}(\mathscr{S}, A)$ that induces a probability distribution equal to μ can be constructively defined?

Observe that any solution to Problems 1 and 2 thereby yields a response to the admissibility objection and the undefinability objection, respectively. And this is precisely what Wald achieves.

8.4.1.1 Wald's Solution to Problem 1

First, in response to Problem 1, Wald proves four theorems[25], the first of which is:

[24] Although Wald appears to be asking for necessary conditions in the statement of Problem 1, as we'll see, the answers he provides to Problem 1 come in the form of sufficient conditions. Nothing of importance appears to hinge on this matter.

[25] The other three theorems of Wald's are more general than this first one: Wald's Theorem II addresses the case that A is infinite (either countably or uncountably infinite), his Theorem III concerns the case that A is a measure space of arbitrary cardinality in terms of which is defined a σ-algebra of measurable sets, and his Theorem IV is a special case of Theorem III in which the measure space in question is n-dimensional Euclidean space equipped with the algebra of Peano-Jordan measurable sets. (Roughly, a set is Peano-Jordan measurable if it can be approximated from the "inside" and the "outside" by finite unions of cubes. See, for instance, [Bur98], pp. 20-22

Theorem (Wald's Theorem I)**.** If A is a finite collection of attributes, $\mathscr{S}$ is a set of countably many place selections, and μ a probability distribution on A, then there are continuum many collectives in $\mathcal{C}(\mathscr{S}, A)$ that induce a probability distribution on A that is identical to μ.[26]

Wald's Theorem I guarantees the existence of a collective relative to this restricted collection, and hence von Mises' account is neither trivial nor inconsistent as long as we are willing to restrict the collection of admissible place selections to some countable collection. But is there any reason to think this restriction is problematic? According to Wald, there is not; in practice, we only apply at most countably many place selections to a given collective.[27]

Further, as von Mises held that a place selection should always be given in terms of an arithmetical rule, this is consistent with Wald's restriction to countable many place selections. For according to Wald, "The concept of mathematical law can only

---

for details.) In each of these cases, the general outcome is the same: for each countable collection of place selections, there is a collective that induces the probability distribution over the attribute space in question.

[26] A similar result was also independently proved by the American probabilist J.L. Doob in 1936, [Doo36], who proved that for every countable collection of place selections $\mathscr{S}$, the set of collectives with respect to $\mathscr{S}$, $\mathcal{C}(\mathscr{S})$ has Lebesgue measure one.

[27] As Wald notes, "[I]t is clear that in each specific task of probability at most countable number of selection rules are actually used." In the original German: "Denn es ist klar, dass in jeder konkreten Aufgabe der Wahrscheinlichkeitsrechnung hchstens abzhlbar viele Auswahlvorschriften tatschlich verwendet werden." ([Wal37], p. 86) In addition, Wald claims that any weakening, presumably to uncountably many place selections, would be "of little interest" ("kaum von Interesse wäre") and elsewhere that it would be "of little importance" ("kaum von Bedeutung wäre"). ([Wal37], p. 86) Von Mises agrees with Wald's conclusion, noting in a discussion of Wald's theorems, "I know of no problem in probability in which a sequence of attributes is subjected to more than an enumerably infinite number of place selections, and I do not know whether this is even possible" ([vM81], p. 92).

be defined meaningfully within a formal logic, and consequently there are evidently only countably many mathematical laws."[28] ([Wal37], p. 86) As Wald argues, since there are at most countably many mathematical laws, and since every admissible place selection is given by a mathematical law, it follows that there are at most countably many admissible place selection rules.[29,30]

8.4.1.2 Wald's Solution to Problem 2

In response to Problem 2, Wald proves two additional theorems, one in the case that the collection A of attributes is finite (Theorem V), and the other in the case

[28] In German: "Der Begriff des mathematischen Gesetzes kann ja nur innerhalb eine formalisierten Logik sinnvoll definiert werden und mithin gibt es offenbar nur abzählbar vielen mathematische Gesetze."

[29] As an example of a class of place selections given by mathematical laws definable in a formal logic, Wald suggests that we can consider the collection of place selections definable in *Principia Mathematica*.

[30] Wald's invocation of logic is particularly noteworthy, for as Martin-Löf writes, "When discussing the restriction to denumerably many selection rules, Wald contributes a very decisive argument in which for the first time is felt the direct influence from mathematical logic" ([ML69b], p. 29). But one might wonder here why Wald doesn't consider also uncountable languages. The reason appears to be that he could not establish the consistency of collectives defined in terms of an uncountable collection of place selections. For after assimilating Wald's insights into his own theory, von Mises' discusses this possibility, writing

> I know of no problem in probability in which a sequence of attributes is subjected to more than an enumerably infinite number of place selections, and I do not know whether this is even possible. Rather, it might be in the spirt of modern logic to maintain that the total number of all the place selections *which can be indicated* is enumerable. Moreover, it has in no way been proved that if a problem should require the application of a continuously infinite number of place selections this would lead to a contradiction. This last question is still an open one ([vM81], p. 93, emphasis in the original).

This question was later answered in the positive by Kamae in [Kam73]. It also follows from work of van Lambalgen that every Martin-Löf random sequence is invariant under the place selections in some uncountable collection of place selections. See, for instance, [vL90].

that the collection is infinite (Theorem VI). As before, we restrict our attention to the finite case.

In proving Theorem V, Wald isolates conditions under which collectives are "constructively definable". For Wald, a sequence is constructively definable if there is a procedure or method ("Verfahren") such that, given the natural number i, computes the ith value of the sequence in a finite number of steps.[31] Additionally, Wald's solution to Problem 2 involves several other constructive defined objects:

- a constructively defined countable collection of place selections, where a countable collection of place selections $\mathscr{S} = \{S_i\}_{i \in \omega}$ is constructively defined if there is some procedure that, given any n and initial segment σ of a sequence X, can calculate $S_n(\sigma)$ in a finite number of steps (thereby telling us whether the value that immediately follows this initial segment σ is to be selected or not); and

- a constructively defined probability distribution, where a probability distribution μ on $A = \{a_0, a_1, \ldots, a_k\}$ is constructively defined if there is a procedure that given i computes the value $\mu(a_i)$ in a finite number of steps.[32]

[31] We shouldn't think that Wald has the formal notion of computability in mind here, but rather the pre-formal notion that was in the air, so to speak, just prior to the notion's formalization by Church and Turing. For instance, Wald uses the same word for 'procedure' as Hilbert does in his statement of his famous Tenth Problem. In Hilbert's words: "Man soll ein Verfahren angeben, nach welchem sich mittles einer endliche Anzahl von Operationen enscheiden lässt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist." ([Hil01], p. 216) Compare this to Wald: "Die Merkmalfolge $\{m_i\}$ ($i = 1, 2, \ldots,$ ad inf.) heisse konstruktiv definiert, falls ein Verfahren vorliegt, das fr jede natrliche Zahl i den Merkmalwert m_i in endlich vielen Schritten tatschlich zu berechnen gestattet" ([Wal37], p. 87).

[32] Actually, Wald defines what it means for a probability distribution to be constructively defined when the collection of attributes A is infinite and then makes the further assumption that every

Wald's Theorem V can thus be stated as follows:

Theorem (Wald's Theorem V)**.** If A is a finite collection of attributes, $\mathscr{S}$ is a constructively defined collection of countably many place selections, and μ is a constructively defined probability distribution on A, then there is a constructively defined collective $X \in \mathcal{C}(\mathscr{S}, A)$ that induces a probability distribution on A that is identical to μ.

Note that this result doesn't entirely address the undefinability objection, for although Wald isolates conditions under which collectives are constructively definable, it only follows that *some* collective is constructively definable. For if the concern is that without a formula for defining a given collective X, then X cannot be used in calculating probabilities, then Wald's result would not, in general, address this concern.

8.4.2 Von Mises' Modified Account

Von Mises apparently found Wald's response to the admissibility objection to be noteworthy, given that he modified modified his theory, replacing the admissibility condition with an alternative requirement of invariance inspired by Wald's Theorem I. In later presentations of von Mises' account, such as those found in later editions

probability distribution on a finite collection of attributes is already constructively defined. As Martin-Löf notes, this is a mistake, since there are probability distributions on a finite set of attributes that assign to some of element of the set a probability that is not constructively definable (and not even computable for that matter, where $x \in \mathbb{R}$ is computable if there is a computable sequence of rational numbers $\{q_i\}_{i \in \omega}$ and a computable function $f : \omega \to \omega$ such that $|x - q_n| \leq 2^{-f(n)}$ for every $n \in \omega$): "Wald erroneously considered any probability distribution over a finite sample space to be constructively defined, forgetting that we must require the probability masses associated with the various outcomes to be computable real numbers" ([ML69b], p. 30).

of *Probability, Statistics, and Truth* and in his textbook *Mathematical Theory of Probability and Statistics*, von Mises drops the condition of admissibility and instead opts for a relative approach: For a countable collection $\mathscr{S}$ of place selections, a sequence X is a collective relative to $\mathscr{S}$ if the limiting relative frequencies of the attributes in X exist and X is invariant under the place selections in $\mathscr{S}$. Moreover, this move is justified by Wald's Theorem I: for each such countable collection $\mathscr{S}$ of place selections, there is some $X \in C(\mathscr{S}, A)$.

While this move addresses the admissibility objection, it seems to raise another problem: which countable collection $\mathscr{S}$ of place selections should we use to define collectives? Von Mises' response is subtle. According to the later version of von Mises' theory, as laid out in his textbook *Mathematical Theory of Probability and Statistics*, for each problem in the probability calculus that we are attempting to solve, there is some countable collection $\mathscr{S}$ of place selections such that we can solve the given problem by means of collectives that are invariant under the place selections in $\mathscr{S}$. Thus, von Mises writes,

> We obtain a concrete idea of the set $[\mathscr{S}]$ of place selections which are supposed not to change the frequency limits if we visualize $[\mathscr{S}]$, for example, as follows: in $[\mathscr{S}]$ are contained all those place selections which present themselves in a particular problem under consideration ([vM64], p. 12).

But this raises two further questions. What role do place selections play in solving problems of the probability calculus? And is there some collection $\mathscr{S}$ that allows us to solve all problems of the probability calculus, or do we instead consider different collections $\mathscr{S}_{\mathcal{P}}$ indexed by different problems $\mathcal{P}$?[33]

[33] We might further ask: What does it mean for a place selection to "present itself"? And what

As a first step in answering these two questions, we must look more closely at the way in which one solves a problem of the probability calculus on von Mises' account.

8.5 Von Mises' Approach to Solving Problems in the Probability Calculus

As discussed in Subsection 8.3.4, on the transformational account of probability, the beginning and end of each problem of the probability calculus are probabilities. Thus, in solving a problem of the probability calculus, we need a method for transforming the initial probabilities into final probabilities, which requires transforming the initial collectives in terms of which the initial probabilities are given into the derived collectives in terms of which the final probabilities are given. To do so, von Mises makes use of four fundamental operations on collectives.

8.5.1 Four Fundamental Operations on Collectives

The four fundamental methods or operations identified by von Mises for deriving new collectives from initial ones are (1) *selection*, (2) *mixing*, (3) *partition*, and (4) *combination*. Let us consider each operation in turn.

8.5.1.1 Selection

The first operation, *selection*, is a familiar one: it is simply the application of an admissible place selection to a collective to extract a new collective. Note, however, that if selection is to transform collectives into further collectives, then we must

role does the problem in question play in this presentation?

require a closure condition on the collection $\mathscr{S}$, namely that for every $S_0, S_1 \in \mathscr{S}$, there is some $S_2 \in \mathscr{S}$ such that for every collective X, $(X_{S_0})_{S_1} = X_{S_2}$.

8.5.1.2 Mixing

The operation of *mixing* involves defining new attributes as the collection of (or disjunction of) certain attributes in the original sequence, resulting in a new collective composed of the new attributes. More precisely, given a collective $X \in A^\omega$, where $A = \{a_0, \ldots, a_n\}$, we define $\widehat{A}$ to be a partition[34] of A, so that $\widehat{A} = \{B_0, B_1, \ldots, B_k\}$, where (i) for each $a_i \in A$ there is a unique $j \leq k$ such that $a_i \in B_j$ and (ii) $\bigcup_{j \leq k} B_j = A$. Then we apply the operation of mixing to X by replacing each a_i by the set B_j to which it belongs, thus transforming $X \in A^\omega$ into a collective $\widehat{X} \in (\widehat{A})^\omega$.[35] As a result of this operation, the distribution of the new sequence is determined by adding the probabilities of the original attributes, so that the probability of B_j in the sequence $\widehat{X}$ is just the sum of the probabilities of the attributes $a_i \in B_j$ in X, thus establishing the addition rule of the probability calculus.[36]

[34]Although we define the operation of mixing by forming a partition of the set A of attributes, this operation should not be confused with the operation of partition.

[35]More precisely, we replace each a_i by a label 'B_j', which corresponds to the set B_j that contains a_i. Henceforth, I will conflate B_j and 'B_j' when I discuss the operation of mixing, but the reader should bear this point in mind.

[36]According to the addition rule, the probability of the finite union of a collection of mutually exclusive events is the sum of the probabilities of those events.

8.5.1.3 Partition

The operation of *partition* is carried out whenever we simply delete from a collective X every occurrence of certain attributes. Formally, we first partition the collection of attributes A into two disjoint subsets A_0 and A_1. Next, given a collective X, we define the collective $X^* \in (A_1)^\omega$ derived by partition to be the collective derived from X by eliminating all occurrences of those attributes in A_0. Von Mises is careful to point out that this operation is not the same as selection; whenever we apply selection to a collective, we select places from that collective in accordance with an admissible place selection, irrespective of the attributes in those places, while whenever we apply partition to a collective, the new collective is derived by selecting all elements belonging to A_1 and none of those belonging to A_0, an operation that is not, in general, given in terms of an admissible place selection. Additionally, used in tandem with mixing, partition can be used to define conditional probability.[37]

8.5.1.4 Combination

Combination is an operation that, unlike the other three operations, derives a new collective from *two* initial collectives. The idea is that given two collectives X and Y, we use one collective, say X, called the *sampling collective*, to extract a subsequence from the other collective Y, which is called the *sampled collective*. To do

[37]Let X be a collective composed of attributes from $A = A_0 \cup A_1$ representing a certain collection of events, where A_0 and A_1 form a partition of A, and let $a \in A_1$ be some fixed attribute. The conditional probability of the occurrence of the event represented by a, given that we know some event represented by an attribute in A_1 has occurred, is the limiting value of the relative frequency of a in the collective $X^* \in (A_1)^\omega$ formed by partition divided by the limiting value of the relative frequency of the attribute A_1 in the collective $\widehat{X} \in \{A_0, A_1\}^\omega$ formed by mixing.

so, we place X and Y into a one-to-one correspondence and then choose an attribute a that appears in X infinitely often to serve as the basis for selecting values in the sampled collective Y. At those positions where a occurs in the sampling collective X, we select the elements from the sampled collective Y, resulting in a new collective, which I'll denote $Y_{X,a}$.

Let us note two consequences of the operation of combination. First, by means of the operation of combination, von Mises defines what it means for two collectives to be independent of one another: Given collectives X and Y, Y is *independent* of X if no matter how we use X to sample a subsequence from Y, the limiting values of the relative frequencies of the derived sequence are the same as those in Y. That is, Y is independent of X if for all $a \in A$ that appears in X infinitely often, we have $\mathsf{relfreq}(Y_{X,a}) = \mathsf{relfreq}(Y)$. This relation of independence is symmetric, a fact that follows from the second consequence of the operation, namely that combination allows von Mises to derive the product rule of the probability calculus, according to which the probability of the conjunction of two independent events is the product of the probabilities of those events.

In order to explain how these operations are used to solve problems in the probability calculus, we need to look more carefully at how problems and their solutions can be formulated on von Mises' account.

8.5.2 Problems and Solutions

Although von Mises' does not give a precise definition of a "problem of the probability calculus", he provides enough clues for us to begin to sharpen the notion.

First, given that the beginning and end of a problem of the probability calculus are probabilities, and given that probabilities are always given in terms of collectives, it appears that a problem of the probability calculus is first given by an initial collective or a collection of initial collectives. Moreover, as von Mises says that the amount of invariance for the collectives is determined by the problem in question, and that the invariance is never determined by more than countably many place selections, we thus have: A problem $\mathcal{P}$ of the probability calculus is first given by

(i) a collection $\mathcal{C}$ of initial collectives, and

(ii) a collection of place selections $\mathscr{S}$ such that every collective in $\mathcal{C}$ is invariant under the place selections in $\mathscr{S}$.

But this can't be all that constitutes a problem of the probability calculus for von Mises; there must be some further demand, such as 'Find the probability that event E occurs under the following conditions'. Thus we include a further component, *the demand of the problem* $\mathcal{P}$, which we will denote $\mathcal{D}$. Put simply, a demand $\mathcal{D}$ is simply a request for the probability of a certain event E or collection of events $\{E_1, E_2, \dots\}$.

To sum up, we will consider a problem $\mathcal{P}$ to be given by a triple $(\mathcal{C}, \mathscr{S}, \mathcal{D})$, where $\mathcal{C}$ is some collection of initial collectives, $\mathscr{S}$ is a collection of place selections such that the collectives in $\mathcal{C}$ are invariant under the every $S \in \mathscr{S}$, and $\mathcal{D}$ is a demand for the probability of a certain event E or collection of events $\{E_1, E_2, \dots\}$ to be determined.[38]

[38]This picture can be made even more precise, for instance, by considering the conditions under which the events in question arise, or how these events are to be related to the attributes in the initial collectives. Still, for our purposes, the account we present here should suffice.

Now the events E_i referenced in the demand $\mathcal{D}$ may belong to a collection of attributes other than those which compose the initial collectives. To solve the problem $\mathcal{P}$, then, one must determine how to arrive at the correct derived or final collective, one that contains each E_i as an attribute, so that one can read off the probability of E_i in the derived collective, thus satisfying the demand of $\mathcal{P}$. Hence, a solution to $\mathcal{P}$ is given by a sequence $O_1, O_2, O_3, \ldots$ of some combination of the four fundamental operations discussed in Subsection 8.5.1 (selection, mixing, partition, and combination) such that by successively applying $O_1, O_2, O_3, \ldots$ to the initial collectives in $\mathcal{C}$, we will arrive at the correct derived collectives, which will contain the probability of the events E_i (in this case, let us say that we have *met the demand* $\mathcal{D}$). Moreover, any use of selection must be from the collection $\mathscr{S}$, so that we have $O_i \in \mathscr{S} \cup \{\text{mixing, partition, combination}\}$ for each i.[39]

8.6 An Ideal of Completeness

We've now addressed the question as to the role that place selections play in solving problems of the probability calculus. However, we have yet to determine whether there is some collection $\mathscr{S}$ that allows us to solve all problems of the probability calculus.

[39]There is a problem with this analysis, however. If we require that the initial collectives in $\mathcal{C}$ are invariant under the selections in $\mathscr{S}$, in applying combination to two initial collectives, it is not apparent that the resulting collective will be invariant under the selections in $\mathscr{S}$. But the worry is that we may need to apply a selection to this combined collective, but we have no guarantee that the combined collective is invariant under the selection we need to apply. To circumvent this problem, it appears that we must require that the collection $\mathscr{S}$ be such that *any* collectives derived from the initial collectives are invariant under the selections in $\mathscr{S}$ (and not just the initial collectives).

8.6.1 Formulating the Resolutory Ideal

As we just discussed, in order to solve certain problems of the probability calculus, we need to apply a sequence of operations to an initial collective. However, unless the initial collective is sufficiently invariant under place selections, we might apply the operation of selection at some stage in our solution and produce a non-collective. In other words, unless our initial collectives are sufficiently random, there will be problems of the probability calculus that we are unable to solve. This, then, is the *resolutory role of randomness*: collectives must be sufficiently random in order to solve certain problems of the probability calculus.[40]

Given that von Mises judges the merits of his account based on its usefulness for scientific purposes, if there are problems of the probability calculus that it cannot solve, this would prove to be a defect of the theory. But according to von Mises, his theory is not defective in this respect. He writes, "The knowledge of the effect of the four fundamental operations on the distribution enables us, in principle, *to solve all problems of the calculus of probabilities*" ([vM81], p. 65, emphasis added).[41] In

[40]In particular, in the course of solving problems, we will often begin with a single collective C, use selection to extract certain subsequences from C, and then use combination to combine these subsequences into a derived sequence the attributes of which are tuples of attributes from C. However, as we discussed in Subsection 8.2.2, the derived sequence produced by the operation of combination will not be a collective unless the original collective C was invariant under the selections described above. In this, selection and combination work in tandem to produce new collectives from initial collectives.

[41]It is peculiar for von Mises to claim that the "*knowledge* of the effect of the four fundamental operations" allows for the solution of all problems of the probability calculus. Why appeal to knowledge of the effect of the operations, rather that just stating that by means of the four operations, all such problems are solvable? One suggestion is this: If we start with what von Mises refers to as a known collective (itself not an unproblematic notion), then if we have knowledge of the effect of the four operations, when we apply any of the four operations, the derived collective will also be known by us. That is, with knowledge of the fundamental operations, we can transform known collectives into known collectives. Of course, to make sense of this suggestion, we need to determine what it

other words, von Mises took his account to be complete, insofar as it allowed for the solution of all problems of the probability calculus, an ideal of completeness I call the *resolutory ideal of completeness*.

But why did von Mises think that his theory could attain this resolutory ideal? The beginning of an answer is found at the beginning of his third lecture in *Probability, Statistics, and Truth*, in which von Mises writes,

> If it were my intention to give a complete course on the theory of probability, I should now demonstrate how new collectives are derived from given ones by more and more complicated combinations of the four fundamental operations, and, on the other hand, how all problems usually treated in probability calculus can be reduced to combinations of this kind. It would, however, be impossible to do this without using mathematical methods out of place in this book ([vM81], p. 66).

Von Mises then directs "[t]hose who are interested in this side of the theory" to several sources, including his lecture notes on probability and statistics [vM46]. Upon searching through these notes (and his later textbook based on these notes), one doesn't find some general theorem showing that his four operations allow him to solve all problems of the probability calculus. Instead, one finds von Mises working out all of the results typically found in a standard text on probability and statistics, except that all of these results are obtained by applying different combinations of his four fundamental operations to certain initial collectives (including those defined over uncountable sets of attributes). Thus, for von Mises, one doesn't show that the resolutory ideal can be attained by proving general meta-theoretic results about the

means to know a collective. Is it just to know the limiting relative frequencies of the attributes that make up the sequence? Or does the knowledge of a collective also involve knowledge of those place selections under which the collective is invariant? No easy answer is forthcoming.

fundamental operations; rather, one just does the hard work of deriving all of the results in standard textbooks on probability and statistics by means of collectives and the four fundamental operations.[42,43]

[42]Von Mises concedes that this can be rather difficult in some cases, due to "complications arising from the accumulation of a great number of elementary operations" ([vM81], p. 65).

[43]There are, however, a number of general questions that might reasonably be asked of von Mises' fundamental operations and their relation to the resolutory ideal that are left unanswered. For instance, von Mises doesn't address why he chose these operations as fundamental, nor does he discuss whether these operations are independent of one another, or whether there are some other operations, or even a proper subset of his four operations, that would allow his theory to attain the resolutory ideal. Nonetheless, one can venture an educated guess as to how von Mises would address these points. First, what is distinctive about the four fundamental operations is that they preserve the condition of randomness. In the beginning of his lecture notes, von Mises describes a procedure for deriving a new sequence from a collective, but one that "destroys the randomness" of the original sequence: Given a binary sequence $X \in 2^\omega$ generated by the tosses of a fair coin (which von Mises takes to be a given collective), one obtains a sequence $Y \in 3^\omega$ by adding the first and second values of X, then the second and third values, the third and fourth values, and so on. As a result of this operation, the sequence Y has the property that no 2 occurring in Y is ever followed by a 0, since a 2 in Y is obtained by adding two consecutive 1s in X, which means that the next value in Y would have to be at least 1. For similar reasons, the consecutive values 010 never appears in Y. Note that Y is thus not a collective for if we select every value that follows a 2, we would obtain a sequence that contains no 0s. It is this preservation of randomness, von Mises claims, that allows for his theory to solve "most of the known problems of probability calculus" ([vM64], p. 15). Concerning the latter two points, on the independence of his operations and whether some proper subset of them or some other set of operations would suffice to attain the resolutory ideal, it seems that what was most important for von Mises was deriving the standard rules of the probability calculus (the addition rule, the division rule used in the definition of conditional probability, and the product rule). But it also seems that von Mises wanted to derive these rules from operations that could be given a reasonable physical interpretation, given that he thinks of collectives as sequences of observations. We take the question as to whether there are some other operations that satisfy these two conditions (derivability of the rules of the probability calculus and a reasonable physical interpretation) to be an open question, but insofar as the fundamental operations satisfy these conditions, it seems they would be appropriate for attempting to achieve out the resolutory ideal, at least from von Mises' perspective.

8.6.2 Attaining the Resolutory Ideal

There is an additional feature of von Mises' account of the resolutory ideal: according to von Mises, the resolutory ideal cannot be attained if we restrict the collection of place selections to some fixed, predetermined countable collection. In particular, von Mises rejected Copeland's suggestion that the collection of collectives be identified with the Bernoulli sequences (those sequences that are collectives with respect to those place selections that select at every index of the form $an+b$ for fixed $a, b \in \omega$), arguing that on such a restricted approach, the resolutory ideal could not be attained. As von Mises observed, a number of problems in the probability calculus can be solved using Bernoulli sequences, but not all of them: "[T]here is [...] no doubt that a number of other meaningful questions would now remain unanswered" ([vM81], p. 90).[44,45] Thus, it is the resolutory ideal of completeness that trumps any benefits provided by restricting the definition of probability to Bernoulli sequences.

[44]Von Mises offers an example of a problem of the probability calculus that cannot be solved using Bernoulli sequences, a modification of a famous problem of Chevalier de Méré, first solved by Fermat. The problem of Chevalier de Méré was this: Suppose we cast a die four times in a row. Which is more probable, for a 6 to occur at least once, or for no 6 to occur at all? By means of Bernoulli sequences (where each attribute is a four-tuple representing the results of tossing the die four times) and the four fundamental operations, one can solve this problem of de Méré. However, as von Mises notes, "What happens, for instance, if a player decides, at the beginning, that he will consider only the first, second, third, fifth, seventh, eleventh,... casts of the die, that is to say, only those whose order number is a prime number? Will this change his chances of winning or not?" ([vM81], p. 90) That is, if given a sequence of four-tuples of tosses of the die, by means of Bernoulli sequences, can we still solve the problem of de Méré in the subsequence selected from the positions whose indices are prime numbers? As von Mises notes, we cannot.

[45]According to Martin-Löf,

> [V]on Mises forcibly urged that sequences like the admissible numbers [of Copeland] cannot be regarded as satisfactory idealizations of sequences obtained by actual coin tossing. Since an admissible number may be defined by a mathematical law, we could, when playing with such a sequence, ensure ourselves an unbroken series of wins ([ML69b], p. 24).

But von Mises makes a much more radical claim:

> If, instead of restricting ourselves to Bernoulli sequences, we consider some differently defined class of sequences, we do not improve the state of affairs. In every case it will be possible to indicate place selections which will fall outside the framework of the class of sequences which we have selected. *It is not possible to build a theory of probability on the assumption that the limiting values of the relative frequencies should remain unchanged only for a certain group of place selections, predetermined once and for all* ([vM81], p. 91, emphasis added).[46]

In other words, if the collection of place selection rules is restricted to some fixed, countable collection, then the resulting theory of probability will not attain the resolutory ideal of completeness; there will be at least one problem of the probability calculus that will not be solvable by this theory.

Let us try to better understand this view by formulating this suggestion in more precise terms. One reasonable suggestion, given in term of the analysis of von Mises' approach to problems and solutions provided in Subection 8.5.2, is the following: For every countable collection of place selections $\mathscr{S}$, there will be a problem $\mathcal{P} = (\mathcal{C}, \mathscr{S}^*, \mathcal{D})$ such that

(i) there is some sequence $O_1^*, O_2^*, O_3^*, \ldots$ of fundamental operations such that (a)

But this does not appear to be right. Based on what we've seen, von Mises wanted a definition of randomness that would ensure the solution of problems in the probability calculus, not one that would provide "satisfactory idealizations of sequences obtained by actual coin tossing". Moreover, such idealizations would not be applicable in those contexts in which the collectives under consideration have more than two attributes, so this doesn't seem to be a line of argument that von Mises would pursue.

[46]This passage gives us strong evidence against one common interpretation of von Mises' definition, that he wanted an absolute definition of randomness. For instance, Martin-Löf writes, "[Von Mises] wanted to define random sequences in an absolute sense, sequences that were to possess all conceivable properties of stochasticity." In the next chapter, however, we will see that this is an appropriate of description of Jean Ville's approach to randomness.

$O_i^* \in \mathscr{S}^* \cup \{\text{mixing, partition, combination}\}$, and (b) successively applying the operations O_i^* to any initial collective in $\mathcal{C}$ will produce a derived collective which will allows us to meet the demand $\mathcal{D}$; and

(ii) for every sequence $O_1, O_2, O_3, \ldots$ of fundamental operations such that $O_i \in \mathscr{S} \cup \{\text{mixing, partition, combination}\}$, by successively applying the operations O_i to any initial collective in $\mathcal{C}$, the resulting sequence will either (a) fail to be a collective, (b) yield one or more incorrect values of the probabilities referenced in the demand $\mathcal{D}$, or (c) will not contain one or more of the attributes referenced in the demand $\mathcal{D}$.

These two conditions clearly imply that $\mathscr{S}^* \neq \mathscr{S}$, and thus we see that while we can solve $\mathcal{P}$ with sequences that are invariant with respect to the place selections in $\mathscr{S}^*$, with sequences that are invariant with respect to the place selections in $\mathscr{S}$, we cannot.

Given these rather intricate conditions, one is left to wonder what could possibly justify von Mises' belief that for any problem $\mathcal{P}$, these two conditions must hold. One might attempt to provide the justification on formal grounds, but the meta-theoretic apparatus that would have to be put in place to even pose the problem in precise enough terms would be extremely powerful: one would need at least third-order arithmetic, given that the fundamental operations map collectives to collectives, which are themselves second-order objects.

8.7 Church on the Resolutory Ideal

Although the resolutory role of randomness and the related resolutory ideal of completeness play a central role in von Mises account, it was seemingly ignored by all of von Mises' contemporaries, the lone exception being Alonzo Church. In his short paper "On the Concept of a Random Sequence", published in 1940, Church offers a restricted version of von Mises' second axiom (VM_2), the principle of the impossibility of a gambling system. Speaking of von Mises' formulation of (VM_2), Church writes, "Grave question is raised whether this requirement, made in vague terms by von Mises, can be satisfactorily represented in an exact definition at all" ([Chu40], p. 132).[47] It is this exact definition that Church attempts to provide.

Given von Mises' view that the resolutory ideal cannot be attained when we restrict the collection of place selections to some predetermined collection, one might conclude that in attempting to provide an "exact definition" of the condition of randomness, Church is either explicitly rejecting this ideal or at least is unaware of it. Surprisingly, on the contrary, Church clearly holds that the resolutory ideal is one that the resulting definition of probability should strive to attain, a view he sets forth in his discussion of the restricted collectives of Copeland and Reichenbach.

First, Church acknowledges the attempts of Copeland and others to provide an exact formulation of the criterion of randomness, writing,

[47] When Church refers to an exact definition, it's not immediately clear what he has in mind, as there are at least two ways to make von Mises' definition exact. First, one could fix, once and for all, one precise class of place selection rules and define collectives solely in terms of this class. Or, one could show how each problem in the probability calculus corresponds to a precise collection of place selections, each of which guarantees enough independence in collectives that can be used to solve that problem. Whereas the second approach is more faithful to von Mises' stated intentions, it is the first approach that Church takes.

> This difficulty [that of providing an exact definition of randomness] may be avoided by abandoning the attempt to define a random sequence and substituting some less restricted class of sequences, such as the admissible numbers of Copeland or the equivalent normal sequences of Reichenbach ([Chu40], p. 132).[48]

But just as von Mises rejected the definitions of Copeland, Reichenbach, and others on the basis of the resolutory ideal of completeness, so too does Church:

> The admissible numbers [of Copeland] have properties which are sufficient to form a basis for a large part of the theory of probability, and they have the important advantage that their existence, for any assigned probability p, can be proved. Their use for this purpose, however, is open to certain objections from the *point of view of completeness of the theory*, as has been forcibly urged by von Mises, and it is therefore desirable to consider further the question of finding a satisfactory form for the definition of a random sequence ([Chu40], p. 133, emphasis added).

Note here that Church also has the worry about the existence of collectives in mind, as he cites as an advantage of Copeland's definition the fact that one can prove the existence of admissible numbers, i.e., collectives with respect to the collection of Bernoulli place selections. Thus, in agreement with von Mises, Church holds that concerns about the resolutory ideal of completeness trump the virtues of Copeland's definition.

This agreement notwithstanding, Church differs with von Mises as to one aspect of the resolutory ideal of completeness, namely, they disagree about the scope

[48] One might be puzzled that whereas von Mises characterizes the work of Copeland and Reichenbach as providing a restricted notion of randomness, Church considers their notion to be *less* restricted. But there is no disagreement here, since when von Mises refers to restricted randomness, he is considering the restriction of the collection of place selections, resulting in a larger collection of collectives, while Church takes this larger collection of sequences to be less restricted, since it would have been smaller had we included more place selections. In short, the more we restrict the condition of randomness, the larger and less restricted the resulting class of collectives will be.

of the problems that must be solvable in our theory of probability in order for it to be considered complete. As this difference in approach to the resolutory ideal follows naturally from the difference between von Mises' definition of randomness and Church's restricted version of this definition, it will be instructive to consider Church's definition.

8.7.1 Church's Restricted Definition of Randomness

In order to provide an exact definition of randomness along the lines suggested by von Mises' definition, Church proposes that we define collectives in terms of the collection of place selections that are effectively calculable. He writes,

> It may be held that the representation of a *Spielsystem* [i.e, a gambling system] by an arbitrary function ϕ is too broad. To a player who would beat the wheel at roulette a system is unusable which corresponds to a mathematical function known to exist but not given by explicit definition; and even the explicit definition is of no use unless it provides a means of calculating the particular values of the function. As a less frivolous example, the scientist concerned with making predictions or probable predictions of some phenomenon must employ an effectively calculable function: if the law of the phenomenon is not approximable by such a function, prediction is impossible. Thus a *Spielsystem* should be represented mathematically, not as a function, or even as a definition of a function, but as an effective algorithm for the calculation of the values of a function ([Chu40], p. 133).

Church's point is that any gambling system or method of prediction that is not effectively calculable, or at least is not effectively approximable, cannot be implemented and thus is, for all practical purposes, useless.[49] Consequently, on Church's ap-

[49]Church elsewhere reiterates this point, writing, "[A] betting system must be based on an effectively calculable method of selection, otherwise it is no betting system at all, in the sense that

proach, the collection of place selection rules is restricted to those that are effectively calculable.

Having provided a justification for restricting to the effectively calculable place selections, Church appeals to his formal definition of effective calculability of λ-definability as adequate "to represent the empirical notion of an effective calculation", citing as support Turing's result of the equivalence of λ-definability and general recursiveness ([Chu40], p. 134). Thus, significantly, one of the earliest applications of Church's Thesis comes in the service of providing a definition of randomness, a notion we shall henceforth refer to as *Church randomness*.[50] Having established that the collection of effectively calculable place selections is a formally definable collection, Church then appeals to Wald's Theorem I to prove the existence of Church random sequences: Since there are countably many computable place selections, it follows from Wald's Theorem I that there are continuum many random sequences.[51]

After concluding that Church random sequences exist, Church compares his notion of randomness with Copeland's notion of admissible numbers. While every Church random sequence is an admissible number, not every admissible number is a Church random sequence. The reason for this is that (i) there is some admissible number $A = a_1a_2a_3\ldots$ such that the function $f(n) = a_n$ is a computable function,

---

no bettor could actually employ and carry out such a system" ([Chu66a], pp. 1-2).

[50]Today, it is common to refer to Church's definition as *Church stochasticity* rather than Church randomness, as Church's definition is generally held today to be an inadequate definition of randomness. The reasons for this will be made clear in the next chapter.

[51]Church also concludes via a theorem of Doob's, discussed in footnote 25, that the collection of binary sequences with limiting relative frequencies equal to 1/2 for each event 0 and 1 forms a set of measure one in 2^ω.

but (ii) no Church random sequence can have this property, since given such a sequence A and a corresponding function f such that $f(n) = a_n$ for every n, the set $S = \{n : f(n) = 0\}$ is a computable set, and thus the place selection that selects the nth bit of a sequence if and only if $n \in S$ is a computable place selection that selects a sequence of all 0s from A. Church thus took his definition to be an improvement over Copeland's, which von Mises rejected on the grounds that it could not attain the resolutory ideal.

8.7.2 Church's Restricted Version of the Resolutory Ideal

But isn't Church's definition vulnerable to the same problem? Interestingly, Church didn't think so. First, he is clear that invariance under the computable place selections is necessary to attain the resolutory ideal, a point Church makes explicitly when he responds to a potential objection to his restricted definition. He writes,

> Use of the above proposed definition of a random sequence as fundamental to the theory of probability is consequently open to the objection that by its means such otherwise apparently combinatorial matters as elementary questions of probability in connection with the tossing of a coin are made to depend on the powerful (and dubious) non-constructive methods of analysis. ([Chu40], p. 135).

In response to this objection, Church writes,

> It is clear, however, that any definition of a random sequence more stringent than this one would have the same disadvantage, and on the other hand that no definition in any respect less stringent could be regarded as even approximately representing von Mises's intention or as being free from such objections as those brought by him against the use of admissible numbers or normal sequences ([Chu40], p. 135).

Thus we have a sort of bottleneck here: On the one hand, we have another definability tradeoff: any definition that is "more stringent", i.e. one that is defined in terms of place selections that form a proper superset of the collection of computable place selections, will be vulnerable to the same non-constructivity objection given above. On the other hand, any definition of randomness that is "less stringent", i.e. one that is defined in terms of place selections that are a proper subset of the collection of computable selections, would result in a definition of probability that is vulnerable to the objections von Mises raised against the definitions of Copeland and Reichenbach, that there are some problems of the probability calculus not solvable by means of their restricted definitions. Hence, according to Church, any definition of probability given in terms of place selections strictly weaker than the computable place selections would not be able to attain the resolutory ideal of completeness.

Did Church think his definition also provided a sufficient amount of invariance to attain the resolutory ideal? In his article, he doesn't say. In correspondence with von Mises' wife, Hilda Geiringer,[52] Church claims something very much along these lines. In his first letter to Geiringer, Church makes a claim that has been made

[52]As far as I can tell, the brief correspondence between Geiringer and Church, which took place in 1966, has not been discussed in the literature on algorithmic randomness. I am grateful to the Department of Rare Books and Special Collections at the Princeton University Library for helping me to obtain copies of these letters from the Alonzo Church papers. This correspondence is also noteworthy because von Mises never explicitly acknowledged Church's contribution: In a review of von Mises' textbook, *Mathematical Theory of Probability and Statistics*, D.V. Lindley writes,

> It appears to your reviewer, though the logical ideas here are outside of his special field of knowledge, that Church's idea avoids many, if not all, of the difficulties [that beset von Mises' definition], and it is surprising that von Mises makes no mention of it. (There is no reference to Church in the book: I am indebted to G.A. Barnard for drawing it to my attention. ([Lin66], p. 749)

several times in the algorithmic randomness literature, that von Mises would have restricted collectives to those invariant under computable place selections if only the definition were available to him:

> [I]t seems to me very plausible to say (though of course no proof of such a proposition can be offered) that the definition of "collective" which results from the approach of this paper is the one which von Mises in some sense actually intended when he wrote in 1931, but that it was impossible for him to make the definition in this way because at that date the precise mathematical definition of effective calculability did not yet exist ([Chu66b], p. 1)

In response to Church's letter, Geiringer asks,

> Could one say that your criterion while clearer and sharper than Wald's L-criterion [according to which we restrict to place selections definable in some logic] is more restrictive (I mean that selections appearing in L would not appear in your scheme). In my opinion, it would be of great interest to possess a clear and sharp criterion which is sufficiently comprehensive ([Gei66], p. 2).

From Church's answer to Geiringer's question, we see that he held that the resolutory ideal *could* be attained with a restricted definition, as long as we were willing to restrict the scope of the problems solvable by the resulting definition. He writes,

> [...] I agree that a good reconstitution of the notion of collective must make use of a class of selections, *not necessarily wide enough to cover all selection that anybody has ever made or claimed to make in a probability problem or probability proof*, but wide enough so that a comprehensive probability theory can be developed on the basis of the notion of collective that results from the class of selections in question with significant gaps neither in the internal logical structure of the theory nor in its applications ([Chu66a], p. 1, emphasis added).

Here we see Church implicitly acknowledge that there may be some problems of the probability calculus that cannot be solved with his restricted definition, but nonetheless, the resulting theory of probability would still be "comprehensive". Interestingly, the reason he thought this is the same reason von Mises offered for why his definition could attain the resolutory ideal, to wit, that one could recover all of classical probability theory using his approach. As Church writes,

> I have every reasonable expectation that the criterion in my paper results in a class of selections for which this is true. But no real guarantee can be obtained except by writing a comprehensive work on probability theory, developing the theory on this basis. And while I suppose this might not be a difficult task, it is certainly a long one, and I have not actually done it ([Chu66a], p. 1).

But what about von Mises' claim that there would always be problems left as unsolvable when the collection of place selections is fixed once and for all? Church was not bothered by this possibility,

> And I would hold that if it is true that no such calculation procedure exists [to implement a given place selection], then the indicated method of selection is an unreasonable one to use in any probability problem. (Has any one ever used it in a probability problem? I don't know, but I would think it unlikely.) ([Chu66a], p. 1)

We thus have two versions of the resolutory ideal of completeness, an unrestricted version that, according to von Mises, cannot be attained by a definition of probability given in terms of a fixed collection of place selections, and a restricted version that, in Church's view, can be attained in a correspondingly restricted sense by a definition of probability given in terms of the computable place selection.

What exactly is at stake with the disagreement over these two versions of the resolutory ideal? Perhaps not much, for as some have claimed, von Mises *would have* opted for restricted version of the resolutory ideal had the definition of computable function been available when he was first formulating his theory.[53] But this is not so clear, especially given von Mises' statement that "[i]t is not possible to build a theory of probability on the assumption that the limiting values of the relative frequencies should remain unchanged only for a certain group of place selections, predetermined once and for all" ([vM81], p. 91).

Yet what we gain from the restricted approach is uniformity: on Church's approach, there is only one fixed collection of place selections to which we must appeal to solve every problem that we encounter in practice. On the unrestricted approach, however, for each problem, we have to determine which place selections should be used in the solution of the problem, and even though von Mises claims that these place selections "present themselves", it's hard to see how this is supposed to work in practice.

Thus, which version of the resolutory ideal is to be preferred depends on whether the uniformity discussed above is to be valued over resolutory completeness. If one were to show, for instance, that solving problems of the probability calculus is more

[53]For instance, Church writes to Geiringer,

> It seems to me very plausible to say (though of course no proof of such a proposition can be offered) that the definition that results from the approach of this paper ["On the Concept of a Random Sequence"] is the one which von Mises in some sense actually intended when he wrote in 1931, but that it was impossible for him to make the definition in this way because at that date the precise mathematical definition of effective calculability did not yet exist ([Chu66b], p. 1).

efficient on the restricted approach than the unrestricted approach, in that one need not be saddled with the problem of determining which place selections should be used to solve a given problem, then this would certainly count in its favor. However, if one could show that too many problems were left unsolved on the restricted approach, then this count in favor of the unrestricted approach.

In the next chapter, we will consider another ideal of completeness associated with definitions of randomness, one that also comes in an unrestricted version and a restricted version, where the restriction is given in terms of computability, just as with Church's restriction of the resolutory ideal. However, unlike the resolutory ideal, this second ideal, which I call the *exemplary ideal of completeness*, played an important role in bringing us to the situation we find in the present day, in which some claim of a given definition of randomness that it captures our commonly-held intuitions about randomness. Given that the exemplary ideal and its role in the development of algorithmic randomness have not been isolated in previous studies, it certainly merits our close attention.

CHAPTER 9

THE EXEMPLARY ROLE OF RANDOMNESS

9.1 Introduction

Whereas von Mises' sought a definition of randomness that would permit the solution of all problems of the probability calculus, another ideal of completeness associated with definitions of randomness was introduced by Jean Ville in his dissertation, *Étude Critique de la Notion de Collective* [Vil39], alongside a number of results important for both the study of collectives and the development of algorithmic randomness. In particular, this ideal was inspired by one of the central results of the *Étude*, which shows that von Mises' collectives satisfy a certain property that many have taken to reveal a fundamental defect in von Mises' definition of randomness, that is counts as random certain sequences satisfying a property not typical of sequences produced at random.

In light of this purported defect, Ville hoped to improve upon von Mises' definition by reformulating the principle of the impossibility of a gambling system in terms of a more general notion of a betting strategy than the one employed by von Mises. Moreover, it was Ville's goal to produce a definition that was free from this purported defect. Towards this end, Ville sought a definition $\mathscr{D}$ with the property that every $\mathscr{D}$-

random sequence satisfies all of the properties that are typical of sequences produced at random, properties I call $\mathcal{R}$*-properties*; that is, $\mathscr{D}$-random sequences, although themselves not necessarily randomly chosen, should *exemplify* all of the $\mathcal{R}$-properties of randomly chosen sequences. Such a definition would thus attain what I call the *exemplary ideal of completeness.*

There is, however, something rather puzzling about the exemplary ideal. For Ville never identifies which properties are to be counted as the $\mathcal{R}$-properties, and moreover, he ultimately concludes that the exemplary ideal is unattainable: "this [...] problem [of finding a definition attaining the exemplary ideal] is considered by us as unsolvable" ([Vil39], p. 93). Yet in his 1966 article "The Definition of Random Sequences" [ML66], Per Martin-Löf presents his definition of randomness, now known as Martin-Löf randomness, as one that attains the exemplary ideal: "Finally, the [collectives] introduced by von Mises obtain a definition which seems to satisfy all intuitive requirements" ([ML66], p. 602).

However, Martin-Löf randomness was not the only candidate for attaining the exemplary ideal, for several years after the publication of Martin-Löf's definition, C.P. Schnorr introduced a slightly weaker definition of randomness,[1] nowadays referred to as Schnorr randomness, which, like Martin-Löf randomness, attained the exemplary ideal, but which further captured "the true concept of randomness". In Schnorr's view,

> Many insufficient approaches have been made [to define randomness] until a definition of random sequences was proposed by Martin-Löf which for

[1] That is, every Martin-Löf random sequence is Schnorr random, but not every Schnorr random sequence is Martin-Löf random.

> the first time included all standard statistical properties of randomness. However, the inverse postulate now seems to have been violated ([Sch71], p. 255).

The violation is that Martin-Löf randomness required *too many properties*, properties that "are of no significance to statistics" and have no "physical meaning". For this reason, Schnorr found Martin-Löf's definition to be unacceptable.

While the importance of the exemplary ideal in the development of algorithmic randomness should be clear enough, as it grew out of concerns with a purported defect of von Mises' definition and later was claimed to be attained by Martin-Löf's definition of randomness, the exemplary ideal is relevant to the larger aims of this study, those of determining the roles of the various definitions of randomness and whether any of these definitions can rightly be claimed to provide a conceptual analysis or explication of the concept of randomness. For there are two salient features of Ville's account that we will encounter in later chapters when we consider extensional adequacy theses such as the MLCT. First, the unclarity in the formulation of the exemplary ideal, given in terms of these nebulous $\mathcal{R}$-properties, is one that persists when we inquire into whether one of the various definitions of randomness is correct. In fact, as I argue, this unclarity proves to be problematic to those who attempt to establish extensional adequacy theses such as the MLCT.

The second salient feature of Ville's account is that when Ville seeks a definition that attains the exemplary ideal, the task of defining randomness appears to be detached from considerations of the purpose that such a definition might play. While it is clear from von Mises' account why he considers a definition of randomness satisfying (VM_1) and (VM_2) to be useful, Ville never indicates why a definition of randomness that attains the exemplary ideal might be of further use. This feature also persists: when we arrive at Martin-Löf's definition, although he claims that his definition appears to satisfy "all intuitive requirements", it's not clear for what purpose these requirements are to be fulfilled. This, too, is problematic for those seeking to establish the MLCT or similar theses.

The main goals of this chapter are (1) to trace the path that took Ville from von Mises' collectives to the formulation of the exemplary ideal of completeness, (2) to consider Ville's reason for holding that the exemplary ideal could not be attained, and (3) to highlight the introduction of Martin-Löf randomness as a candidate for attaining the exemplary ideal.

The remainder of the chapter will proceed as follows. First, in Section 9.2, I explain Ville's putative counterexample to von Mises' definition of randomness. This counterexample led Ville to provide an alternative formalization of a betting strategy that was intended to be an improvement over von Mises' definition of a betting strategy, as I'll discuss in Section 9.3. Next, in Section 9.4, I introduce the exemplary ideal of completeness, highlight its central features, and explain Ville's rationale for holding that it could not be attained by any definition of randomness. In Section 9.5, I discuss Martin-Löf's definition of randomness and his presentation of this definition as one fulfilling the exemplary ideal. In Section 9.6, I discuss Schnorr's definition of randomness and his arguments that his definition, and not Martin-Löf's, captures the "true concept of randomness", and then in Section 9.7, I conclude by discussing the problematic nature of the exemplary ideal.

9.2 Ville's Putative Counterexample to von Mises' Definition

The first major contribution in Ville's *Étude* is often referred to as Ville's Theorem, which purportedly shows a fundamental defect in von Mises' definition of randomness. Before I lay out the details of Ville's Theorem, let me say a brief word on the notation used throughout the ensuing discussion. Hereafter, we will restrict our attention to collectives in 2^ω, and instead of merely referring to $C(\mathscr{S})$, the collection of collectives in 2^ω with respect to the place selections in $\mathscr{S}$ as I have done in the previous chapter, here I will make reference to $C(\mathscr{S}, p)$, the collection of $\mathscr{S}$-collectives in 2^ω that have the property that the limiting relative frequency of 1 is equal to $p \in (0, 1)$.

9.2.1 Ville's Theorem

As motivation for his main theorem, Ville states the following result:[2]

Theorem. Let $p \in (0,1)$ and $\mathcal{A} \subseteq 2^\omega$. Then there is a countable collection of place selections $\mathscr{S}$ such that $C(\mathscr{S}, p) \subseteq 2^\omega \setminus \mathcal{A}$ only if $\mathcal{A}$ has p-measure zero.[3]

In other words, given a set $\mathcal{A} \subseteq 2^\omega$, we can ensure by some choice of place selections $\mathscr{S}$ that no collective $X \in C(\mathscr{S}, p)$ belongs to $\mathcal{A}$ *only if* $\lambda_p(\mathcal{A}) = 0$ (where λ_p is the

[2]This result is essentially a more general version of Doob's Theorem, discussed in footnote 25 in Chapter 8.

[3]The p-measure of a set $\mathcal{X} \subseteq 2^\omega$ is also referred to as Bernoulli p-measure, a generalization of the Lebesgue measure. Given $\sigma \in 2^{<\omega}$, let $\#_0(\sigma)$ and $\#_1(\sigma)$ denote the number of 0s and 1s in σ, respectively. Then the p-measure of the basic open set determined by σ, $[\![\sigma]\!] = \{X \in 2^\omega : \sigma \prec X\}$, is defined to be

$$\lambda_p([\![\sigma]\!]) = p^{\#_0(\sigma)}(1-p)^{\#_1(\sigma)},$$

which reduces to the Lebesgue measure when $p = \frac{1}{2}$.

p-measure defined in footnote 3 below). It is natural to ask whether the converse holds: can every set of p-measure zero be covered by the complement of a fixed set of collectives? Ville's Theorem shows us that the answer is "no". That is, Ville shows that for some $p \in (0,1)$, there is a set $\mathcal{G}$ of p-measure zero such that for *any* countable collection $\mathscr{S}$ of place selections, $C(\mathscr{S}, p) \cap \mathcal{G} \neq \emptyset$. In particular, he shows,

Theorem (Ville's Theorem)**.** For any countable collection $\mathscr{S}$ of place selections, there is $X = x_1x_2x_3 \ldots \in C(\mathscr{S}, \frac{1}{2})$ such that for every n,

$$\frac{\#\{i < n : x_i = 1\}}{n} \geq \frac{1}{2}.$$

Thus, for every countable collection $\mathscr{S}$ of place selections, we can always find a sequence $X \in 2^\omega$ that is a collective relative to $\mathscr{S}$, but has the property that every one of its initial segments contains more 1s than 0s. Now, if we define

$$\mathcal{G} := \left\{ X \in 2^\omega : (\forall n) \frac{\#\{i < n : x_i = 1\}}{n} \geq \frac{1}{2} \right\}$$

one can show that $\lambda(\mathcal{G}) = 0$, which follows from a classic result known as the Law of the Iterated Logarithm, first proven by A.Y. Khinchin in 1924, and later, in a different form, by A.N. Kolmogorov in 1929.[4] Thus, Ville identifies a collection $\mathcal{G}$ such that, for every collection of place selections $\mathscr{S}$, there will be a collective $X \in C(\mathscr{S}, \frac{1}{2})$ that belongs to $\mathcal{G}$.

[4]For details of the Law of the Iterated Logarithm, see, for instance, Chapter VIII.5 of [Fel68].

9.2.2 Consequences of Ville's Theorem?

It has been claimed by a number of individuals, including Fréchet in his address at the 1937 Geneva conference, that Ville's Theorem reveals a fundamental defect in von Mises' definition. But what exactly is this defect? In other words, to what statement is Ville's counterexample a counterexample? These questions are especially pressing, given that von Mises was aware of Ville's Theorem, but was entirely unmoved by it.[5]

In the subsequent literature on the implications of Ville's Theorem for von Mises' theory of probability, one finds at least three different answers to these questions:

1. Ville's Theorem shows that von Mises' definition does not capture our commonly-held intuitions of randomness.

2. Ville's Theorem show that von Mises' definition cannot serve as the foundation for measure-theoretic probability.

3. Ville's Theorem shows that von Mises' definition is given in terms of inadequate formalization of a betting strategy.

Let's consider each answer in turn.

9.2.2.1 Response 1: On commonly-held intuitions of randomness

This first response is summed up well by Rod Downey and Denis Hirschfeldt in their recently published monograph *Algorithmic Randomness and Complexity*

[5]He even says as much in his response published in the proceedings of the 1937 Geneva conference: "I accept this theorem, but I do not see in it an objection, and I do not find in it a reason to modify the theory of collectives." In the original French: "J'accepte ce théorème, mais je n'y vois pas une objection et n'y trouve pas une raison pour modifier la théorie du collectif" ([vM38], p. 66).

[DH10]. They write of any sequence contained in the set $\mathcal{G}$ defined above, "Such a sequence is clearly not random. After all, if we were to flip a supposedly fair coin many times and never at any point observed an excess of tails over heads, we would surely begin to suspect that something was amiss" ([DH10], p. 230).[6] This is perhaps the most commonly offered response to the question as to whether Ville's Theorem shows von Mises' definition of randomness to be defective. While most would agree that a definition of randomness that is intended to capture our commonly-held intuitions of randomness should not include as random any sequence in the set $\mathcal{G}$, von Mises would not be moved by this objection to his definition because his goal was not to capture some absolute definition of randomness. Rather, as was detailed in the previous chapter, von Mises wanted a definition of randomness that would enable him to solve problems in the probability calculus, and there's no indication in his works that he thought that a definition that captured all of the commonly-held intuitions of randomness would thereby permit the solution of all problems of the probability calculus.[7]

[6] A related response is summarized by van Lambalgen, who writes,

> [Collectives] do not necessarily satisfy all asymptotic properties proved by measure theoretic methods and since the type of behaviour exemplified by [the property $\frac{\#\{i<n:x_i=1\}}{n} \geq \frac{1}{2}$ holding of a sequence for every n] will not occur in practice (when tossing a fair coin), [collectives] are not satisfactory models of random phenomena ([Lam87], p. 52).

[7] This holds true of the later formulation of his theory, for as we discussed in Section 8.4.2, on the later approach the amount of invariance required to hold of collectives was determined by the problem one was trying to solve.

9.2.2.2 Response 2: On the foundation of measure-theoretic probability

This second response is summed up by Glenn Shafer, who writes,

> The main point of [Ville's] critique of collectives was that von Mises' and Wald's idea of selecting subsequences was inadequate as a foundation for classical probability theory. [...] It was also inadequate for representing the classical idea of ruling out events of probability zero, because the relative frequency of 1s in a sequence of 0s and 1s can converge to a number p and yet do so in a way that has probability zero [...] and this behavior cannot be ruled out by specifying a countable number of subsequences on which the limiting frequency must also converge to p ([Sha09], p. 37).

The key claim here is that if von Mises' definition is to serve as a foundation for classical probability, it should count as having p-measure zero those events that are counted as having p-measure zero according to classical probability. In defense of this claim, one might argue that there are problems of measure-theoretic probability that cannot be solved in von Mises' framework (such as showing the set $\mathcal{G}$ to have Lebesgue measure zero), and thus if we want a definition of probability that permits us to solve all problems of measure-theoretic probability, von Mises' definition will not suffice.

Again, von Mises would not be threatened by this objection, given that he was not attempting to develop a foundation for measure-theoretic probability. In fact, von Mises explicitly contrasts his frequency-theoretic approach to probability with the measure-theoretic approach, which takes Kolmogorov's axioms of probability as a starting point. While on von Mises' approach, collectives are the primitive objects in terms of which one defines probability, on the approach given in terms of Kolmogorov's axioms, one takes probability as the primitive term. Further, on von

Mises' approach, calculations in the probability calculus are rooted in operations on collectives, which guarantee that the standard axioms of probability hold. By contrast, on the measure-theoretic approach, the situation is inverted, as the axioms of probability are the starting point of our calculations. In general, von Mises did not think his theory was bound by the standards of measure-theoretic probability.

9.2.2.3 Response 3: An inadequate formalization of a betting strategy

This last response was offered by Ville, who held that his theorem shows von Mises' formalization of a gambling system to be defective, insofar as it is inadequate for representing the idea of the impossibility of a gambling system. In particular, given a sequence in $\mathcal{G}$ that has more 1s than 0s in every initial segment, a gambler should be able to make use of the knowledge to make arbitrarily large amounts of money. The problem with von Mises' formalization of a gambling system is that it only allows the gambler to choose the trials on which he will bet, whereas in reality, a gambler can do more than merely choose trials; he can also vary the amount of his bet. But the gambler who employs a von-Mises- gambling system cannot capitalize on the knowledge that the sequence on which he is betting is biased in all of its initial segments. As Ville puts it in his *Étude*,

> There is a very important restriction there; von Mises does not try to translate in a completely general manner the conduct of a player who tries to modify his chances. He only speaks of a systematic choice of moves, and does not try to express the fact that one can equally try to modify his chances by a systematic distribution of his bets. We shall see that the restriction made by von Mises prevents it from attaining, at least in our view, the end that he seems to be proposing in writing the

passage cited.[8,9] ([Vil39], p. 89)

From von Mises' point of view, this perceived defect is not truly a defect. While the notion of a gambling system plays a useful role in motivating his second axiom (VM_2), the principle of the impossibility of a gambling system, the bottom line for von Mises is always solving problems in the probability calculus. Does the fact that his formalization of a gambling system is not sufficiently general result in there being certain problems that are not solvable? Given his view that his definition can attain the tory ideal, von Mises surely would not accept this. In the end, thinking of place selections as gambling systems may be a useful heuristic, but for the purposes of solving problems, it is entirely dispensable.

Thus, as far as von Mises was concerned, Ville's Theorem had no implications for his theory of probability. Invariance of limiting relative frequencies under place selection guarantees that calculations could be carried out: even with the property given by the set $\mathcal{G}$ above, the possibility of carrying out the requisite calculations is

[8] In the original French:

> Il y a là une restriction très importante; M. de Misès ne cherche pas à traduire d'une manière tout à fait générale la conduite du joueur qui cherche à modifier ses chances. Il ne parle que d'un choix systématique des coups, et ne cherche pas à exprimer le fait que l'on peut également chercher à modifier ses chances par une répartition systématique des mises. Nous verrons (p. 83) que la restriction faite par M. des Misès l'empêche d'atteindre pleinement, à notre avis du moins, le but qu'il semble s'être proposé en écrivant le passage cité.

[9] The passage to which Ville refers is from von Mises' 1931 volume *Vorlesungen aus dem Geiete der angewandten Mathematik* (*Lectures from the Field of Applied Mathematics*) [Mis45]. Ville quotes a French version of this passage: "Le fait que l'on ne peut modifier ses chances par un choix systématique des coups [sur lesquels on mise] compte parmi les conceptions les plus essentielles qui sont, à nos yeux, indissolublement lièes à la notion de ¡¡hasard¿¿ et de jeu de hasard." ("The fact that we cannot modify our chances by a systematic choice of trials on which to place our bets is among the most essential conceptions which are, in our eyes, indissolubly linked to the notion of 'chance' and of gambling" ([Vil39], p. 89).

not jeopardized.[10]

9.3 Ville's Alternative Formalization of a Betting Strategy

In the beginning of the fourth chapter of his *Étude*, "Criteria of Irregularity Based on the Notion of Martingale",[11] Ville writes, "In this chapter we propose a new manner of expressing the axiom of irregularity of von Mises [i.e. (VM_2)], which still conforms to his general idea [...], but which is not equivalent to the mathematical translation that he has given it."[12] ([Vil39], p. 85) Thus we see that Ville does not entirely reject von Mises' approach; rather Ville just seeks to give an improved formulation of von Mises' second axiom of collectives that isn't vulnerable to the purported defect exposed by Ville's Theorem.

To this end, Ville considers the following general situation. Suppose a gambler is

[10] I don't think the book is closed on this matter, for there is an additional way in which Ville's Theorem might pose a problem for von Mises' account. Glossing over some important details, the idea is this: Von Mises expresses the view, thoroughly spelled out by van Lambalgen in his dissertation, that in order for a theorem of measure-theoretic probability to be physically meaningful (whatever this is supposed to mean), it must be translatable into a theorem about collectives. Thus, given that the Law of the Iterated Logarithm cannot be shown to hold in terms of collectives, it is not a physically meaningful result. The problem is that other features of von Mises' account commit him to the claim that the passage from results about finite sequences of attribute to their infinitary analogues preserves physical meaning (so that the limit of physically meaning finitary properties is a physically meaningful infinitary property).What's more, one can show that a finite version of the Law of the Iterated Logarithm holds of collectives, which, in the limit, yields the standard Law of Iterated Logarithm for infinite sequences. Thus von Mises faces a dilemma: either he has to deny that the passage from finitary properties to their infinitary analogues preserves physical meaning (which would undercut the supposed empirical basis of his theory) or he has to concede that his account does not have the inside track with respect to determining the physical meaning of statements of measure-theoretic probability.

[11] In the original French: "Critères d'Irrégularité Fondés sur la Notion de Martingale".

[12] In the original French: "Nous allons dans ce chapitre proposer une nouvelle manière d'exprimer l'axiome d'irrégularité de M. de Misès, qui reste conforme à son idée générale [...], mais non équivalente à la traduction mathématique qu'il en a donnée."

betting on a sequence X of which he knows

$$(*) \text{ The limiting relative frequency of } X \text{ is not equal to } p.$$

Equipped with this piece of information $(*)$, but no other information about the specific values of X, nor the actual limiting relative frequency of X, how can the gambler win arbitrarily large amounts of money in betting on X?[13]

While there is, in general, no method for a gambler to exploiting a bias in a sequence by means of place selections[14] (or, more precisely, in the case that the gambler knows that there is a bias, but does not know what it is), Ville's was eventually

[13]We should note that the bias problem was not merely a theoretical question for Ville, as it was inspired by conversations with a relative of his, who happened to be a professional gambler. In a letter to the French mathematician Pierre Crepel shortly before Ville's death, Ville recounts these conversations:

> But I had an acquaintance, a relative of the woman who became my wife, who claimed to make a (modest) living by gambling, which he pursued like a drudge, "working" for hours recording and counting the outcomes of boule or roulette spins, and then betting according to a calculation that he kept secret. His name was Mr. Parcot. I claimed that it was impossible for him to win. The probability calculus showed that for simple martingales, everything ended up a loss. Because the calculus was applied to martingales one by one, the layman was left with the impression that one could find a crack in the armor and slip though. Mr. Parcot claimed to have found a crack. I did not try to convince him; I don't even remember now if I had an opportunity to do so. I knew that there was a general refutation of the possibility of winning for sure, but I went no farther. Mr. Parcot's continuing profits simply made me think. Why not? I knew that a certain role was played by confusions between infinitely small and zero, and between actual and virtual infinity, nothing more. I did not doubt Mr. Parcot's good faith, and there was something that pointed out a path. [...] I studied the probability calculus in Laplace, and I found there a way to win in heads and tails if you know the coin is asymmetric without knowing which side is favored. From this, I concluded that Parcot had perhaps discovered and taken advantage of a flaw in the roulette wheel. Taking advantage of a known flaw in a roulette wheel is child's play, but taking advantage of the fact, for example, that the spins are not independent, without knowing exactly how they are dependent, is another matter. So this is where I was, say in 1932 ([Cré09], pp. 14-15).

[14]The reason for this is that von Mises' notion of a gambling system doesn't take into account the winnings of a gambler.

able to solve the bias problem by means of martingales.

9.3.1 The Definition of Martingale

For Ville, a martingale is a formalization of a game in which a gambler bets on the values of a sequence. Let A be a gambler , let $X = x_1x_2x_3\ldots \in 2^\omega$ be some sequence, and let $p, q \in (0,1)$ be such that $p+q=1$. The details of the game are as follows: At the start of the game, A has capital $s_0 = 1$. A then wagers some portion of his capital λ_0 that the first value of X will be a 0, and some portion μ_0 that the first value will be a 1, where $\lambda_0 + \mu_0 \leq 1$.[15] If a 0 occurs, A receives $\frac{\lambda_0}{p}$ (and loses $\frac{\mu_0}{q}$), and if a 1 occurs, he receives $\frac{\mu_0}{q}$ (and loses $\frac{\lambda_0}{p}$). Thus, A's capital at the end of the first round of the game will be either be

$$1 + \frac{\lambda_0}{p} - \frac{\mu_0}{q},$$

in the event that a 0 occurs, or

$$1 + \frac{\mu_0}{q} - \frac{\lambda_0}{p},$$

in the event that a 1 occurs.

More generally, suppose that A has capital s_n at the end of the nth round of the game, having already bet on $x_1x_2\ldots x_n$. Then we define two functions λ_n and μ_n as follows: $\lambda_n(x_1, x_2, \ldots, x_n)$ is the portion of A's capital that he will bet that the

[15] The fact that $\lambda_0 + \mu_0 \leq 1$, rather than $\lambda_0 + \mu_0 = 1$, means that A can choose to bet only some of his capital at a given round. In fact, A even has the option of setting $\lambda_0 = \mu_0 = 0$, which is to say, A has the option of not betting any money on either outcome.

$(n+1)$st value of X, x_{n+1}, is a 0, while $\mu_n(x_1, x_2, \ldots, x_n)$ is the portion of A's capital that he will bet that x_{n+1} is a 1. That is, A bets the amount $\lambda_n s_n$ on the occurrence of a 0, and the amount $\mu_n s_n$ on the occurrence of a 1. As before, if a 0 occurs, he will receive $\frac{\lambda_n}{p} s_n$, and if a 1 occurs, he will receive $\frac{\mu_n}{q} s_n$. Thus we have

$$s_{n+1} = \begin{cases} \frac{\lambda_n}{p} s_n + (1 - \lambda_n - \mu_n) s_n & \text{if } x_n = 0 \\ \frac{\mu_n}{q} s_n + (1 - \lambda_n - \mu_n) s_n & \text{if } x_n = 1 \end{cases}$$

where the second summand $(1 - \lambda_n - \mu_n)s_n$ in each of the two above expressions is the amount of capital that A decides to save at the $(n+1)$st round (that is, he need not bet his entire capital at each round). More explicitly, we have

$$s_{n+1}(x_1, x_2, \ldots, x_n, 0) = \frac{\lambda_n}{p} s_n(x_1, x_2, \ldots, x_n) + (1 - \lambda_n - \mu_n) s_n(x_1, x_2, \ldots, x_n) \quad (9.1)$$

and

$$s_{n+1}(x_1, x_2, \ldots, x_n, 1) = \frac{\mu_n}{q} s_n(x_1, x_2, \ldots, x_n) + (1 - \lambda_n - \mu_n) s_n(x_1, x_2, \ldots, x_n). \quad (9.2)$$

Combining equations (9.1) and (9.2), rearranging terms, and applying a few straight-foward algebraic operations, we get

$$s_n(x_1, x_2, \ldots, x_n) = p s_{n+1}(x_1, x_2, \ldots, x_n, 0) + q s_{n+1}(x_1, x_2, \ldots, x_n, 1). \quad (9.3)$$

After developing all of this machinery, Ville defines a martingale to be the system of games corresponding to the functions $\{s_n\}_{n\in\omega}$ that satisfy (9.3).[16]

Let us make a few clarifying remarks about this definition. First, the role that the values p and q play in the above definition is that they represent the odds that determine the payoffs made by the house to those who make winning bets. In the case that $p = q = \frac{1}{2}$, the payoff for the game is double-or-nothing. That is, if A successfully predicts the next value of the sequence, he receives a payment equal to double the bet that he placed on that outcome; if he fails to predict the next value, he loses his entire bet. Second, given a sequence X that has limiting relative frequencies of p and q for the attributes 0 and 1, respectively, the game will be fair if the payoff for successfully predicting a 0, given a bet α, is $\frac{\alpha}{p}$ (and thus the payoff for successfully predicting a 1 after a bet of α is $\frac{\alpha}{q}$). For instance, if a 1 occurs only $\frac{1}{10}$th of the time in a sequence, then a fair payoff would be to give A ten times her original bet whenever she successfully predicts the occurrence of a 1. Third, in current work in algorithmic randomness, a martingale is simply defined as a function $M : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ such that

$$2M(\sigma) = M(\sigma 0) + M(\sigma 1), \tag{9.4}$$

where the payoff is given by $p = q = \frac{1}{2}$.[17] That is, the term 'martingale' is usually reserved for those functions that satisfy the equation (9.4). To avoid ambiguity, I

[16] Unfortunately, Ville isn't entirely clear about what it means for a martingale to be the system of games corresponding to the functions $\{s_n\}_{n\in\omega}$ that satisfy (9.3), for he only provides a high-level description of these games. The standard approach nowadays is to identify the martingale with the functions $\{s_n\}_{n\in\omega}$ (actually, just one function that agrees with each s_n on strings of length n, given by (9.4) below).

[17] More generally, if the underlying measure is a Bernoulli p-measure, λ_p, where a 1 occurs with

will henceforth use the term "p-martingale" to refer to any function $M : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ that satisfies

$$M(\sigma) = pM(\sigma 0) + qM(\sigma 1)$$

(where $p + q = 1$) and the term "martingale" to refer to a p-martingale for some $p \in (0, 1)$.

Although we've defined Ville's notion of a betting strategy, we haven't defined what it means for a gambler to succeed when employing a given strategy. The idea is straightforward: a martingale M succeeds on a sequence X if for every N, there is some initial segment $X{\upharpoonright}n$ of X such that $M(X{\upharpoonright}n) \geq N$; that is, a gambler using M will win arbitrarily much capital as she bets on the initial segments of X. But in Ville's formulation, this doesn't mean that gambler's winnings simply grow without bound. Instead, her winnings can dip down to levels arbitrarily close to 0 infinitely often, as long as infinitely often her winnings grow larger and larger. Thus, a p-martingale M succeeds on $X \in 2^\omega$ if

$$\limsup_{n\to\infty} M(X{\upharpoonright}n) = \infty.$$

For a fixed p-martingale M, let $\mathcal{S}_{M,p}$, the *success set* of M, be the collection of $X \in 2^\omega$

---

probability p, then a λ_p-martingale is a function $M : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ such that

$$M(\sigma) = pM(\sigma 0) + qM(\sigma 1).$$

In the most general case, which Ville doesn't consider, for a probability measure μ on 2^ω, a μ-martingale M is a function $M : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ that satisfies

$$M(\sigma) = \mu(\sigma 0|\sigma)M(\sigma 0) + \mu(\sigma 1|\sigma)M(\sigma 1),$$

where $\mu(\sigma i|\sigma)$ is the conditional probability $\frac{\mu(\sigma i)}{\mu(\sigma)}$ for $i \in \{0, 1\}$.

on which M succeeds. Thus, M succeeds on $X \in 2^\omega$ if and only if $X \in \mathcal{S}_{M,p}$.[18]

Having established his definition of martingale, Ville next turns to the task of extending von Mises' axiom of irregularity by means of the notion of martingale.

9.3.2 Ville's Correspondence between Martingales and Null Sets

Ville opens his next section, "Extension of the axiom of irregularity in the sense of von Mises-Wald by means of the notion of martingale",[19] by writing, "It is clear from von Mises' account that the notion of selection and of the invariance of frequency was intended in his mind to express the impossibility of indefinitely winning a fair game"([Vil39], p. 89).[20] As noted above, Ville held that in von Mises' formulation of the principle of the impossibility of a gambling system, an important aspect of a gambler's behavior was not captured: the ability of the gambler to vary his bets. Ville, by contrast, was able to show on his approach to betting strategies, a bias could be exploited by a gambler to win arbitrarily much capital, for he proved the following:

Theorem 9.1. *For every set $\mathcal{U} \subseteq 2^\omega$ of p-measure zero, there is a p-martingale M*

[18]One might worry here that this definition involves a standard of success that is too idealized, as the gambler who employs it must have unbounded resources of time at her disposal. While one can slightly weaken the criterion of success by replacing the lim sup in the definition of success with a limit, if we weaken the criterion of success so that a gambler can succeed on a sequence only after finitely many steps, then whether a martingale succeeds on a sequence will be determined by its finite initial segments, a consequence hardly appropriate for defining random sequences.

[19]In the original French, "Élargissement à l'aide de la notion de martingale de l'axiome d'irrégularité au sens de de Misès-Wald."

[20]In the original French: "Il ressort de l'exposé de M. de Misès que la notion de sélection et d'invariance de la fréquence était destinée dans son esprit á exprimer l'impossibilité de gagner indéfiniment à un jeu équitable."

such that $\mathcal{U} \subseteq \mathcal{S}_{M,p}$.

In other words, for every set $\mathcal{U}$ of p-measure zero, there is a p-martingale that succeeds on every sequence in $\mathcal{U}$. In particular, if we let $\mathcal{U}$ be the collection of sequences X that have limiting relative frequency not equal to p, then since $\mathcal{U}$ is a set of p-measure zero, it follows from Theorem 9.1 that there is a p-martingale M such that $\mathcal{U} \subseteq \mathcal{S}_{M,p}$.[21]

Another consequence of Theorem 9.1 is that Ville's formulation of gambling system extends that of von Mises:

Corollary 9.2. *For every countable collection of place selections* $\mathscr{S}$ *and every* $p \in (0,1)$, *there is a* p*-martingale* M *such that* $2^\omega \setminus C(\mathscr{S}, p) \subseteq \mathcal{S}_{M,p}$.

For every countable collection $\mathscr{S}$ of place selections, there is a martingale M that succeeds on the set of sequences that are not invariant under the selections in $\mathscr{S}$. Thus, Ville's formulation of a gambling system is an extension of that of von Mises: for every von-Mises-gambling system, there is a Ville-gambling-system that succeeds on every sequence on which the von-Mises-gambling-system succeeds. Further, Ville's Theorem shows is that the converse of the above corollary doesn't hold: If we let

$$\mathcal{G} := \left\{ X \in 2^\omega : (\forall n) \frac{\#\{i < n : x_i = 1\}}{n} \geq \frac{1}{2} \right\}$$

as in the proof of Ville's Theorem, then there is a $\frac{1}{2}$-martingale M such that $\mathcal{G} \subseteq \mathcal{S}_{M,\frac{1}{2}}$,

[21] Note that Ville doesn't tell us explicitly how to compute the martingale M. An explicit construction can be found in Laurent Bienvenu's his dissertation, *Game-theoretic Characterizations of Randomness: Unpredictability and Stochasticity* [Bie08]. There he shows that for any $\delta > 0$, one explicitly define a $\frac{1}{2}$-martingale M_δ that succeeds on any sequence for which the limiting relative frequency of 1s is greater that $\frac{1}{2} + \delta$.

but there is no collection of collectives $\mathscr{S}$ such that $C(\mathscr{S}, \frac{1}{2}) \subseteq 2^\omega \setminus \mathcal{S}_{M,\frac{1}{2}}$. For otherwise we would have

$$\mathcal{G} \subseteq \mathcal{S}_{M,\frac{1}{2}} \subseteq 2^\omega \setminus C(\mathscr{S}, \frac{1}{2}),$$

which is impossible by Ville's Theorem.

It should be mentioned that Ville also proved the converse of Theorem 9.1, a result that will be of interest shortly.

Theorem 9.3. *For a fixed p-martingale M, the success set of M, $\mathcal{S}_M, p$ has p-measure zero.*

Thus, Ville establishes a one-to-one correspondence between martingales and null sets.

9.4 The Exemplary ideal of Completeness

9.4.1 In Search of an Improved Axiom of Irregularity

Having provided his definition of martingale and established the correspondence between martingales and null sets, Ville turns his attention to isolating those properties in terms of which he can define randomness, or as Ville puts it, irregularity. To motivate his account of irregularity, Ville writes in the first section of his fourth chapter,

> We will see that one can give conditions of irregularity (more strict than those of Wald[22]), forming a system $\mathcal{C}$ of conditions C such that if one

[22]Throughout his dissertation, Ville often refers to "Wald's conditions of irregularity" or "the

> also denotes by C the set of points x such that the sequence x satisfies the condition C, one has the property:
>
> (Γ). *Any set in the complement of a set C can be enclosed in a set of p-measure zero, and any set of p-measure zero can be enclosed in the complement of a set C* ([Vil39], p. 86).[23]

In other words, Ville intended for his conditions of irregularity to be invulnerable to the deficiency of von Mises' approach as revealed by Ville's Theorem. That is, Ville's expresses the goal of providing conditions of irregularity, subsets of 2^ω that I will refer to as *irregular sets*, so that not only is the complement of every irregular set (which I will call a *regular set*) contained in a set of measure zero, but every set of measure zero is contained in a regular set.

But which sets should be counted as the irregular sets? Ville considers one suggestion, "There is an immediate way to answer the question [as to which sets are the irregular sets]; it is to take $\mathcal{C}$ to be all the sets of p-measure equal to 1, or even only the F_σ[24] sets of p-measure equal to 1" ([Vil39], p. 87).[25] However, the astute

von Mises-Wald conditions", but to be consistent with my usage in earlier chapters, I will refer to these conditions as "von Mises' conditions of irregularity" (although Wald is certain worthy of credit for sharpening von Mises' definition).

[23]In the original French:

> Nous allons voir que l'on peut donner des conditions d'irrégularité (plus strictes que celle de M. Wald), formant un système $\mathcal{C}$ de conditions C telles que si l'on désigne également par C l'ensemble des points x tels que la suite x satisfasse á la condition C (p. 49), on ait la propriété: (Γ). *Tout ensemble complémentaire d'un ensemble C peut être enfermé dans un ensemble de p-mesure nulle, et tout ensemble de p-mesure nulle peut être enfermé dans un complémentaire d'ensemble C.*

[24]A set $\mathcal{S} \subseteq 2^\omega$ is F_σ if it is the countable union of closed subsets of 2^ω.

[25]In the original French: "Il y a une manière immédiate de résoudre la question, c'est de prendre pour les C tous les ensembles de p-mesure égale à 1, ou même seulement les F_σ de p-mesure égale à 1."

reader will recognize that there is a serious problem with this approach: If we were to take the conditions of irregularity $\mathcal{C}$ to be the collection of all sets of measure one, then the result would be a vacuous notion of irregularity, as the intersection of all such sets would be empty, for the simple reason that for each $X \in 2^\omega$, the set $2^\omega \setminus \{X\}$ has measure one. This is just another version of the admissibility objection raised against von Mises' definition of collectives, the upshot of which is that if we define collectives in terms of *every* place selection, we end up with a trivial definition of randomness.

Ville recognized that the collection of irregular sets must be restricted in some way, and he was aware of this version of the admissibility objection,[26] but he didn't justify the claim by appealing to this objection. Instead, he gives a surprisingly different reason, writing, "But these conditions of irregularity *no longer present any intuitive character*" ([Vil39], p. 87, emphasis added).[27] But what is this intuitive character, and why should it act as a constraint on our choice of the irregular sets? At this point of his discussion in the *Étude*, Ville doesn't say, but instead presses on to determine which sets should be taken to the be irregular sets. Let's press on with him, returning to these questions about the intuitive character shortly.

Now, given the correspondence between martingales and sets of null sets that we discussed at the end of Section 9.3, the question

> "Which conditions should be used to characterize the axiom of irregularity?"

[26]In particular, he references the admissibility objection in passage on Fréchet and Lévy quoted below on page 249.

[27]In the original French, "Mais ces conditions d'irrégularité ne présentent plus aucun caractère intuitif."

can be reformulated as

> "Which martingales should be used to define the axiom of irregularity?"

Although one might hope that with this recast question, we are in a better position to characterize irregularity in terms of martingales instead of place selections, Ville concedes that even on his more general account, there is no reasonable answer to be given. He writes,

> But the condition of irregularity given in terms of martingales is relative; it assumes a prior choice of properties (of probability zero) to exclude. If, in a certain sense, it solves the question of irregularity more completely than the condition of Wald, *it fails to give an arithmetical model of a sequence with all the characteristics of a sequence taken at random*; this last problem is considered by us as unsolvable, and we submit ourselves on this point to the opinion of numerous mathematicians, among them E. Borel, Fréchet, P. Lévy ([Vil39], p. 93, emphasis added).[28]

This is a telling passage, for it clearly indicates what Ville was aiming for in identifying the conditions of irregularity: Ville wanted a definition of irregularity that would provide "an arithmetical model of a sequence with all the characteristics of a sequence taken at random". But not just any model will do: Ville sought a complete definition, one satisfying *all* the properties typically held by sequences chosen at random, properties I referred to in the introduction as $\mathcal{R}$-properties. This, then is the

[28]In the original French:

> Mais la condition d'irrégularité par la martingale est relative; elle suppose un choix préalable des propriétés (de probabilité nulle) à exclure. Si, dans un certain sens, elle résout la question de l'irrégularité plus complètement que la condition de M. Wald, elle ne parvient pas à donner un modèle arithmétique d'une suite présentant *tous* les caractères d'une suite prise au hasard; ce dernier problème est considéré par nous comme insoluble, et nous nous soumettons sur ce point à l'opinion de nombreux mathématiciens, parmi lesquels MM. E. Borel, Fréchet, P. Lévy.

exemplary ideal of completeness: a definition $\mathscr{D}$ of randomness is complete if and only if all $\mathscr{D}$-random sequences, being paradigmatic instances of randomly chosen sequences, satisfy all $\mathcal{R}$-properties.

9.4.2 The Exemplary Ideal as Unattainable

But why exactly did Ville hold that the exemplary ideal was unattainable? From the quote given above, the answer appears to be that on the one hand, a definition of irregularity given in terms of martingales is always relative to a choice of martingales, but on the other hand, in Ville's view, only an absolute definition, one not determined by a choice of martingales, could attain the exemplary ideal.

Further elaborating on this relativity in the concluding chapter of the *Étude*, Ville writes,

> We concede that the definition [of irregularity] can only be relative, and here is why: as noted by Fréchet, if we want to clarify how we might recognize in the classical theory the sequences that do not have the character of incidental sequences, one is led to consider as non-incidental the sequences that are logically possible, but which are not encountered in practice, that is to say, which have a property of which the probability is zero in the modernized classical theory. Or, as noted by P. Lévy, if one excludes all the sequences presenting a property of probability zero, one excludes all the logically possible sequences. It is then necessary to make a choice among the properties of probability zero. Once this choice is made, the sequences that do not possess these properties will be considered as irregular relative to these properties ([Vil39], pp. 135-136).[29]

[29]In the original French,

> Nous concédons que la définition ne peut ètre que relative, et voici pourquoi: comme l'a fait remarquer M. Fréchet, si l'on veut préciser comment on pourrait reconnaître dans la théorie classique les suites qui n'offrent pas le caractère des suites fortuites, on est amené à considérer comme non fortuites les suites logiquement possibles, mais qu'on

In Fréchet's view, to define the irregular/incidental sequences, we should take as regular/non-incidental those sequences that are not encountered in practice (though it is not impossible for them to occur); for instance, the infinite sequence of all heads produced by the tosses of a fair coin is such a sequence.[30] But according to Lévy, there is no *non-arbitrary* choice of sets of measure zero in terms of which we can define these irregular sequences. Thus, if both Fréchet and Lévy are right, we have no alternative but to define irregularity as a relative notion.

Ville reiterates this conclusion in the paragraph that follows the passage about Fréchet and Lévy given above:

> In this manner, the notion of irregularity is defined in a relative manner: the systems of properties of probability zero is chosen arbitrarily. If the system chosen is denumerable, there will exist surely sequences irregular with respect to this system. We thus obtain a relative, but coherent, definition of irregularity. We can consider this solution as incomplete, but the solution of this question, in our view, cannot be made such. All we can do is seek to make the solution less abstract ([Vil39], p. 136).[31]

> ne rencontre pas pratiquement, c'est-á-dire qui ont une propriété dont la probabilité est nulle dans la théorie classique modernisée. Or, comme l'a fait observer M. P. Lévy, si l'on exclut toutes les suites présentant une propriété de probabilité nulle, on exclut toutes les suites logiquement possible. Il faut donc faire un choix parmi les propriétés de probabilité nulle. Une fois ce choix fait, les suites qui ne posséderont pas ces proprietétés seront considérées comme irrégulières *relativement à ces propriétés.*

[30]More precisely, sufficiently long initial segments of this sequence are not encountered in practice.

[31]In the original French,

> De cette maniére, la notion d'irrégularité est définie de manière relative; ce qui est choisi arbitrairement, c'est le système de propriétés de probabilité nulle. Si le système choisi est dénombrable, il existera sûrement des suites irrégulières relativement à ce système. Nous obtenons ainsi une définition relative, mais cohèrente, de l'irrégularité. On peut considérer cette solution comme incomplète, mais la solution de cette question, à notre avis, ne peut être que telle. Tout ce que l'on peut faire, c'est de chercher à rendre la

Here Ville repeats that the claim that the exemplary ideal of completeness cannot be attained; not only is his solution to the irregularity problem incomplete, but *no* solution can be complete. But he adds an interesting twist that does not appear in his earlier discussion of the unattainability the exemplary ideal: "All we can do is seek to make the solution less abstract." What does it mean of a solution to be abstract, and what does it mean for one solution to be less abstract than other?

Ville provides a first clue when he later writes, "The definition of irregularity by selections, proposed by Wald, has the advantage of having a concrete sense; but it is not equivalent to the above abstract definition, although it is also a relative solution" ([Vil39], p. 136).[32] Later, he sheds more light on the concreteness of the von Mises-Wald definition, writing, Ville continues,

> On the contrary, to give a countable system of selections and to exclude the sequences where the total frequency is destroyed or modified by application of selections from the system, this is a proposition that has an immediate sense and one we can place at the beginning of a theory, as did M. Wald; similarly, to give a martingale and to exclude the sequences where a player, in applying this martingale, wins indefinitely, this has a very concrete sense, independent of all measure theory. We will consider these last criteria as satisfying from an intuitive point of view ([Vil39], p. 138).[33]

---

solution moins abstraite.

[32]In the original French, "La définition de l'irrégularité par les sélections, proposée par M. Wald, a l'avantage d'avoir un sens concret; mais elle n'est pas équivalente à la définition abstraite ci-dessus, bien qu'elle aussi soit une solution relative."

[33]In the original French,

> Au contraire, se donner un système dénombrable de sélections et exclure les suites où la fréquence totale est détruite ou modifiée par application d'une des sélections dy système, c'est là une proposition qui a un sense immédiat, et que l'on peut placer au début d'une théorie, comme l'a fait M. Wald; de même, se donner une martingale et exclure les

Thus, some characterizations of irregularity are concrete, immediate, and intuitive. What is distinctive of these characterizations is that they are not given in purely extensional terms but are given by some appropriate description (involving intuitive notions such as betting strategies and games of chance), unlike the "abstract conditions of irregularity", the phrase Ville uses to refer to the sets of measure one. Ville is clear on this point, saying of these abstract conditions that they do not have a "concrete and intuitive meaning" ([Vil39], p. 137). In support of this claim, Ville observes that the complement of the Cantor set[34] is included among these abstract conditions of irregularity, but taking the complement of the Cantor set to be a criterion of irregularity is, in his view, "a proposition that will be considered as reasonable only by someone who has pushed far enough in the study of probability, and which, separated from measure theory and the theory of denumerable probabilities, seems arbitrary" ([Vil39], p. 138).[35,36] In this respect, the choice to include the complement of the Cantor set among the irregular sets is not intuitive, requiring additional, specialized knowledge to justify the inclusion.

Thus, we have concrete and abstract characterizations of irregularity, and con-

suites où un joueur, en appliquant cette martingale, s'enrichit indéfiniment, cela a un sens très concret, indépendant de toute théorie de la mesure. Nous considérerons ces derniers critères comme satisfaisants au point de vue intuitif.

[34]The middle third Cantor set is a subset of [0,1] of measure zero, formed by removing the middle third from [0,1], then the middle thirds of the remaining intervals $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$, and so on.

[35]"une proposition qui ne sera considérée comme raisonnable que par quelqu'un qui a poussé assez loin l'étude des probabilités, et qui, séparée de la théorie de la mesure et de celles des probabilités dénombrables, paraît arbitraire."

[36]Ville adds, "We consider such a criterion as not having an immediate subjective meaning" ("Nous considèrons un tel critére comme n'ayant pas de signification subjective immédiate.") ([Vil39], p. 138).

crete ones are to be preferred to the abstract ones. But further, according to Ville, certain concrete characterizations of irregularity are more complete than others. But how do we measure the degree of completeness of a concrete characterization of irregularity?

The answer is to consider the relation of implication that might hold between a concrete condition of irregularity and an abstract condition of irregularity. Given two sets of concrete conditions of irregularity $\{C_i^1\}_{i\in I}$ and $\{C_i^2\}_{i\in J}$,[37] we'll say that $\{C_i^1\}_{i\in I}$ is *more complete* than $\{C_i^2\}_{i\in J}$ if for every abstract condition of irregularity A, if there is some concrete condition C_k^2 from the latter collection such that

$$C_k^2(X) \Rightarrow A(X)$$

for all $X \in 2^\omega$, there is some concrete condition C_ℓ^1 from the former collection such that

$$C_\ell^1(X) \Rightarrow A(X).$$

From this it follows that Ville's collection of conditions of irregularity, given in terms of martingales, is more complete than the collection of conditions of irregularity given in terms of von Mises' place selections. On von Mises approach, "we obtain concrete criteria, but the irregularity is not defined in a manner as complete as possible", given that there are sets of measure zero that we cannot show to have measure zero by means of place selections (namely, the collection of sequences that fail to satisfy the Law of the Iterated Logarithm). However, defining irregularity in terms

[37]Here I and J are index sets that need not be countable.

of martingales provides a more complete solution, as "the criteria of irregularity using the notion of martingale have both a concrete sense and do not give rise to the gap that we have indicated" ([Vil39], p. 138).[38]

In fact, given the correspondence between martingales and null sets established by Ville, the conditions of irregularity given by martingales are the *most* complete, since for every abstract condition of irregularity (i.e. set of measure zero), there is a concrete condition of irregularity, defined in terms of a martingale, that implies it. Nonetheless, this completeness does n't imply the attainment of the exemplary ideal. As Ville puts it,

> It can be naturally possible to find criteria of irregularity that have another expression; but the practical solution, which should give criteria having an intuitive sense, can never go beyond the abstract solution: there is a limitation that is due to the nature of the problem ([Vil39], p. 138).[39]

This view, that it is due to the nature of the problem of irregularity that no intuitive definition of irregularity (and thus no definition of randomness) can attain the exemplary ideal, did not prevent others from seeking an intuitive definition of randomness that could attain the ideal. In particular, the view was neither shared by Per Martin-Löf nor C.P. Schnorr, who each produced definitions of randomness

[38] In the original French, "[L]es critères d'irrégularité utilisant la notion de martingale ont à la fois un sens concret et ne donnent pas lieu à la lacune que nous venons de signaler."

[39] In the original French,

> Il peut être naturellement possible de trouver des critères d'irrégularités ayant une autre expression; mais la solution pratique, qui doit donner des critères ayant un sens intuitif, ne peut jamais aller plus loin que la solution abstraite: il y a là une limitation qui tient à la nature du probeème.

that they took to attain some version of the exemplary ideal (at least initially, in the case of Martin-Löf).

9.5 Martin-Löf's Definition of Randomness

In 1966, Per Martin-Löf published the paper "The Definition of Random Sequences", which contained a definition of algorithmic randomness for infinite sequences with the remarkable property of extending Church's definition of randomness while at the same time avoiding the problem posed by Ville's Theorem, as every sequence random according to Martin-Löf's definition satisfies the Law of the Iterated Logarithm. Thus for the first time after the publication of Ville's *Étude*, a legitimate contender for a definition of randomness satisfying the exemplary ideal of completeness was put forward: "Finally, the [collectives] introduced by von Mises obtain a definition which seems to satisfy all intuitive requirements" ([ML66], p. 602). However, this remark raises two questions that we will have to address here: What did Martin-Löf understand these intuitive requirements to be, and why did he think that his definition satisfied them?

A partial answer to this question is provided by Martin-Löf's discussion of the definition of randomness for finite strings developed by A.A. Kolmogorov, his dissertation supervisor. In fact, Martin-Löf intended his own definition for infinite sequences to be an extension of the definition given by Kolmogorov. Let us briefly review this definition.

9.5.1 Kolmogorov's Definition of Random Finite Strings

The main intuition behind Kolmogorov's definition of randomness for finite strings is that intuitively, an object such as a sufficiently long finite binary string or an infinite binary sequence is judged to be random if it contains no discernible pattern.[40] This phrase "judged to be random" is typically taken to mean something like "judged to be of random origin", for instance by being produced by the tosses of a fair coin. For example, the string

$$11111111111111111111111111111111111111111111111111 \tag{9.5}$$

will be judged by any competent thinker to be non-random, insofar as it is highly unlikely that this string was produced by the tosses of a fair coin. But one might object that for any string of the same length as the above string (9.5) (50 bits, to be precise), the probability of producing that string is the same, namely 2^{-50}. But there is an important difference between the string (9.5) and the other $2^{50} - 1$ binary strings of length 50: (9.5) is the only binary string of length 50 that contains no 0s, and thus it is highly improbable that we obtain a string that contains no 0s by the tosses of a fair coin. In fact, it is already quite probable (about a 66% chance) that a binary string of length 50 produced by the tosses of a fair coin will contain between 22 and 28 0s. However, the string

$$01010101010101010101010101010101010101010101010101 \tag{9.6}$$

[40]For the moment, I am being deliberately vague about what I mean by a 'discernible pattern', but I'll be more specific about this shortly.

certainly satisfies the property of having between 22 and 28 0s. Nonetheless, it is highly unlikely that (9.6) has been produced by the tosses of a fair coin (since the vast majority strings of length 50 contain the substrings 00 and 11), and thus we judge (9.6) to be non-random.

We can continue to consider more and more strings of length 50, ruling them out as non-random on the basis of some property or other, but it would be nice to have some fixed collection of properties $\mathcal{P}$, which we might consider as the collection of detectable patterns, so that for any string of length 50 (or any sufficiently long string, for that matter) that has a property in $\mathcal{P}$, we are justified in concluding that it is non-random.

Kolmogorov's insight is that any string τ instantiating such a detectable pattern will thereby be compressible: there is some Turing machine M and another string σ such that $M(\sigma) = \tau$ and $|\sigma| < |\tau|$. Thus, we might call a string τ (M, c)-random if for every σ such that $M(\sigma) = \tau$, we have $|\sigma| \geq |\tau| - c$; in other words, a finite string is random if it cannot be compressed very much.

More precisely, we define

$$C_M(\tau) = \min\{|\sigma| : M(\sigma) = \tau\}$$

to be the plain Kolmogorov complexity of τ relative to the machine M. However, if we fix a universal Turing machine U, we have for every Turing machine M there is a constant d_M such that

$$C_U(\sigma) \leq C_M(\sigma) + d_M$$

for every finite string σ, and thus we can define

$$C(\sigma) := C_U(\sigma)$$

for every σ (as changing the choice of universal machine only changes the values given by the associated complexity measure by at most an additive constant).

With this definition of complexity, we can now define what it means for a string to be random. As noted above, random strings should not be compressible, and so the formal analogue of this is to require the Kolmogorov complexity of a random string to be not much lower than the length of the string itself. Thus, we say that $\sigma \in 2^{<\omega}$ is c-incompressible if

$$C(\sigma) \geq |\sigma| - c.$$

Of course, c-incompressibility only provides a reasonable notion of randomness for strings that are sufficiently long.

Before we return to Martin-Löf's contribution, there is a conditional variant of Kolmogorov complexity that bears mentioning. On the conditional approach, the Turing machines in the definition now take two strings as input, $M(\sigma, \rho)$. Again, fixing a Turing machine U that is universal for this collection of machines, we define

$$C_U(\tau|\rho) = \min\{|\sigma| : U(\sigma, \rho) = \tau\},$$

the idea being that this measures the amount of algorithmic information that ρ has about τ. Lastly, as above, we can define a string $\sigma \in 2^{<\omega}$ to be c-incompressible relative to $\rho \in 2^{<\omega}$ if

$$C_U(\sigma|\rho) \geq |\sigma| - c.^{41,42}$$

9.5.2 Martin-Löf on Kolmogorov's Definition

As mentioned above, Martin-Löf intended to provide a definition of randomness for infinite sequences extending Kolmogorov's definition for finite strings. Martin-Löf certainly held this definition in high esteem, as evidenced by statements such as the following two:

> Kolmogorov has proposed a definition of randomness for which strong arguments can be given that it is coextensive with our corresponding intuitive concept ([ML69a], p. 265).

> The thesis has been put forward by Kolmogorov that this provides an adequate formalization of our intuitive notion of randomness ([ML66], pp. 603-604).

Martin-Löf further supports these claims by proving, in his words, "that the random elements as defined by Kolmogorov possess all conceivable statistical properties of randomness" ([ML66], p. 602). Martin-Löf is particularly bold in referring here to "all conceivable statistical properties" rather than, say, "all statistical properties" or

[41] There are many questions that one can raise concerning Kolmogorov's definition, the notion of algorithmic information, problems concerning the choice of universal machine in terms of which we define Kolmogorov complexity, etc., but there is no space to even begin to raise these questions here.

[42] For more information on Kolmogorov complexity, see, for instance, Kolmogorov's original paper on his measure of complexity, "Three Approaches to the Definition of the Notion of Amount of Information" [Kol65], or Li and Vitanyi's textbook, *An Introduction to Kolmogorov Complexity and Its Applications* [LV97].

"all perceivable statistical properties" is somewhat surprising, but Martin-Löf does have some reason for this stronger claim, which we will discuss shortly. The key point is that in Martin-Löf's view, the exemplary ideal can be attained, at least by a definition of random finite strings. Later, he adds, "In order to justify the proposed definition of randomness [that is, Kolmogorov's definition] we have to show that the sequences, which are random in the stated sense, possess the various properties of stochasticity with which we are acquainted in the theory of probability" ([ML66], p. 604).

To prove that Kolmogorov's definition attains the exemplary ideal, Martin-Löf offers a statistical definition of randomness for finite strings. To motivate this statistical definition of randomness, Martin-Löf directs his reader to consider the statistical test that tests for the hypothesis that a given string has a relative frequencies of 1s that is sufficiently close to 1/2. In implementing such a test, we reject the hypothesis that the string is of random origin when the relative frequency of 1s is sufficiently far from $\frac{1}{2}$. But how far is sufficiently far?

Martin-Löf answers this question by appealing to levels of significance. In standard statistical testing, given a null hypothesis H_0 (such as "the string σ is of random origin"), we reject the hypothesis at a significance level $\alpha \in (0, 1)$ if the probability of rejecting a true null hypothesis is at most α.[43] To reflect this aspect of statistical practice Martin-Löf considers levels of significance $\epsilon = \frac{1}{2,}\frac{1}{4}, \frac{1}{8}, \cdots$, stipulates that the null hypothesis, that the string $\xi_1\xi_2 \ldots \xi_k$ is random, is to be rejected at level $\epsilon = 2^{-m}$

[43] The rejection of a true null hypothesis is referred to as a type I error.

if

$$|2s_n - n| > f(m, n), \tag{9.7}$$

where $s_n = \sum_{i=1}^{n} \xi_i$ and $f(m, n)$ is the least value such that the inequality (9.7) holds for no more than 2^{n-m} strings of length n. That is, we reject the null hypothesis at level $\epsilon = 2^{-m}$ on the basis of the condition (9.7), which holds for $\frac{1}{2^m}$th of all strings of length n.

In the more general setting, Martin-Löf considers sequential tests, which are "prescriptions" such that, given a level of significance ϵ, specifies a collection of strings for which we should reject the null hypothesis. This collection of strings is referred to as a critical region for the significance level ϵ. More precisely, a sequential test is a set $U \subseteq \omega \times 2^{<\omega}$ such that

(i) $U_n := \{\sigma : (n, \sigma) \in U\}$ is the critical region for the significance level $\epsilon = 2^{-n}$;

(ii) $|U_m \cap 2^n| \leq 2^{n-m}$; and

(iii) $U_n \supseteq U_{n+1}$ for every $n \in \omega$.

Moreover, this prescription needs to be given effectively, which amounts to requiring that U be a computably enumerable set of pairs (n, σ). Of this last requirement, Martin-Löf writes, "This is the weakest requirement we can imagine, and, in fact, all the tests of use in statistical practice are even of a much simpler type." ([ML66], p. 605) Already, we see Martin-Löf assert the strength of his definition: *any* test that we might employ in statistical practice is included among the sequential tests.

Martin-Löf proves something even stronger: there is one single test U, such that for every other sequential test V, there is some constant c such that

$$V_{m+c} \subseteq U_m.$$

That is, U is *universal* for all computably enumerable sequential tests.

Given this definition of an effective sequential test for finite strings, two questions naturally arise: How can this notion be used to produce a definition of randomness for finite strings? And given such a definition of randomness, what relation does it have to Kolmogorov's definition?

To answer the first question, we need to consider the notion of a critical level of a string σ, which is the greatest m such that $\sigma \in U_m$, i.e. it is the greatest index of a critical region containing σ, or, as Martin-Löf puts it, "the smallest level of significance on which the hypothesis is rejected." ([ML66], p. 606) Thus we define

$$m_U(\sigma) = \max\{m : \sigma \in U_m\}.[44]$$

By setting $U_0 = 2^{<\omega}$, it follows that $0 \leq m_U(\sigma) \leq |\sigma|$. This suggests that the higher the critical level of a string, the more random it is. That is, we only reject the null hypothesis at a very small level of significance.

This measure $m_U(\sigma)$ of the level of significance of a string with respect to the test U also allows us to answer the second question, concerning the relationship of

[44]Strictly speaking, the critical level not the significance level 2^{-m}, but rather the value m. To be precise, we can define the critical level to be the significance level, and then just define the function m to be minus the logarithm of the critical level.

this definition of randomness with Kolmogorov's, for Martin-Löf proves:

Theorem. There is some constant c such that for each n and each $x \in 2^{<\omega}$ of length n,

$$|(n - C(x|n)) - m_U(x)| \leq c$$

Here's how we should understand this result: on Kolmogorov's conditional definition of randomness, a string that is non-random will be such that $C(x|n) \leq n - c$ for some $c \in \omega$. What Martin-Löf shows is that the universal test U is such that the level at which we reject the hypothesis of randomness is within a constant of the amount that a string can be compressed. That is, for each compressible string, we reject the hypothesis of randomness at a critical level that is determined, within a constant, by the amount the string can be compressed.

Of course, the constant in question depends on our choice of the asymptotically optimal machine, as well as our choice of universal test, but we shouldn't be surprised by this; in the context of finite strings, it appears that we can do no better than establishing results up to an additive constant. Nonetheless, this result is an important one, for it shows a connection between the compressibility of a string and its having certain statistical regularities. Furthermore, this result shows why Martin-Löf holds that strings random according to Kolmogorov's definition "possess all conceivable statistical properties of randomness": as long as each conceivable property for randomness can be tested for by some computably enumerable statistical test, Martin-Löf's claim follows.[45]

[45] Of course, one might reply to Martin-Löf that there are *conceivable* properties of randomness that cannot be tested for by any computably enumerable statistical test. But as long as one holds that a conceivable property of randomness is one that can be tested by a statistical test that is

9.5.3 The Definition of Martin-Löf Randomness

Next Martin-Löf shows how one can extend his definition of randomness for finite strings to one for infinite sequences.[46] In order to motivate this definition, he writes,

> Imagine a random device, such as the tossing of a coin, capable of delivering a potentially infinite binary sequences $\xi_1\xi_2\cdots\xi_n\cdots$. To conform with our intuitive conception of randomness, such a sequence has to satisfy for example the law of large numbers [...], or, requiring more, the law of the iterated logarithm ([ML66], p. 609).

There are several remarks to make about this passage. First, Martin-Löf motivates his definition by considering the properties that are typically held by sequences generated at random. This clearly hearkens back to Ville's desire for a definition of random sequence given in terms of the properties typically held by sequences chosen at random.[47] Second, it is noteworthy that the so-called "intuitive conception of randomness" has now apparently assimilated both the Law of Large Numbers (von

humanly implementable, then this reply is not a very forceful one.

[46]Martin-Löf's transition from the discussion of random finite strings to random infinite strings is remarkable, given both the modesty in how he characterizes his work on random finite strings and the recognition that his definition of random infinite sequences answers a long standing open question:

> In the case of finite binary sequences the introduction of the universal test led to nothing but a useful reformulation of what could have been established by means of the complexity measure of Kolmogorov. We shall now see that by defining in a similar way a universal sequential test we obtain a natural definition of infinite random sequences. Such a definition has so far not been obtained by other methods ([ML66], p. 608).

[47]It is reasonable to ask whether there is any significant different between *generating* a sequence at random and *choosing* a sequence at random. This depends on what we take to be the methods of sequence generation and the methods of sequence selection. Although there may be some interesting cases in which the properties typical of randomly generated sequences differ from the properties typical of randomly selected sequences, for our purposes, consideration of these cases would detract from the present discussion.

Mises' first axiom of collectives) and the Law of the Iterated Logarithm. But are we to understand the satisfaction of these properties as the "intuitive requirements" satisfied by Martin-Löf's definition?

This is not altogether clear. For Martin-Löf writes,

> Wald did away with all purely mathematical objections against von Mises' [collectives]. The only doubt there could remain was whether the precisely delimited mathematical concept constitutes an adequate idealization of our intuitive notion of a random sequence. Have we the right to assert that the [collectives] possess in some sense all possible properties of randomness? ([ML69b], p. 30)

Martin-Löf thus appears to hold that a definition of randomness satisfies all intuitive requirements, yielding an "adequate idealization of our intuitive notion of a random sequence", if and only if it counts as random those sequences that "possess in some sense all possible properties of randomness".

But an unclarity still persists: In what sense do sequences that are random according to an exemplarily complete definition possess all possible properties of randomness? To answer this question, Martin-Löf extends his definition of a sequential test for finite strings to a definition of sequential test for infinite sequences.

The idea is familiar: We begin with a uniform collection U of sets of finite strings, but "in the spirit of constructive analysis", we can think of these strings as defining a constructive open set in 2^ω, where an open set $\mathcal{U} \subseteq 2^\omega$ is constructively open if the collection of strings σ such that $[\![\sigma]\!] \subseteq \mathcal{U}$ is a computably enumerable set.[48]

[48] Recall that $[\![\sigma]\!] = \{X \in 2^\omega : \sigma \prec X\}$.

Next, Martin-Löf defines a constructive null set (or Martin-Löf test) to be a set $\mathcal{A} \subseteq 2^\omega$ such that $\mathcal{A} \subseteq \mathcal{U}_i$ for every $\mathcal{U}_i$ in a constructive sequence $\mathcal{U}_1, \mathcal{U}_2 \ldots$ of constructively open sets with the property that $\lambda(\mathcal{U}_i)$, the Lebesgue measure of $\mathcal{U}_i$, constructively approaches 0 as i approaches infinity. That is, there is some computable function $f : \omega \to \omega$ such that $\lambda(\mathcal{U}_m) \leq 2^{-k}$ for all $m \geq f(k)$. With a bit of work, one can show that this requirement is equivalent[49] to the requirement that $\lambda(\mathcal{U}_k) \leq 2^{-k}$ for every k.[50]

With this definition of a constructive null set, Martin-Löf then defines a sequence X to be random (henceforth *Martin-Löf random*) if and only if for every constructive null set$\{\mathcal{U}_i\}_{i\in\omega}$, $X \notin \bigcap_{i\in\omega} \mathcal{U}_i$. Given that there are constructive null sets that contain every sequence that fails to satisfy the Law of Large Numbers or the Law of the Iterated Logarithm, it follows that every Martin-Löf sequence satisfies these properties.

More generally, as with the definition of a sequential test for finite strings, Martin-Löf claims that his notion of a test for infinite sequences is so general that "[a]ny sequential test of present or future use in statistics is given by an explicit prescription, which, for every level of significance $\epsilon = \frac{1}{2}, \frac{1}{4}, \cdots$, tells us for what sequences the hypothesis is to be rejected" ([ML66], p. 609). Thus, Martin-Löf appears to hold that a property $\mathcal{P}$ is a "possible property of randomness" if and only if there is some computably enumerable statistical test $\mathcal{T}$ such that for all sequences X,

[49] These requirements are equivalent in the sense that both definitions yield the same class of constructive null sets.

[50] As with the definition of a Martin-Löf test for finite strings, Martin-Löf also proves the existence of a universal test of randomness for infinite sequences.

$$X \text{ passes } \mathcal{T} \text{ if and only if } X \text{ has } \mathcal{P}.$$

Martin-Löf's approach bypasses the problem of enumerating natural language descriptions of the properties of randomness by giving merely a syntactical characterization: any property that is satisfied by all and only those sequences passing a given c.e. sequential test can be defined by a certain Σ^0_2 predicate.[51] Hence, if Martin-Löf is right in holding that the possible properties of randomness, those properties typical of sequences generated at random, are to be identified with those properties that are testable via c.e. sequential tests, it follows that his definition attains the exemplary ideal.[52]

[51]That is, given a Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$, there is a Σ^0_2 predicate Φ such that

$$\Phi(X) \text{ if and only if } X \notin \bigcap_{i\in\omega} \mathcal{U}_i.$$

[52]Incidentally, Martin-Löf was not able to prove that his definition of randomness for infinite sequences could be formulated in terms of Kolmogorov complexity. While he was able to show that for $X \in 2^\omega$, if there is some $c \in \omega$ such that

$$(\exists^\infty n) C(X{\restriction}n) \geq n - c \tag{9.8}$$

then X is Martin-Löf random, he could not prove that the converse holds. Moreover, he showed that the set of sequences satisfying

$$(\exists c)(\forall n) C(X{\restriction}n) \geq n - c$$

is empty. See [ML71] for the details. Several years later, Schnorr and Levin showed independently that a necessary and sufficient condition for Martin-Löf randomness is given by

$$(\exists c)(\forall n) K(X{\restriction}n) \geq n - c,$$

where K is a modification of Kolmogorov complexity known as *prefix-free complexity* (discussed in Section 10.3.1), thus vindicating Martin-Löf's attempt to extend Kolmogorov's definition. Interestingly, it was only in the last decade that the condition 9.8 is necessary and sufficient for a sequence to be Martin-Löf random *relative to the halting set*, which means roughly that such sequences pass all statistical tests that come equipped with the halting set as an oracle.

9.6 Schnorr's Alternative Definition

Why should we identify the possible properties of randomness with those properties testable by a c.e. statistical test? This question was raised by Schnorr, who argued that these possible properties of randomness should *not* be identified with those properties testable by the Martin-Löf tests, but rather a proper subclass of the collection of these tests. Let us consider Schnorr's definition of randomness.

9.6.1 The Rationale for Schnorr Randomness

At face value, there isn't much of a difference between Martin-Löf's definition of randomness and Schnorr's definition. If we take a Martin-Löf test to be a computable sequence $\{\mathcal{U}_i\}_{i\in\omega}$ of effectively open subsets of 2^ω such that $\lambda(\mathcal{U}_i) \leq 2^{-i}$ for every $i \in \omega$, then the definition of a Schnorr test is nearly identical: A Schnorr test is simply a Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$ such that $\lambda(\mathcal{U}_i) = 2^{-i}$ for every $i \in \omega$. Then a sequence is Schnorr random if and only if it is not contained in any Schnorr test. Could such a slight modification possibly make a difference?

One immediate difference is that the collection of Martin-Löf random sequences is not coextensive with the collection of Schnorr random sequences: while every Martin-Löf random sequence is Schnorr random (since every Schnorr test is a Martin-Löf test), Schnorr proved that the converse does not hold. But according to Schnorr, there is a further significant difference between these two types of tests. As he writes,

> The deficiency residing in the previous concepts of randomness [Martin-Löf randomness and several other definitions] is, in our opinion, that properties of random sequences are postulated which are of no significance to statistics. Many insufficient approaches have been made until

> a definition of random sequences was proposed by Martin-Löf which for the first time included all standard statistical properties of randomness. However, the inverse postulate now seems to have been violated ([Sch71], p. 255).

Here Schnorr acknowledges that the collection of Martin-Löf tests accounts for every property that one would encounter in actual statistical practice, just as Martin-Löf claimed when he first offered his definition. Yet, in Schnorr's view, the Martin-Löf tests include *too many* tests; "all standard statistical properties of randomness" are guaranteed to hold of Martin-Löf random sequences, but also certain properties that "are of no significance to statistics." Which properties are these? Schnorr continues,

> The acceptable definition of random sequences cannot be any formulation of recursive function theory which contains all relevant statistical properties of randomness, but *it has to be precisely a characterization of all those properties of randomness that have a physical meaning.* These are intuitively those properties that can be established by statistical experience. This means that a sequence fails to be random in this sense if and only if there is an effective process in which this failure becomes evident. On the other hand, it is clear that if there is no effective process in which the failure of the sequence to be random appears, then the sequence behaves like a random sequence ([Sch71], p. 255, emphasis mine).

Thus, Schnorr "the acceptable definition of random sequences" must require that random sequences satisfy all and only the physically meaningful properties. But what counts as a physically meaningful property? In Schnorr's view, a property $\mathcal{P}$ is physically meaningful only if there is an effective procedure such that, for every sequence X, if X does not satisfy $\mathcal{P}$, the effective procedure indicates that X does not satisfy $\mathcal{P}$. Let us call such a property *effectively detectable.*[53]

[53]Strictly speaking, it's not the property, but rather its absence, that is effectively detectable.

For instance, thinking of a martingale d as a test for randomness (as Schnorr does), so that a sequence fails the test if d succeeds on it, Schnorr writes, "[I]f the sequence $d(X \upharpoonright n)$ increases so slowly that no one working with effective methods only would observe its growth, then the sequence X behaves as if it withstands the test F" ([Sch71], p. 256). That is, even if it fails the test, and thus should be counted as non-random, unless this failure can be effectively detected, this sequence is indistinguishable from a "truly" random sequence, i.e., one that does not fail any such test.

9.6.2 An Alternative Formulation of the Exemplary Ideal?

Recall that Church offered a version of von Mises' definition, restricting the collection of place selections to the computable place selections. Moreover, Church suggested that his definition need not produce a theory of probability attaining the resolutory ideal of completeness as understood by von Mises, solving all problems of the probability calculus. For while there are many problems that cannot be solved by means of computable place selections, Church was doubtful whether such problems ever arise in actual practice. Thus, only a restricted version of the resolutory ideal should be attained, and Church at least suspected that the theory of probability given by the computable place selections could attain this ideal.

We now find ourselves in a similar situation with Schnorr's criticism of Martin-Löf's definition. Schnorr acknowledges that although Martin-Löf random sequences satisfy "all standard statistical properties of randomness", but these sequences are required to satisfy too many properties, many of which do not arise in actual statistical practice. The remedy, then, is to restrict to those properties the failure of which can be effectively determined. Corresponding to this restriction, one might formulate a restricted version of the exemplary ideal: A definition of randomness attains the restricted exemplary ideal if and only if it satisfies all effectively detectable properties that are typical of sequences generated at random, i.e. it satisfies all effectively detectable $\mathcal{R}$-properties.

Schnorr, in fact, held that his definition attained this restricted exemplary ideal, referring to Schnorr randomness as "the true concept of randomness" ([Sch71], p. 216). In fact, Schnorr explicitly asserts:

> THESIS [...]: A sequence behaves within all effective procedures like a p-random sequence iff it is Schnorr p-random ([Sch77], p. 198).[54]

Thus, surprisingly, before the explicit formulation of the Martin-Löf-Chaitin Thesis, Schnorr formulated his own thesis, which I will henceforth refer to as Schnorr's Thesis (which I'll abbreviate as ST).

9.6.3 Martin-Löf's Response to Schnorr's Objection?

Was Martin-Löf moved by Schnorr's objection? Although Martin-Löf doesn't explicitly address Schnorr's objections in print, one indication that he wasn't convinced by Schnorr's objection is that in 1970, he published an alternative definition of randomness which is even *stronger* than Martin-Löf randomness. He writes,

[54] Here p is some computable probability measure, as Schnorr is considering randomness with respect to various measures. In this study, we restrict our attention primarily to randomness with respect to the Lebesgue measure, although I briefly discuss randomness for different measures in Chapter 12.

> The practice in probability theory is the following. As soon as we have proved that a certain property, such as the law of large numbers [...] or the law of the iterated logarithm [...] has probability one, we say that this is a *property of randomness*. However, if we try, within the classical mathematical framework, to define a sequence to be random if it possesses all properties of randomness, we are led to a vacuous notion ([ML70], p. 74).

Martin-Löf's response here is not to define the properties of randomness in terms of c.e. statistical tests, but in terms of *hyperarithmetical* tests. He writes,

> It is proposed to avoid this paradox, born of the classical conception of the totality of all sets of probability one, by restricting our attention to hyperarithmetical sets or, equivalently, to properties expressible in the constructive infinitary propositional calculus. This may be regarded as a constructive version of the restriction to Borel sets which is usually accepted in probability theory. Actually, the specific Borel sets considered there are always obtained by applying the Borelian operations to recursive sequences of previously defined sets, which means precisely that they are hyperarithmetical ([ML70], p. 74).

To pursue the details of this definition would take us far from the task at hand, but the point worth making is that Martin-Löf was not ultimately content with his claim that his definition attained the exemplary ideal. Contra Schnorr, random sequences must satisfy even more conditions than those enforced by Martin-Löf tests, not fewer.

9.7 Summing Up

There are a few important points to take away from this discussion, as we move to the task of evaluating the MLCT and related theses. As we've seen, the exemplary ideal is a problematic ideal of completeness for definitions of randomness, for several reasons. First, there is no clear identification of the $\mathcal{R}$-properties that are to be satisfied by a sequence that is random according to an exemplarily complete definition. Yet without an identification of the $\mathcal{R}$-properties, we have no way to evaluate the claim that a given definition attains the exemplary ideal.

A second problem is that, in our discussion of the exemplary ideal, there is a notable lack of any talk of why we might want an exemplarily complete definition in the first place. That is, neither Ville, Martin-Löf, nor Schnorr tell us what we really gain by identifying an exemplarily complete definition. Although in the accounts of Ville, Martin-Löf, and Schnorr, reference is made to "statistically meaningful properties" and the practice of statistics and probability theory, but it is far from clear what role an exemplarily complete definition of randomness would play in those disciplines. One begins to wonder: Is this merely idealization for the sake of idealization? This is an important question to keep in mind as we turn towards understanding and evaluating the MLCT.

CHAPTER 10

THE STATUS OF THE MARTIN-LÖF-CHAITIN THESIS

10.1 Introduction

Martin-Löf's claim that his definition of randomness for infinite sequences "appears to satisfy all intuitive requirements", which we encountered in the latter half of the previous chapter, has been repeated a number of times in the years following the publication of his paper "The Definition of Random Sequences". In particular, some have held that in Martin-Löf randomness, we find a definition that captures what Martin-Löf referred to as the "intuitive conception of randomness", just as the formal definition of a computable number-theoretic function is held to capture the intuitive conception of an effectively calculable number-theoretic function. This statement, introduced in Chapter 7 as the Martin-Löf-Chaitin Thesis, can be expressed concisely as

> MLCT: An infinite binary sequence is Martin-Löf random if and only if it is intuitively random.

My goal in this chapter is to inquire into the status of the MLCT. In particular, I want to focus my attention on the claim that Martin-Löf randomness can serve as the basis of a conceptual analysis of the notion of randomness, thereby filling the conceptual-analytic role.

Given that extensional adequacy is a necessary condition for a successful conceptual analysis, Martin-Löf randomness successfully fills the conceptual-analytic role only if the MLCT is true. But what reasons are there for accepting the MLCT? Moreover, are there any good reasons for rejecting it? To answer these questions, I survey and evaluate the main arguments that have been given support of the MLCT, as well as the arguments given against it, many of which have been offered by those who hold that some alternative definition of randomness captures the so-called intuitive conception of randomness.

There are two primary reasons for carefully attending to the details concerning the status of the MLCT. First, a systematic treatment of the various arguments given for and against the MLCT has not been provided in the philosophical literature on algorithmic randomness, and so the catalogue I provide here fills an important void. Second, the issues that face the advocate of the MLCT, the MLR-*advocate*, are the same issues that face the advocate of the claim that definition $\mathscr{D}$ captures the so-called intuitive conception of randomness (to whom I have been referring as the $\mathscr{D}$-advocate), where $\mathscr{D}$ is any currently available definition of randomness. Further, these issues will feature prominently in my discussion of the Justificatory Challenge that the $\mathscr{D}$-advocate must meet, as laid out in the next chapter.

The rest of the chapter will proceed as follows. In Section 10.2, I carry out some initial ground-clearing by making several clarificatory remarks, which in turn will facilitate our discussion of the MLCT. Next, in Section 10.3, I lay out the key pieces of evidence that have been offered in support of the MLCT, namely a number of convergence results, theorems that show that two intensionally non-equivalent definitions of randomness are nonetheless extensionally equivalent. As there is good reason to hold that these convergence results alone do not suffice to establish the MLCT, I suggest several ways that the MLR-advocate might supplement these convergence results. In the course of discussing severals ways of supplementing these convergence results, we will begin to see the challenge facing the MLR-advocate to carry our this justificatory task. In Sections 10.4 and 10.5, I discuss the challenge posed to the MLCT by two alternative definitions of randomness, Schnorr randomness and weak 2-randomness, and I further discuss how the MLR-advocate might respond to this challenge. Lastly, in Section 10.6, in anticipation of my discussion of the Justificatory Challenge, I close with some general reflections.

10.2 Some Clarificatory Remarks

Before I outline the evidence given in support of the MLCT and discuss the challenge that other definitions of algorithmic randomness pose to the MLCT, several clarificatory remarks are in order.

10.2.1 The So-Called Intuitive Conception of Randomness?

Our first order of business is to clarify what is meant by the phrase "the intuitive conception of randomness", which regularly appears in the literature on algorithmic randomness. The most immediate concern one might have with this phrase is the definite article: why think there is just one intuitive conception of randomness and not many? For if we hold that a conception of randomness serves as the basis of our judgments of randomness, and that an intuitive conception of randomness is an informal, pre-theoretic collection of beliefs on the basis of which one makes judgments of randomness, then there seems to me to be no reason to restrict to one single conception as *the* intuitive conception.

Nonetheless, as it is not unreasonable to hold that there are widely held beliefs about the notion of randomness that serve as the basis for many common attributions of randomness, I will hereafter refer to this collection of beliefs as "the prevailing intuitive conception of randomness" or to "our commonly-held intuitions of randomness". However, I don't want to commit to there even being a unique prevailing intuitive conception of randomness, nor do I want to commit to the claim that commonly-held intuitions about randomness should have the final say in determining the adequacy of our definitions of randomness. Still, for dialectical purposes, let us assume that "the prevailing intuitive conception of randomness" is not an empty expression, and that this conception serves as a constraint in the task of evaluating the various definitions of randomness.

10.2.2 "Capturing" the Prevailing Intuitive Conception of Randomness?

Next, we need to clarify the relation that holds between a formal definition of randomness and the prevailing intuitive conception of randomness, a relation I've been referring to as the "capturing" relation. From the discussions of the MLCT, one can gather that a minimum, a definition $\mathscr{D}$ of randomness captures the prevailing intuitive conception of randomness only if the extension of the definition $\mathscr{D}$ is equal to the extension of the prevailing intuitive conception. For our purposes, we can take this to be necessary *and sufficient* for the capturing relation to hold between a formal definition of randomness and the prevailing intuitive conception of randomness.[1]

But a bit more can be said here. Given that there are multiple non-equivalent definitions of randomness that are, in a sense, competing for the role of capturing the prevailing intuitive conception of randomness, it will be useful to precisely define the "capturing" relation. Taking $\mathscr{D}(X)$ to stand for "X is $\mathscr{D}$-random" and $\mathsf{IR}(X)$ to stand for "X is intuitively random" (i.e. X is random according to the prevailing intuitive conception), let us say that a given definition $\mathscr{D}$ of algorithmic randomness

(i) *licenses the attribution of intuitive randomness to a sequence* X if

$$\mathscr{D}(X) \Rightarrow \mathsf{IR}(X);$$

(ii) *licenses the attribution of intuitive non-randomness to a sequence* X if

$$\neg\mathscr{D}(X) \Rightarrow \neg\mathsf{IR}(X);$$ [2]

(iii) *captures the intuitive notion of randomness* if for every X, either $\mathscr{D}$ licenses the attribution of intuitive randomness to X or licenses the attribution of intuitive

[1]Perhaps one might require additional conditions that also guarantee intensional adequacy of the correct definition of randomness. However, I think it is reasonable to restrict to the notion of extensional adequacy at this stage of the game, given that the question as to the extensional adequacy of various definitions of algorithmic randomness has not received a satisfactory treatment.

[2]Here I am assuming that a sequence is intuitively non-random if and only if it is not intuitively random. Some might balk at this assumption, but given that the theses we consider here are true only if the collection of intuitively random sequences has a definite extension, this assumption is not inappropriate.

non-randomness to X, i.e., for every X,

$$\mathscr{D}(X) \Leftrightarrow \mathsf{IR}(X).$$

Hereafter, I will also refer to a definition that captures the prevailing intuitive notion of randomness as *correct*.

With this terminology in mind, let us consider the ways in which a definition of randomness can fail to capture the prevailing intuitive notion of randomness. In particular, there are two circumstances in which this failure can occur. First, $\mathscr{D}$ may falsely license the attribution of intuitive non-randomness to some sequence X:

(I) $\mathscr{D}(X)$ is false but $\mathsf{IR}(X)$ is true.

Second, $\mathscr{D}$ may falsely license the attribution of intuitive randomness to some sequence X:

(II) $\mathscr{D}(X)$ is true but $\mathsf{IR}(X)$ is false.

Borrowing standard terminology from statistical practice, let us say of any definition $\mathscr{D}$ for which (I) holds that it is *Type I defective* and of any definition $\mathscr{D}$ for which (II) holds that it is *Type II defective*.[3]

From the claim that a definition $\mathscr{D}$ of randomness captures the intuitive conception of randomness, it follows that all other definitions are either Type I defective or Type II defective, or both. This observation will be central to my discussion of the

[3]Note that these two types of defectiveness are analogous to the two main types of errors in hypothesis testing in statistics: A Type I error is committed when one rejects a given null hypothesis when it is in fact true, while a Type II error is committed when one fails to reject the null hypothesis when in fact the alternative hypothesis is true.

status of the MLCT beginning in the next section. Before we turn to this discussion, however, there is one further matter of clarification.

10.2.3 Establishing the Defectiveness of a Definition of Randomness

If the $\mathscr{D}$-advocate (i.e., the advocate of the claim that definition $\mathscr{D}$ of randomness captures the intuitive notion of randomness) is going to justify her claim, it appears that she must show that all definitions of randomness that are not extensionally equivalent to $\mathscr{D}$ are either Type I defective or Type II defective. But how does one even show that a *single* definition is Type I defective or Type II defective? Let us consider these two in reverse order.

10.2.3.1 Type II Defectiveness

The standard approach to showing that a given definition of randomness $\mathscr{D}$ is Type II defective is to identify some property $\mathcal{P}$ such that (i) $\mathcal{P}$ is satisfied by a $\mathscr{D}$-random sequence but (ii) no intuitively random sequence has $\mathcal{P}$. I call such a property a *disqualifying property.* We will see a number of such properties throughout the latter half of the chapter, but before we consider these examples, we need to ask: What makes a property a disqualifying property?

First, the disqualifying properties that we'll consider shortly all bear a close connection to what I call *hallmarks of randomness.* To understand what makes a property a disqualifying property, we first need to discuss these hallmarks of randomness.

Hallmarks of randomness are properties on the basis of which one attributes randomness to a sequence; in general, these are the properties associated with random phenomena, even in everyday, non-scientific settings. While I don't intend to provide a full catalogue of the various hallmarks of randomness, there are four that have been isolated as significant in the algorithmic randomness literature: statistical typicality, unpredictability, incompressibility, and independence. The first three hallmarks play a particularly important role in the general theory of algorithmic randomness, for one often finds definitions of randomness as falling into one of three paradigms: the "typicality paradigm", the "unpredictability paradigm", and the "incompressibility paradigm".[4,5]

Let us consider each of these four hallmarks in turn.[6]

- *Typicality*: According to "typicality paradigm" in the algorithmic randomness literature, an intuitively random sequence is typical if it satisfies all "statistical properties of randomness", to use Martin-Löf's phrase that we discussed at the end of the previous chapter. Some examples of these statistical properties of randomness, which should now be quite familiar to us, are the Law of Large

[4]For example, the opening paragraph of the chapter on Martin-Löf randomness in the recent Downey/Hirschfeldt volume [DH10] lists these paradigms as "the computational paradigm", "the measure-theoretic paradigm", and "the unpredictability paradigm". See ([DH10], p. 226) for more details.

[5]Unlike the other three hallmarks, there is no "independence paradigm" for definitions of algorithmic randomness, but this not detract from the importance of this hallmark.

[6]One might worry that the hallmarks as described here are rather loosely characterized. This is largely by design. For one thing, in the algorithmic randomness literature, not much of substance is said about these hallmarks, as one usually passes quickly from very broad descriptions of a hallmark to a formal definitions that is based, in some way, upon that hallmark. Of course, this doesn't mean that we shouldn't try to be more precise in characterizing these hallmarks, and as we will see in this chapter and the next, these hallmarks can be made more precise. Nonetheless, as crudely as these hallmarks might be formulated, they are still useful for my purposes.

Numbers and the Law of the Iterated Logarithm.

- *Unpredictability*: According to the "unpredictability paradigm", the bits of an intuitively random sequence should be unpredictable. Here the prediction need not be absolute; if an effective method correctly predicts the values of a sequence, say, three-fourths of the time, then the sequence is still counted as predictable.

- *Incompressibility*: According to the "incompressibility paradigm", an intuitively random sequence should be incompressible, in the sense that its initial segments cannot be compressed. The common refrain in the algorithmic randomness literature is that a finite string is incompressible if its shortest description is not much shorter than the string itself.[7]

- *Independence*: Lastly, the bits of an intuitively effective random sequence should be independent of one another. One might think that unpredictability guarantees a degree of independence: if the $(n+1)$st value of a sequence cannot be predicted from the previous n values, then we might say that the $(n+1)$st value is independent of the previous n values. But one might consider more

[7]This characterization relies on a very specific account of what counts as a description (and thus I explicitly avoided the use of the term "description" in explaining Kolmogorov complexity in the previous chapter). For instance, if one takes the admissible descriptions to be given by computer programs, then a string σ cannot be compressed if the shortest computer program that outputs σ is roughly as long as σ.

general sorts of independence; for instance, one should not be able to use the even-indexed bits of a random sequence to effectively generate or predict the odd-indexed bits of that sequence, and vice versa.

So what do these hallmarks of randomness have to do with disqualifying properties? If we hold that a disqualifying property is a property in virtue of which one attributes non-randomness to a sequence, then the connection should be clear: A property in virtue of which a sequence is a statistically atypical, predictable, or compressible is usually counted as a disqualifying property. For instance, a sequence that does not contain the subword 11 is highly atypical, and thus should not be counted as intuitively random. Similarly, if every 5th bit of a sequence is a 0, then there is an effective method of prediction that correctly predicts every 5th bit of the sequence; again, this sequence would not be counted as intuitively random.

It is not my intention to provide necessary or sufficient conditions for a property to count as a disqualifying property, for as I discuss in the next chapter, it's not clear to me that such a set of conditions can be identified. For now we'll just consider disqualifying properties on a case-by-case basis, provisionally accepting some properties as disqualifying properties.

To sum up, the standard way to show that a definition of randomness $\mathscr{D}$ is Type II defective is to establish that some $\mathscr{D}$-random sequence satisfies a disqualifying property.

10.2.3.2 Type I Defectiveness

How do we show that a definition of randomness is Type I defective? It is much less clear how one is to establish that a definition $\mathscr{D}$ of randomness is Type I defective. One suggestion is to take the following approach:

(i) identify some property $\mathcal{H}$ that can be instantiated by sequences and is a hallmark of randomness;

(ii) show that some sequence X that is not $\mathscr{D}$-random bears $\mathcal{H}$; and

(iii) establish that X bears no disqualifying properties.

But this approach is problematic. Suppose we would like to show that a definition $\mathscr{D}$ is Type I defective, as it fails to count as random a sequence X that is intuitively random. Thus, we have to carry out the above steps (ii) and (iii) for this sequence X. But if we are able to dispute with the $\mathscr{D}$-advocate whether X is intuitively random, then X must be, in some sense *accessible* to us, where, following a definition suggested by Borel, a number is accessible if it is possible to define it in such a way that "any two mathematicians will be certain that they are speaking about one and the same entity" ([Bag53], p. 407).[8] More generally, suppose we are not merely disagreeing with the $\mathscr{D}$-advocate about a single sequence, but rather some collection $\mathcal{S}$ of sequences. Again, this collection $\mathcal{S}$ must be accessible to us, where, extending Borel's definition, a collection of sequences is accessible if it is possible to define it in such a way that any two mathematicians will be certain that they are talking about

[8] See also [Bor52].

one and the same collection. In either of these two cases, we face a problem: Why shouldn't the property that defines X or the property that defines $\mathcal{S}$ be counted as a disqualifying property? This is a serious question that we must answer if an argument for the Type I defectiveness of $\mathscr{D}$ like the one given above is to succeed.

One alternative approach to arguing for the Type I defectiveness of a given definition $\mathscr{D}$ is to argue that some of the properties that are necessary for $\mathscr{D}$-randomness are not necessary for intuitive randomness. One example of this approach is provided by Schnorr, who argued that Martin-Löf's definition is Type I defective because it requires random sequences to pass statistical tests that have no "physical meaning".

I'm not so much interested, at this point of our discussion, in the particulars of Schnorr's argument, but rather Schnorr's argumentative strategy:[9] Instead of arguing about the individual sequences that are Schnorr random but not Martin-Löf random, Schnorr argues about the properties that are necessary and sufficient for randomness. That is, Schnorr argues for one collection of properties to be considered as the properties of randomness (those that can be given in terms of Schnorr tests), rejecting the collection identified by Martin-Löf (those properties that can be given in terms of Martin-Löf tests). While this appears to be a more promising approach than the other approach to demonstrating Type I defectiveness, the downside is that for this approach to succeed, one must successfully identify the properties of randomness, which is a rather difficult task, as I argue in the next chapter.

We will revisit the problems that come with establishing that a definition of randomness is Type I or Type II defective later in this chapter as well as in the next

[9]However, below in Section 10.4.1, I consider the particulars of Schnorr's argument.

chapter. My goal at this point of the discussion is merely to point out that there are such problems without indicating just how problematic they are. Later, I will argue that they are *deeply* problematic, so much so that it is doubtful whether one can successfully establish that one definition of algorithmic randomness is correct while all others are defective.

We now turn to the heart of the chapter, in which I lay out and evaluate the evidence given in support of the MLCT and consider the threat posed to the MLCT by certain alternative definitions of randomness.

10.3 On the Evidence for the MLCT

The standard argument for the MLCT is that a number of intensionally different definitions of algorithmic randomness prove to be extensionally equivalent to Martin-Löf randomness; hereafter I will refer to this argument as the *Appeal to Convergence Results*.[10] In what follows, I will highlight the convergence results to which the advocates of the MLCT appeal in support of their view. Next, I will argue that the Appeal to Convergence Results, taken in isolation, fails to establish Martin-Löf randomness as the correct definition of randomness. Lastly, in light of this failure,

[10]It should be noted that other arguments given in support of the MLCT of varying quality are discussed in Jean-Paul Delahaye's article "The Martin-Löf-Chaitin Thesis: The Identification by Recursion Theory of the Mathematical Notion of Random Sequence" [Del11]. In his article, Delahaye classifies a number of types of arguments that can be given in support of the claim that a given formal notion is coextensive with a specific informal, pre-theoretic notion and then compares the strength of two particular instances of each argument-type, the argument-instance in support of the CTT and one argument-instance in support of the MLCT. Most of the argument-types in Delahaye's classification do not give rise to particularly strong argument-instances in favor of the MLCT, as most establish that Martin-Löf random sequences have some property that is shared by a number of non-equivalent definitions of randomness, and so we will not consider these argument-types here.

I suggest several ways to supplement the appeal to convergence results to establish the truth of the MLCT, namely by means of a squeezing argument or by what Alan Turing calls a "direct appeal to intuition".

10.3.1 Three Convergent Definitions

One widely held view on the justification of the Church-Turing Thesis (henceforth, the CTT) is that the convergence of multiple formal definitions provides the strongest evidence in support of the CTT. As evidence of the widespread acceptance of this view, one merely needs to consult any of the standard textbooks on computability theory.[11] Similarly, some have claimed that multiple definitions of randomness are equivalent provides a similarly strong justification for the MLCT. Antony Eagle summarizes this position well, writing,

> Different intuitive starting points have generated the same set of random sequences. This has been taken to be evidence that ML-randomness [...] is really the intuitive notion of randomness, in much the same way as the coincidence of Turing machines, Post machines, and recursive functions was taken to be evidence for Church's Thesis, the claim that any one of these notions captures the intuitive notion of effective computability ([Eag10]).

For instance, in their textbook on Kolmogorov complexity, Paul Vitanyi and Ming Yi write,

> The fact that such different effective formalizations of infinite random sequences turn out to define the same mathematical object constitutes evidence that our intuitive notion of infinite sequences that are effectively random coincides with the precise notion of Martin-Löf random infinite sequences ([LV97], p. 222).

[11] See, for example, [Rog87], pp. 18-21, [Coo04], pp. 42-43, or [Soa87] p. 14.

More forcefully, A. Dasgupta writes,

> Perhaps the strongest evidence for the Martin-Löf-Chaitin thesis available so far is Schnorr's theorem, which establishes the equivalence between a naturally formulated typicality definition (Martin-Löf randomness) and a naturally formulated "incompressibility definition" (Kolmgorov-Chaitin randomness) ([Das11], p. 707).

Let us briefly consider the most well-known of these convergence results.[12] There are three definitions of randomness in particular that are the most widely studied, each of which is intended to formalize a different hallmark of randomness: typicality, unpredictability, and incompressibility:

Randomness as typicality: $X \in 2^\omega$ is random$_1$ if and only if for every Martin-Löf test $\{\mathcal{U}_i\}_{i\in\omega}$, $X \notin \bigcap \mathcal{U}_i$ (i.e., X is Martin-Löf random).

Randomness as unpredictability: X is random$_2$ if and only if no computably enumerable martingale succeeds on X.[13]

Randomness as incompressibility: X is random$_3$ if and only if the initial segment complexity of X is sufficiently high, i.e.,

$$(\exists c)(\forall n)K(X{\upharpoonright}n) \geq n - c.^{14}$$

[12]The reader interested in the full details of these different definitions of randomness, and proofs that they are equivalent, should consult Section 2.5.1.

[13]A martingale d is computably enumerable (c.e.) if the range values of d are uniformly left-c.e. real numbers, which means that the left cuts corresponding the to the values $d(\sigma)$ for every $\sigma \in 2^{<\omega}$ are uniformly c.e. sets of rational numbers. That is, there is some computable function $f : 2^{<\omega} \to \omega$ such that for each $\sigma \in 2^{<\omega}$, the left cut of d is equal to the c.e. set of rationals, $W_{f(\sigma)}$. For more details, see Section 2.5.1 .

[14]Here $K(\sigma)$ is the *prefix-free* Kolmogorov complexity of σ, defined in terms of a prefix-free universal Turing machine U, one for which $U(\sigma)\downarrow$ and $\sigma \prec \sigma'$ implies that $U(\sigma')\uparrow$. It is necessary to

While each of these definitions appears to be well-motivated, as each is given in terms of a different hallmark of randomness, on the basis of which we make attributions of randomness in both scientific and non-scientific settings, it is striking that each of these definitions picks out the same extension of random sequences. Thus we have, for every $X \in 2^\omega$,

$$X \text{ is random}_1 \Leftrightarrow X \text{ is random}_2 \Leftrightarrow X \text{ is random}_3.$$ [15]

Recasting each of these definitions of randomness in terms of the hallmarks of randomness they are purported to formalize, we have, for every $X \in 2^\omega$

$$X \text{ is typical} \Leftrightarrow X \text{ is incompressible} \Leftrightarrow X \text{ is unpredictable.}$$

Thus, three intensionally non-equivalent definitions are extensionally equivalent.

10.3.2 An Initial Worry

Now, these convergence results are interesting and even quite surprising, for they appear to indicate that the various hallmarks of randomness are more closely related than we might have otherwise suspected. But how much evidential support do the

use some restricted version of Kolmogorov complexity to extend Kolmogorov's definition of random finite string to the infinite case, since the collection of $X \in 2^\omega$ such that

$$(\exists c)(\forall n)C(X{\upharpoonright}n) \geq n - c$$

is empty, a result due to Martin-Löf [ML71]. The idea to overcome this obstacle by considering some restricted class of Turing machines is due independently to Chaitin, Levin, and Schnorr. See Section 2.5.1 for more details.

[15] The first equivalence was proved by Schnorr, who gave an effective version of Ville's proof of the correspondence between martingales and sets of measure zero, while variants of the second were proved independently by Schnorr [Sch73] and Levin [Lev73]. See also [Cha75].

above convergence results really provide for the MLCT? More generally, how much evidential support do convergence results provide for a given thesis? Related to this second question, we might further ask: Does this evidential support come in degrees, and if so, what determines the degree of support? Is this degree merely determined the number of intensionally different definitions that are extensionally equivalent, or does the *quality* of the convergence results factor into this degree of support in some way?

Answering these last two questions could potentially help in answering the first question about the amount of evidential support the convergence results provide for the MLCT. The semblance of an answer to these questions can be gleaned from the literature on algorithmic randomness. First, some of those who appeal to convergence results in support of the MLCT appear to hold that there are varying degrees of evidential support provided by convergence results. Let's distinguish between two different, though not incompatible, approaches to measuring the degree of evidential support of convergence results. According to the *quantitative approach* to convergence results, the more convergent definitions there are, the more evidence we have for the associated thesis. An example of this approach is provided by Jean-Paul Delahaye, who, in comparing the strength of the evidential support of the Appeal to Convergence Results in support of the CTT with the Appeal to Convergence Results in support of the MLCT, writes,

> Here we have a clear advantage in favour of the Church-Turing Thesis over the Martin-Löf-Chaitin Thesis for the first one is supported by several hundred equivalent definitions, meanwhile the second one has only several equivalent definitions [. . .] ([Del11], p. 130)

Next, according to the *qualitative approach* to convergence results, not all convergence results have the same evidential strength; some lend more evidence to the relevant thesis than do others. This appears to be the approach Dasgupta takes when he claims that the Levin-Schnorr theorem (the equivalence of randomness$_1$, given in terms of typicality, and randomness$_3$, given in terms of incompressibility) provides "perhaps the strongest evidence for the [MLCT]". For if all convergence results have the same evidential strength, then presumably the equivalence of the "typicality definition" and the "unpredictability definition" should have just as much evidential support as the Levin-Schnorr Theorem, the equivalence of the "typicality definition" and the "incompressibility definition".[16]

These two approaches to convergence results must be further developed and clarified before we can say with any confidence that the Appeal to Convergence Results really does provide good evidence for the MLCT, especially in the face of alternative definitions of randomness. However, there is good reason to doubt that the Appeal to Convergence Results provides much evidential support for the MLCT in the first place.

10.3.3 Kreisel's Concern

The problem of determining the evidential support provided by convergence results is eclipsed by a more serious concern. In particular, Georg Kreisel raised a

[16]It is not clear what would account for some convergent results having more evidential strength than others on this approach. Is the evidential strength of the Levin-Schnorr theorem due to its being unexpected or surprising? Does a convergence result that is unexpected or surprising offer more evidential support than one that is neither unexpected nor surprising? If not, what else could account for the strength of a single convergence result?

worry that these types of convergence results do not exclude the possibility of systematic error. We might have a collection of definitions that converge to the wrong extension. Concerning whether one can generalize the CTT to abstract structures in higher computability theory, Kreisel writes, "Equivalence results do not play a special role, simply because one good reason is better than 20 bad ones, which may be all equivalent because of systematic error" ([Kre71], p. 144). The fact that a number of formal definitions have the same extension doesn't guarantee that they also have the same extension as some notion they are intended to capture. Let's call this worry *Kreisel's Concern*.

Kreisel's Concern, then, is that we may have a number of definitions that converge in extension, but which converge to the *wrong* extension. How are we to rule out this possibility? Perhaps the quantitative approach to the Appeal to Convergence Results might address Kreisel's Concern. For instance, given that there are hundreds of definitions of the class of computable number-theoretic functions that are equivalent to one another, and no serious contenders that are not equivalent to these definitions, as long as we keep piling on more and more equivalent definitions, the chance of systematic error gets smaller and smaller.

But this is not a line of argument available to the defender of the MLCT, for just as there are a number of definitions that converge to the extension of Martin-Löf randomness, there are *also* multiple definitions of randomness that converge to the extension of Schnorr randomness. In fact, one can find definitions of algorithmic randomness equivalent to Schnorr randomness that are given in terms of the same three hallmarks of randomness discussed above:

Randomness as typicality: $X \in 2^\omega$ is random$_4$ if and only if for every Schnorr test $\{\mathcal{U}_i\}_{i\in\omega}$, $X \notin \bigcap \mathcal{U}_i$ (i.e., X is Schnorr random).

Randomness as unpredictability: $X \in 2^\omega$ is random$_5$ if and only if for every computable martingale d and every unbounded, non-decreasing computable function h, it is not the case that

$$d(X{\upharpoonright}n) \geq h(n)$$

for infinitely many n.[17]

Randomness as incompressibility: $X \in 2^\omega$ is random$_6$ if and only if for every computable measure machine M,

$$(\exists c)(\forall n)K_M(X{\upharpoonright}n) \geq n - c,$$

where a computable measure machine M is a prefix-free machine such that the Lebesgue measure of the collection of infinite extensions of the strings in the domain of M is a computable real number.

Thus we have for every $X \in 2^\omega$,

$$X \text{ is random}_4 \Leftrightarrow X \text{ is random}_5 \Leftrightarrow X \text{ is random}_6.^{18}$$

[17] A martingale d is computable if if the range values of d are uniformly computable real numbers, which means that the left cuts corresponding the to the values $d(\sigma)$ for every $\sigma \in 2^{<\omega}$ are uniformly *computable* sets of rational numbers.

[18] The first equivalence was also proved by Schnorr [Sch71]. The second was proven only recently by Rod Downey and Evan Griffiths in [DG04].

If the Appeal to Convergence Results provides evidential support to the MLCT, it also provides evidential support for ST. In fact, Schnorr explicitly presents the Appeal to Convergence Results argument in support of ST, writing, "[An] important argument for our thsis which proposes [Schnorr randomness] as the 'really true' concept of randomness is that some different approaches lead to an equivalent definition" ([Sch71], p. 257). So clearly convergence results alone cannot privilege Martin-Löf randomness over Schnorr randomness with respect to capturing the prevailing intuitive conception of randomness.[19]

In light of these equivalent definitions that converge to the extension of Schnorr randomness, the quantitative approach to convergence results isn't going to get us very far in deciding between the MLCT and ST.[20] Further, given that there is an analogue of the Levin-Schnorr theorem for Schnorr randomness,[21] the qualitative approach, such as the one Dasgupta takes, is also threatened.[22] Thus, in this case,

[19]There is one qualitative difference between Martin-Löf randomness and Schnorr randomness that some have held shows Martin-Löf randomness to be the superior definition: while there is a universal Martin-Löf test, there is no universal Schnorr test. While this existence of a universal test has certain technical advantages, it's hard to see why a definition $\mathscr{D}$ of randomness should be disqualified from capturing the prevailing intuitive conception of randomness because these is no universal test for $\mathscr{D}$-randomness.

[20]For even if we were to count the number of definitions equivalent to Martin-Löf randomness and the to Schnorr randomness, how useful would that be? An educated guess is that the numbers wouldn't differ by too much.

[21]Downey and Griffiths showed the equivalence of randomness$_4$ and randomness$_6$ in [DG04].

[22]In fact, this isn't the only result showing the equivalence of a definition given in terms of typicality with a definition in terms of incompressibility. As discussed in footnote 52 in the previous chapter, X is 2-random, i.e., Martin-Löf random relative to the halting set $\emptyset'$, if and only if

$$(\exists c)(\exists^{\infty} n)C(X\restriction n) \geq n - c.$$

Kreisel's concern is legitimate: if there can be only one correct definition of randomness, then a systematic error is occurring in at least one of these two cases.

10.3.4 Supplementing the Convergence Results

Clearly, then, the convergence results are not enough to help us adjudicate between MLCT and ST, and in particular, taken in isolation, they certainly don't show that MLCT is to be preferred to ST. However, one might hope to settle the issue by supplementing the convergence results with additional evidence. In what follows, I will discuss two ways one might carry this out: by offering a squeezing argument, or by making a "direct appeal to intuition".

10.3.4.1 Squeezing Arguments

First, one might supplement the convergence results by means of a *squeezing argument*.[23] The idea of a squeezing argument, as formulated by Kreisel in [Kre72], is the following: Suppose we have an intuitive notion $\mathscr{I}$, the extension of which we are trying to capture by some formal definition. If there is

(a) a formal definition $\mathscr{D}_1$ such that every object that satisfies $\mathscr{D}_1$ also satisfies $\mathscr{I}$, and

(b) another formal definition $\mathscr{D}_2$ such that every object that satisfies $\mathscr{I}$ also satisfies $\mathscr{D}_2$,

[23] For a particularly helpful discussion of squeezing arguments, see Peter Smith's recent "Squeezing Arguments" [Smi11].

it follows that

$$\mathsf{ext}(\mathscr{D}_1) \subseteq \mathsf{ext}(\mathscr{I}) \subseteq \mathsf{ext}(\mathscr{D}_2),$$

where $\mathsf{ext}(\cdot)$ denotes the extension of the given notion. Now if we can show that an object satisfies $\mathscr{D}_1$ if and only if it satisfies $\mathscr{D}_2$, we will thereby have shown that

$$\mathsf{ext}(\mathscr{D}_1) = \mathsf{ext}(\mathscr{I}) = \mathsf{ext}(\mathscr{D}_2).$$

That is, we will have "squeezed" the extension $\mathscr{I}$ between the extensions of the definitions $\mathscr{D}_1$ and $\mathscr{D}_2$, or to put the matter differently, $\mathscr{D}_1$ and $\mathscr{D}_2$ will converge "around" $\mathscr{I}$. If we can thus show that the extension of the intuitively random sequences is squeezed between the extensions of two coextensive definitions of randomness, on this approach, we can conclude that we've captured the intuitive notion.

Can we use a squeezing argument to establish the MLCT? First recall that $\mathsf{MLR} \subsetneq \mathsf{SR}$. This has a significant consequence for our question: Since MLR is sufficient for SR, if the advocate of MLCT can successfully establish that MLR is not Type I defective (that is, that MLR is necessary for intuitive randomness), that is, for every X,

$$\mathsf{IR}(X) \;\Rightarrow\; X \in \mathsf{MLR}, \tag{10.1}$$

this would imply that *no* $X \in \mathsf{SR} \setminus \mathsf{MLR}$ is intuitively random.[24] Thus, if the MLR-advocate is going to squeeze the extension of the intuitively random sequences between the extensions of two equivalent definitions of MLR, he will have to establish

[24]Note, however that if the advocate of MLCT establishes that MLR is not Type II defective (that is, that MLR is sufficient for intuitive randomness), this would not rule out the possibility that SR is also sufficient for intuitive randomness.

(10.1). That is, if a squeezing argument is going establish the MLCT, we need to determine whether or not the sequences in $\mathsf{SR} \setminus \mathsf{MLR}$ are intuitively random.

Assuming that the extension of intuitive randomness can be captured by some formal definition, either

(i) no $X \in \mathsf{SR} \setminus \mathsf{MLR}$ is intuitively random;

(ii) some $X \in \mathsf{SR} \setminus \mathsf{MLR}$ is intuitively random, but some $Y \in \mathsf{SR} \setminus \mathsf{MLR}$ is not intuitively random; or

(iii) every $X \in \mathsf{SR} \setminus \mathsf{MLR}$ is intuitively effectively random.

Clearly, (ii) and (iii) both imply that the MLCT is false, and thus the MLR-advocate will want to reject (ii) and (iii) while establishing (i) (which I'll henceforth abbreviate as $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$).

Now, to show that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$, one reasonable strategy is to identify one or more properties that hold of every $X \in \mathsf{SR} \setminus \mathsf{MLR}$ and argue that this property is incompatible with the intuitive conception of effective randomness. In Subsection 10.4.2, I'll discuss some attempts to carry this out. But the main point is this: Even if we supplement the convergence results with a squeezing argument in order to establish the MLCT, we still have an additional task to carry out, namely, showing that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$.

10.3.4.2 A "Direct Appeal to Intuition"

One might also supplement the convergence results by making what Turing refers to in his famous 1937 paper "On Computable Numbers, with an Application to the

Entscheidungsproblem" [Tur38] as a "direct appeal to intuition". That is, one can attempt to sharpen the prevailing intuitive conception of randomness, distilling certain features that are individually necessary and jointly sufficient for intuitive randomness and arguing that some formal definition of randomness has each of these features. This is the approach that Turing takes in arguing that every effectively calculable number-theoretic function is Turing computable in "On Computable Numbers". Specifically, what Turing carries out is an analysis of the defining features of human computors (i.e. those who compute number-theoretic functions with pencil and paper), features such as (i) being capable of observing only finitely many symbols at a given time, (ii) having one's behavior determined entirely by the symbols he is observing and the state that he is in, (iii) being in one of only a finite number of possible states, and so on. Having provided this analysis, Turing then argues that every computation that can be carried out by a computor, i.e. an individual with the features he has identified, can be carried out by a Turing machine.[25]

How might one justify the MLCT by a direct appeal to intuition? Following Turing's lead, one would have to identify a list of features of the prevailing intuitive conception of randomness that are individually necessary and jointly sufficient for intuitive randomness. But even in one were to identify such a list of features, there would still be work to be do to establish the MLCT. In particular, one would have to further argue that these features are individually necessary and jointly sufficient for Martin-Löf randomness. But here's the rub: it's hard to see how either of these

[25]Some have disputed the claim that what Turing carries out is, properly speaking, a *conceptual analysis* of the intuitive notion of numerical computability; see, for instance, Michael Rescorla's "Church's Thesis and the Conceptual Analysis of Computability" [Res07].

steps can be carried out without showing that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$. For if this "direct appeal to intuition" is going to allow us to establish the MLCT, the list of necessary and jointly sufficient conditions for intuitive randomness must be incompatible with the sequences in $\mathsf{SR} \setminus \mathsf{MLR}$. But to provide such a set of conditions for intuitive randomness that is incompatible with the sequences in $\mathsf{SR} \setminus \mathsf{MLR}$ is precisely to show that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$.

The upshot is this: on both of the suggested ways to supplement the convergence results, in order to establish the MLCT, it appears that one still must establish $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$. Can the MLR-advocate carry out this task? More generally, how much of a threat does Schnorr randomness pose to the MLCT?

10.4 The Challenge of Schnorr Randomness

We've already seen that the Appeal to Convergence Results does not establish the MLCT, given that Schnorr made the Appeal to Convergence Results in support of ST. As I argued above, the MLR-advocate thus needs to establish the Type II defectiveness of Schnorr randomness as a necessary step in establishing the MLCT. Yet from the point of view of the SR-advocate, Martin-Löf randomness is Type I defective. Does the SR-advocate have any good reason to hold this?

10.4.1 Is Martin-Löf Randomness Type I Defective?

We now turn to the particulars of Schnorr's argument for the Type I defectiveness of Martin-Löf randomness. As mentioned above, instead of arguing that there is some sequence that is (i) Schnorr random, (ii) not Martin-Löf random, but (iii) should be counted as intuitively random, Schnorr argued that Martin-Löf randomness is Type I Defective on the grounds that it is based on an inadequate formalization of the notion of a statistical test.

Recall that Schnorr acknowledged that Martin-Löf's definition "included all standard statistical properties of randomness". But it also included additional properties, properties that, in Schnorr's view, lack "physical meaning" and are of no significance for statistics. In particular, some Martin-Löf tests are such that there is no effective procedure that indicates when a given sequence has failed the test. Thus, there are sequences that fail to pass some Martin-Löf tests, but there is no way to effectively verify this failure. From Schnorr's point of view, these sequences should still be considered as effectively random.

Should the MLR-advocate take the MLCT to be threatened by Schnorr's argument? I think not, as there are a number of issues with Schnorr's argument on which the MLR-advocate could press Schnorr. First, he might question the extent to which definitions of randomness should be constrained by actual statistical practice. Moreover, it's far from clear that Schnorr's restricted definition is any more faithful to statistical practice than Martin-Löf's definition is. For instance, why think that effective detectable properties are the only ones that are of significance for statistics?[26]

[26] Antony Eagle raises a related point:

> In general, one might be worried (for good reason) about Schnorr's *operationalist* idea that a non-random sequence must be one that can be effectively determined to be non-random: we don't normally take the existence of good evidence for a property to be equivalent to the presence of the property ([Eag10]).

Second, the MLR-advocate might call into question the notion of physical meaning to which Schnorr appeals, asking, for instance, why it is that a statistical test must have this extra verifiability condition in order to count as physically meaningful. A test that lacks an effective method for verifying when a sequence passes the test may very well fail to be adequate for actual statistical practice, but it doesn't follow that the test lacks physical meaning.

Lastly, the MLR-advocate might question whether Schnorr's definition is any more physically meaningful than Martin-Löf's definition, in light of the fact that both definitions are already highly idealized and detached from everyday statistical experience (as both involve infinite sequences of events, both involve unbounded running time for enumerating tests, and so on).

In general, if our goal is to produce a definition that will be of use in actual statistical practice, then we must determine exactly which conditions must be fulfilled in order for a definition of randomness to fulfill this purpose. Further, if one of these conditions is that we must be able to effectively verify that non-random sequences do not pass various tests for randomness, then, by all means, we should add this as a constraint. But in the absence of a clear reason for such a requirement, it seems hard to justify, and Schnorr provides no such reason. At best, then, Schnorr's argument is in need of further development.

The MLR-advocate need not merely play defense, rebutting Schnorr's claim that MLR is Type I defective; the MLR-advocate also has grounds for holding that Schnorr randomness is Type II defective. Specifically, there are several putative disqualifying properties to which the MLR-advocate can appeal in support of the claim that Schnorr randomness is Type II defective.

10.4.2 Putative Counterexamples to ST

The standard approach to showing that Schnorr randomness is Type II defective consists of identifying Schnorr random sequences that are not intuitively random.[27] Three such purported counterexamples, along with the standard justification given to support the claim that these are legitimate counterexamples, are the following:

> **Putative Disqualifying Property 1:** There exists a Schnorr random sequence $R = R_0 \oplus R_1$ (where R_0 consists of the even indexed bits of R, while R_1 consists of the odd indexed bits) such that R_0 is not Schnorr random relative to R_1, and vice versa (in fact, one can use R_0 to compute R_1, and vice versa).[28] This constitutes a failure of independence, and since intuitively random sequences are such that their bit values are independent of one another, it follows that such a sequence cannot be intuitively random.

This is not an unreasonable argument for the MLR-advocate to make, as Martin-Löf randomness does not have this problem; if $R = R_0 \oplus R_1$ is Martin-Löf random, then R_0 is Martin-Löf random relative to R_1, and vice versa.[29] Thus, crucially, by

[27] By "standard approach", I mean that this is the approach one commonly finds in the algorithmic randomness literature.

[28] Note that every definition of algorithmic randomness can be relativized to an oracle $X \in 2^\omega$. For instance, X-Martin-Löf randomness is defined in terms of X-Martin-Löf tests, where each component of such a test is given by an X-c.e. collection of finite sequences.

[29] This result, known as van Lambalgen's Theorem, is a powerful tool in the study of algorithmic randomness. See, for instance, [vL90].

appealing to this counterexample, the MLR-advocate does not undermine the claim of the correctness of MLR.

Another putative disqualifying property that has been attributed to some Schnorr random sequences involves incompressibility.

> **Putative Disqualifying Property 2:** There are Schnorr random sequences that are highly compressible, where this compressibility is understood in terms of prefix-free Kolmogorov complexity.[30] But since intuitively random sequences are not compressible, it follows that these sequences are not intuitively random.

Again, the MLR-advocate does not face this problem, as the initial segments of a Martin-Löf random sequence have high Kolmogorov complexity and are thus incompressible. Thus, as with Putative Disqualifying Property 1, this counterexample also doesn't undermine the MLCT.

Third, we have a disqualifying property that is somewhat surprising:

> **Putative Disqualifying Property 3:** There is a Schnorr random sequence X and a computable place selection S such that there limiting relative frequency of 0s in X is not equal to the limiting relative frequency of 0s in the selected

[30]Even though no Schnorr random is compressible by means of a computable measure machine (as touched on briefly in Subsection 10.3.3) there are some Schnorr random sequences that highly compressible by any universal prefix-free machine (which is not a computable measure machine): There is a Schnorr random sequence $X \in 2^\omega$ such that for every computable, non-decreasing, unbounded function $g : \omega \to \omega$,

$$K(X{\upharpoonright}n|n) \leq h(n)$$

for almost every $n \in \omega$ (where $K(\cdot|\cdot)$ denotes prefix-free conditional complexity).

subsequence X_S. In other words, X is not random according to Church's definition.[31]

Given that Schnorr's definition is "more effective" than Martin-Löf's, requiring of the statistical properties of randomness that they be effectively detectable, it is surprising that although every Martin-Löf random sequence is random in Church's sense, some Schnorr random sequences are not.

But even if we were to accept that intuitively random sequences cannot have these disqualifying properties, there is still more work to be done if the MLCT is going to be established. For these two counterexamples, if successful, only show that SR is Type II defective, but they don't imply that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR})$. In particular, even if there are some sequences in $\mathsf{SR} \setminus \mathsf{MLR}$ that are not intuitively random, this doesn't rule out the possibility that intuitive randomness is captured by some definition the extension of which lies strictly *between* MLR and SR.[32] What the MLR-advocate needs is to define some property that is incompatible with intuitive randomness but which is satisfied by every sequence in $\mathsf{SR} \setminus \mathsf{MLR}$. Here's one suggestion:

[31] See [Wan96], Theorem 3.3.5.

[32] To date, researchers in algorithmic randomness have identified at least *five* non-equivalent definitions strictly between MLR and SR, and potentially a sixth, pending the solution of the most important and longest standing open problem in the field, whether MLR is equivalent to a definition of AR known as Kolmogorov-Loveland randomness. [MMN$^+$06] and [BHKM10] provide all the relevant details.

Putative Disqualifying Property 4: Each sequence is $\mathsf{SR} \setminus \mathsf{MLR}$ computes a function that dominates (or grows faster than) every computable function.[33] But the collection of sequences that compute such a fast-growing function is a set of measure zero, and thus such sequences are atypical. Since intuitively random sequences are should not have any atypical computational properties, it follows that intuitively random sequences cannot compute any function that dominates every computable function.

Given this latter collection of counterexamples, showing that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR})$ initially looks promising. If the MLR-advocate can successfully establish that Putative Disqualifying Property 4 is a legitimate disqualifying property, he will not only show that Schnorr randomness is Type II defective, but he will show that *every* $X \in \mathsf{SR} \notin \mathsf{MLR}$ fails to be intuitively random.

However, this line of argument is not available to the MLR-advocate, as the conclusion comes at a cost that the MLR-advocate cannot afford: some Martin-Löf random sequences *also* have this property of computing sequences that grow faster than every computable function.[34] It follows that if we are to appeal to Disqualifying Property 4 in order to show that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$, we thereby would show that MLR is Type II defective. Thus, we have a way of sharpening the notion of typicality

[33] That is, for each such sequence X there is a function f that is Turing computable in X (denoted $f \leq_T X$) such that for every computable function g,

$$(\exists m)(\forall n \geq m)[g(n) \leq f(n)].$$

[34] This is due to the fact that some Martin-Löf random sequences are *Turing complete*; that is, using such a sequence as an oracle, one can solve Turing's halting problem. I'll discuss this further in Subsection 10.5.1.

that seems to rule out every $X \in \mathsf{SR} \setminus \mathsf{MLR}$ from being intuitively random, but unfortunately for the MLR-advocate, some sequences in MLR are thus ruled out as well.

In sum, there's still unfinished business for the MLR-advocate, as there is currently no other putative disqualifying property for him to invoke in order to show that $(\mathsf{SR} \setminus \mathsf{MLR}) \Rightarrow \neg\mathsf{IR}$. To make matters worse, Schnorr randomness is not the only definition of randomness that poses a threat to Martin-Löf randomness.

10.5 The Challenge of Weak 2-randomness

Not only does the MLR-advocate have to face the challenge of the SR-advocate, who claims that Martin-Löf randomness is Type I defective, but he also has to face the challenge posed by the advocate of weak 2-randomness (denoted W2R), a strictly stronger definition of randomness.[35] Whereas a sequence is Martin-Löf random if it avoids all Martin-Löf tests, a sequence is weakly 2-random if it avoids all *generalized* Martin-Löf tests, where a generalized Martin-Löf test is a uniformly computable sequence of Σ^0_1 classes $\{\mathcal{U}_i\}_{i\in\omega}$ such that

$$\lim_{n\to\infty} \lambda(U_i) = 0.$$

Equivalently, a sequence is weakly 2-random if it is not contained in any Π^0_2 classes of measure 0. Thus it follows that every weakly 2-random sequence is Martin-Löf random (since the intersection of each Martin-Löf test is a Π^0_2 class of measure 0) ,

[35]This definition was originally formulated in [Kur81] and further studied in [GS82] and [Kau91].

but below we will see that the converse does not hold.

Perhaps the most striking characterization of weak 2-randomness is given by a recent algorithmic randomness, according to which $X \in 2^\omega$ is weak 2-random if and only if X is Martin-Löf random and forms a minimal pair with the halting set $\emptyset'$ in the Turing degrees.[36] This latter condition means that any set that is computable from X and computable from $\emptyset'$ must be a computable set; that is, there are no non-computable sets computable from both X and from $\emptyset'$.[37]

This result has important consequences: The collection of weakly 2-random sequences is the largest subcollection of the Martin-Löf random sequences that avoids certain properties that some have argued are incompatible with intuitive randomness. Consequently, the MLR-advocate has to address a number of putative disqualifying properties offered by the $\mathsf{W2R}$-advocate in support of the claim that Martin-Löf randomness is Type II defective.

10.5.1 Putative Counterexamples to the MLCT

Compared to the other putative disqualifying properties, this first putative disqualifying property has received the most attention in the philosophical literature on algorithmic randomness. However, it's not entirely clear *why* this property should disqualify a sequence instantiating it from being counted as intuitively random.

[36]One direction is shown in [DNWY06], while the other direction was proven by Hirschfeldt and Miller (unpublished). The full proof can be found in Section 2.5.3.1.

[37]In the language of the Turing degrees, if $X \in \mathbf{x}$, then $\mathbf{x} \vee \mathbf{0}' = \mathbf{0}$.

Putative Disqualifying Property 5: Some Martin-Löf random sequences are Δ^0_2, or equivalently, are decidable by a trial-and-error predicate.[38]

Why might such sequences be counterexamples to the MLCT? There are several ways that this has been argued. For instance, Panu Raatikainen writes of Chaitin's Ω, a well-known and much studied Martin-Löf random sequence that is Δ^0_2 (to be defined shortly),

> The important aspect that matters here is that a trial and error procedure is still completely deterministic; the machine described above proceeds in a perfectly determinate manner. This means in particular that Ω, although not recursively enumerable, can still be generated by a completely deterministic procedure. And this, in turn, should raise some doubts about the genuine randomness of Ω, and more generally, about the plausibility of a definition of randomness that counts such sequences as random ([Raa00], p. 221).

If we hold that no intuitively random sequence can be "generated by a completely deterministic procedure", and we accept that sequences such as Ω can be generated by a completely deterministic procedure, then Raatikainen's conclusion would follow. However, there is some cause for concern, given that the procedure that one can use to generate Ω is not a Turing-computable procedure, but is rather a *hypercomputational* procedure.[39] Should a sequence be disqualifed from being counted as intuitively

[38]A predicate $P \subseteq \omega$ is a trial-and-error predicate if and only if there is a total computable function $f : \omega^2 \to \omega$ such that (i) $P(x)$ holds if and only if $\lim_{y\to\infty} f(x, y) = 1$ and (ii) $\neg P(x)$ holds if and only if $\lim_{y\to\infty} f(x, y) = 0$, where $\lim_{y\to\infty} f(x, y) = n$ if there is some stage z after which $f(x, y)$ has stabilized; that is,

$$\lim_{y\to\infty} f(x, y) = n \Leftrightarrow (\exists z)(\forall y \geq z)f(x, y) = n.$$

[39]A hypercomputational procedure is one that computes functions or sets of natural numbers that are not Turing computable.

random if it can be generated by a hypercomputational procedure? Perhaps not. For on a sufficiently broad reading of "hypercomputational procedure", *every* sequence can be generated by a hypercomputational procedure.[40] For Raatikainen's argument to succeed, then, he would have to identify a proper subcollection of the collection of all hypercomputational procedures such that any sequence that can be generated by one of these procedures cannot be intuitively random.[41]

An alternative explanation as to why being Δ^0_2 is a disqualifying property, which doesn't face the same problems as Raatikainen's explanation, is given by Daniel Osherson and Scott Weinstein, who describe the "tension" between randomness and decidability by a trial-and-error in the following way:

> Consider a physical process that, if suitably idealized, generates an indefinite sequence of independent random bits. One such process might be radioactive decay of a lump of uranium whose mass is kept at just the level needed to ensure that the probability is one-half that no alpha particle is emitted in the nth microsecond of the experiment. Let us think of the bits as drawn from {0,1} and denote the resulting sequence by x with coordinates $x_0, x_1, \ldots$. Now wouldn't it be odd if there were a computer program P with the following property?
>
> **1.** *For any input i, P enters a nonterminating routine that writes a nonempty, finite sequence $b_1, \ldots, b_m$ with $b_m = x_i$ (m depends on i).*
>
> The program will not, in general, allow prediction of x_i inasmuch as there is no requirement that the ultimate bit b_m written by $P(i)$ be marked as final. Nonetheless, shouldn't randomness exclude any computational

[40]For instance, any Turing machine equipped with a sufficiently strong oracle is a hypercomputational procedure, and since every sequence is computable by a Turing machine equipped with some sufficiently strong oracle, my claim follows. Or, for a given sequence X, one can cook up a more exotic hypercomputational procedure that computes X.

[41]This is just another variant of the admissibility objection as discussed in Chapter 8.

process from having the kind of intimate knowledge of x_i described in 1? ([OW08], p. 56)

There certainly appears to be some tension here, for property 1 is a fairly strong computational property: although a sequence satisfying property 1 is not decidable, it *is* decidable in the limit. But Osherson and Weinstein fail to say why it is that randomness *should* exclude any computational process from deciding the sequence in the limit.[42] Nonetheless, the main improvement of this argument over Raatikainen's is that Osherson and Weinstein offer an alternative definition of randomness that does not have this property, namely weak 2-randomness (which they refer to as *strong randomness*).[43] Thus, the suggestion is that in the presence of a definition of randomness that is stronger than Martin-Löf randomness and avoids this property, this latter definition should be preferred as capturing the prevailing intuitive conception of randomness.

While the W2R-advocate can appeal to Martin-Löf random sequences that are Turing below $\emptyset'$ (i.e., the Δ^0_2 sequences) in support of the claim that Martin-Löf randomess is Type II defective, she can also appeal to the Martin-Löf random sequences that are Turing *above* $\emptyset'$:

[42] Here are two suggestions. First, property 1 can be seen as a kind of predictability; consequently, sequences with property 1 are not unpredictable, and thus are not intuitively random. Second, the collection of sequences satisfying property 1 is a set of measure zero (as only countably many sequences have this property), and so sequences with property 1 are atypical, and hence not intuitively random.

[43] No weakly 2-random sequence is Δ^0_2, since every Δ^0_2 sequence is a Π^0_2 singleton (and thus is contained in a Π^0_2 subset of 2^ω of measure 0). Alternatively, this also follows from the fact that being Δ^0_2 is equivalent to being Turing computable from $\emptyset'$, and thus no non-computable Δ^0_2 sequence can form a minimal pair with $\emptyset'$ (for if A is non-computable and Δ^0_2, then if $A \in \mathbf{a}$, $\mathbf{a} \vee \mathbf{0'} = \mathbf{a} \neq \mathbf{0}$).

Disqualifying Property 6: Every Turing degree above (and including) the Turing degree of the halting problem $\emptyset'$ contains a Martin-Löf random sequence.

This result, known as the Kučera-Gács theorem,[44] has a very strong consequence: For every sequence $X \in 2^\omega$, there is some Martin-Löf random sequence such that $X \leq_T A$. That is, every set, no matter how complicated, is Turing reducible to some Martin-Löf random sequence.[45] Thus, for instance, we can find a Martin-Löf random sequence from which the first-order theory of true arithmetic (i.e., every sentence in the language of first-order arithmetic that is true in $(\mathbb{N}, +, \times, 0, 1)$) can be effectively determined.

Why should this count as a disqualifying property? Again, sequences with this property are atypical, as all are contained in $\mathcal{S} = \{X : X \geq_T \emptyset'\}$, a set of measure zero. Further, $\mathcal{S}$ contains no weakly 2-random sequences (since no sequence in $\mathcal{S}$ forms a minimal pair with $\emptyset'$[46]), so it's not unreasonable for the W2R-advocate to suggest this property as a putative disqualifying property.

One last putative disqualifying property that the W2R-advocate can offer in support of the claim that Martin-Löf randomness is Type II defective involves Chaitin's Ω, already referenced in the discussion of Disqualifying Property 5. If we let U be a

[44]This result was proven independently by Kučera in [Kuč85] and by Gács in [Gác86].

[45]Given $X \in 2^\omega$, the Turing degree of $X \oplus \emptyset'$ is above the Turing degree of $\emptyset'$, and thus $X \oplus \emptyset'$ isTuring equivalent to some Martin-Löf random sequence.

[46]For $X \in \mathcal{S}$, if $\mathbf{x} = deg_T(X)$, then $\mathbf{x} \vee \mathbf{0'} = \mathbf{0'}$.

universal, prefix-free Turing machine, then we can define Ω as:

$$\Omega := \lambda(\mathrm{dom}(U)) = \sum_{\sigma \in \mathrm{dom}(U)} 2^{-|\sigma|}.^{47}$$

There are a number of very strong properties satisfied by Ω but not satisfied by any weakly 2-random sequence. Thus, from the point of view of the W2R-advocate, we have yet another disqualifying property:

> **Putative Disqualifying Property 7:** Martin-Löf randomness counts Chaitin's Ω among the random sequences.

Why think that the Type II defectiveness of Martin-Löf randomness follows from the fact that Ω is Martin-Löf random? Here are several properties of Ω that might lead one to think that it is not intuitively non-random:

1. For every $n \in \Omega$, from the first n bits of Ω, $\Omega \upharpoonright n$, we can determine whether $U(\sigma)\downarrow$ for all $\sigma \in 2^{<\omega}$ such that $|\sigma| \leq n$. From this it follows that Ω is Turing complete.[48]

2. Ω is a left-c.e. real, i.e., it is the limit of a computable, non-decreasing sequence of rational numbers.[49]

[47]Since there are infinitely many universal prefix-free Turing machines, we should technically write Ω as Ω_U and refer to Ω-numbers, not one single number Ω.

[48]This is due to Chaitin, who first defined Ω in [Cha75].

[49]This simply follows from the definition of Ω, given above.

3. Ω is a Solovay complete, which means that for any left-c.e. real α, there is some constant c and some computable function f such that for any rational number $q < \alpha$,

$$c(\Omega - q) > \alpha - f(q).$$

Thus for any sequence of rationals numbers effectively converging to Ω can be effectively transformed to a sequence of rationals effectively converging to α at roughly the same rate.[50]

These latter two properties taken together yield a particular noteworthy consequence, for being a left-c.e. Solovay complete real is a *sufficient condition* for being Martin-Löf random (and thus by (2) and (3) above, it is necessary and sufficient for being equal to Ω_U for some universal prefix-free Turing machine U).[51] Clearly there is something strange going on here. For here we have a property that prima facie has nothing to do with the prevailing intuitive conception of randomness (and moreover, appears to be non-random according to the prevailing intuitive concept of randomness) and yet the satisfaction of this property is a sufficient condition for being counted as random.

A rough explanation of what is going on here is this: one can exploit the computational power of Ω to diagonalize into the collection of Martin-Löf random sequences. For instance, consider the proof that Ω is incompressible. The general idea of the

[50]This was first shown by Solovay in his unpublished notes [Sol75].

[51]In [CHKW01], Calude, Hertling, Khoussainov, and Wang show that if a sequence is a Solovay complete left-c.e. real then it is an Ω-number. Later, Kučera and Slaman in [KS01] show that any left-c.e. Martin-Löf random real is already Solovay complete, and is thus an Ω-number.

proof is as follows:[52]

(a) We consider an approximation of Ω given in terms of some increasing sequence of rationals that converge to it.

(b) For a fixed stage s, we consider in parallel the Kolmogorov complexity of initial segments of the stage s approximation of Ω.

(c) If we ever see a potential witness that one such initial segment can be compressed, we define a new computation in terms of some fixed prefix-free machine M.

(d) Because the prefix-free machine in terms of which Ω is defined is universal, this new computation must be simulated by the universal machine, thereby forcing the stage s approximation of Ω to change.

(e) By a careful choice of both M (by means of the recursion theorem) and the string involved in the new computation, the change in the stage s approximation of Ω occurs at a place that invalidates the potential witness of compressibility.

(f) Every witness to a compressible initial segment of some approximation of Ω is eventually thwarted, and thus Ω is incompressible.

This is a fascinating phenomenon, and one that surely raises a number of questions about the relationship between Martin-Löf randomness and the prevailing intuitive conception of randomness. At a minimum, there is certainly tension here, which is

[52]Here I am following the proof as given in [DH10], p. 228.

recognized by many of those working in the field of algorithmic randomness. For instance, Downey and Hirschfeldt write, “It is important to note that Ω is a somewhat misleading example of [Martin-Löf randomness], as it is rather computationally powerful” ([DH10], p. 228). Shortly thereafter they add, “Thus one should keep in mind that, while Ω is certainly the most celebrated example of [Martin-Löf randomness], it is not ‘typically [Martin-Löf random]’ ([DH10], p. 229).

10.5.2 The Response of the MLR-Advocate?

Before concluding this section, let's consider whether there anything that the MLR-advocate can say in response to the W2R-advocate. Here are two suggestions: First, the MLR-advocate can raise Putative Disqualifying Property 1 again, as there are weakly 2-random sequences $R = R_0 \oplus R_1$ such that R_0 is not weakly 2-random relative to R_1, and vice versa.[53]

But a more interesting suggestion is this: the MLR-advocate can concede that the putative disqualifying properties offered by W2R-advocate are legitimate disqualifying properties, but instead of further conceding that weak 2-randomness is the correct definition of randomness, he can opt for an even stronger definition, Martin-Löf randomness relative to the halting problem $\emptyset'$, also referred to as 2-randomness (which I'll write as 2MLR). This might be an attractive possibility for a number of reasons: first, the convergence results discussed above still hold, in relativized form, for 2-randomness.[54] Further, from the point of view of the 2MLR-advocate,

[53] This was shown by Kautz in [Kau91], and independently by Barmpalias, Downey, and Ng in [BDN11].

[54] As mentioned in footnote 29, there is at least another convergence result that might support

weak 2-randomness is now Type II defective, for every 2-random sequence is weakly 2-random, but the converse doesn't hold.

However, this is a dangerous game, for one can define a version of Chaitin's Ω relative to $\emptyset'$, $\Omega^{\emptyset'}$, a sequence that can shown to be 2-random by almost exactly the same proof that shows that Ω is Martin-Löf random that I outlined above. Moreover, although $\Omega^{\emptyset'}$ cannot compute $\emptyset''$, when the information in $\emptyset'$ is also available, one can compute $\emptyset''$.[55] Lastly, $\Omega^{\emptyset'}$ is left-c.e. relative to $\emptyset'$ and Solovay complete relative to $\emptyset'$, and, as before, these the satisfaction of these two properties is sufficient for 2-randomness.[56]

Surely, the advocate of weak *3*-randomness[57] will offer these as putative disqualifying properties, but the advocate of 3-randomness[58] can trump these concerns, and so on. This phenomenon should give us pause as we ponder the possibility of a single definition that captures the prevailing intuitive conception of randomness.

---

the cause of the 2MLR-advocate, the equivalence of 2-randomness with a notion of incompressibility given in terms of plain Kolmogorov complexity.

[55]That is $\Omega^{\emptyset'} \oplus \emptyset'$ is Turing equivalent to $\emptyset''$.

[56]In fact, for any $A \in 2^\omega$, one can define, Ω^A, Chaitin's Ω relative to A, which is Martin-Löf random relative to A, left-c.e. relative to A, and Solovay complete relative to A (where these two conditions are necessary and jointly sufficient for a sequence being an Ω-number relative to A. These results were first shown in [DHMN05].

[57]$X \in 2^\omega$ is weakly 3-random if and only if X is not contained in any Π^0_3 subset of 2^ω of measure zero.

[58]X is 3-random if and only if X is not Martin-Löf random relative to $\emptyset''$.

10.6 The Next Step

My intention in this chapter was not to settle any of the specific issues that I raised here (such as determining which putative disqualifying properties should be considered as legitimate disqualifying properties, or how to supplement the convergence results so as to establish the MLCT), but rather to lay out the key challenges that faces the MLR-advocate, as well as the advocates of alternative definitions such as Schnorr randomness, weak 2-randomness, and 2-randomness. In so doing, I have provided much of the raw material necessary to understand the Justificatory Challenge facing the $\mathscr{D}$-advocate, which I discuss in the next chapter.

CHAPTER 11

THE JUSTIFICATORY CHALLENGE

11.1 Introduction

Over the last two chapters, we've seen that carrying out the task of identifying the "intuitive requirements" that an acceptable definition of randomness should satisfy is far from straightforward, particularly in the absence of a clear indication as to what role such a definition is intended to play. Even if we identify this role as the role of providing a conceptual analysis of the notion of random sequence, this does not really make the task any easier; the advocate of the claim that a given definition of randomness $\mathscr{D}$ successfully fills the conceptual-analytic role still must face the burden of picking out the properties that are necessary and sufficient for a sequence to be intuitively random.

Why do I characterize this task as a burden? Answering this question is the central goal of the present chapter. The main claim for which I argue is that the $\mathscr{D}$-advocate must overcome what I call the *Justificatory Challenge* in order to establish that $\mathscr{D}$ captures the prevailing intuitive conception of randomness. This challenge for the $\mathscr{D}$-advocate is as follows:

Justificatory Challenge: Provide a sharpening of the prevailing intuitive conception of randomness that is precise enough to block the claims of extensional adequacy made concerning alternative definitions of randomness without undermining the claim of the extensional adequacy of $\mathscr{D}$.

Why think that the $\mathscr{D}$-advocate must face this particular challenge? This follows from three pieces of data, which indicate (1) that the $\mathscr{D}$-advocate must establish the Type II defectiveness of weaker definitions of randomness while deflecting the claim, made by the advocates of stronger definitions of randomness, that $\mathscr{D}$ is Type II defective, (2) that she must do so by making recourse to the so-called prevailing intuitive conception of randomness, and (3) that this prevailing intuitive conception is not precise enough to underwrite the claims that alternative definitions of randomness are Type II defective and $\mathscr{D}$-randomness is not.

After arguing that the $\mathscr{D}$-advocate must meet the Justificatory Challenge, I will argue that there are no compelling reasons to hold that this challenge can be met. This alone is not particularly strong evidence that the $\mathscr{D}$-advocate cannot establish her claim of the extensional adequacy of $\mathscr{D}$. That is, we do not yet have sufficient grounds for holding the No-Thesis Thesis, the claim that no definition of randomness that has a definite, well-defined extension can capture the prevailing intuitive conception of randomness. In the next chapter, I supplement the argument given here with the goal of providing these grounds.

There is one serious objection that one can raise to my claim that the $\mathscr{D}$-advocate must face the Justificatory Challenge: If the $\mathscr{D}$-advocate must face this burden, why shouldn't we also require the advocate of the CTT to meet a similar Justificatory Challenge? That is, one might argue that I am unfairly demanding the $\mathscr{D}$-advocate to meet a challenge that is not imposed on the advocates of other claims of extensional adequacy. This is an important matter to clarify, and so I will devote the latter part of the chapter to isolating the salient differences between the task of justifying the CTT and the task of justifying the MLCT. As I argue, these differences are significant, so much so that it is not problematic to hold both the view that the MLR-advocate must the Justificatory Challenge and the view that the advocate of the CTT does not.

In the next section, Section 11.2, I explain why the $\mathscr{D}$-advocate must face the Justificatory Challenge, laying out the three pieces of data discussed above. In Section 11.3, I address two questions raised by my discussion of the Justificatory Challenge: I outline what it means to offer a sharpening of the prevailing intuitive conception of randomness, and then I explain just how burdensome the Justificatory Challenge should be for the $\mathscr{D}$-advocate. Lastly, in Section 11.4, I respond to the objection discussed in the previous paragraph, namely that I unfairly pose the Justificatory Challenge to the $\mathscr{D}$-advocate but not to, say, the adherent of the CTT.

11.2 The Basis of the Justificatory Challenge

In what follows, let $\mathscr{D}$ be some currently available of definition of algorithmic randomness, potentially the subject of the claim asserting that $\mathscr{D}$-randomness fills the conceptual-analytic. The first step in my argument is to explain why it is the that $\mathscr{D}$-advocate must face the Justificatory Challenge. Towards this end, I appeal to three pieces of data.

11.2.1 Datum 1: Weaker and Stronger Definitions of Randomness

The first datum to which I appeal in support of the claim that the $\mathscr{D}$-advocate must address the Justificatory Challenge is the fact that every definition of randomness can be "sandwiched" between two definitions of randomness that are not extensionally equivalent to $\mathscr{D}$:

> **Datum 1:** For every currently available definition $\mathscr{D}$ of algorithmic randomness, there are definitions $\mathscr{D}_1$ and $\mathscr{D}_2$ of algorithmic randomness such that
>
> $$\mathsf{ext}(\mathscr{D}_1) \subsetneq \mathsf{ext}(\mathscr{D}) \subsetneq \mathsf{ext}(\mathscr{D}_2), \tag{†}$$
>
> where $\mathsf{ext}(\mathscr{D})$ is the extension of the definition $\mathscr{D}$.[1]

This property (†) follows from a more general phenomenon: for every currently available definition $\mathscr{D}$ of algorithmic randomness, there is some definition $\mathscr{D}_1$ of algorithmic randomness and a parameter-free[2] formula Θ of second-order arithmetic,

[1] As formulated, this statement is not quite true: there is at least one exception to this property (†), namely Kurtz randomness (or at least two, if we count the collection of normal sequences as a definition of randomness). However, to the best of my knowledge, no one claims that weak randomness or normality captures the intuitive conception of randomness, for reasons that will be discussed in Sections ? and ?. It would thus be more accurate for Datum 1 to begin with "For every currently available definition of algorithmic randomness that is a serious candidate for capturing the prevailing intuitive conception of randomness. . . "

[2] It's important that we require Θ to be parameter-free, for as soon as we allow parameters, then for each parameter A, the formula $X \neq A$ defines a set of measure one that excludes only the sequence A. As we learned from the admissibility objection, discussed in Chapter 8, a definition of typicality that includes sets definable in terms of such formulas is entirely too restrictive, as every sequence comes out as atypical.

such that

(i) $\mathsf{ext}(\mathscr{D}_1) \subsetneq \mathsf{ext}(\mathscr{D})$,

(ii) $\Theta(2^\omega) := \{X \in 2^\omega : \Theta(X)\}$ is a null set,

(iii) $\Theta(X)$ holds for some $X \in \mathsf{ext}(\mathscr{D})$, and

(iv) $\neg\Theta(X)$ holds for *every* $X \in \mathsf{ext}(\mathscr{D}_1)$.

Such a formula Θ defines what I referred to in the previous chapter as a putative disqualifying property, for from the point of view of the $\mathscr{D}_1$-advocate, $\mathscr{D}$ is disqualified from capturing the intuitive conception of randomness by dint of there being some $\mathscr{D}$-random sequence that satisfies Θ. In short, the $\mathscr{D}_1$-advocate holds that $\mathscr{D}$ is Type II defective.

11.2.1.1 Instantiations of the Property (†)

We have already encountered one instance of this property (†) in the previous chapter, as there we saw that the collection of Martin-Löf random sequences lies strictly between the collection of weak 2-random sequences and the collection of Schnorr random sequences:

$$\mathsf{W2R} \subsetneq \mathsf{MLR} \subsetneq \mathsf{SR}$$

Schnorr randomness also satisfies the property (†), sandwiched strictly between two definitions, computable randomness[3] (denoted CR) and Kurtz randomness[4] (denoted KR), definitions we have yet to discuss but which will come up later in the chapter:

$$\mathsf{CR} \subsetneq \mathsf{SR} \subsetneq \mathsf{KR}$$

Further, the strongest definition of randomness we've encountered thus far in our discussion, 2-randomness,[5] is strictly stronger than weak 2-randomness, but in fact, there are infinitely many definitions that are stronger:

$$\ldots \subsetneq \mathsf{3MLR} \subsetneq \mathsf{W3R} \subsetneq \mathsf{2MLR} \subsetneq \mathsf{W2R}$$

That is, in general, we have

$$\mathsf{W(n+1)R} \subsetneq \mathsf{nMLR} \subsetneq \mathsf{WnR}$$

for every $n \in \omega$.[6]

[3]As formulated by Schnorr, a sequence $X \in 2^\omega$ is computably random if and only if no computable martingale succeeds on X. This is strictly stronger than Schnorr randomness, which counts as non-random those sequences on which a computable martingale succeeds with computable verification, as discussed in the previous chapter.

[4]A sequence X is Kurtz random if and only if $X \notin \mathcal{P}$ for every Π^0_1 class $\mathcal{P}$ of measure 0. Kurtz randomness, sometimes referred to as weak randomness, is named after Kurtz, who formulated the definition in his dissertation [Kur81].

[5]Recall that 2-randomness is Martin-Löf randomness relativized to the halting problem, $\emptyset' = \{x : \phi_x(x)\downarrow\}$, so that a sequence is 2-random if and only if it passes every Martin-Löf test that comes equipped with $\emptyset'$ as an oracle.

[6](†) also holds for many other definitions of randomness, most of which are not the subject of any claims of extensional adequacy: resource-bounded notions of randomness, hyperarithmetical

Given that $\mathscr{D}$ satisfies the property (†), in order for the $\mathscr{D}$-advocate to justify the claim that $\mathscr{D}$ captures the prevailing intuitive conception of randomness, she must establish that $\mathscr{D}_2$ is Type II defective while blocking the claims made by the $\mathscr{D}_1$-advocate that $\mathscr{D}$ is Type II defective. But in fact, she has to show even more, namely that every sequence in $\mathsf{ext}(\mathscr{D}) \setminus \mathsf{ext}(\mathscr{D}_1)$ is intuitively random and that no sequence in $\mathsf{ext}(\mathscr{D}_2) \setminus \mathsf{ext}(\mathscr{D})$ is intuitively random. Let us consider how she might carry this out.

11.2.2 Datum 2: Grounds for Adjudication

Unlike Datum 1, Datum 2 is not a straightforward claim about the various mathematical definitions of randomness, but rather it is a claim about the grounds according to which one can adjudicate between the various definitions of randomness.

> **Datum 2:** To establish the correctness of $\mathscr{D}$, the $\mathscr{D}$-advocate must provide grounds for adjudicating between the various definitions of randomness, but she can only do so by making recourse to the prevailing intuitive conception of randomness.

I take these grounds of adjudication to be principled reasons to hold that one definition of randomness is adequate while all of the others are not. Clearly, then, the $\mathscr{D}$-advocate needs to identify the grounds of adjudication on the basis of which she can conclude that (i) every $\mathscr{D}$-random sequence is intuitively random (so that $\mathscr{D}$ is

notions of randomness, and set-theoretic definitions of randomness. It is beyond the scope of this project to consider any of these definitions in detail, but I will discuss them briefly in the next chapter, as such definitions lend themselves to the approach that I develop there.

not Type II defective) and (ii) no non-$\mathscr{D}$-random sequence is intuitively random. Let us thus consider the ways that one might adjudicate between the various definitions of randomness (some of which we already discussed in the previous chapter).

11.2.2.1 Adjudication via Paradigm Instances

First, one might adjudicate between the various definitions of randomness by appealing to paradigm instances of random and non-random sequences. But there is are several problems with this approach. The first problem is that the paradigm instances of non-random sequences are already counted as non-random by each of the definitions of randomness that are taken to be candidates for capturing the prevailing intuitive conception of randomness.[7] Thus, appealing to paradigm instances of non-random sequences is of little help to us in adjudicating between the various definitions of randomness.

The second problem is that it is far from clear what should be considered a paradigm instances of a random sequence. In fact, one might reasonably question whether there are *any* paradigm instances of random sequences. For suppose that we identify some sequence X as a paradigm instance of a random sequence. In order to adjudicate between various definitions by checking which definitions count X as random and which ones do not, we need some way to make reference to X, but in such a way that guarantees that we are referring to X and not some other sequence. That is, there is some linguistic expression that refers to X such that any competent

[7] For example, each of these definitions counts among the non-random sequences all computable sequences, all biased sequences (that fail the Law of Large Numbers), and all sequences that fail the Law of the Iterated Logarithm.

user of that expression can recognize that the expression refers to X and not to some other sequence; that is, such a sequence satisfies Borel's definition of an accessible sequence, which we briefly discussed in Section 10.2.3.2. Herein lies the worry: If X is accessible, why should it be counted as a paradigm instance of randomness? Moreover, why *shouldn't* X be counted as a paradigm instance of non-randomness?

I don't intend to offer answers to these two questions, and for my purposes, answers to these questions are not required, since the fact is that no one attempts to adjudicate between the various definitions by appealing to paradigm instances of randomness. Such instances just aren't available.

11.2.2.2 Adjudication via Disqualifying Properties

Clearly, we cannot adjudicate between the various definitions of randomness by appealing to paradigm instances of random and non-random sequences. An alternative approach is to adjudicate between the various definitions of randomness by appealing to certain putative disqualifying properties. This is an approach that we have already discussed at length in the previous chapter. For instance, we've seen that the MLR-advocate invokes certain putative disqualifying properties in support of the claim that Schnorr randomness is Type II defective, while the $\mathsf{W2R}$-advocate invokes other putative disqualifying properties in support of the claim that Martin-Löf randomness is Type II defective.

Let's consider the general form of the arguments for Type II defectiveness that are given in terms of disqualifying properties. In what follows, $\mathscr{D}^*$ is some definition of randomness that is claimed to be Type II defective.

(P1) There is some $\mathscr{D}^*$-random sequence X that satisfies property $\mathcal{P}$.

(P2) No sequence satisfying $\mathcal{P}$ is intuitively random.

(P3) If there is a $\mathscr{D}^*$-random sequence that is not intuitively random, then $\mathscr{D}^*$ fails to capture the prevailing intuitive conception of randomness.

(C) Therefore, $\mathscr{D}^*$ fails to capture the prevailing intuitive conception of randomness.

For each instance of this argument schema, the property $\mathcal{P}$ in (P1) and (P2) is filled in by a putative disqualifying property. While (P1) is straightforward and (P3) is relatively uncontroversial,[8] instances of (P2) are the subject of some dispute. For how does one show that a putative disqualifying property is a legitimate one? That is, how can we establish of a given property $\mathcal{P}$ that it is incompatible with the prevailing intuitive conception of randomness, in the sense that no intuitively random sequence can satisfy it? An appeal to the prevailing intuitive conception of randomness is clearly necessary in this case. As for the form such an appeal is to take, let us bracket this until we discuss the third datum below.

11.2.2.3 Adjudication via Properties of Randomness

In addition to adjudicating between the various definitions of randomness by appealing properties that intuitively random sequences should not satisfy, one might also try to adjudicate between the various definitions of randomness by appealing to properties that random sequences *should* satisfy, the properties that Martin-Löf

[8]Rather, (P3) should be relatively uncontroversial among those who seek a definition of randomness that captures the prevailing intuitive conception of randomness.

referred to as "properties of randomness".

Of course, this approach will only succeed if we can determine which properties are the properties of randomness. From our discussion of the exemplary ideal in Chapter 9, the difficulty of identifying these properties of randomness should be clear. Ville considered this to be an unsolvable problem, while Martin-Löf made a reasonable attempt to associate the properties of randomness with those properties that are testable by means of c.e. sequential tests. However, Schnorr argued that not every c.e. sequential test corresponds to a property of randomness; only those sequential tests that test for effectively detectable properties (so that we can effectively determine whether a sequence has passed the test) correspond to properties of randomness.

In general, for any definition of algorithmic randomness $\mathscr{D}$, there is a countable collection of properties $\{\Phi_i^{\mathscr{D}}\}_{i\in\omega}$ such that

$$\mathsf{ext}(\mathscr{D}) = \bigcap_{i\in\omega}\{X : \Phi_i^{\mathscr{D}}(X)\}.$$

The task of adjudicating between the various definitions of randomness via properties of randomness thus boils down to the task of determining which collection of formulae $\{\Phi_i^{\mathscr{D}}\}_{i\in\omega}$ define the properties of randomness. This collection would then yield the individually necessary and jointly sufficient conditions for a sequence to be intuitively random.

Consequently, the $\mathscr{D}$-advocate will need to argue that $\{\Phi_i^{\mathscr{D}}\}_{i\in\omega}$ is the desired collection of properties, but in order to do so, she must successfully establish that all and only the intuitively random sequences satisfy all of the properties in $\{\Phi_i^{\mathscr{D}}\}_{i\in\omega}$. Clearly, then, as was the case with adjudication via disqualifying properties, recourse to the prevailing intuitive conception of randomness is necessary to carry out this task.[9]

We have seen that among the grounds for adjudication available to the $\mathscr{D}$-advocate are to appeal to disqualifying properties and to properties of randomness. But there is another possibility: the $\mathscr{D}$-advocate can appeal to certain purposes that an extensionally adequate definition of randomness might fulfill.

11.2.2.4 Adjudication via Purposes

How might one adjudicate between the various definitions of randomness by appealing to the purposes that such definitions might fulfill? In particular, how can the $\mathscr{D}$-advocate appeal to these purposes to establish the correctness of $\mathscr{D}$? Here's one approach:

> **Step One:** Identify a purpose that can only be fulfilled by a definition of randomness that captures the prevailing intuitive conception of randomness.
>
> **Step Two:** Establish that $\mathscr{D}$ fulfills this purpose.

[9]Note, however, that this task cannot be carried out independently of the task of identifying which putative disqualifying properties should be counted as legitimate. Moreover, it's not clear that these two tasks are even *distinct*. Since a disqualifying property $\mathcal{Q}$ is a property that is satisfied by only measure zero many sequences (for otherwise, this would conflict with the universally accepted view that the collection of random sequences should have measure one), it follows that the collection of sequences not satisfying $\mathcal{Q}$ has measure one. Thus if we hold that no intuitively random sequence can satisfy $\mathcal{Q}$, then it follows that every intuitively random sequence should satisfy the property $\mathcal{P}$ such that for each $X \in 2^\omega$,

$$\mathcal{P}(X) \text{ if and only if } \neg\mathcal{Q}(X).$$

That is, $\mathcal{Q}$ should be considered as a legitimate disqualifying property if and only if $\mathcal{P}$ is considered as a property of randomness.

But to what purpose could the $\mathscr{D}$-advocate appeal here? Obviously, it can't be the purpose of capturing the prevailing intuitive conception of randomness, as that is the claim the $\mathscr{D}$-advocate is ultimately trying to establish. Thus, she must identify some *other* purpose that only can be fulfilled by a correct definition of randomness. Perhaps that purpose is to serve as a replacement for imprecise uses of the phrase "random sequence" in certain scientific contexts. If so, the $\mathscr{D}$-advocate should identify these uses and these contexts and explain why "$\mathscr{D}$-random sequence" can serve as a suitable replacement for uses of the phrase "random sequence" in those contexts.

There is some risk in this approach, however: if the $\mathscr{D}$-advocate merely identifies one or more contexts in which "$\mathscr{D}$-random sequence" is a suitable replacement for the use of the phrase "random sequence" in those contexts, this leaves open the possibility that there is some other context and some other definition of randomness $\mathscr{D}_0$ such that "$\mathscr{D}_0$-random sequence" is a suitable replacement for the uses of "random sequence" in this other context.[10] To rule of this possibility, the $\mathscr{D}$-advocate needs to provide an account as to why the contexts she has identified require a definition of randomness that captures the prevailing intuitive conception of randomness to serve as a suitable replacement for the informal notion of randomness used in those contexts. Further, she most argue that any other contexts not included among those she identified do *not* require a definition that captures the prevailing intuitive conception in order to serve as a suitable replacement of the informal notion of randomness used in those contexts.

[10]As I argue in the next chapter, this is precisely the situation we find with many of the definitions of algorithmic randomness.

Of course, the considerations are only applicable if the $\mathscr{D}$-advocate can identify a purpose that can only be fulfilled by an extensionally adequate definition of randomness. Yet as I will discuss in Section 11.4, such a purpose has yet to be identified in the literature on algorithmic randomness. Since there are no readily identifiable purposes that can only be fulfilled by an extensionally adequate definition of randomness, hereafter we will restrict our attention to adjudication via disqualifying properties and properties of randomness.

To sum up, it appears that the only grounds for adjudication available to the $\mathscr{D}$-advocate are disqualifying properties and properties of randomness. In each of these cases, however, she must make recourse to the prevailing intuition conception of randomness to identify which putative disqualifying properties are legitimate and which properties are the properties of randomness.

11.2.3 Datum 3: The Lack of Precision

We've identified the grounds of adjudication to which the $\mathscr{D}$-advocate can reasonably appeal. But there is a further problem: although the prevailing intuitive conception of randomness must serve as a constraint in determining which disqualifying properties and properties of randomness are the legitimate ones that determine the correct definition of randomness, common formulations of this intuitive conception are not very precise, and certainly not precise enough to meet the needs of $\mathscr{D}$-advocate:

> **Datum 3:** The prevailing intuitive conception of randomness, as commonly formulated, is not precise enough to permit justified adjudication

between the various definitions of randomness.

Before I offer evidence in support of Datum 3, let me say a word about this datum. First, in presenting Datum 3, I am not denying that the prevailing intuitive conception of randomness allows *some* justified adjudication between the various definitions. For instance, most (if not all) informed individuals would hold that any sequence that contains more 0s than 1s in each of its initial segments should not be counted as random, in the sense that such a sequence is distinguishable from the sequences that typically result from the repeated tosses of a fair coin. My claim is that we cannot *fully* adjudicate between the various definitions of randomness if the prevailing intuitive conception of randomness is to be the final arbiter in these matters, as the prevailing intuitive conception lacks the precision necessary to underwrite the judgments needed for such adjudication.

To verify Datum 3, let us consider both disqualifying properties and properties of randomness as grounds of adjudication. First, I claim that the prevailing intuitive conception of randomness does not help to resolve the question as to which putative disqualifying properties are legitimate. If we look back at the disqualifying properties considered in the previous chapter, it's hard to see how the prevailing intuitive conception of randomness can help us decide which of these are legitimate. For instance, should a sequence that is decidable in the limit be counted as not intuitively random? What about a sequence that is compressible by some prefix-free machine but not by any computable measure machine?[11] What about a sequence from which

[11]While every Schnorr random sequence has high Kolmogorov complexity when measured in terms of any computable measure machine, some Schnorr random sequences have extremely low prefix-free Kolmogorov complexity.

we can compute the halting problem? If we're going to answer these and related questions by appealing to the prevailing intuitive conception of randomness, a more precise account of this conception is needed.

Do we fare any better when considering properties of randomness? Perhaps. For in the previous chapter, I highlighted four hallmarks of randomness, properties on the basis of which one attributes randomness to a sequence: typicality, unpredictability, incompressibility, and independence. Each of these can be taken to provide some sharpening of the prevailing intuitive conception of randomness. Do these help us to determine which properties are to be included as the properties of randomness? Unfortunately for the $\mathscr{D}$-advocate, they do not.

In support of this claim, recall that as there are definitions of randomness based on the first three of the these four hallmarks of randomness that converge to the extension of Martin-Löf randomness, there are also similar definitions that converge to the extension of Schnorr randomness. Additionally, one can find many other non-equivalent definitions of randomness given in terms of these hallmarks. The general point here is that the $\mathscr{D}$-advocate cannot simply assert the correctness of $\mathscr{D}$-randomness on the grounds that $\mathscr{D}$ is a formalization of some hallmark of randomness (or is equivalent to a number of definitions of randomness each of which is a formalization of some hallmark of randomness), at least if we understand these hallmarks of randomness as they are commonly formulated (such as the way they are formulated in the literature on algorithmic randomness).

But there is another problem for the $\mathscr{D}$-advocate involving the various hallmarks of randomness, one that is closely related to the problem with identifying which putative disqualifying properties should be counted as legitimate. Specifically, for each hallmark of randomness, there are ways of further sharpening the hallmark that might undermine the claim of the correctness of $\mathscr{D}$-randomness. For instance, the MLR-advocate can argue that the MLCT is supported by the convergence of Martin-Löf randomness with a definition of randomness based on the hallmark of unpredictability. However, as Δ^0_2 sequences are predictable-in-the-limit, and given that there are Δ^0_2 Martin-Löf random sequences, it follows that some Martin-Löf sequences are, in some sense, predictable. So we have a hallmark of randomness, unpredictability, that understood in one respect (in terms of computably enumerable martingales) lends support to the MLCT, while understanding it in another respect (in terms of predictions-in-the-limit) militates against it.

In sum, if the $\mathscr{D}$-advocate is going to establish the correctness of $\mathscr{D}$-randomness by appealing to disqualifying properties and properties of randomness, she must offer a precise enough account of the prevailing intuitive conception of randomness to justify her choice of disqualifying properties and properties of randomness.

11.2.4 Upshot of the Data

Recall the statement of the Justificatory Challenge as given in Section 11.1:

> **Justificatory Challenge:** Provide a sharpening of the prevailing intuitive conception of randomness that is precise enough to block the claims of extensional adequacy made concerning alternative definitions of randomness without undermining the claim of the extensional adequacy of $\mathscr{D}$.

In light of our three pieces of data, that the $\mathscr{D}$-advocate must face the Justificatory Challenge is immediate. By Datum 1, the definition $\mathscr{D}$ satisfies the property (†), i.e., there are definitions of randomness $\mathscr{D}_1$ and $\mathscr{D}_2$ such that

$$\mathsf{ext}(\mathscr{D}_1) \subsetneq \mathsf{ext}(\mathscr{D}) \subsetneq \mathsf{ext}(\mathscr{D}_2).$$

Thus, if the $\mathscr{D}$-advocate is going to establish the correctness of $\mathscr{D}$-randomness, she must show that (i) every sequence in $\mathsf{ext}(\mathscr{D}) \setminus \mathsf{ext}(\mathscr{D}_1)$ is intuitively random and (ii) no sequence in $\mathsf{ext}(\mathscr{D}_2) \setminus \mathsf{ext}(\mathscr{D})$ is intuitively random. By Datum 2, to carry out this step, the $\mathscr{D}$-advocate needs to provide grounds for adjudicating between the various definitions of randomness. Moreover, these grounds must make recourse to the prevailing intuitive conception of randomness. By Datum 3, however, this prevailing intuitive conception that plays a necessary role in adjudication is not precise enough to underwrite the claims that the $\mathscr{D}$-advocate needs to secure to establish the correctness of $\mathscr{D}$.

It appears, then, that the only option available to the $\mathscr{D}$-advocate is to offer a more precise account of the intuitive conception of randomness, what I'm referring to as a *sharpening* of the intuitive conception. That is, the $\mathscr{D}$-advocate must meet out the Justificatory Challenge.[12]

[12]One approach that the $\mathscr{D}$-advocate might pursue in order to avoid meeting the Justificatory Challenge is to fall back to the position that $\mathscr{D}$ provides not a conceptual analysis of randomness, but an *explication* of randomness. According to Carnap, the key features of an explication are

> 1. The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.
>
> 2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an exact form, so as to introduce the

11.3 Several Questions

We've established that the $\mathscr{D}$-advocate must meet the Justificatory Challenge. But there are several pressing questions to be addressed. First, what exactly is a sharpening of the prevailing intuitive conception of randomness? Second, just how burdensome is this Justificatory Challenge?

11.3.1 What is a Sharpening of the Prevailing Intuitive Conception of Randomness?

As a first pass at answering this question, let's briefly consider Peter Smith's discussion of the passage from a pre-theoretic concept to a precise formalization of that concept, as found in his article "Squeezing Arguments" [Smi11]. According to Smith, there are three conceptual levels involved in this passage:

> **Level One:** First, we have "initial, inchoate, 'unrefined', ideas" of the concept in question.

> explicatum into a well-connected system of scientific concepts.
>
> 3. The explicatum is to be a fruitful concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).
>
> 4. The explicatum should be as simple as possible; this means as simple as the more important requirements (1), (2), and (3) permits ([Car50], p. 7).

One benefit of this approach for the $\mathscr{D}$-advocate is that $\mathscr{D}$-randomness can be a successful explicatum of intuitive randomness even if not every $\mathscr{D}$-random sequence is intuitively random. However, by taking this approach, the $\mathscr{D}$-advocate does not avoid having to meet the Justificatory Challenge, for now she needs to provide grounds for holding that $\mathscr{D}$, and not some alternative definition, provides an explication of the concept of randomness, since it can be reasonably claimed of a number of definitions of randomness that they satisfy 1-3 above. Thus, this approach faces the same problems as those faced by the advocate of the claim that $\mathscr{D}$ provides a conceptual analysis of the concept of randomness.

Level Two: Next, we have an "idealized though still informal and vaguely framed notion", having sharpened the Level One ideas in certain respects.

Level Three: Lastly, we arrive at "crisply defined notions", i.e. one or more formal concepts ([Smi11], pp. 28, 29).

Of particular interest for the present discussion is the passage from Level One to Level Two, concerning which Smith writes,

> The move from the first to the second level involves a certain exercise in conceptual sharpening. And there is no doubt a very interesting story to be told about the conceptual dynamics involved in such a reduction in the amount of 'open-texture', as we get rid of some of the imprecision in our initial inchoate ideas and privilege some strands over others — for this exercise isn't an *arbitrary* one ([Smi11], p. 29).

Moreover, what is distinctive about the passage from Level One to Level Two is that "having done *this* much informal tidying, although on the face of it we've still left things rather vague and unspecific, *in fact* we've done enough to fix a determinate extension for the notion" in question ([Smi11], p. 28, emphasis in the original). Thus, on Smith's account, we don't have a single intuitive concept at Level One, but only a jumble of informal, intuitive ideas; it is only at Levels Two and Three that we consider concepts, informal ones at Level Two and formal ones at Level Three. Thus, when one shows that an intuitive notion is coextensive with some formal concept, she shows a Level Two concept to be coextensive with a Level Three concept, at least on Smith's account.

From the above discussion, we can isolate two kinds of sharpening: the passage from Level One to Level Two, which I'll call $sharpening_{1\to2}$ and the passage from Level

Two to Level Three, which I'll call *sharpening*$_{2\to3}$. It is the former kind of sharpening, sharpening$_{1\to2}$, of the intuitive conception of randomness that I claim the $\mathscr{D}$-advocate must provide. But although Smith's account gives as a good start in understanding sharpening$_{1\to2}$, we need to be more clear about what this sharpening$_{1\to2}$ is, so as not to burden the $\mathscr{D}$-advocate with some indeterminate task.

There are several distinctive features of sharpening$_{1\to2}$ that are readily identifiable. First, Smith's account suggests that we can think of sharpening$_{1\to2}$ as a process that takes an input and produces some output, the input being a jumble of pre-formal, intuitive ideas and the output being an informal concept that has a definite extension. But there is an additional constraint: the informal concept that is the output of the sharpening$_{1\to2}$ must bear some sort of conceptual connection to the informal ideas that were the input of the sharpening$_{1\to2}$. Of course, now we're saddled with the task of explaining what this conceptual connection could be, itself a tall order.[13]

However, we might get around the problem of identifying this conceptual connection by considering some paradigm examples of sharpening$_{1\to2}$. Smith offers two

[13]What we take this conceptual connection to be will be determined, in part, by how we understand the Level One ideas that serves as our input. Are they ideas concerning certain properties $P_1, \ldots, P_n$? If so, then we might require of output concept C that it satisfy

$$C(x) \Rightarrow P_i(x)$$

for at least one $i \leq n$, where x is the relevant sort of object (it might be too much to require that $C(x)$ implies $P_i(x)$ for *every* $i \leq n$, since there's no guarantee that $P_1, \ldots, P_n$ are all satisfiable by one single object x). What if we take these ideas to be certain propositions? Then we might require some sort of implicational relationship between certain propositions about the concept C and these Level One propositions. Clearly, there is much to fill in here, a task that would take us beyond the scope of this study.

such examples, the passage from Level One ideas of validity to the Level Two informal concept of validity-in-virtue of form, and the passage from Level One ideas of computability to the Level Two informal concept of effective calculability. I won't consider the details of these examples here (although I did discuss the latter example in the previous chapter under the guise of Turing's "direct appeal to intuition"), but in both cases, it is clear that the Level One ideas and the associated Level Two concept are closely linked, so that, say, a competent user of the Level Two concept will readily recognize that the Level One ideas apply to the same objects to which the Level Two concept applies.

There's a lot more to say about the notion of sharpening.[14] At a minimum, for the $\mathscr{D}$-advocate to provide a sharpening of the intuitive conception of randomness requires her to provide an informal Level Two concept of randomness that has a definite extension. This leads us to the second question raised above: Just how burdensome is this task?

11.3.2 How Burdensome is the $\mathscr{D}$-advocate's Burden?

Let's recap what the supposed burden is that the $\mathscr{D}$-advocate must address. The $\mathscr{D}$-advocate needs to provide a sharpening of the intuitive conception of randomness so that

(i) every sequence in $\mathsf{ext}(\mathscr{D}) \setminus \mathsf{ext}(\mathscr{D}_1)$ is intuitively random according to this suf-

[14] I haven't addressed how one actually engages in this activity of conceptual sharpening. To answer this question, we need some way of gauging the degree of precision or sharpness of a characterization of a concept. Unfortunately, no account of the degrees of precision of various characterizations of a given concept is available, and to develop such an account would itself be a major undertaking.

ficiently precise account, and

(ii) no sequence in $\mathsf{ext}(\mathscr{D}_2) \setminus \mathsf{ext}(\mathscr{D})$ is intuitively random according to this sufficiently precise account.

What's to prevent the $\mathscr{D}$-advocate from actually carrying this out? One reason to be doubtful that the $\mathscr{D}$-advocate can bear this burden is there just aren't any reasonable options available to her. In the previous chapter, I laid out all of the currently available lines of argument for the MLR-advocate to pursue just to show that no sequence in $\mathsf{SR} \setminus \mathsf{MLR}$ is intuitively random, and none were successful. But more pressing, it's far from clear how the MLR-advocate should address the putative disqualifying properties offered by the W2R-advocate. And the situation is no better for both the SR-advocate and the W2R-advocate.

This should remind us of the predicament facing Ville as he tried to identify the conditions of irregularity that would yield a definition of randomness that attained the exemplary ideal. The problem was that every choice of conditions was an arbitrary choice. It may have seemed that there was no need to make an arbitrary choice when Martin-Löf posed his definition in 1966; at the time, Martin-Löf randomness may have appeared to be a candidate for attaining the exemplary ideal.

Since then, however, many definitions of randomness have accumulated, and to adjudicate between them, it appears that we have to make arbitrary choices about which putative disqualifying properties should be taken to be legitimate disqualifying properties and which properties are to be counted as properties of randomness. That is, there is no *principled* account offered as to why this property and not that one is a property of randomness, or why this putative disqualifying property is legitimate while that one is not.

At this stage of my argument, I don't take this to be particularly strong evidence: the inability to discern what such a principled account could be may just be due to a lack of vision on my part. But I think there are additional reasons to hold that these apparently arbitrary choices made in support of various definitions of randomness *must* be arbitrary; namely, an alternative account of the roles that these definitions are to play, according to which none of the various definitions of randomness fully captures all of the significant truths about randomness, but that multiple definitions nonetheless provide insights into certain mathematical uses of the concept of randomness. I discuss these roles in the next chapter.

11.4 A Serious Objection

Before we turn to the next chapter, there is one serious objection that should be addressed here.[15] The objection is this: If the $\mathscr{D}$-advocate must face this burden, why shouldn't we also require the advocate of the CTT to meet a similar Justificatory

[15] There is also a noteworthy but less serious objection to consider. According to this objection, my account is incomplete, as I've neglected to consider the possibility of a correct *disjunctive* definition of randomness, formed by taking the disjunction of the various definitions of algorithmic randomness, such as MLR, SR, CR, and so on. This isn't much of an objection, however, for the following reason: the definitions under consideration form a partial order (under containment of the extensions of the definitions), and thus the extension of a disjunctive definition of randomness is simply the union of all of the extensions of the definitions that are the disjuncts. Moreover in the case that one definition $\mathscr{D}$ in the union is below all the others in the ordering given by $\subseteq$, the union of the definitions in the disjunction is simply coextensive with $\mathscr{D}$. And this is precisely the situation in which we find ourselves here: For instance, the union of the extensions of the definitions we've considered in this chapter is

$$\mathsf{W2R} \cup \mathsf{MLR} \cup \mathsf{CR} \cup \mathsf{SR} \cup \mathsf{KR} = \mathsf{KR}.$$

Therefore, nothing is gained in considering disjunctive definitions of randomness.

Challenge? But given that the CTT is widely accepted even though the corresponding Justificatory Challenge has not been met, why not hold that the $\mathscr{D}$-advocate can establish the correctness of $\mathscr{D}$ without meeting the Justificatory Challenge as laid out in this chapter? More generally, one might worry that the standards that I've set forth to justify the claim of the correctness of a formal definition of an intuitive notion are so stringent as to rule out the possibility of justifying *any* claim that one has provided a successful conceptual analysis of a concept such as randomness or computability.

To be clear, although I've introduced the randomness-theoretic theses such as the MLCT as analogues of the CTT, my intention is not to call into question extensional adequacy theses *in general*; in fact, my intention is to present an argument for the No-Thesis Thesis the acceptance of which does not require one to reject the CTT. Towards this end, I will highlight several important differences between the task of justifying the CTT (henceforth, the CTT-JT (for *justificatory task*)) and the task of justifying the MLCT (henceforth, the MLCT-JT).

In particular, there are three differences between the CTT-JT and the MLCT-JT that I want to emphasize here:

(1) the challenge of viable alternatives,

(2) the sharpening of the associated intuitive notion, and

(3) the role of extensional adequacy.

11.4.1 Difference 1: Viable Alternatives

The first difference between the CTT-JT and the MLCT-JT concerns the challenge of viable alternatives to the relevant theses. Unlike the case with the MLCT, there is a lack of viable alternatives to the CTT. While this certainly doesn't mean that the CTT is true by default, it does mean that the CTT-JT is not saddled with the same problems that beset the MLCT-JT.[16]

But might the various alternative definitions of randomness be more akin to the various models of computability than I am letting on? For just as there are various relativized notions of randomness, there are also various relativized notions of computability, and just as there are notion of randomness weaker than Martin-Löf randomness, in the sense that they are less restrictive (for instance, computable randomness, Schnorr randomness, and weak randomness), there are models of computability weaker than Turing computability, in the sense that they compute fewer functions (for instance, the primitive recursive functions, those functions computable by a pushdown automaton, those computable by a finite state automaton, those computable at various levels of the subrecursive hierarchies, and so on).

Do any of these alternative models of computability give rise to a thesis that is a viable alternative to the CTT? First, there is a very good reason to think that these "weaker" definitions do not capture all intuitively computable functions. Each such collection of functions, such as the collection of primitive recursive functions, is effectively enumerable, so if we let $\{\psi_e\}_{e\in\omega}$ be the collection of primitive recursive functions, then setting $\theta(e) = \psi_e(e) + 1$ yields an intuitively computable function

[16] Of course, this is not to say that the CTT-JT might face problems that do not beset the MLCT-JT.

that is not primitive recursive, so that we have effectively diagonalized out of the class of primitive recursive functions. Similarly, one can effectively diagonalize in this way out of any class of functions weaker than the collection of Turing computable functions.

But the same cannot be said for the class of Turing computable functions, since the collection of total computable functions is not effectively enumerable. In fact, the fact that one cannot diagonalize out of the class of computable functions convinced Kleene of the truth of the CTT:

> When Church proposed this thesis [i.e. Church's Thesis], I sat down to disprove it by diagonalizing out of the class of λ-definable functions. But, quickly realizing that the diagonalization cannot be done effectively, I became overnight a supporter of the thesis ([Kle81], p. 59).

But what about stronger alternative definitions of computability, which count more functions among the intuitively computable functions? As we will now see, the fact that the intuitive concept of computability can be made sufficiently precise allows us to conclude that strong alternative definitions of computability are too strong.

11.4.2 Difference 2: Sharpening the Intuitive Notion

The second difference between the CTT-JT and the MLCT-JT that I want to highlight here concerns the notion of sharpening that I discussed in Section 11.3.1. Whereas the MLR-advocate is burdened with the task of sharpening the intuitive conception of randomness in a unique way, a task that I have argued is unlikely to be carried out successfully, the task of sharpening the intuitive conception of human computability appears to have been carried out successfully by Turing, as we discussed in the previous chapter (see Section 10.3.4.2). Gödel, among many others, recognized this, stating that "the correct definition of mechanical computability was established beyond any doubt by Turing", supporting this claim by describing Turing's so-called "direct appeal to intuition" ([Göd93], p. 168).

One consequence of the sharpening that Turing offers is that by isolating the distinctive features of a computation carried out by a human computor, we can rule out those computational procedures involving access to an oracle as surpassing the notion of human computability. By contrast, if we consider the notion of randomness, there appears to be no reasonable sharpening of the intuitive conception of randomness that allows us to conclude that definitions of randomness stronger than Martin-Löf randomness, such as weak 2-randomness surpass the intuitive conception of randomness.

But there is more that one can say here. Recall the three conceptual levels involved in the passage from a pre-theoretic, intuitive conception to a corresponding formal concept, as identified by Smith and discussed in Section 11.3.1:

> **Level One:** First, we have "initial, inchoate, 'unrefined', ideas" of the concept in question.
>
> **Level Two:** Next, we have an "idealized though still informal and vaguely framed notion", having sharpened the Level One ideas in certain respects.
>
> **Level Three:** Lastly, we arrive at "crisply defined notions", i.e. one or more formal concepts ([Smi11], pp. 28, 29).

Now what do these levels look like when we consider the passage from initial, inchoate ideas of randomness to formal definitions of randomness? First, at Level One, we have something along the lines of untutored, folk intuitions of randomness, intuitions that inform everyday attributions of randomness to, say, events, sequences of events, configurations of objects, and so on. In identifying the various hallmarks of randomness, on the basis of which we make attributions of randomness, we've already taken a step to make these Level One ideas more precise. For instance, if we hold that a sequence is random only if it is typical, this is certainly a sharpening of the Level One ideas of randomness, but one not sharp enough to yield a definite extension of sequences. Suppose we further sharpen the notion of typicality by formulating it in terms of the avoidance of certain sets of measure zero—at this point, we've moved well beyond merely 'intuitive' characterizations of randomness, as such a sharpening will likely be informed by measure-theoretic considerations. But observe that we still haven't arrived at a concept with a definite extension, and so, according to Smith's account, we haven't properly arrived at a Level Two concept, as there are many choices available to us as to which sets of measure zero should be included in our definition of typicality.

The crucial observation to make is this: as soon as we identify the sets of measure zero in terms of which typicality is defined, we arrive at a Level Three, formal concept. More importantly, *at no point in this passage do we arrive at an informal concept with a definite extension.* Further, our informal, Level Two schema, prior to the identification of these sets of measure zero, is consistent with multiple, non-equivalent extensions. For instance, once we choose to characterize typicality in terms of a certain collection of sets of measure zero, if we specify that a sequence is typical if it avoids all Π^0_1 subsets of 2^ω, we arrive at the definition of Kurtz randomness; if instead of Π^0_1 classes of measure zero, we consider the collection of Π^0_2 classes $\mathcal{F}$ with the restriction that each such $\mathcal{F}$ must have the form $\mathcal{F} = \bigcap_i \mathcal{U}_i$ with $\lambda(\mathcal{U}_i) = 2^{-i}$ for each i, then we arrive at the definition of Schnorr randomness. Further, if we consider those Π^0_2 classes $\mathcal{F}$ of the form $\mathcal{F} = \bigcap_i \mathcal{U}_i$ with $\lambda(\mathcal{U}_i) \leq 2^{-i}$ for each i, then we arrive at the definition of Martin-Löf randomness; lastly, if we consider *all* Π^0_2 classes of measure zero, then we arrive at the definition of weak 2-randomness. Importantly, prior to the specification of the class of measure zero sets, each of these definitions was compatible with the above characterization of random sequences as typical (i.e., that random sequences avoid certain sets of measure zero).

Smith's model of the three conceptual levels involved in the passage from pre-theoretic ideas of a notion to a precise formal concept thus does not appear to be the appropriate model for considering the passage from informal, intuitive ideas of randomness to formal definitions of randomness.[17] In particular, since Smith requires of Level Two informal concepts that they have a definite extension, it's not clear that the standard informal characterizations of randomness, such as those given in terms of the various hallmarks of randomness, can be properly seen as Level Two informal concepts of randomness. Instead, it appears that these informal characterizations yield Level Two concept *schemata*. Prior to instantiating each schema, we don't have a concept with a definition extension, but once we instantiate the schema by specifying the relevant collection of objects, we arrive at a Level Three formal concept.

[17] I don't intend to suggest here that Smith took his model to hold of all instances of the phenomenon of passing from pre-theoretic ideas of a notion to a precise formal concept.

The crucial observation to make here is this: *it is the multitude of ways of instantiating this schema that correspond to various sharpenings of the prevailing intuitive conception of randomness.* Perhaps this should come as no surprise. For in looking back to the initial discussion of the various hallmarks of randomness in Chapter 4, one will find that each of the three hallmarks of randomness in terms of which the various definitions of randomness are given (typicality, unpredictability, and incompressibility) can be viewed schematically. Let's compare the hallmarks of randomness as formulated in Chapter 4 with schematic versions of those same hallmarks:

(1a) **Non-schematic approach to typicality:** A sequence is typical if it is passes every statistical test for randomness, and so is not detected as non-random by any such tests.

(1b) **Schematic approach to typicality:** A sequence is typical with respect to a collection $\mathcal{C}$ of statistical tests if it passes every test in $\mathcal{C}$, and so is not detected as non-random by any of the tests in $\mathcal{C}$.

(2a) **Non-schematic approach to unpredictability:** A sequence is unpredictable if its bits cannot be predicted by any method of prediction.

(2b) **Schematic approach to unpredictability:** A sequence is unpredictable with respect to a collection $\mathcal{C}$ of methods of prediction if its bits cannot be predicted by any method of prediction in $\mathcal{C}$.

(3a) **Non-schematic approach to incompressibility:** A sequence is incompressible if its initial segments cannot be compressed by any mode of compression.

(3b) **Schematic approach to incompressibility:** A sequence is incompressible with respect to a collection $\mathcal{C}$ of modes of compression if its initial segments cannot be compressed by any mode of compression in $\mathcal{C}$.

Thus, we can view (1b), (2b), and (3b) as Level Two schemas: unless we fill in some appropriate collection of objects $\mathcal{C}$ in each case, these don't yield a concept with a definite extension. But what's more, it appears that the collection $\mathcal{C}$ is not the only variable component of (1b), (2b), and (3b), for there are several ways to fix the meaning of "passing a test", "being predicted", and "being compressed". So for instance, we can hold that a sequence is predictable if there is some computable martingale that succeeds on it, or we can further require that there is a computable procedure that verifies this success.[18]

On this schematic approach, we can catalogue the various definitions of randomness in the algorithmic randomness literature, as formulated in terms of

1. a hallmark of randomness, which I'll call the *motif* of a definition,
2. a collection of resources that are appropriate for the given motif (a certain class of sets of measure zero for the typicality motif, certain methods of prediction for the unpredictability motif, and certain modes of compression for the incompressibility motif), and
3. a criterion of success, a condition or set of conditions that must be satisfied in order for a sequence to be counted as non-random according to the definition.

[18] Recall that the definition that results from the former option is computable randomness, while Schnorr randomness results from the latter.

Further, by varying the resources and the criterion of success of a definition of randomness, the result will be a range of definitions. More exactly, for a fixed definition $\mathscr{D}$ with a given motif, as we vary the resources of the definition $\mathscr{D}$, the corresponding extension of the resulting definition changes—all other things being equal, if we increase the resources of $\mathscr{D}$, the resulting definition $\mathscr{D}'$ will be a stronger definition, counting fewer sequences as random. Similarly, if we decrease the resources of $\mathscr{D}$, the resulting definition $\mathscr{D}'$ will be a weaker definition, counting more sequences as random. If instead of varying the resources of $\mathscr{D}$ we vary the criterion of success, then the same phenomenon occurs; a more demanding criterion of success results in a smaller extension of sequences that are counted as random, while a less demanding criterion of success results in a larger extension of sequences that are counted as random.

Thus, when we sharpen the prevailing intuitive notion of randomness, we arrive at various schemata, which, depending on they are is filled in, give different formal definitions of randomness. Yet when Turing sharpened the intuitive notion of human computability, he arrived at a concept with a definite extension (at least according to the standard account of Turing's analysis, which I am not challenging here).[19]

[19] One might object here that we can give a schema of definitions of computability given in terms of, say, oracle computability or the subrecursive hierarchies. But such a schema did not arise in the course of attempting to make more precise our informal notions of computability *with the aim of justifying a claim of extensional adequacy.* This is the crucial difference. Of course, the $\mathscr{D}$-advocate might be able to provide a sharpening of our informal ideas of randomness that does not yield a schema, but rather a concept with a fixed extension. In light of the data I present in the next chapter, I think that such a sharpening is not forthcoming.

11.4.2.1 Difference 3: The Role of Extensional Adequacy

The last difference between the CTT-JT and the MLCT-JT concerns the purposes an extensionally adequate definition might serve. First, let's consider the purposes an extensionally adequate definition of computability can serve. Following the terminology of Boolos, Burgess, and Jeffrey, we can distinguish between two primary uses of the CTT, the "unavoidable" uses of the CTT and the "lazy" uses of the CTT.[20]

The paradigm example of an unavoidable use of the CTT is given by Turing (and Church as well). In order to show that the Entscheidungsproblem, the problem of finding an effective procedure that would determine whether a given first-order formula is derivable from the axioms of first-order logic, has a negative answer, Turing showed that there is no Turing computable procedure that could decide the truth of statements of first-order logic. Then, appealing to Turing's thesis, he concluded that no effective procedure could decide the truth of statements of first-order logic. In general, an unavoidable use of the CTT involves the replacement of "Ψ is not Turing computable" with "Ψ is not effectively calculable", where Ψ stands for some predicate, set, procedure, etc.

The lazy uses of the CTT allow one to pass from an informal description of an effective procedure to an index for a partial computable function. That is, instead of deriving the desired partial computable function from, say, the initial functions, the task of verifying that our informal description yields a partial computable function is bypassed. The CTT thus functions here as an informal rule of inference in a proof: One defines an intuitively computable function, and thus by the CTT, it follows that

[20] See [BBJ02], p. 136.

this function is a partial computable function. Thus, there is some e such that ϕ_e that computes the desired function, and ϕ_e is then used in the next steps of the proof.

What about the purposes that an extensionally adequate definition of randomness might serve? As it currently stands, the general theory of algorithmic randomness does not make any "unavoidable uses" of any definition of randomness that is comparable to the use of, say, Turing computability in the proof of the negative solution of the Entscheidungsproblem.[21] That is, there is no proof in the general theory of algorithmic randomness in which one shows that no Martin-Löf random sequence satisfies some property and then further concludes that no intuitively random sequence satisfies that property (nor is there any proof in which one shows that every Martin-Löf random sequence satisfies some property and further concludes that every intuitively random sequences satisfies that property).

What about "lazy uses" of some definition of randomness such as Martin-Löf randomness? Here, too, there does not appear to be a discernible loss. For instance, there is no rule of inference that allows one to pass from an intuitively random sequence to, say, a Martin-Löf random sequence, and reasonably so, for if my account

[21] There is perhaps one exception: Chaitin proves that no reasonable theory of arithmetic (i.e. on that is computably axiomatizable and extends Robinson's Q) can prove more than finitely many statements asserting true values of Ω (of the form "$\Omega(n) = 1$" or "$\Omega(n) = 0$"). Chaitin claims of this result that it shows there are truths of arithmetic that are "true for no reason", the rationale being something like this: since Ω is random, its bit values are akin to the tosses of an unbiased coin. Moreover, almost all of these truths are unprovable in arithmetic, and therefore these are seemingly random facts that are true for no reason. Admittedly, there are many worrisome steps in this argument (and it's far from clear how to make this argument respectable), but it appears that one step that is needed is to pass from the Martin-Löf randomness of Ω to the intuitive randomness of Ω, so that we can conclude that "its bit values are akin to the tosses of a fair coin" or something like that. Thus the only unavoidable use of the MLCT that I can find comes in an argument that is already deeply problematic apart from this implicit use of the MLCT.

is correct, such a rule of inference would not be completely reliable. For as I argue in the next chapter, there are some contexts in which Martin-Löf randomness licenses all of the attributions of randomness and non-randomness, and in these contexts, this rule of inference would be completely reliable, but outside of such contexts, reliability is not guaranteed.

CHAPTER 12

THE CALIBRATIVE AND LIMITATIVE ROLES OF RANDOMNESS

12.1 Introduction

Thus far we have established that the $\mathscr{D}$-advocate must face the Justificatory Challenge: she must provide a sharpening of the prevailing intuitive conception of randomness in order to establish the correctness of the definition $\mathscr{D}$. In addition, in the previous chapter I argued for two points that are pertinent to the present discussion. First, I argued that it is doubtful whether the $\mathscr{D}$-advocate can meet this challenge, as there seems to be no principled choice of the legitimate disqualifying properties and the properties of randomness. Second, I argued that there are no discernible purposes for a definition of randomness that can only be fulfilled by an extensionally adequate definition. While these two points shouldn't be taken to provide particularly strong evidence for the No-Thesis Thesis, the claim that no definition of randomness that has a definite, well-defined extension can capture the prevailing intuitive conception of randomness, they do provide an important first step towards the No-Thesis Thesis.

This is not the fully story on the No-Thesis Thesis, however. For in this chapter, I discuss two roles for definitions of randomness, each of which can be successfully filled by multiple definitions of randomness. Further, as I argue, the fact that multiple definitions of randomness can successfully fill these two particular roles gives us good reason to hold not only that the $\mathscr{D}$-advocate *cannot* meet the Justificatory Challenge, but also to accept the No-Thesis Thesis.

The two roles that I discuss here are what I call the *calibrative role of randomness* and the *limitative role of randomness*. Roughly, a definition $\mathscr{D}$ fills the calibrative role if and only if there is some notion of "almost-everywhere" typicality $\mathcal{T}$ occurring in classical mathematics such that the $\mathscr{D}$-randomness of a sequence is necessary and sufficient for that sequence to be $\mathcal{T}$-typical (where these notions of typicality are specified in the next section). Significantly, there is no single definition of randomness that is coextensive with all the various notions of almost-everywhere typicality that one encounters in classical mathematics. This already strongly suggests that no single definition can capture everything that mathematicians have held to be significant about the notion of randomness.

This is further confirmed by the fact that many definitions of randomness fill what I call the *limitative role of randomness*; such definitions illuminate an interesting phenomenon, the *indefinite contractibility of the notion of absolute randomness*. Broadly speaking, that the notion of absolute randomness is indefinitely contractible means that for every extension $\mathcal{E}$ of sequences that purportedly contains all absolutely random sequences, there is some $X \in \mathcal{E}$ that is not absolutely random, where, following a suggestion of Myhill's, a sequence is absolutely random if and only if it satisfies no property that is (i) satisfied by only measure zero many sequences and

(ii) is definable without parameters in the language of set theory.[1] By means of the various definitions of randomness, we can systematically study this phenomenon of contractibility. In fact, this phenomenon allows us to diagnose the source of the various putative disqualifying properties discussed in the previous two chapters.

An important feature of these two roles is that a definition can successfully fill one of these roles (or both) without capturing the prevailing intuitive conception of randomness. What's more, the fact that multiple definitions of randomness can successfully fill these roles gives us reason to accept the No-Thesis Thesis, or so I argue.

The remainder of the chapter will proceed as follows. In Section 12.2, I introduce the calibrative role of randomness, while in Section 12.3, I consider two objections concerning the significance of the calibrative role as I've characterized it. Next, in Section 12.4 I introduce the limitative role of randomness. Lastly, in Section 12.5, I present my full argument for the No-Thesis Thesis, based on the that the fact that both the calibrative and the limitative roles are successfully filled by multiple definitions of randomness.

12.2 The Calibrative Role

The first role that we will consider in this chapter is what I call the *calibrative role of randomness*. Roughly, a definition $\mathscr{D}$ fills the calibrative role if and only if there is some notion of "almost-everywhere typicality" (hereafter *a.e.-typicality*)

[1] We cannot formally define absolute randomness, but we can study the properties of a predicate $\mathcal{R}$ satisfying certain general conditions that are necessary for absolute randomness, as discussed in Section 12.4.

occurring in classical mathematics such that all and only the $\mathscr{D}$-random sequences instantiate this notion of a.e.-typicality. To understand what this means, we have a bit to unpack here. What is a notion of a.e.-typicality? What does it mean for a sequence to instantiate a notion of a.e.-typicality? And what does it mean for a notion of a.e.-typicality to occur in classical mathematics?

As a first pass at answering these questions, let us consider several examples of the phenomenon in question. In particular, I invoke recent results concerning the relationship between (i) a number of definitions of randomness and (ii) theorems involving certain properties studied in classical mathematics that hold on a set of measure one (theorems from such areas as analysis, ergodic theory, information theory, and probability theory).[2]

Consider the following theorems of classical analysis, each of which concerns some property that holds of almost every real number in [0,1], or equivalently, in a subset of [0,1] of measure one:[3,4]

Theorem 1. For every non-decreasing real-valued function $f : [0,1] \to \mathbb{R}$, f is differentiable almost everywhere.

Theorem 2. For every real-valued function $f : [0,1] \to \mathbb{R}$ of bounded variation, f

[2]Most of the results discussed below are recent, with the exception of one, proved by Demuth in [Dem75]. However, only recently has the significance of Demuth's result been realized.

[3]Here we rely on the fact that every sequence in 2^ω corresponds to a unique point in [0,1], while every non-rational point in [0,1] corresponds to a unique member of 2^ω.

[4]These theorems can be found in any standard textbook on real analysis, such as [Rud87] or [Fol99].

is differentiable almost everywhere.[5]

Theorem 3. For every integrable function $f : [0,1] \to \mathbb{R}$,

$$\lim_{r\to\infty} \frac{1}{\lambda(B_r(x))} \int_{B_r(x)} f(x)dx \tag{12.1}$$

converges to $f(x)$ for almost every x.[6]

Each of these theorems has the form

$$(\forall f \in \mathscr{C})(\forall^{\mathsf{a.e.}} x \in [0,1])\Phi(x,f),$$

where $\mathscr{C}$ is some collection of functions from $[0,1]$ to $\mathbb{R}$, $\forall^{\mathsf{a.e.}}$ is the almost-everywhere quantifier, and $\Phi(x,f)$ is some formula of second-order arithmetic. Further, in each of these cases, the set of real numbers for which the theorem holds is a set of measure one; this is precisely what it means for a result to hold for almost every $x \in [0,1]$. Such results are commonly glossed as follows: If we choose a point $x \in [0,1]$ at random, then with probability one, the relevant property will hold for x; we might also say that it is the typical behavior of points $x \in [0,1]$ for the each of the above properties to hold at x, or that these properties hold of the random member of $[0,1]$.[7]

[5] $f : [0,1] \to \mathbb{R}$ is of bounded variation if its total variation is finite. That is, the supremum of

$$\sum_{i=0}^{n_P} |f(x_{i+1}) - f(x_i)|$$

over all partitions $P = \{x_0, x_1, \ldots, x_{n_P}\}$ of [0,1] is finite.

[6] This result is known as the Lebesgue differentiation theorem.

[7] I should emphasize that these glosses here are not, strictly speaking, formal, insofar as the notions of "typicality" or "random point" are not given precise definitions.

Such behavior is what I am referring to as *almost-everywhere behavior*.

Let us now make some further restrictions on the three instances of almost-everywhere behavior given by Theorems 1-3. Instead of considering

$\mathscr{C}_1 = \{f : [0,1] \to \mathbb{R} : f \text{ is non-decreasing}\}$,

$\mathscr{C}_2 = \{f : [0,1] \to \mathbb{R} : f \text{ is of bounded variation}\}$, and

$\mathscr{C}_3 = \{f : [0,1] \to \mathbb{R} : f \text{ is integrable}\}$

in each of the above three theorems, let us restrict each collection $\mathscr{C}_i$ to just the *computable* functions in each collection,[8] then we have the following analogues of Theorems 1-3:

Theorem 4 (Brattka, Miller, Nies, [BMN11])**.** *$z \in [0,1]$ is computably random if and only if z is a point of differentiability of every non-decreasing, computable, real-valued function $f : [0,1] \to \mathbb{R}$.*[9]

Theorem 5 (Demuth [Dem75]; Brattka, Miller, Nies, [BMN11])**.** *$z \in [0,1]$ is Martin-Löf random if and only if z is a point of differentiability of every computable, real-valued function $f : [0,1] \to \mathbb{R}$ of bounded variation.*

[8] $f : [0,1] \to \mathbb{R}$ is computable if (i) for every computable sequence of real numbers $(x_k)_{k\in\omega}$, the sequence $f(x_k)_{k\in\omega}$ is computable, and (ii) there is a computable function $g : \omega \to \omega$ such that for every $x, y \in [0,1]$ and every $n \in \omega$,

$$|x - y| \leq 2^{-g(n)} \Rightarrow |f(x) - f(y)| \leq 2^{-n}.$$

[9] The definition of computable randomness can be found in Section 2.5.2.1.

Theorem 6 (Pathak, Rojas, Simpson, [PRS11]; Rute, [Rut11])**.** $z \in [0,1]$ is Schnorr random if and only if for every L_1-computable function $f : [0,1] \to \mathbb{R}$,

$$\lim_{r\to\infty} \frac{1}{\lambda(B_r(z))} \int_{B_r(z)} f(x)dx$$

converges to $f(z)$.

More concisely, given a collection $\mathscr{C}$ of real-valued functions $f : [0,1] \to \mathbb{R}$, if $\mathscr{C}^*$ is the subset of $\mathscr{C}$ consisting of the *computable* members of $\mathscr{C}$, $\Phi(z, f)$ says "f is differentiable at z", $\Phi'(z, f)$ expresses the relation given by (12.1) above, and $\mathscr{C}_i$ for $i = 1, 2, 3$ are as above, then we have:

Theorem 4*: $(\forall f \in \mathscr{C}_1^*)\Phi(z, f)$ if and only if $z \in \mathsf{CR}$.

Theorem 5*: $(\forall f \in \mathscr{C}_2^*)\Phi(z, f)$ if and only if $z \in \mathsf{MLR}$.

Theorem 6*: $(\forall f \in \mathscr{C}_3^*)\Phi'(z, f)$ if and only if $z \in \mathsf{SR}$.

These are quite surprising results, for the restriction of each of the classes of functions in Theorems 1-3 to the computable real-valued functions yields not just three properties that hold almost-everywhere, but three properties each of which is coextensive with some definition of algorithmic randomness. Thus we have three distinct notions of what we might call *almost-everywhere typicality* on our hands:

z is a.e.-typical$_1$ if and only if $(\forall f \in \mathscr{C}_1^*)\Phi(z, f)$;

z is a.e.-typical$_2$ if and only if $(\forall f \in \mathscr{C}_2^*)\Phi(z, f)$;

z is a.e.-typical$_3$ if and only if $(\forall f \in \mathscr{C}_3^*)\Phi'(z, f)$.

Combining the above three statements with Theorems 4*, 5*, and 6* yields

$$z \text{ is a.e.-typical}_1 \text{ if and only if } z \in \mathsf{CR};$$

$$z \text{ is a.e.-typical}_2 \text{ if and only if } z \in \mathsf{MLR};$$

$$z \text{ is a.e.-typical}_3 \text{ if and only if } z \in \mathsf{SR}.$$

There are a number of additional results along these lines,[10] and it is plausible to hold that even more will be established for other notions of a.e.-typicality that are not equivalent to these three. But what is the significance of these results?

To answer this question, let us revisit the exemplary role of randomness. Recall that in Chapter 9 we discussed the exemplary role of randomness, first formulated by Jean Ville, who wanted to find a definition of randomness that would satisfy all of the properties typically held by a sequence chosen at random (properties I referred to as the $\mathcal{R}$-properties). Based on my argument in Chapter 11, there does not appear to be a single collection of properties that falls under the description "the properties typically held by a sequence chosen at random", as there seems to be no principled

[10] There is an even stronger result along the lines of Theorems 4^*, 5^*, and 6^* involving weak 2-randomness:

Theorem (Brattka, Miller, Nies [BMN11])**.** $z \in [0,1]$ is weakly 2-random if and only if every almost everywhere differentiable computable function $f : [0,1] \to \mathbb{R}$ is differentiable at z.

Recently, Miyabe has shown a number of similar results for weak randomness (defined above in footnote 6); for instance:

Theorem (Miyabe [Miy12])**.** $z \in [0,1]$ is weakly random if and only if $f(z) < \infty$ for every non-negative extended computable function $f : [0,1] \to \mathbb{R} \cup \{-\infty, \infty\}$ such that $f(x) < \infty$ almost everywhere.

There are a number of other almost-everywhere results from areas outside of analysis that pick out classes of random sequences. For instance, an effective version of Birkhoff's ergodic theorem holds at all and only the Martin-Löf random points, as recently shown in [BDH$^+$11] and [FGMN11]. Similarly, Hoyrup [Hoy12] has recently show that an effective version of the Shannon-Breimann-McMillman Theorem from information theory holds for all and only the Martin-Löf random sequences.

way to define such a collection of $\mathcal{R}$-properties, a conclusion that led Ville to conclude that no definition could be given that satisfies the exemplary ideal.

We now have another reason to hold that there no single collection of properties falling under the description "the properties typically held by a sequence chosen at random": there are multiple choices of such properties that yield definitions of randomness that capture certain instances of a.e.-typicality occurring in classical mathematics. We thus have something along the lines of a *restricted* exemplary role. For each definition $\mathscr{D}$ that is coextensive with a notion of a.e.-typicality, $\mathscr{D}$-randomness exemplifies the degree of randomness needed for a given almost-everywhere property to hold. Thus, while we cannot specify a complete set of $\mathcal{R}$-properties, properties that hold of all and only the sequences chosen at random, for a given notion of a.e.-typicality $\mathcal{T}$, there is some collection of properties, the $\mathcal{R}_{\mathcal{T}}$ properties, such that X exemplifies $\mathcal{T}$-a.e.-typicality if and only if X satisfies all the $\mathcal{R}_{\mathcal{T}}$-properties.

This phenomenon suggests a research program that is already under way, namely to classify various notions of a.e.-typicality in classical mathematics according to the corresponding definition of randomness. While a number of results along these lines have been established (as discussed above), there is still much work to be done in this direction—for each result in classical mathematics involving an almost-everywhere quantifier (where the notion of almost-everywhere is in understood in the sense of measure, and not, say, in the sense of Baire category), one can investigate the degree of randomness necessary and sufficient for the result to hold effectively. Thus we have an analogue of reverse mathematics, a program that one might refer to as *reverse* *randomness*.[11,12]

12.3 Two Objections

Before we consider the limitative role, there are two objections to the calibrative role as I've formulated it above. The first objection is directed towards the notions of a.e.-typicality that the various definitions purportedly capture, while the second objection involves a worry that the various definitions of randomness, and the corresponding notions of a.e.-typicality, are artificial, and thus these definitions do not capture what mathematicians consider to be significant truths about randomness.

12.3.1 Objection 1: Bad Company for Notions of a.e.-Typicality

Let $\Psi(X)$ be the formula $X = X$, which holds for almost every sequence in 2^ω, because it holds for *every* sequence in 2^ω. Why isn't the collection of sequences satisfying Ψ, namely 2^ω, counted as fulfilling the calibrative role? Specifically, the objection is that I haven't said enough about what counts as an *admissible* notion of a.e.-typicality (i.e. one that is potentially coextensive with a definition of randomness), and without some constraint on the notion of a.e.-typicality, my account runs the risk of counting all sorts of pathological notions as admissible.

This objection can be seen as an analogue of the Bad Company Objection to the

[11] This designation is due to Laurent Bienvenu.

[12] I think it is reasonable to ask whether the correspondence between notions of a.e.-typicality and definitions of randomness can be formalized within a sufficiently rich formal system, such as one equipped with a generalized quantifier that can be interpreted as quantifying over all and only the $\mathscr{D}$-random sequences for different definitions of randomness $\mathscr{D}$. There is the potential for some interesting formal work to be done here.

use of abstraction principles such as Hume's Principle to establish that all arithmetical truths are conceptual truths. Just as one might object that Hume's Principle cannot underwrite the introduction of new concepts due to the fact that it is too closely related to certain unacceptable abstraction principles, one might object that the notions of a.e.-typicality cannot be used to secure the various definitions of randomness as capturing significant truths about the concept of randomness since these notions of a.e.-typicality are too closely related to, say, the notion of typicality given by Ψ, a notion that clearly captures no significant truths about the concept of randomness.

One initial response is to rule out such formulas as $X = X$ from defining a notion of a.e.-typicality on the grounds that a notion of typicality that counts every sequence as typical (so that the corresponding notion of atypicality counts no sequence as atypical) is an odd notion of typicality. But what exactly is odd about such a notion, and why should the oddness of a notion of typicality give us grounds for its disqualification?

Perhaps one might further appeal to intuitive considerations. Consider everyday, informal uses of the predicate "x is typical": "the typical American family", "the typical income", "the typical interview questions", etc. We don't think that the typical American family is just *any* American family, that the typical income is just *any* income, and that the typical interview questions include all questions that have been or might be asked at an interview. That is, in each of these cases, the predicate "x is typical" picks out a proper subcollection of the domain in question.

But surely this doesn't imply that *every* use of "x is typical" picks out a proper subcollection of the domain in question. For instance, if all of the members of a given domain $\mathcal{D}$ are indistinguishable relative to some property (or collection of properties) that is shared by all members of the domain, then "the typical x" would apply to all members of $\mathcal{D}$. It's not clear what is problematic in such a case.

Let us consider a third, more promising approach. For certain purposes, a notion of typicality that is satisfied by every object in a given domain might be interesting or even useful. But for many mathematical purposes, a notion of typicality that counts all sequences as typical is neither interesting nor useful; there are many properties that are of interest to mathematicians but which do not hold of every object in a given domain, but only of every sequence in a large subset of the domain. In some cases, this notion of largeness is given in terms of Lebesgue measure (i.e., the property of having Lebesgue measure one), while in other cases, it is given in terms of Baire category (i.e., the property of being comeager).[13] Further, it is the former variety of typicality, the measure-theoretic variety, that are of studied in classical probability theory and statistics.

The notions of a.e.-typicality discussed in 12.2 are all of the former variety: each holds for measure one many sequences, and fails to hold for uncountably many sequences. *These* are the notions of typicality that are relevance to the calibrative role; not even notions of typicality given in terms of Baire category are relevant to the calibrative role of randomness. This is because the most basic laws of probability, such as the Law of Large Numbers, don't hold of sequences that are typical in the

[13]Recall that a subset of a topological space with a countable dense set is meager if it is the union of nowhere dense sets, and a set is comeager if its complement is meager.

sense of Baire category.

Note that even with this restriction to notions of typicality that fall under the purview of classical probability theory and statistics, one can still cook up a number of gerrymandered sets of measure one (even ones definable in second-order arithmetic without parameters) that should not be considered legitimate notions of a.e.-typicality. To rule out these gerrymandered notions of a.e.-typicality, one might further require that admissible notions of a.e.-typicality must exclude all non-normal sequences, where a sequence X is normal if for every finite binary sequence σ, the limit of the relative frequency of the occurrence of σ in X as a subword is $2^{-|\sigma|}$.[14] This too is not an unreasonable requirement, given that the collection of normal sequences has a nice algorithmic characterization, as the collection of sequences that are invariant under all place selections computable by a finite state automaton.[15] Additionally, nearly all of the currently available definitions of algorithmic randomness satisfy this requirement.

One notable exception is Kurtz randomness, introduced back in Section 11.2. The problem is that the collection of Kurtz random sequences includes all weakly 1-generic sequences,[16] which fail to satisfy even the most basic laws of probability,

[14]Following the notation of [MR06], for a finite sequences $\sigma, \tau \in 2^{<\omega}$, $\mathsf{occ}_\sigma(\tau)$ is the number of times that σ occurs as a subword of τ, and $\mathsf{freq}_\sigma(\tau) := \mathsf{occ}_\sigma(\tau)/|\tau|$. Then $X \in 2^\omega$ is *normal* if for every $\sigma \in 2^{<\omega}$,

$$\lim_{n\to\infty} \mathsf{freq}_\sigma(X{\restriction}n) = 2^{-|\sigma|}.$$

[15]This remarkable result is one of the earlier results in the theory of algorithmic randomness, proved by V.N. Agafonov in [Aga68]. One might even hold that this is the first result in the project of reverse randomness: a notion of algorithmic randomness, the collection of sequences invariant (in von Mises' sense) under selections computable by finite state automata, is necessary and sufficient for a notion of a.e.-typicality occurring in classical mathematics, namely, normality.

[16]A sequence $X \in 2^\omega$ is weakly 1-generic if and only if for every X meets every dense Σ^0_1 subset

such the Law of Large Numbers. Kurtz randomness presents an odd case, since it can be seen as the melding of two notions of typicality discussed above, typicality in the sense of Lebesgue measure and typicality in the sense of Baire category.

But we should exercise further caution. Including normality as a necessary condition for any admissible notion of a.e.-typicality is only a viable option if we're interested in formalizing *unbiased* randomness (so that we consider our sequences as produced by the repeated tosses of an *unbiased* coin), as there are many notions of *biased* random sequences, some much more well-behaved than others, but most of which are incompatible with normality.[17] These are still measure-theoretic notions of typicality; they just aren't given in terms of the Lebesgue measure.

For the purposes of studying the calibrative role of randomness, one need not precisely delimit the notions of a.e.-typicality that are admissible for the purposes of calibrating the amount of randomness necessary and sufficient for such typicality to be instantiated. We've already seen a number of notions of typicality that aren't admissible (notions that count every sequence at typical, notions given in terms of Baire category) and a number of notions of typicality that are (each of which is a measure-theoretic notion of typicality).

The mathematician is not left without direction as to how to study this calibrative role; he only needs to grab an analysis textbook or a treatise on probability to find notions of a.e.-typicality that can be calibrated by various notions of random-

of 2^ω. Since the complement of every Π^0_1 subset of 2^ω is a dense and Σ^0_1, it follows that every weakly 1-generic sequence is Kurtz random.

[17]For more on notions of randomness with respect to biased measures, see the mathematical portion of this dissertation.

ness. What's more, the examples of this phenomenon that we already have are not jeopardized by gerrymandered notions of typical; Martin-Löf randomness, Schnorr randomness, and other definitions really do appear to capture significant truths about randomness.

12.3.2 Objection Two: The Artificiality Problem

The second objection that we will consider here is that these notions of randomness, and the associated notions of a.e.-typicality, are ad hoc, artificial notions and as such, they are not relevant to understanding the notion of randomness.

An objection along these lines is provided by Michiel van Lambalgen in his dissertation, "Random Sequences". Articulating a view consonant with my own regarding the possibility of a correct definition of randomness, he writes,

> As regards the interpretation of statistical tests, the very generality of Martin-Löf's definition presents a problem. There is a glaring contrast between the careful, piecemeal discussion of statistical tests in the literature [...] and Martin-Löf's sweeping generalisation. It seems to me that there is no use in trying to establish once and for all all properties of random sequences if we cannot survey this totality and if there are no general arguments for the choice of a particular class of properties. In this case, these arguments would have to be supplied by recursion theory. Now the prospects for such general arguments look bleak: without too much effort we could devise several alternatives to the definitions proposed by Martin-Löf and Schnorr ([Lam87], p. 92).

While I mostly agree with van Lambalgen here,[18] this is the extent of our agreement

[18] One point of disagreement: it's unclear to me that van Lambalgen's supposed ability to devise alternatives to Martin-Löf randomness and Schnorr randomness has much of a bearing on the possibility of identifying the correct definition of randomness. That is, it's not the mere presence of multiple definitions of randomness that should be troubling to the MLR-advocate (or the $\mathscr{D}$-advocate for any definition $\mathscr{D}$). Thus I take my argument to be an improvement of van Lambalgen's insofar

on the matter. For van Lambalgen continues,

> If these general arguments do not exist, the use of recursion theory may be rather inessential here. After the discovery of a statistical law which should be true of random sequences, we may determine its recursion theoretic structure; but this structure seems to be rather accidental. *It is open to doubt whether there really exists such an intimate connection between randomness and recursion theory* ([Lam87], p. 92, emphasis added).

If there really is no intimate connection between randomness and recursion theory, then to study the notions of a.e.-typicality captured out by the different definitions of randomness may be misguided, if not foolhardy. That is, without such an intimate connection, there's no reason to think that the definitions of randomness pick out classes sequences that are of interest to anyone but the computability theorists who were interested in these definitions to begin with. Let's call this problem the *artificiality problem*.

In response to this artificiality problem, I think it's reasonable to hold that the notions of a.e.-typicality discussed in this chapter are *not* artificial, as they result from restricting certain theorems from classical mathematics to an effective subset of the relevant set of objections. Moreover, it's not as if we considered restricted versions of these theorems and only then did we cook up definitions of randomness that are necessary and sufficient for the notion of typicality referenced in these theorems. That our definitions prove to be equivalent to these naturally-occurring notions of typicality suggests that our definitions are far from artificial, and should be of interest to mathematicians outside of the community of computability theorists (and even

as I account for why this multiplicity proves to be problematic for advocates of a given extensional adequacy thesis.

outside of the community of mathematical logicians).

In response to this, the "artificiality objector" might claim that I'm just brushing the artificiality problem under the rug, so to speak, for the restriction of the a.e. results to computable instances is itself an artificial restriction. For these definitions of randomness, he might claim, don't capture naturally occurring notions of typicality, but rather, these notions of typicality only result from artificially restricting certain almost-everywhere results from classical mathematics. Moreover, prior to the restriction, there is no notion of typicality picked out by these results. For example, as we saw, a point $x \in [0,1]$ is computably random if and only if x is a point of differentiability of every computable, non-decreasing real-valued function. But in the unrestricted setting, for every $x \in [0,1]$, we can always find a non-decreasing real-valued function so that f is not differentiable at x. So if we consider the entire class of non-decreasing real-valued functions, the points of differentiability for all such functions is empty.[19]

It is true that the different notions of randomness are not explicitly identifiable when we consider the almost-everywhere results in their unrestricted form. But to hold that these definitions are therefore artificial due to the fact that they are only

[19]The "artificiality objector" might further object that the restriction to computable real-valued functions is artificial because every computable real-valued function $f : \mathbb{R} \to [0,1]$ is already uniformly continuous. But it turns out that the restriction to computable real-valued functions is not the only restrict one can make. Recently, Hoyrup and Rojas have studied the class of *layerwise computable* functions, a notion defined explicitly in terms of a universal Martin-Löf tests. These functions are not necessarily uniformly continuous, but more importantly, Rojas and Hoyrup ([HR09b], [HR09a]) show that the collection of layerwise computable functions on a computable probability space is coextensive with the collection *effectively measurable* real-valued functions on that space (an effectivization of a very natural notion from analysis, that of a measurable function). Connections between randomness and other notions of computability for real-valued functions have been and continue to be explored.

become relevant once we restrict the class of properties that hold almost everywhere to some nicely definable class is to dismiss as artificial many of the insights provided by the study of the effective content of classical mathematics over the last sixty years.[20] For much fruitful research in computable algebra, computable model theory, computable analysis, and varieties of constructive mathematics has been carried out by considering theorems of classical mathematics, restricting the relevant objects to some nicely definable class, and studying the extent to which the theorems still hold true in these restricted settings. In many cases, in this restricted setting, additional information about the objects in question is uncovered, information which likely would not have been uncovered in the unrestricted setting.[21]

I submit that the notions of typicality associated with almost-everywhere theorems, which are only apparent when we consider restricted versions of these theorems, should be counted among the additional information that is uncovered by restricting to the effective setting. Taken in the context of this larger project of unearthing the effective content of classical mathematical theorems, this seems to be far from artificial. In fact, Kolmogorov seemed to anticipate this very role, writing,

> The notions of [algorithmic randomness] in their application to infinite sequences make possible some very interesting research that, although it is not necessary from the point of view of the foundations of probability, may have a certain significance in the study of the algorithmic aspect of mathematics as a whole ([Kol83], p. 217).

This is precisely what I take the calibrative role to deliver for us: to help us better understand the effective content of almost-everywhere theorems from classical

[20] For a brief survey of the history of work in effective mathematics, see [Har98], pp. 5-7.

[21] For examples of this phenomenon, see the many selections in [EGN$^+$98a] and [EGN$^+$98b].

mathematics. In this respect, the definitions are far from artificial.

12.4 The Limitative Role

The second role of randomness that is successfully filled by multiple definitions of randomness is the limitative role of randomness. To understand this role, we need to consider an approach to defining randomness suggested by John Myhill in a letter he wrote to Arthur Kruse in 1963.[22]

In his letter, Myhill suggests an intensional approach to randomness, one "not definable in usual mathematical terms" (Kruse p. 321). Towards this end, Myhill provides the following axioms of the notion of randomness, where λ is the Lebesgue measure on 2^ω and $\mathcal{R}(X)$ is a predicate on 2^ω with the intended interpretation "X is random":

($\mathcal{R}_1$) $\lambda(\{X : \mathcal{R}(X)\} = 1$;

($\mathcal{R}_2$) If $\lambda(\{X : \Phi(X)\} = 1$ then $(\forall X)(\mathcal{R}(X) \to \Phi(X))$, where Φ is a parameter-free formula in the language of set theory with one free variable X.

Moreover, Myhill allows the formulas Φ in ($\mathcal{R}_2$) to contain the predicate $\mathcal{R}$, he writes the cryptic remark, "The 'circularity' of the schema above with [$\mathcal{R}$] allowed to appear in Φ is quite justified if we are convinced that [$\mathcal{R}$] belongs to a new order of ideas, entirely outside the set-theoretic order" ([Kru67], p. 321).[23]

[22]Kruse provides the text of Myhill's letter in his article, "Some Notions of Random Sequence and Their Set-Theoretic Foundations" [Kru67].

[23]Myhill further notes that there are strong consequences to allowing $\mathcal{R}$ to occur in the formulas Φ in ($\mathcal{R}_2$): For instance, if there is some predicate $\mathcal{R}$ satisfying these two axioms, then there can

Thus, Myhill's axioms can be considered as axioms for the notion of *absolute randomness*, a notion that lies "entirely outside the set-theoretic order".[24] But how is this notion of absolute randomness related to the notions of randomness that we've considered here? My claim is that the various definitions of randomness that we've considered here, as well as many others that we haven't considered, illustrate the *indefinite contractibility of the notion of absolute randomness.* But what does it mean to say that the notion of absolute randomness is indefinitely contractible?

12.4.1 Indefinite Contractibility of the Concept of Absolute Randomness

To understand the claim that the concept of absolute randomness is indefinitely contractible, we first need to consider what it means for a concept to be indefinitely extensible, a phenomenon that has been well-studied in the philosophy of mathematics. As a first pass, let's consider the characterization of indefinitely extensible concepts provided by Michael Dummett, who writes,

> [an] indefinitely extensible concept is one such that, if we can form a definite conception of a totality all of whose members fall under the concept, we can, by reference to that totality, characterize a larger totality all of whose members fall under it ([Dum96], p. 441).

be no definable well-ordering of the continuum, and thus $V \neq L$. For if there is a definable well-ordering of the continuum, then let X be the least random sequence in this well-ordering. The set $(0,1) \setminus \{X\}$ is a definable set of measure one, and thus by $(\mathcal{R}_2)$, it contains every random sequence, contradicting the fact that X is not in this set. Concerning this result, Myhill writes, "This strikes me as very pleasant; it may have the reverse effect on you" ([Kru67], p. 321).

[24]This isn't the only approach to defining absolute randomness. For instance, one can study a notion of absolute randomness given in terms of ordinal-definable sets of measure one, a notion shown to be consistent by Solovay in [Sol70]. See also [DKUV03].

That is, the collection of objects falling under an indefinitely extensible concept cannot be formed into one single totality.

To make Dummett's characterization more precise, let us distinguish between two kinds of indefinitely extensible concepts, weakly indefinitely extensible concepts and strongly indefinitely extensible concepts.[25]

- ◦ **Weak indefinite extensibility:** Let $\mathcal{O}$ be a collection of objects, and let C be a concept such that the objects that fall under C are members of $\mathcal{O}$. The concept C is *weakly indefinitely extensible* if for every set S consisting entirely of objects falling under C, there is some $x \in \mathcal{O}$ such that (i) $x \notin S$ and (ii) x falls under C.

Note that we obtain a larger totality S' of objects falling under C simply by adding x to S. But there is one ingredient from Dummett's characterization that is missing from this definition: we don't require that the object x be obtained "by reference to the totality" S. According to Dummett, one passes from the original totality S to the object x, and hence to the larger totality S' by means of a "principle of extension", or what Russell calls a "self-productive process" (references). One way to cash this out is to require that there be a function f that maps S to some object x_S not contained in S but which falls under C. Then $S' = S \cup \{f(S)\}$ is the larger totality, all of whose members fall under C. Thus we have:

- ◦ **Strong indefinite extensibility:** Let $\mathcal{O}$ be a collection of objects, and let C be a concept such that the objects that fall under C are members of $\mathcal{O}$. The concept C is *strongly indefinitely extensible* if there is some function $f : \mathcal{P}(\mathcal{O}) \to \mathcal{O}$

[25] I've borrowed this definition from [SW06], p. 266, fn. 8.

such that for every set S consisting entirely of objects falling under C, f maps S to some $x \in \mathcal{O}$ such that (i) $x \notin S$ and (ii) x falls under C.

Let us consider several examples.[26]

The paradigm example of an indefinitely extensible concept is the concept of ordinal number. To see that the concept of ordinal number is indefinitely extensible, let S be a definite collection of ordinal numbers, and let S' be the downward closure of S under $\leq$, so that $\alpha \in S'$ if and only if there is some $\beta \in S$ such that $\alpha \leq \beta$. Clearly, S' is an ordinal. Now if we let γ be the order type of S', it follows that $S' \cup \{\gamma\}$ is an ordinal; let γ' be its order type. It follows that $\gamma' \notin S'$. Other standard examples of indefinitely extensible concepts are the concept of non-self-membered set (since by the proof of Russell's paradox, for any collection S of non-self-membered sets, we can define a non-self-membered set, namely S itself),[27] and the concept of cardinality (since given a set S of cardinal numbers, we can define a new set S' by replacing each $\kappa \in S$ with a set of size κ, and then forming the powerset of S', thus yielding a set with cardinality larger than any cardinal in S).

Now, let us consider indefinitely contractible concepts. In the introduction of this chapter, I loosely characterized the indefinite contractibility of a concept C as follows: for each definite extension $\mathcal{E}$ that purportedly contains all of the objects falling under C, there is some object $X \in \mathcal{E}$ such that there are grounds for holding that X should not be counted among the objects falling under C. Let us now attempt

[26] These examples, and others, can be found in [SW06] and [Wri10].

[27] As noted by Crispin Wright, this example also shows that the concept of set is also indefinitely contractible, since for any set S, we can consider the set of non-self-membered sets in S, which is not a member of S.

to make this more precise. As with the indefinite extensibility of a concept, we will distinguish between weak and strong forms of indefinite contractibility.

- **Weak indefinite contractibility:** Let $\mathcal{O}$ be a collection of objects, and let C be a concept such that the objects that fall under C are members of $\mathcal{O}$ (but not all members of $\mathcal{O}$ necessarily fall under C). Then the concept C is *weakly indefinitely contractible* if for every set S containing all of the objects falling under C, there is some $x \in \mathcal{O}$ such that (i) $x \in S$ and (ii) x does not fall under C.

Just as a concept is weakly indefinitely extensible but not strongly indefinitely extensible in the absence of a principle of extension, which allows one to define a new totality in terms of the initial totality, a weakly indefinitely contractible concept need not have a principle of *contraction*. Such a principle allows one to pass from a totality containing all of the objects falling under C to a sub-totality containing objects falling under C. As with principles of extension, we will think of a principle of contraction as being in terms of a function:

- **Strong indefinite contractibility:** Let $\mathcal{O}$ be a collection of objects, and let C be a concept such that the objects that fall under C are members of $\mathcal{O}$ (but not all members of $\mathcal{O}$ necessarily fall under C). Then the concept C is *strongly indefinitely contractible* if there is some function $f : \mathcal{P}(\mathcal{O}) \to \mathcal{O}$ such that for every set S containing all of the object falling under C, f maps S to some $x \in \mathcal{O}$ such that (i) $x \in S$ and (ii) x does not fall under C.

The salient difference between indefinitely extensible concepts and indefinitely contractible sequences is this: if we are trying to define a totality of all and only those objects that fall under a concept C, whereas if C is indefinitely extensible, then we will be systematically prevented from defining a definite totality containing *all* of the objects falling under C, if C is indefinitely contractible, then we will be systematically prevented from defining a definite totality containing *only* the objects falling under that concept.

Now, why should we think that the concept of absolute randomness is indefinitely contractible? Let $\mathscr{D}$ be a definition of randomness such that $\mathsf{ext}(\mathscr{D})$ is a definite collection of sequences that purportedly contains all of the absolutely random sequences. If we were to show that there is some set $\mathcal{S} \subseteq 2^\omega$ of measure zero, definable in a parameter-free way in some language that is interpretable in ZFC, such that X is contained in $\mathcal{S}$, we would thereby establish that X not absolutely random. Further, if we could show this for every set of measure one that purportedly contains all absolutely random sequences, we would thereby show that the concept of absolute randomness is weakly indefinitely contractible. Still further, if we could show that for every $\mathcal{E} \subseteq 2^\omega$ of measure one, one can define a sequence $X \notin \mathcal{E}$ *in terms of* $\mathcal{E}$, we would thereby establish the strong indefinite contractibility of the concept of randomness.

Unfortunately, it is well beyond the scope of this project to address the general case for an arbitrary set of measure one. However, we can still consider a sufficiently general case. Before we do so, let us first consider those measure one sets that have been discussed in previous chapters, namely those measure one sets that are the extensions of the various definitions of randomness.

12.4.2 Indefinite Contractibility and Definitions of Algorithmic Randomness

Recall that in the course of laying out the Justificatory Challenge, I stated that for every definition $\mathscr{D}$ of algorithmic randomness, there is some non-equivalent definition $\mathscr{D}_1$ and a parameter-free formula Θ of second-order arithmetic such that

(i) $\mathsf{ext}(\mathscr{D}_1) \subsetneq \mathsf{ext}(\mathscr{D})$,

(ii) $\Theta(2^\omega) := \{X \in 2^\omega : \Theta(X)\}$ is a null set,

(iii) $\Theta(X)$ holds for some $X \in \mathsf{ext}(\mathscr{D})$, and

(iv) $\neg\Theta(X)$ holds for *every* $X \in \mathsf{ext}(\mathscr{D}_1)$.

Observe that by the definition of absolute randomness given in the introductory remarks of this chapter, it follows that the sequence $X \in \mathsf{ext}(\mathscr{D})$ such that $\Theta(X)$ holds is not absolutely random. Although we obtain a contraction of the notion of absolute randomness simply by considering $\mathsf{ext}(\mathscr{D}) \setminus \{X\}$, in many cases, the formula Θ does not merely hold of a single sequence, but potentially infinitely many such sequences (even uncountably many in some cases). Thus, by (iv), $\mathscr{D}_1$ yields a contraction of the notion of absolute randomness that is incompatible with the property defined by Θ; that is, no $\mathscr{D}_1$-random sequence satisfies Θ.

But what exactly is this formula Θ? The answer to this question is not straightforward, as it depends, in part, on which definition $\mathscr{D}$ is under consideration. But here is the important point: for each such formula Θ that one encounters in the literature on algorithmic randomness, one can show that there is a $\mathscr{D}$-random sequence

X such that $\Theta(X)$ holds either by explicitly or indirectly construction. I refer to this phenomenon as the *systematic generation of unruly instances*.

To understand how unruly instances can be systematically generated, let's first consider how one might explicitly construct an algorithmically random sequence. The key fact here is that each definition $\mathscr{D}$ of algorithmic randomness is such that, in order for a sequence X to be counted as $\mathscr{D}$-random, X must satisfy a countable collection of requirements, where these requirements can often be specified computably or given as a collection of computably enumerable conditions. Thus, by means of an elaborate construction (sometimes using the priority method, but often using techniques specific to the given requirements), one can explicitly build a sequence X satisfying the collection of requirements, thereby guaranteeing that X is counted as $\mathscr{D}$-random.[28] Moreover, in many cases, one can satisfy the collection of requirements while simultaneously satisfying a formula Θ that has the properties described above.[29]

Unruly instances can also be produced by an indirect construction, where an indirect of a random sequence satisfying a given putative disqualifying property involves applying a general procedure to a specific collection of random sequences, with the result of this procedure being a single sequence satisfying the property in question. Allow me to illustrate this with several examples.

[28] This is precisely the approach that Wald took to prove the consistency of collectives and that certain collectives are constructively definable, as discussed in Chapter 8.

[29] One such technique is discussed at length in Section 7.4 of [Nie09], entitled "How to build a computably random set". In addition, Ville's Theorem is proved in a similar way: one builds a sequence that is invariant under a countable collection of place selections while simultaneously ensuring that it fails to satisfy the Law of the Iterated Logarithm.

In Chapter 9, I mentioned that Martin-Löf proved the existence of a universal Martin-Löf test, a uniformly computable collection $\{\widehat{\mathcal{U}}\}_{i\in\omega}$ of effectively open subsets of 2^ω such that $X \in 2^\omega$ is Martin-Löf random if and only if $X \notin \bigcap_{i\in\omega}\widehat{\mathcal{U}}_i$. Consequently, every sequence in the complement of one of the $\widehat{\mathcal{U}}_i$'s is a Martin-Löf random sequence. Further, the complement of each $\widehat{\mathcal{U}}_i$ is a Π^0_1 class. One particularly useful fact about Π^0_1 classes is that certain *basis theorems* are true of Π^0_1 classes, where a basis theorem is a given by identifying some property Φ and showing that for every Π^0_1 class $\mathcal{P}$, there is some $X \in \mathcal{P}$ such that $\Phi(X)$ holds. For instance, we have:

The Low Basis Theorem: Every Π^0_1 class contains a sequence of low Turing degree.[30]

The Hyperimmune-Free Basis Theorem: Every Π^0_1 class contains a sequence has hyperimmune-free Turing degree.[31]

The Kreisel Basis Theorem: Every Π^0_1 class contains a sequence of c.e. Turing degree.[32]

Now since there is a Π^0_1 class that consists entirely of Martin-Löf random sequences, it follows that there is

- a low Martin-Löf random sequence,

[30]A sequence A has low Turing degree if the halting problem relative to A is Turing equivalent to the halting problem: $A' \equiv_T \emptyset'$. We also say that such sequences are low.

[31]A sequence A has hyperimmune-free Turing degree if every function computable from A is dominated by a computable function.

[32]A sequence A has c.e. degree if there is some c.e. set W such that $A \equiv_T W$.

- a Martin-Löf random sequence of hyperimmune-free degree, and

- a Martin-Löf random sequence of c.e. degree.

In each of these cases, the property in question only holds for a measure zero collection of sequences, and thus it follows that such sequences are not absolutely random.

But what's more interesting for our purposes is that each of these properties arises by means of a forcing construction using Π^0_1 subclasses of the original Π^0_1 we started with.[33] That is, we apply a certain general procedure (given by the forcing construction) to a specific collection of random sequences (here a Π^0_1 class of random sequences), and the result is an individual sequence that satisfies a putative disqualifying property.

Thus far, we've seen behavior that is consistent with weak contractibility of the concept of absolute randomness. But it's not clear that there is a uniform procedure for producing unruly instances that is applicable for all currently available definitions of randomness. That is, there appears to be no general principle of contraction, but rather a variety of ways of contracting the extension of a given definition of randomness.

This fact notwithstanding, there are instances in which a uniform procedure is available for an interesting subcollection of definitions. One particularly interesting example is provided by Chaitin's Ω, discussed in Section 10.5.1. As discussed there, one proves that Ω is incompressible by means of a diagonal construction: every possible witness to the compressibility of Ω is thwarted by a clever use of the Recursion Theorem. But this result relativizes nicely: For every $A \in 2^\omega$, we can construct a

[33]For details, see [Nie09], pp. 57-60 or [DH10] pp. 77-82.

version of Chaitin's Ω that is relative to A, denoted Ω^A, and by the same proof that Ω is incompressible, we can show that Ω^A is incompressible by Turing machines that come equipped with A as an oracle.

Here's the relevance of this result: If we start with the collection of Martin-Löf randomness, which can be characterized as the collection of sequences that are not compressible by a universal prefix-free Turing machine U, then we can define Ω in terms of U:

$$\Omega = \sum_{U(\sigma)\downarrow} 2^{-|\sigma|}.$$

Further, $\Omega \in \mathsf{MLR}$, as shown by the proof via diagonalization that Ω is incompressible. But Ω is not absolutely random, since, for instance, it satisfies the property of being Turing equivalent to the halting problem $\emptyset'$, a property only satisfied by countably many sequences. Thus, to rule out all such sequences, one natural option is to consider the collection of sequences that are Martin-Löf random relative to $\emptyset'$, i.e. the collection of 2-random sequences. It follows that $\Omega \notin \mathsf{2MLR}$, and hence we've contracted the notion of absolute randomness. However, $\mathsf{2MLR}$ can be characterized as those sequences that are not compressible by a universal prefix-free Turing machine equipped with $\emptyset'$ as an oracle. Again, we define

$$\Omega^{\emptyset'} = \sum_{U^{\emptyset'}(\sigma)\downarrow} 2^{-|\sigma|},$$

and it follows that $\Omega^{\emptyset'} \in \mathsf{2MLR}$. Thus we can contract the class of sequences, relativizing our definition to $\emptyset''$, thus yielding $\mathsf{3MLR}$. But $\Omega^{\emptyset''} \in \mathsf{3MLR}$, and so we can further contract the collection of sequences.[34] Thus, for every n, we can carry out this procedure for nMLR, thus repeatedly contracting definitions of randomness that are defined along levels of the arithmetical hierarchy.[35]

---

[34]Here are two ways to carry this out formally. For each n, we have $\mathsf{nMLR} \cap \Delta^0_{n+1} \neq \emptyset$ but $(\mathsf{n}+1)\mathsf{MLR} \cap \Delta^0_{n+1} = \emptyset$, and thus it follows that the Δ^0_{n+1} sequences is a parameter-free collection of sequences of measure zero, and thus contains no absolutely random sequences. However, we can form by iteration the sequence $\Omega, \Omega^{\Omega}, \Omega^{\Omega^{\Omega}}, \ldots$, we have:

$$\begin{gathered} \Omega \in \Delta^0_2 \\ \Omega^{\Omega} \in \Delta^0_3 \\ \Omega^{\Omega^{\Omega}} \in \Delta^0_4 \\ \vdots \\ \Omega^{\Omega^{\cdot^{\cdot^{\cdot^{\Omega}}}}} \in \Delta^0_{n+1} \\ \vdots \end{gathered}$$

Similarly, for each n, we have

$$\mathsf{nMLR} \cap \{X : X \equiv_T \emptyset^{(n)}\} \neq \emptyset$$

but

$$(\mathsf{n}+1)\mathsf{MLR} \cap \{X : X \equiv_T \emptyset^{(n)}\} = \emptyset,$$

and thus it follows that the collection of sequences such that $\{X : X \equiv_T \emptyset^{(n)}\}$ is a parameter-free collection of sequences of measure zero, and thus contains no absolutely random sequences. Then using the facts that $\Omega^A \oplus A \equiv_T A'$ and $\Omega^A \oplus A \in \mathsf{MLR}$ if and only if $A \in \mathsf{MLR}$, we can form by iteration the sequence $\Omega, \Omega^{\Omega} \oplus \Omega, \Omega^{(\Omega^{\Omega} \oplus \Omega)} \oplus (\Omega^{\Omega} \oplus \Omega), \ldots$, and thus, if we set

$$\begin{gathered} S_0 := \Omega \\ S_{n+1} := \Omega^{S_n} \oplus S_n \end{gathered}$$

it follows that $S_n \equiv_T \emptyset^{(n+1)}$ for every $n \in \omega$. Thus, we have two cases of the uniform generation of unruly instances for a collection of putative disqualifying properties (being Δ^0_n for some $n \in \omega$ or being Turing-equivalent to $\emptyset^{(n)}$ for some $n \in \omega$).

[35]Note that we have another difference between the definition of Turing computability and the definition of Martin-Löf randomness. Whereas the collection of Turing computable functions is not diagonalizable (as discussed in Section 11.4.1), there is a sense in which the collection of Martin-Löf random sequences *is* diagonalizable. Recall that in Section 10.5.1 I outlined the proof that Chaitin's Ω is incompressible, which proceeds roughly as follows: every possible witness to the compressibility of an initial segment of Ω is eventually thwarted, using the Recursion Theorem in tandem with the fact that Ω encodes the information about all prefix-free Turing machines. Thus, the incompressibility of Ω is secured by diagonalizing against all possible computations that witness the compressibility of Ω. Thus, by means of this diagonalization, Ω is included among the random sequences.

This behavior, namely that we can contract measure one sets defined at various levels of the arithmetical hierarchy, suggests that this contractibility can be uniformly carried out for a collection of definitions of randomness defined in terms of resources given by the levels of some hierarchy.

12.4.3 A More General Approach to Indefinite Contractibility

Suppose we consider a family of definition given in terms of some hierarchy that occurs in mathematical logic, such as one of the various subrecursive hierarchies, the arithmetical hierarchy, the hyperarithmetical hierarchy, the constructible hierarchy, and so on. Let us consider a family of definitions of randomness with a fixed motif and criterion of success, but vary the resources along one of these hierarchies. Suppose, for instance, that our hierarchy $\mathscr{H}$ is indexed by some set I on which we have a natural ordering $\leq$. Then we can write $\mathscr{H} = \{\mathscr{H}_\alpha\}_{\alpha \in I}$, where $\mathscr{H}_\alpha$ is the αth level of $\mathscr{H}$. Further, we can define a family of definitions $\{\mathscr{D}_\alpha\}_{\alpha \in I}$, where for each $\alpha \in I$, $\mathscr{D}_\alpha$ uses as its resources all objects in $\mathscr{H}$. Then one will be able to prove the following five conditions, which I call the *contractibility conditions*:

(C1) For each $\alpha, \beta \in I$, $\alpha \leq \beta$ implies that $\mathsf{ext}(\mathscr{D}_\beta) \subseteq \mathsf{ext}(\mathscr{D}_\alpha)$.

(C2) For each $\alpha, \beta \in I$, $\alpha \leq \beta$ implies that $\lambda(\mathsf{ext}(\mathscr{D}_\alpha) \setminus \mathsf{ext}(\mathscr{D}_\beta)) = 0$.

(C3) For each $\alpha \in I$, no $\mathscr{D}_\alpha$-random sequence is definable in $\mathscr{H}_\alpha$.

(C4) For each $\beta \in I$ and for each $\alpha < \beta$, there is some $\mathscr{D}_\alpha$-random sequence that is definable in $\mathscr{H}_\beta$.

(C5) For each $\alpha, \beta \in I$, $\alpha \leq \beta$ implies that the set $\{X : X \in \mathsf{ext}(\mathscr{D}_\alpha)\}$ is definable with parameters from $\mathscr{H}_\beta$.

From these five conditions, it follows that we can contract the extensions of the various definitions defined at the various levels of the hierarchy $\mathscr{H}$: Given $\alpha \in I$, if $\mathsf{ext}(\mathscr{D}_\alpha)$ purportedly contains all absolutely random sequences, then for any $\beta > \alpha$, by (C4), there is some $\mathscr{D}_\alpha$-random sequence X that is definable in $\mathscr{H}_\beta$. However, by (C3), X cannot be $\mathscr{D}_\beta$-random, and hence $X \in \mathsf{ext}(\mathscr{D}_\alpha) \setminus \mathsf{ext}(\mathscr{D}_\beta)$. By (C5), the set $\mathcal{S} = \{X : X \in \mathsf{ext}(\mathscr{D}_\alpha) \setminus \mathsf{ext}(\mathscr{D}_\beta)\}$ is definable in $\mathscr{H}_\beta$, but by (C2), $\lambda(\mathcal{S}) = 0$. Thus it follows that X is not absolutely random. Lastly, by (C1) and the fact that X is not $\mathscr{D}_\beta$-random, we have $\mathsf{ext}(\mathscr{D}_\beta) \subseteq (\mathsf{ext}(\mathscr{D}_\alpha) \setminus \{X\})$, yielding the desired contraction.

Thus we have a plethora of examples illustrating the indefinite contractibility of the concept of absolute randomness. Moreover, we have seen that any family of definitions of randomness defined in terms of resources from some hierarchy that satisfies the contractibility conditions (C1)-(C5) above will thus be contractible.[36]

Summing up, we have seen that the various definitions of randomness fill the limitative role by showing the extent to which every definition of randomness with a definite, well-defined extension will inevitably count as random a sequence that is ruled out as non-random by stronger definitions of randomness. This phenomenon is not well-understood; there is much formal work to be done to better understand the limitative role. That is, we can formally study the systematical generation of un-

[36] There is one last possibility to consider: we let $\mathscr{D}$-randomness be given by the intersection of all of the extensions of the definitions $\mathscr{D}_\alpha$ given in terms of $\mathscr{H} = \{\mathscr{H}_\alpha\}_{\alpha \in I}$. The problem is that if $\mathscr{H}$ is some hierarchy that is definable in the von Neumann hierarchy, we can find stronger definitions of randomness in terms of the von Neumann hierarchy. Further, if $\mathscr{H}$ is the von Neumann hierarchy, then the intersection of the extensions of the definitions $\mathscr{D}_\alpha$ will not be a set, and thus the resulting definition will not be yield a well-defined, definite extension of random sequences.

ruly instances, classifying the various ways in which one can produce such instances. Moreover, the contractibility of a family of definitions the resources of which are given by some hierarchy of resources has not been systematically studied. For instance, the details of this phenomenon in various hierarchies, from the subrecursive up through the set-theoretic, merits careful attention. The hope is that in studying this phenomenon, we might better understand the very hierarchies in terms of which we can define randomness.

12.5 Towards the No-Thesis Thesis

Let us conclude this chapter by returning to the No-Thesis Thesis. In the previous chapter, we determined that the $\mathscr{D}$-advocate must meet the Justificatory Challenge: she must provide a sharpening of the prevailing intuitive conception of randomness that is precise enough to block the claims of extensional adequacy made concerning alternative definitions of randomness without undermining the claim of the extensional adequacy of $\mathscr{D}$. Further, I argued that there is no reason to hold that the $\mathscr{D}$-advocate can meet this challenge. In light of the data we've considered in this chapter, I claim that we now have reason to hold that the $\mathscr{D}$-advocate *cannot* meet this challenge.

Recall that the $\mathscr{D}$-advocate's goal is to establish the correctness of $\mathscr{D}$, that a sequence is $\mathscr{D}$-random if and only if it is intuitively random. If she were to establish this, then we would be able to replace uses of "intuitively random sequence" with "$\mathscr{D}$-random sequence", just as the CTT permits us to replace uses of "effectively calculable number-theoretic function" with "Turing computable number-theoretic function". However, one problem with this approach is that the collection of intuitively random sequences does not appear to be a well-defined, definite collection of sequences, or at least the $\mathscr{D}$-advocate hasn't given us any reason to think that this is the case. To further compound matters, in our discussion of the calibrative role, we have seen that there are multiple definitions of randomness $\mathscr{D}_1, \mathscr{D}_2, \ldots$ such that, for each $\mathscr{D}_i$, there are certain contexts in which we are justified in replacing uses of the informal predicate "X is random" with the formal predicate "X is $\mathscr{D}_i$-random". There is no single definition that can capture each of these notions of a.e.-typicality. Thus, it appears that no single definition can capture all of what mathematicians have considered to be significant truths concerning the concept of randomness.

In addition, the fact that the various definitions of randomness all successfully fill the limitative role of randomness further confirms the view that no single definition of randomness is adequate for all purposes for which we might develop a definition of randomness that yields a well-defined, definite collection of sequences. In particular, each definition of randomness that we've considered, and in fact, *any* definition $\mathscr{D}$ that is part of a family of definitions satisfying the contractibility conditions, can be "contracted": we will also be able to find a putative disqualifying property $\mathcal{P}$ that is satisfied by some $\mathscr{D}$-random sequence, and a stronger definition $\mathscr{D}'$ such that no $\mathscr{D}'$-random sequence has $\mathcal{P}$.

In spite of this evidence, one might continue to hold on to the hope that we might one day establish some single definition of randomness as the *correct* one. However, I think we have good reason to accept the alternative account that I have offered here. There are specific contexts in which one definition of randomness $\mathscr{D}$

is appropriate for the task at hand, while all others are disqualified. But there are contexts in which $\mathscr{D}$ is *not* the appropriate definition for the task at hand; rather, some other, non-equivalent definition $\mathscr{D}'$ is the appropriate one. Further, on this approach, we can bring our formal tools to bear on the analysis of the ways in which one definition is appropriate in certain contexts (by studying how that definition fills the calibrative role) and not appropriate for others (by studying how that definition fills the limitative role). Thus on this approach, it is the multiplicity of definitions of randomness, and not one single definition of randomness, that captures much of what mathematicians have taken to be significant about the notion of randomness.

BIBLIOGRAPHY

[Aga68] V. N. Agafonov. Normal sequences and finite automata. *Dokl. Akad. Nauk SSSR*, 179:255–256, 1968.

[Bag53] F. Bagemihl. Book Review: Les nombres inaccessibles. *Bull. Amer. Math. Soc.*, 59(4):406–409, 1953.

[BBJ02] George S. Boolos, John P. Burgess, and Richard C. Jeffrey. *Computability and logic*. Cambridge University Press, Cambridge, fourth edition, 2002.

[BDH+11] Laurent Bienvenu, Adam Day, Mathieu Hoyrup, Ilya Mezhirov, and Alexander Shen. A constructive version of Birkhoff's ergodic theorem for Martin-Löf random points. *Information and Control*, 2011. forthcoming.

[BDN11] George Barmpalias, Rod Downey, and Keng Meng Ng. Jump inversions inside effectively closed sets and applications to randomness. *J. Symbolic Logic*, 76(2):491–518, 2011.

[BHKM10] Laurent Bienvenu, Rupert Hölzl, Thorsten Kräling, and Wolfgang Merkle. Separations of non-monotonic randomness notions. *Journal of Logic and Computation*, 2010.

[Bie08] Laurent Bienvenu. *Game-theoretic characterizations of randomness: unpredictability and stochasticity*. PhD thesis, Université de Provence, 2008.

[Bin03] Stephen Binns. *The Medvedev and Muchnik lattices of Pi(0/1) classes*. ProQuest LLC, Ann Arbor, MI, 2003. Thesis (Ph.D.)–The Pennsylvania State University.

[Bin08] Stephen Binns. Π^0_1 classes with complex elements. *J. Symbolic Logic*, 73(4):1341–1353, 2008.

[BM09] Laurent Bienvenu and Wolfgang Merkle. Constructive equivalence relations on computable probability measures. *Ann. Pure Appl. Logic*, 160(3):238–254, 2009.

[BMN11] Vasco Brattka, Joseph S. Miller, and André Nies. Randomness and differentiability. *Preprint*, 2011.

[Bor52] Émile Borel. *Les nombres inaccessibles. Avec une note de Daniel Dugué*. Gauthier-Villars, Paris, 1952.

[Bur98] Frank Burk. *Lebesgue measure and integration: an introduction*. Wiley-IEEE, 1998.

[Car50] Rudolf Carnap. *Logical Foundations of Probability*. University of Chicago Press, 1950.

[Cha75] Gregory J. Chaitin. A theory of program size formally identical to information theory. *J. Assoc. Comput. Mach.*, 22:329–340, 1975.

[CHKW01] Cristian S. Calude, Peter H. Hertling, Bakhadyr Khoussainov, and Yongge Wang. Recursively enumerable reals and Chaitin Ω numbers. *Theoret. Comput. Sci.*, 255(1-2):125–149, 2001.

[Chu40] Alonzo Church. On the concept of a random sequence. *Bull. Amer. Math. Soc.*, 46:130–135, 1940.

[Chu66a] Alonzo Church. Letter to Hilda Geiringer von Mises, 11 june 1966. Alonzo Church Papers, Box 22, Folder 21; Department of Rare Books and Special Collections, Princeton University Library, June 1966.

[Chu66b] Alonzo Church. Letter to Hilda Geiringer von Mises, 16 mar. 1966. Alonzo Church Papers, Box 22, Folder 21; Department of Rare Books and Special Collections, Princeton University Library, March 1966.

[Coo04] S. Barry Cooper. *Computability theory*. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[Cop28] Arthur H. Copeland. Admissible Numbers in the Theory of Probability. *Amer. J. Math.*, 50(4):535–552, 1928.

[Cré09] Pierre Crépel. Jean Ville's recollections, in 1984 and 1985, concerning his work on martingales. *J. Électron. Hist. Probab. Stat.*, 5(1):18, 2009. Reported and with an introduction by Pierre Crépel, Translated from the French by Glenn Shafer.

[Das11] Abhijit Dasgupta. Mathematical foundations of randomness. In Prasanta S. Bandyopadhyay, Malcolm R. Forster, and Dov M. Gabbay, editors, *Handbook of the Philosophy of Statistics*, pages 641–712. North Holland, 2011.

[Day10] Adam Day. Process and truth-table characterisations of randomness. *Submitted*, 2010.

[Del11] J.-P. Delahaye. The Martin-Löf-Chaitin Thesis: The identification by recursion theory of the mathematical notion of random sequenceartin-löf-chaitin thesis: The identification by recursion theory of the mathematical notion of random sequence. In H. Zenil, editor, *Randomness Through Computation: Some Answers, More Questions*. World Scientific Pub Co Inc, 2011.

[Dem75] O. Demuth. The differentiability of constructive functions of weakly bounded variation on pseudo numbers. *Comment. Math. Univ. Carolinae*, 16(3):583–599, 1975.

[Dem88] Osvald Demuth. Remarks on the structure of tt-degrees based on constructive measure theory. *Comment. Math. Univ. Carolin.*, 29(2):233–247, 1988.

[DG04] Rodney G. Downey and Evan J. Griffiths. Schnorr randomness. *J. Symbolic Logic*, 69(2):533–554, 2004.

[DH10] Rodney G. Downey and Denis R. Hirschfeldt. *Algorithmic randomness and complexity*. Theory and Applications of Computability. Springer, New York, 2010.

[DHMN05] Rod Downey, Denis R. Hirschfeldt, Joseph S. Miller, and André Nies. Relativizing Chaitin's halting probability. *J. Math. Log.*, 5(2):167–192, 2005.

[DKUV03] Bruno Durand, Vladimir Kanovei, Vladimir A. Uspensky, and Nikolai Vereshchagin. Do stronger definitions of randomness exist? *Theoret. Comput. Sci.*, 290(3):1987–1996, 2003.

[DNWY06] Rod Downey, Andre Nies, Rebecca Weber, and Liang Yu. Lowness and Π^0_2 nullsets. *J. Symbolic Logic*, 71(3):1044–1052, 2006.

[Doo36] J. L. Doob. Note on probability. *Ann. of Math. (2)*, 37(2):363–367, 1936.

[Dum96] M. Dummett. *The Seas of Language*. Oxford scholarship online. Oxford University Press, 1996.

[Eag10] Antony Eagle. Chance vs. randomness. In Edward N. Zalta, editor, *Stanford Encyclopedia of Philosophy*. 2010.

[EGN+98a] Yu. L. Ershov, S. S. Goncharov, A. Nerode, J. B. Remmel, and V. W. Marek, editors. *Handbook of recursive mathematics. Vol. 1*, volume 138 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1998. Recursive model theory.

[EGN+98b] Yu. L. Ershov, S. S. Goncharov, A. Nerode, J. B. Remmel, and V. W. Marek, editors. *Handbook of recursive mathematics. Vol. 2*, volume 139 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 1998. Recursive algebra, analysis and combinatorics.

[Fel68] William Feller. *An introduction to probability theory and its applications. Vol. I.* Third edition. John Wiley & Sons Inc., New York, 1968.

[FGMN11] Johanna N.Y. Franklin, Noam Greenberg, Joseph Miller, and Keng Meng Ng. Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. *Proceedings of the American Mathematical Society*, 2011. forthcoming.

[Fol99] Gerald B. Folland. *Real Analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons Inc., New York, second edition, 1999. Modern techniques and their applications, A Wiley-Interscience Publication.

[Fra08] Johanna Franklin. Hyperimmune-free degrees and Schnorr triviality. *Journal of Symbolic Logic*, 73:999–1008, 2008.

[Fré38] Maurice Fréchet. Exposé et discussion de quelques recherches récentes sur les fondéments du calcul des probabilités. In *Colloque consacré au calcul des probabilités*, volume 735 of *Actualités Scientifiques et Industrielles*, pages 22–55. Hermann, 1938.

[Gác86] Péter Gács. Every sequence is reducible to a random one. *Inform. and Control*, 70(2-3):186–192, 1986.

[Gei66] Hilda Geiringer. Letter to Alonzo Church, 13 apr. 1966. Alonzo Church Papers, Box 22, Folder 21; Department of Rare Books and Special Collections, Princeton University Library, April 1966.

[Göd93] K. Gödel. Undecidable diophantine propositions. In *Collected Works, Volume III*. Oxford University Press, 193?

[Göd46] K. Gödel. Remarks before the princeton bicentennial conference on problems in mathematics. In *Collected Works, Volume II*. Oxford University Press, 1946.

[GS82] Haim Gaifman and Marc Snir. Probabilities over rich languages, testing and randomness. *J. Symbolic Logic*, 47(3):495–548, 1982.

[Har98] Valentina S. Harizanov. Pure computable model theory. In *Handbook of recursive mathematics, Vol. 1*, volume 138 of *Stud. Logic Found. Math.*, pages 3–114. North-Holland, Amsterdam, 1998.

[Hil01] David Hilbert. Mathematische probleme. *Archiv der Mathematik und Physik*, 3(1):44–63, 213–237, 1901.

[Hoy12] Mathieu Hoyrup. The dimension of ergodic random sequences. *Preprint*, 2012.

[HR09a] Mathieu Hoyrup and Cristóbal Rojas. An application of Martin-Löf randomness to effective probability theory. In *Mathematical theory and computational practice*, volume 5635 of *Lecture Notes in Comput. Sci.*, pages 260–269. Springer, Berlin, 2009.

[HR09b] Mathieu Hoyrup and Cristóbal Rojas. Applications of effective probability theory to Martin-Löf randomness. In *Automata, languages and programming. Part I*, volume 5555 of *Lecture Notes in Comput. Sci.*, pages 549–561. Springer, Berlin, 2009.

[Kam32] Erich Kamke. *Einführung in die Wahrscheinlichkeitstheorie*. S. Hirzel, Leipzig, 1932.

[Kam33] Erich Kamke. Über neuere begründungen der wahrscheinlichkeitsrechnung. *Jahresbericht Deutsche Mathematiker Vereinigung*, 42:14–27, 1933.

[Kam73] Teturo Kamae. Subsequences of normal sequences. *Israel J. Math.*, 16:121–149, 1973.

[Kau91] Steven Kautz. *Degrees of Random Sets*. PhD thesis, Cornell University, 1991.

[KHMS11] Bjørn Kjos-Hanssen, Wolfgang Merkle, and Frank Stephan. Kolmogorov complexity and the recursion theorem. *Transactions of the American Mathematical Society*, 363(10):5465–5480, 2011.

[Kle81] Stephen C. Kleene. Origins of recursive function theory. *Ann. Hist. Comput.*, 3(1):52–67, 1981.

[Kol65] A. N. Kolmogorov. Three approaches to the definition of the notion of amount of information. *Problemy Peredači Informacii*, pages 184–193, 1965. Reprinted in Selected works of A. N. Kolmogorov. Vol. III: Information Theory and the Theory of Algorithms, Shiryayev, A.N. Kluwer Academic Publishers Group.

[Kol83] A. N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. *Uspekhi Mat. Nauk*, 38(4(232)):27–36, 1983.

[Kre71] Georg Kreisel. Some reasons for generalizing recursion theory. In *Logic Colloquium '69 (Proc. Summer School and Colloq., Manchester, 1969)*, pages 139–198. North-Holland, Amsterdam, 1971.

[Kre72] Georg Kreisel. Informal rigour and completeness proofs. In *Problems in the Philosophy of Mathematics*. North-Holland, 1972.

[Kru67] Arthur H. Kruse. Some notions of random sequence and their set-theoretic foundations. *Z. Math. Logik Grundlagen Math.*, 13:299–322, 1967.

[KS01] Antonín Kučera and Theodore A. Slaman. Randomness and recursive enumerability. *SIAM J. Comput.*, 31(1):199–211 (electronic), 2001.

[Kuč85] Antonín Kučera. Measure, Π^0_1-classes and complete extensions of PA. In *Recursion theory week (Oberwolfach, 1984)*, volume 1141 of *Lecture Notes in Math.*, pages 245–259. Springer, Berlin, 1985.

[Kur81] Stuart Kurtz. *Randomness and Genericity in the Degrees of Unsolvability*. PhD thesis, University of Illinois at Urbana-Champaign, 1981.

[Lam87] Michiel Van Lambalgen. *Random Sequences*. PhD thesis, University of Amsterdam, 1987.

[Lev73] L. A. Levin. The concept of a random sequence. *Dokl. Akad. Nauk SSSR*, 212:548–550, 1973.

[Lev10] Leonid A. Levin. Some theorems on the algorithmic approach to probability theory and information theory (1971 dissertation directed by A. N. Kolmogorov). *Ann. Pure Appl. Logic*, 162(3):224–235, 2010. Translated from the Russian original.

[Lin66] D.V. Lindley. Review: Richard von Mises, Hilda Geiringer, *Mathematical Theory of Probability and Statistics*. *Ann. Math. Statistics*, 37(3):747–754, 1966.

[LV97] Ming Li and Paul Vitányi. *An introduction to Kolmogorov complexity and its applications*. Graduate Texts in Computer Science. Springer-Verlag, New York, second edition, 1997.

[Mer03] Wolfgang Merkle. The Kolmogorov-Loveland stochastic sequences are not closed under selecting subsequences. *J. Symbolic Logic*, 68(4):1362–1376, 2003.

[Mil04] Joseph S. Miller. Every 2-random real is Kolmogorov random. *J. Symbolic Logic*, 69(3):907–913, 2004.

[Mis45] R.V. Mises. *Vorlesungen aus dem Gebiete der angewandten Mathematik*. Number v. 1 in Vorlesungen aus dem Gebiete der angewandten Mathematik. M. S. Rosenberg, 1945- , c1939, 1945.

[Mis51] R.V. Mises. *Positivism: a study in human understanding trans. from the German with the collaboration of the author*. s.n., 1951.

[Miy12] Kenshi Miyabe. Characterization of Kurtz randomness by a differentiation theorem. *Preprint*, 2012.

[ML66] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[ML69a] P. Martin-Löf. Algorithms and randomness. *Revue de l'Institut International de Statistique / Review of the International Statistical Institute*, 37(3):pp. 265–272, 1969.

[ML69b] Per Martin-Löf. The literature on von Mises' Kollektivs revisited. *Theoria*, 35:12–37, 1969.

[ML70] Per Martin-Löf. On the notion of randomness. In *Intuitionism and Proof Theory (Proc. Conf., Buffalo, N.Y., 1968)*, pages 73–78. North-Holland, Amsterdam, 1970.

[ML71] Per Martin-Löf. Complexity oscillations in infinite binary sequences. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 19:225–230, 1971.

[MM68] Webb Miller and D. A. Martin. The degrees of hyperimmune sets. *Z. Math. Logik Grundlagen Math.*, 14:159–166, 1968.

[MMN$^+$06] Wolfgang Merkle, Joseph S. Miller, André Nies, Jan Reimann, and Frank Stephan. Kolmogorov-Loveland randomness and stochasticity. *Ann. Pure Appl. Logic*, 138(1-3):183–210, 2006.

[MR06] Wolfgang Merkle and Jan Reimann. Selection functions that do not preserve normality. *Theory Comput. Syst.*, 39(5):685–697, 2006.

[MSU98] Andrei A. Muchnik, Alexei L. Semenov, and Vladimir A. Uspensky. Mathematical metaphysics of randomness. *Theoret. Comput. Sci.*, 207(2):263–317, 1998.

[Nie05] André Nies. Lowness properties and randomness. *Adv. Math.*, 197(1):274–305, 2005.

[Nie09] André Nies. *Computability and randomness*. Oxford Logic Guides. Oxford University Press, 2009.

[NST05] André Nies, Frank Stephan, and Sebastiaan A. Terwijn. Randomness, relativization and Turing degrees. *J. Symbolic Logic*, 70(2):515–535, 2005.

[OW08] Daniel Osherson and Scott Weinstein. Recognizing strong random reals. *The Review of Symbolic Logic*, 1(01):56–63, 2008.

[Pop35] Karl Popper. *Logik der Forschung*. Julius Springer Verlag, Vienna, 1935.

[Por12] Christopher P. Porter. Kolmogorov on the role of randomness in probability theory (forthcoming). *Mathematical Structures in Computer Science*, 2012.

[PRS11] Noopur Pathak, Cristóbal Rojas Rojas, and Stephen G. Simpson. Schnorr randomness and the Lebesgue differentiation theorem. *Preprint*, 2011.

[Raa00] Panu Raatikainen. Algorithmic information theory and undecidability. *Synthese*, 123(2):217–225, 2000.

[Rei32] Hans Reichenbach. Die logischen grundlagen des wahrscheinlichkeitsbegriffs. *Erkenntnis*, 3(1), 1932.

[Res07] Michael Rescorla. Church's thesis and the conceptual analysis of computability. *Notre Dame J. Formal Logic*, 48(2):253–280 (electronic), 2007.

[Rog87] Hartley Rogers, Jr. *Theory of recursive functions and effective computability.* MIT Press, Cambridge, MA, second edition, 1987.

[RS07] Jan Reimann and Theodore A. Slaman. Probability measures and effective randomness. *Preprint*, 2007.

[RS08] Jan Reimann and Theodore A. Slaman. Measures and their random reals. *Preprint*, 2008.

[Rud87] Walter Rudin. *Real and complex analysis.* McGraw-Hill Book Co., New York, third edition, 1987.

[Rut11] Jason Rute. Randomness and the Lebesgue differentation theorem. 2011.

[Sch71] C.-P. Schnorr. A unified approach to the definition of random sequences. *Math. Systems Theory*, 5:246–258, 1971.

[Sch73] C.-P. Schnorr. Process complexity and effective random tests. *J. Comput. System Sci.*, 7:376–388, 1973. Fourth Annual ACM Symposium on the Theory of Computing (Denver, Colo., 1972).

[Sch77] C.-P. Schnorr. A survey of the theory of random sequences. In *Basic problems in methodology and linguistics (Proc. Fifth Internat. Congr. Logic, Methodology and Philos. of Sci., Part III, Univ. Western Ontario,* *London, Ont., 1975)*, pages 193–211. Univ. Western Ontario Ser. Philos. Sci., Vol. 11. Reidel, Dordrecht, 1977.

[SF77] C.-P. Schnorr and P. Fuchs. General random sequences and learnable sequences. *J. Symbolic Logic*, 42(3):329–340, 1977.

[Sha09] Glenn Shafer. The education of Jean André Ville. *J. Électron. Hist. Probab. Stat.*, 5(1):50, 2009.

[Smi11] Peter Smith. Squeezing arguments. *Analysis*, 71(1):22–30, 2011.

[Soa87] Robert I. Soare. *Recursively enumerable sets and degrees*. Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1987. A study of computable functions and computably generated sets.

[Sol70] Robert M. Solovay. A model of set-theory in which every set of reals is Lebesgue measurable. *Ann. of Math. (2)*, 92:1–56, 1970.

[Sol75] R.M. Solovay. Draft of paper (or series of papers) on chaitin's work. May 1975.

[SW06] Stewart Shapiro and Crispin Wright. All things indefinitely extensible. In Agustín Rayo and Gabriel Uzquiano, editors, *Absolute Generality*. Oxford University Press, 2006.

[Tur38] A. M. Turing. On computable numbers, with an application to the Entscheidungsproblem. a correction. *Proceedings of the London Mathematical Society*, s2-43(1):544–546, 1938.

[Ven66] John Venn. *The logic of chance: an essay on the foundations and province of the theory of probability, with especial reference to its logical bearings and its application to moral and social science, and to statistics*. Macmillan and Co, London and Cambridge, 1866.

[Vil39] Jean Ville. *Étude critique de la notion de collectif*. Gauthier-Villars, 1939.

[vL90] Michiel van Lambalgen. The axiomatization of randomness. *J. Symbolic Logic*, 55(3):1143–1167, 1990.

[vM19] Richard von Mises. Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.*, 5(1-2):52–99, 1919.

[vM31] Richard von Mises. *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*. Mary S. Rosenberg, 1931.

[vM38] Richard von Mises. Quelques rémarques sur les fondements du calcul des probabilités. In *Colloque consacré au calcul des probabilités*, number 735 in Actualités Scientifiques et Industrielles, pages 57–66. Hermann, 1938.

[vM41] R. von Mises. On the foundations of probability and statistics. *Ann. Math. Statistics*, 12:191–205, 1941.

[vM46] Richard von Mises. Notes on mathematical theory of probability and statistics. Technical report, Harvard University, 1946.

[vM64] Richard von Mises. *Mathematical theory of probability and statistics*. Edited and Complemented by Hilda Geiringer. Academic Press, New York, 1964.

[vM81] Richard von Mises. *Probability, statistics and truth*. Dover Publications Inc., New York, english edition, 1981.

[Wal37] Abraham Wald. Die widerspruchsfreiheit des kollektivbegriffes der warhrscheinlichkeitsrechnung. *Ergebnisse eines Mathematischen Kolloquiums*, 8, 1937.

[Wal38] Abraham Wald. Die widerspruchsfreiheit des kollektivbegriffes. In *Colloque consacré au calcul des probabilités*, volume 735 of *Actualités Scientifiques et Industrielles*, pages 79–99. Hermann, 1938.

[Wan96] Yongge Wang. *Randomness and Complexity*. PhD thesis, University of Heidelberg, 1996.

[Wri10] Crispin Wright. Whence the paradox? Axiom v and indefinite extensibility. For 4th International Lauener Symposium on Analytical Philosophy, Bern, May 27 2010, In Honour of Professor Sir Michael Dummett, 2010.

[ZL70] A. K. Zvonkin and L. A. Levin. The complexity of finite objects and the basing of the concepts of information and randomness on the theory of algorithms. *Uspehi Mat. Nauk*, 25(6(156)):85–127, 1970.

This document was prepared & typeset with pdfLaTeX*, and formatted with* NDdiss2$_\varepsilon$ *classfile (v3.0[2005/07/27]) provided by Sameer Vijay.*