ON ANALOGUES OF THE CHURCH-TURING THESIS IN ALGORITHMIC RANDOMNESS

CHRISTOPHER P. PORTER

Abstract. In this article, I consider the status of several statements analogous to the Church-Turing thesis that assert that some definition of algorithmic randomness captures the intuitive conception of randomness. I argue that we should not only reject the theses that have appeared in the algorithmic randomness literature, but more generally that we ought not evaluate the adequacy of a definition of randomness on the basis of whether it captures the so-called intuitive conception of randomness to begin with. Instead, I argue that a more promising alternative is to evaluate the adequacy of a definition of randomness on the basis of whether it captures what I refer to as a "notion of almost everywhere typicality." In support of my main claims, I will appeal to recent work in showing the connection between definitions of algorithmic randomness and certain "almost everywhere" theorems from classical mathematics.

1. Introduction

Among computability theorists and mathematical logicians more generally, there is a broad consensus that the formal notion of a Turing computable number-theoretic function provides an adequate formalization of the intuitive, pre-theoretic notion of an effectively calculable number-theoretic function. This commitment is more precisely expressed as the thesis that a number-theoretic function is Turing computable if and only if it is effectively calculable, a thesis commonly referred to as the *Church-Turing thesis*.[1]

One key datum that lends credence to the Church-Turing thesis is the lack of alternative definitions of computable function that seriously contend to capture the intuitive notion of effective calculability. That is, every definition that has been offered as a formalization of the intuitive notion of effectively calculable number-theoretic function has been proven to be extensionally equivalent to the definition of a Turing computable function.

But what if matters had been otherwise and there were multiple, non-equivalent definitions of computable function, each of which having a reasonable claim to capturing the intuitive, pre-theoretic notion of effectively calculable function? In such a scenario, how would our understanding of the concept of computability have differed from how it is commonly understood?[2]

[1]Properly speaking, this thesis should be referred to as *Turing's thesis*, but given the equivalence of the notion of Turing computable function and Church's notion of lambda-calculable function, it has become common to take "the Church-Turing thesis" to refer to both Church's thesis and Turing's thesis.

[2]According to some, this state of affairs obtains in the actual world. For instance, some hold that various models of hypercomputation, which transcend the so-called Turing barrier (i.e., compute functions that are not Turing computable), also capture the intuitive notion of computability. To be clear, in the discussion here I take the target notion "effectively calculable function" to be synonymous with "a function effectively calculable by a human following a finite list of instructions, carrying out a finite number of steps in any given halting computation," a notion that involves no infinitary operations. This is not to say that models of hypercomputation do not capture anything

In the theory of algorithmic randomness, we are confronted with precisely such a scenario. In particular, there are multiple, non-equivalent definitions of algorithmically random sequence (in fact, infinitely many such definitions), and moreover, several of these definitions have been claimed to capture the intuitive, pre-theoretic conception of randomness. More specifically, in the literature on algorithmic randomness, one can find multiple analogues of the Church-Turing thesis, each instantiating the following schema, which I'll henceforth refer to as the *randomness thesis schema*:

> An infinite sequence is intuitively random if and only if it is $\mathscr{D}$-random.

Here $\mathscr{D}$ is a definition of randomness for infinite sequences,[3] and the $\mathscr{D}$-random sequences are those that fall under the definition $\mathscr{D}$.

This paper is concerned with the status of three specific instances of the above randomness thesis schema, instantiated by definitions of randomness that I will henceforth refer to as *randomness thesis candidates*. The theses, which will be explained in Section §2 below, are as follows:

> **The Martin-Löf-Chaitin thesis**: An infinite sequence is intuitively random if and only if it is Martin-Löf random.[4]

> **Schnorr's thesis**: An infinite sequence is intuitively random if and only if it is Schnorr random.[5]

> **The Weak 2-Randomness thesis**: An infinite sequence is intuitively random if and only if it is weakly 2-random.[6]

Setting aside for the moment the details of these formal notions of randomness and of the notion of intuitive randomness referenced in each thesis, at this point of the discussion it suffices to observe that each thesis is formulated in terms of a notion of algorithmic randomness for infinite sequences that is extensionally distinct from the other two. Given (i) that each of the three instances of the randomness thesis schema implies, at a minimum, a claim of extensional adequacy, i.e., that the relevant formal definition of random sequence has the same extension as the informal, intuitive notion of random sequence, and (ii) that no two of the randomness thesis candidates are extensionally equivalent, it follows that at most one of these three theses can be true.

In this paper, I argue that we should not accept *any* of these three theses. But my view is not that these definitions of randomness simply miss the mark and that we should thus look elsewhere for a definition that captures the intuitive, pre-theoretic conception of random sequence. Rather, my more general goal, towards which the present study is just one step, is to establish that it is a mistake to judge the various definitions of randomness according to the following criterion of adequacy:

---

significant about the general notion of computation, just that they do not capture the notion of human computability that Turing and others originally had in mind.

[3] These sequences are usually taken to be binary, but one can consider more generally infinite sequences formed from any finite alphabet of symbols without affecting the resulting theory of randomness. Hereafter, the sequences under consideration will always be binary.

[4] See, for instance, [Del11] and [Das11].

[5] See [Sch71a] and [Sch77].

[6] See [OW08].

> $Adeq_1$: A definition of randomness $\mathscr{D}$ is adequate if and only if it captures the intuitive, pre-theoretic conception of randomness.[7]

This raises an immediate question: If we set aside $Adeq_1$ as a criterion by which the various definitions of algorithmic randomness should be judged, what condition or conditions determine the adequacy of a given definition of randomness?

Here I present a partial answer to this question by defending the following criterion of adequacy, which provides a *sufficient* condition for the adequacy of a definition of randomness:

> $Adeq_2$: A definition of randomness $\mathscr{D}$ is adequate if it captures some notion of *almost everywhere typicality.*

Of course, the viability of this criterion hinges on what I mean by "almost everywhere typicality," which I will clarify shortly. Having done so, I will present $Adeq_2$ is as a more reasonable alternative to $Adeq_1$. As we will see, the three randomness thesis candidates mentioned above, as well as several other non-equivalent definitions of algorithmic randomness that have not been suggested as randomness thesis candidates, fare equally well with respect to this criterion $Adeq_2$.

The heart of my argument is the recent study, from the point of view of algorithmic randomness, of certain "almost everywhere" theorems found in fields such as analysis, probability theory, ergodic theory, and information theory (see, for instance, [BMN11], [Rut12], [BDH$^+$11], [FGMN11], and [Hoy11]). These are theorems of the form

$$\text{(for almost every } x\text{) } \Phi(x),$$

where the "almost every" quantifier means "for every x in a set of Lebesgue measure one" and Φ is some property that holds of elements of the relevant domain. In this paper, we will consider such theorems from classical analysis, each of which involves the behavior of some class of real-valued functions at almost every real number in [0,1].

What is the connection of such results to algorithmic randomness? The answer is this: For each of the theorems in question, if we restrict the class of functions referenced in the theorem to a specific definable subclass, then the collection of real numbers for which the restricted version of the theorem holds will be coextensive with a collection of algorithmically random sequences (considered as the binary representations of real numbers). That is, if Φ^* is a suitable restriction of the formula Φ given above, then we have the following:

$$\Phi^*(x) \text{ if and only if } x \text{ is } \mathscr{D}\text{-random.}$$

That is, each Φ^* picks out what I referred to above as a notion of *almost everywhere typicality* (or *a.e. typicality*, for short), and moreover, this a.e. typicality is coextensive with $\mathscr{D}$-randomness for some definition $\mathscr{D}$ of algorithmic randomness. Furthermore, different notions of a.e. typicality correspond to different definitions of algorithmic randomness, including each of the three randomness theses candidates introduced above.

[7] I will briefly discuss what one might take "the intuitive, pre-theoretic conception of randomness" to refer to in Subsection 2.2 below.

In light of the correspondence between notions of a.e. typicality and definitions of algorithmic randomness, I will argue that it is not viable to judge the adequacy of definitions according to $Adeq_1$, from which it follows that we ought not accept any of the three main randomness theses. And in addition to suggesting the untenability of judging the adequacy of definitions of randomness according to $Adeq_1$, the correspondence between notions of a.e. typicality and definitions of algorithmic randomness suggests that $Adeq_2$ is a more appropriate criterion of adequacy by which the various definitions of randomness can be judged, or so I shall argue.

The upshot of my account is that we cannot justifiably say of any of these three randomness thesis candidates what Gödel said of the definition of computability, that we have an "absolute definition of an interesting epistemological notion, i.e., one not depending on the formalism chosen" [Göd46, p. 150]. None of these definitions of randomness captures everything that mathematicians have taken to be significant concerning the concept of randomness. Rather, we have a family of definitions of an interesting epistemological notion, many of which provide insight into certain mathematically significant instances of typicality.

The main claims of this paper may come as no surprise to current researchers in algorithmic randomness, many of whom have likely grown accustomed to the multiplicity of interesting and informative definitions of randomness. Despite this, there are a number of reasons that such a study is necessary. First, the main claims of the present study have yet to be clearly articulated and defended elsewhere, and thus this study fills a gap in the philosophical literature on algorithmic randomness. Second, the search for a definition of randomness that captures the so-called intuitive conception was a central impetus in the development of algorithmic randomness;[8] here I account for why this search could not be successfully carried out. Lastly, although theses such as the Martin-Löf-Chaitin thesis do not currently play a role in the theory of algorithmic randomness (unlike the case of the Church-Turing thesis in computability theory[9]), in some recent work, both the Martin-Löf-Chaitin thesis ([Del11], [Das11]) and the Weak 2-Randomness thesis ([OW08]) have been defended. Thus, the view against which I am arguing is not without its adherents.

The outline of the remainder of this paper is as follows. In §2, I provide the necessary technical background on the definitions of algorithmic randomness that are randomness thesis candidates, as well as several others. I also discuss the so-called intuitive conception of randomness. In §3, I discuss the notions of a.e. typicality that can be derived from the restricted almost everywhere theorems, while in §4 I present my argument for the claim that we should not accept any of the three main randomness theses and my argument for the claim that $Adeq_2$ is more appropriate than $Adeq_1$ as a criterion of adequacy for the various definitions of randomness. Lastly, in §5 I consider two objections to my account.

2. Definitions of Algorithmic Randomness

Before I survey the main definitions of algorithmic randomness that we will consider here, I will first try to motivate the general idea behind these definitions. Then I will briefly discuss the

[8] I do not attempt to defend this claim here, but see, for instance, chapter 9 of [Por12].

[9] See, for instance, Section 7.2 of [BBJ07] for a brief discussion of the distinction between avoidable and unavoidable uses of the Church-Turing thesis.

so-called intuitive conception of randomness that certain definitions of algorithmic randomness are claimed to capture.

2.1. **The logical approach to defining randomness.** Each of the definitions of randomness that we discuss in this paper are instances of what one might call the *logical approach* to randomness. The distinctive feature of definitions formulated according to the logical approach is that for each such definition $\mathscr{D}$,

(i) there is a countable collection of properties $\{\Phi_i\}_{i\in\omega}$ expressible in some formal language $\mathcal{L}$ (usually the language of first- or second-order arithmetic) such that for each $i \in \omega$, the set $\{x \in 2^\omega : \Phi_i(x)\}$ has Lebesgue measure one, and
(ii) a sequence $x \in 2^\omega$ is $\mathscr{D}$-random if and only if $\Phi_i(x)$ for every $i \in \omega$.

One immediate consequence of this approach is that for each such definition $\mathscr{D}$, assuming that there is some x that fails to satisfy some Φ_i, 2^ω can be partitioned into a non-empty collection consisting of all $\mathscr{D}$-random sequences and a non-empty collection consisting of all non-$\mathscr{D}$-random sequences.

But a natural question arises: Which properties $\{\Phi_i\}_{i\in\omega}$ should be chosen in formulating a definition of randomness? There is no obvious answer to this question. In fact, much of the early work in the development of algorithmic randomness was motivated by answering this very question.[10] The aim of this early work was to identify those properties that are satisfied by the typical outcomes of some fixed random process such as the repeated tosses of an unbiased coin. For instance, it was held that such properties should at least guarantee that any sequence counted as random satisfies the law of large numbers and the law of the iterated logarithm.[11]

But this approach is far from satisfying, as it is not at all clear that "those properties satisfied by the typical outcomes of some fixed random process" picks out a unique collection of properties. In fact, if the argument I provide in §4 is successful, we should reject the claim that there is any such unique collection of properties. Of course, this is not to say that *no* choice of the properties $\{\Phi_i\}_{i\in\omega}$ gives us a reasonable definition of randomness. As I will argue, there are choices of the properties $\{\Phi_i\}_{i\in\omega}$ that yield definitions of randomness that are adequate according to the criterion $Adeq_2$ (though there is no unique such collection, in my view).

2.2. **The so-called intuitive conception of randomness.** When confronted with a randomness thesis that asserts that some formal definition of randomness captures the intuitive, pre-theoretic conception of randomness, one might reasonably question what this intuitive conception could be. The literature on algorithmic randomness (both mathematical and philosophical) is largely silent on this matter.[12]

What authors who appeal to the intuitive conception of randomness seem to have in mind is clearly something like the prevailing, commonly-held ideas concerning the notion of randomness. Of course, this raises a number of questions, such as: Given that humans appear to be unreliable at

[10] This is precisely the question Martin-Löf took himself to be answering in his 1966 paper [ML66].
[11] That is, the limiting relative frequencies of the 0s and 1s in such sequences should converge to 1/2, and the relative frequencies of 0s and 1s in initial segments of these seqeunces should oscillate above and below 1/2 infinitely often.
[12] This should be contrasted, for instance, with Turing's discussion of the characteristic features of the intuitive notion of human computability in Section 9 of his famous 1936 paper [Tur36].

both detecting randomness and producing randomness,[13] surely our definitions of randomness ought not answer to these purportedly faulty intuitions. But which ideas are to be included among the "prevailing, commonly-held ideas of randomness"? And why think there is only one such collection of prevailing ideas to begin with?

In light of the first question, we might instead hold that the intuitive conception of randomness consists of the prevailing commonly-held ideas concerning the notion of randomness *among the statistically literate*, where this includes those who recognize that among equiprobable events occurring in a sufficiently large number of random trials, there can be both local irregularities (such as the unpredictability of individual outcomes) but global regularity (such as a near-equal distribution of outcomes for each event in the long run).[14]

Regarding the second question, for dialectical purposes, I will simply assume that there is some single such intuitive conception. Under this assumption, presumably held by the advocates of the various randomness thesis, I will then argue that each of the theses are untenable. I should emphasize, however, that nothing in my argument hinges on there being a single intuitive conception of randomness. Moreover, it is beyond the scope of this study to further investigate what this intuitive conception could be or whether there is a single such conception (or more precisely, whether the data I present indicates that there is no single intuitive conception of randomness).

With this background in mind, let us now discuss the three main definitions of randomness that have been proposed as randomness thesis candidates, as well as two other definitions that will be pertinent for the discussion of a.e. typicality in §3.

2.3. **Martin-Löf randomness.** Prior to the publication of Martin-Löf's 1966 paper, "The Definition of Random Sequences" [ML66], there was a general consensus that no satisfactory definition of randomness for infinite sequences had been developed, as all prior definitions of randomness were considered, for one reason or another, to be defective.[15] The definition of randomness offered in the 1966 paper, nowadays referred to as Martin-Löf randomness, was held by Martin-Löf to be a "natural definition of random infinite sequences" that "seems to satisfy all intuitive requirements" [ML66, pp. 602, 608]. Given the failure of previous attempts at defining random infinite sequences, this was a significant breakthrough.[16]

The basic idea behind Martin-Löf's definition is that random sequences, such as those typically produced by the repeated tosses of an unbiased coin, are those sequences that are typical with respect to a certain kind of statistical test. For sake of illustration, let us consider the case of a putatively random *finite* string of 0s and 1s. Suppose we have a device D that putatively produces random strings of a fixed length n (for some sufficiently large n). We can test the null hypothesis that D produces random strings by taking as a sample an output string produced by D and

[13]There is a substantial literature on this topic. For a good starting point, see [BHW91] or [FK97].

[14]Here I have in mind, say, statisticians, professional gamblers, and members of the scientific community at large.

[15]For instance, Richard von Mises' definition of randomness was criticized for failing to be well-defined, and later modifications of the definition, such as those offered by Abraham Wald and Alonzo Church, were shown by Jean Ville to include as random certain sequences that fail to satisfy the law of the iterated logarithm. For more details, see, for instance, [vL87] or [Das11].

[16]Martin-Löf explicitly notes this, writing, "Such a definition has so far not been obtained by other methods" [ML66, p. 608].

checking whether σ bears a given property that is typical of strings produced by, say, the tosses of an unbiased coin. For instance, we might consider a statistical test that checks the number of 0s and 1s of all strings with the same length as that of σ, rejecting the null hypothesis of randomness if it is observed that the ratio of 0s and 1s in the string strays too far from the value 1. For such a test, we set a level of significance $\alpha \in (0,1)$, with a corresponding critical region in the distribution of all possible strings of length n. If our sample string is found in the critical region, we reject the hypothesis that D produces random strings at level α, where α is the probability of rejecting a true null hypothesis.

But how do we carry out such a test when the device in question outputs a potentially infinite sequence? Martin-Löf's key insight was to test the sample *at* all levels of significance of the form $\alpha = 2^{-1}, 2^{-2}, 2^{-3}, \ldots, 2^{-n}, \ldots$. Moreover, if there is some level 2^{-n} such that the sample avoids the corresponding critical region, then we do not reject the null hypothesis. Thus, we reject the null hypothesis only if some initial segment of the sample is contained in each of the critical regions corresponding to arbitrarily small levels of significance. Lastly, if we further generalize by viewing such putative outputs of D as completed totalities, i.e., infinite binary sequences, then such a hypothesis test applied to an infinite sequence x amounts to testing finite initial segments of x.

But which statistical tests should we employ to test for randomness? Another key insight of Martin-Löf's was that we should include all and only those tests that are computably enumerable. As Martin-Löf observes, "Any sequential test of present or future use in statistics is given by an explicit prescription, which, for every level of significance ..., tells us for what sequences the hypothesis is to be rejected" [ML66, p. 609]. From such an explicit prescription, we can actually enumerate the critical regions of our test; this amounts to requiring that the enumeration be given computably. And by proceeding at this level of generality, Martin-Löf took himself to be providing a definition that would never count as random a sequence that would be counted as non-random in statistical practice, even in principle.

Formally, a *Martin-Löf test* is a sequence $(U_i)_{i\in\omega}$ of uniformly computably enumerable sets of finite strings $2^{<\omega}$ such that for each $U_i = \{\sigma_1, \sigma_2, \ldots\}$ we have

$$\sum_{\sigma_j \in U_i} 2^{-|\sigma_j|} \leq 2^{-i},$$

where, for any $\sigma \in 2^{<\omega}$, $|\sigma|$ is the length of σ. Equivalently, if we set

$$[\![U_i]\!] = \{x : (\exists \sigma \prec x)\ \sigma \in U_i\}$$

(where $\sigma \prec x$ means that σ is an initial segment of x), then $[\![U_i]\!]$ is an effectively open subset of 2^ω with Lebesgue measure no greater than 2^{-i}.[17] That is, $[\![U_i]\!]$ is the critical region corresponding to the significance level 2^{-i}.

[17]The Lebesgue measure λ on 2^ω is defined as follows. For a finite sequence $\sigma \in 2^{<\omega}$, the open set $[\![\sigma]\!] = \{x : \sigma \prec x\}$ has Lebesgue measure $\lambda([\![\sigma]\!]) = 2^{-|\sigma|}$, and then λ can be extended in the standard way to the Borel subsets of 2^ω via Carathéodory's theorem.

Next, x *passes the test* $(U_i)_{i\in\omega}$ if $x \notin \bigcap_{i\in\omega} [\![U_i]\!]$, i.e., there is some critical region $[\![U_i]\!]$ such that no initial segment of x is contained in U_i. Lastly, a sequence x is *Martin-Löf random*, denoted $x \in \mathsf{MLR}$, if x passes every Martin-Löf test $(U_i)_{i\in\omega}$.

The definition of Martin-Löf randomness has proven to be extremely useful, finding recent applications in such areas as ergodic theory, information theory, geometric measure theory, reverse mathematics, and computable analysis. Such a wide range of applicability might be considered as grounds for the thesis that Martin-Löf randomness captures the intuitive conception of randomness (referred to above as the Martin-Löf-Chaitin thesis). But prior to the discovery of these applications, the Martin-Löf-Chaitin thesis was accepted primarily on the basis of the confluence of multiple intensionally distinct definitions of randomness, with the idea being that just as the confluence of definitions supports the Church-Turing thesis, so too should it be counted as evidence for the Martin-Löf-Chaitin thesis.[18]

There are a number of problems with this approach to justifying the Martin-Löf-Chaitin thesis.[19] However, it is beyond the scope of this paper to consider these problems, as our goal here is not to consider the strength of the evidence offered in support of the various randomness-theoretic theses, but rather to argue directly against the tenability of the theses.

We now turn to Schnorr randomness, which was explicitly formulated as an alternative to Martin-Löf randomness.

2.4. **Schnorr randomness.** Several years after the publication of Martin-Löf's definition of randomness, Schnorr offered a constructive variant of Martin-Löf's definition of randomness.[20] In Schnorr's view, Martin-Löf's definition counts too many sequences among the random sequences, as certain Martin-Löf tests lack physical meaning and thus should not be included as part of an adequate definition of randomness. According to Schnorr,

> 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. [Sch71a, p. 255]

[18]This point is summed up nicely by Antony Eagle, who writes,

> Different intuitive starting points have generated the same set of random sequences. This has been taken to be evidence that [Martin-Löf]-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]

[19]For instance, there are definitions of algorithmic randomness $\mathscr{D}_1$ and $\mathscr{D}_2$ such that (i) $\mathscr{D}_1$ and $\mathscr{D}_2$ are not extensionally equivalent but (ii) each of $\mathscr{D}_1$ and $\mathscr{D}_2$ are extensionally equivalent to multiple intensionally distinct definitions of randomness. This problem is discussed at length in "The confluence of definitions of algorithmic randomness" (in preparation).

[20]See, for instance, [Sch71a], [Sch71b], or [Sch77].

According to Schnorr, in order for a statistical test to be physically meaningful, there must an explicit example of a sequence passing this test.[21] Furthermore, such an explicit example, on Schnorr's account, is given by a computable sequence.

Not every Martin-Löf test has this property. That is, there is a Martin-Löf test such that no computable sequence passes this test (for instance, there is a universal Martin-Löf test, so that a sequence is Martin-Löf random if and only if it passes this one test). Schnorr thus modified Martin-Löf's definition to exclude those Martin-Löf tests that have the property that no computable sequence passes any such test. To do so, instead of considering tests $(U_i)_{i\in\omega}$ with the property that each set $[\![U_i]\!]$ has measure no greater than 2^{-i} for each i, Schnorr restricted his attention to those tests $(U_i)_{i\in\omega}$ where the measure of each set $[\![U_i]\!]$ is *equal* to 2^{-i} (or, more generally, equal to some computable real number, which gives rise to an equivalent class of tests). Having made this modification, Schnorr proved that for every test $(U_i)_{i\in\omega}$ such that $\lambda([\![U_i]\!]) = 2^{-i}$ for every i, there is some computable sequence x such that $x \notin \bigcap_{i\in\omega}[\![U_i]\!]$. Such tests, which are called *Schnorr tests*, thus satisfy Schnorr's condition of physical meaningfulness.

Defining randomness in terms of Schnorr tests instead of Martin-Löf tests yields a alternative definition of randomness: a sequence x is *Schnorr random*, denoted $x \in \mathsf{SR}$, if and only if x passes every Schnorr test. Moreover, Schnorr proved that this restriction of tests yields a notion of randomness that is strictly weaker than Martin-Löf randomness, in the sense that more sequences are counted as random: every Martin-Löf random sequence is Schnorr random,[22] but there are Schnorr random sequences that are not Martin-Löf random.

Significantly, Schnorr is very clear that he takes his definition to capture "the true concept of randomness" [Sch71a, p. 255]. He even explicitly formulates this claim as a thesis (see [Sch77, p. 198]), also appealing to the confluence of multiple intensionally distinct definitions of randomness that are extensionally equivalent to Schnorr randomness in support of his thesis. Regardless of the merits of Schnorr's position, as early as 1970, there were already two competing candidates on the table for a definition that captures the intuitive conception of randomness.

2.5. **Weak 2-randomness.** Another alternative to Martin-Löf randomness that one finds in the algorithmic randomness literature is known as weak 2-randomness, a notion of randomness that is not obtained by restricting the class of Martin-Löf tests but rather by extending it. Originally introduced in [Kur82] (and independently in [GS82]), weak 2-randomness was not originally formulated as an alternative to Martin-Löf randomness that fares better with respect to capturing the intuitive conception of randomness. However, recently weak 2-randomness has been offered as an alternative randomness thesis candidate, as will be noted below.

Whereas Schnorr randomness is a notion of randomness that is weaker than Martin-Löf randomness, weak 2-randomness is stronger than Martin-Löf randomness, since it requires a sequence to pass more tests to be counted as random. As discussed above, the measures of the critical regions given by a Martin-Löf test $(U_i)_{i\in\omega}$ are effectively bounded: $\lambda([\![U_i]\!]) \leq 2^{-i}$ for each $i \in \omega$. If instead we require merely that the measures of our critical regions converge to zero, but not necessarily

[21] See [Sch71b, p. 35].

[22] This follows from the fact that every Schnorr test is a Martin-Löf test.

effectively,[23] then we have a more general notion of test than that of the Martin-Löf test. Thus, a *generalized Martin-Löf test* is a sequence $(U_i)_{i\in\omega}$ of uniformly computably enumerable subsets of $2^{<\omega}$ such that $\lim_{n\to\infty}\lambda([\![U_n]\!])=0$. In addition, we say that $x\in 2^\omega$ is *weakly 2-random*, denoted $x\in \mathsf{W2R}$, if $x\notin\bigcap_{i\in\omega}[\![U_i]\!]$ for every generalized Martin-Löf test $(U_i)_{i\in\omega}$. Equivalently, $x\in 2^\omega$ is weakly 2-random if and only if x is not contained in any Π^0_2 class of measure zero.[24]

Given that every Martin-Löf test is a generalized Martin-Löf test, it follows that every weakly 2-random sequence is Martin-Löf random. However, there are some Martin-Löf random sequences that are not weakly 2-random, particularly all Δ^0_2 Martin-Löf random sequences.[25] In fact, that some Martin-Löf random sequences are Δ^0_2 has been taken by some to be grounds for rejecting the Martin-Löf-Chaitin thesis (see, for instance, [Raa00] and [OW08]). The objection here is that Δ^0_2 random sequences, though not computable, are nonetheless decidable in the limit. But this appears to be a property incompatible with randomness, since it implies a close relationship between a sequence x and the computable procedure used to approximate the values of x.

Again, we cannot weigh merits of this argument for the claim that Δ^0_2 sequences should not be counted as random, but it is worth pointing out that the weakly 2-random sequences are precisely the Martin-Löf random sequences that cannot compute any non-computable Δ^0_2 sequence (in the sense of Turing reducibility), from which it follows that no weak 2-random sequence can be Δ^0_2. Moreover, Osherson and Weinstein claim that this latter fact "gives reason to identify the intuitive conception of randomness with [weak 2-randomness]" [OW08, p. 58].

2.6. **Two additional definitions.** For our purposes, it will be helpful to introduce two additional definitions of algorithmic randomness into the discussion, namely *computable randomness* and *Kurtz randomness*, neither of which has been proposed as a randomness thesis candidate.

Originally introduced by Schnorr in his study of alternatives to Martin-Löf randomness, computable randomness is defined in terms of a certain collection of betting strategies known as martingales.[26] The formal details are not so important for the ensuing discussion, so I will only give a rough sketch of the definition of martingale.

The basic idea behind these betting strategies is this: a game is played in which, for a fixed sequence $x\in 2^\omega$, a bettor attempts to predict the values of x one bit at a time, placing a bet at the n^{th} round on the value of the n^{th} bit of the sequence, having seen the values of the previous $n-1$ bits. The bettor has some initial capital and at each round he can bet some proportion p of his capital that the next value of the sequence will be a 0 (so that the proportion $1-p$ of his capital is bet that the next value of the sequence will be a 1). The payoff is fair, as the winnings for correctly guessing the next value are twice the amount of the bet placed on that value, while all capital placed on the incorrect value is lost. The bettor wins the game is his capital is not bounded

[23] That is, $\lambda([\![U_i]\!])\to 0$ as $i\to\infty$, but there need not be a computable function $f:\omega\to\omega$ such that $\lambda([\![U_i]\!])\leq 2^{-f(n)}$ for every n.

[24] A subset of 2^ω is Π^0_2 if it is the intersection of a sequence of uniformly effectively open subsets of 2^ω.

[25] A sequence is Δ^0_2 if there is some computable function $f:\omega^2\to\omega$ such that (i) $x(n)=\lim_{s\to\infty}f(n,s)$ and (ii) $f(n,0),f(n,1),f(n,2),\dots$ consists of only finitely many distinct values for each n.

[26] These martingales, defined by Ville in his dissertation [Vil39], should not be confused with the martingales one finds in modern probability theory, which are due independently to Doob and Lévy.

by any finite number (although this capital need not increase monotonically). Lastly, a sequence x is *computably random* if no bettor can win this game when betting on the values of x using a computable betting strategy.[27] The collection of computably random sequences will be denoted CR.

Schnorr was able to prove that every Martin-Löf random sequence is computably random and that every computably random sequence is Schnorr random (and thus that there are computably random sequences that are not Martin-Löf random). Computable randomness and Schnorr randomness were only separated much later.[28]

Kurtz randomness, also known as weak randomness, was introduced by Kurtz in his dissertation [Kur82]. Whereas a sequence is weakly 2-random if it is not contained in any Π^0_2 subclass of 2^ω of measure zero, a sequence is weakly 1-random if it is not contained in any Π^0_1 subclass of 2^ω of measure zero. This yields a particularly weak notion of randomness (one even weaker than Schnorr randomness), as the Kurtz random sequences include all sequences that are generic in a very weak sense,[29] sequences that do not even satisfy the law of large numbers. The collection of Kurtz random sequences will be denoted KR.

In sum, we have five non-equivalent definitions of randomness that stand in the following relationship:

$$\mathsf{W2R} \subsetneq \mathsf{MLR} \subsetneq \mathsf{CR} \subsetneq \mathsf{SR} \subsetneq \mathsf{KR}.^{30}$$

With this background in mind, we now consider the primary data to which I appeal in my argument against the three randomness theoretic theses.

3. The data: "Almost everywhere" theorems

In this section, I present the instances of almost everywhere typicality that serve the dual purpose of (1) showing the untenability of the three randomness theses under consideration and (2) motivating the criterion of adequacy $Adeq_2$ (that a definition of randomness $\mathscr{D}$ is adequate if it captures some notion of a.e. typicality) as an alternative to the criterion $Adeq_1$ (that a definition of randomness $\mathscr{D}$ is adequate if and only if it captures the intuitive, pre-theoretic conception of randomness).

3.1. **A.e. typicality.** In classical analysis, it is very common to encounter theorems that hold of almost every member of some fixed domain of objects, for instance, the set of the real numbers or the unit interval $[0,1]$. Hereafter, let us restrict our attention to results where the domain in question is $[0,1]$. A number of these results involve some collection $\mathscr{C}$ of real-valued functions $f : [0,1] \to \mathbb{R}$ and have the form

$$(\forall f \in \mathscr{C})(\forall^{\mathsf{a.e.}} x \in [0,1])\ \Phi(x,f),$$

[27]Roughly, a computable betting strategy is given by a computable function that determines the amounts of capital to be bet on each outcome at each round (amounts that are themselves computable real numbers).

[28]See [Wan99].

[29]Such sequences, called *weakly 1-generic*, are contained in every dense effectively open subset of 2^ω.

[30]This gives the false impression that the collection of definitions of algorithmic randomness is linearly ordered, but there are definitions of randomness incomparable with one another in the ordering given by $\subseteq$.

where $\forall^{\text{a.e.}}$ is the almost everywhere quantifier (so that $(\forall^{\text{a.e}} x \in [0,1])\ \Phi(x)$ means that the set $\{x : \Phi(x)\}$ has Lebesgue measure one), and $\Phi(x, f)$ is some predicate such as "f is differentiable at x."

Such results are commonly glossed as follows: For each f, if we choose a point $x \in [0,1]$ at random, then with probability one, the property $\Phi(\cdot, f)$ will hold at x. Alternatively, one can say that it is the *typical* behavior of points $x \in [0,1]$ for the each of the above properties $\Phi(\cdot, f)$ to hold at x, or that $\Phi(\cdot, f)$ holds of the *Lebesgue-typical* member of [0,1], or that these properties hold of the random member of $[0,1]$.[31] As stated in the introduction, such typical behavior will be referred to as *a.e. typicality.*

In the contexts of computable mathematics or constructive mathematics, these kinds of almost everywhere results can be studied for a restricted class of functions. One common restriction is to consider only *computable* real-valued functions.[32] It was originally observed by Demuth in [Dem75] that the points at which these restricted almost everywhere results hold correspond precisely to various classes of algorithmically random sequences. This correspondence was recently rediscovered by Brattka, Miller, and Nies in [BMN11], and has led to a fruitful strand of research in algorithmic randomness. In the next subsection, I provide the details of the correspondence between a.e. typicality and randomness.

3.2. **Some examples.** Let us first consider two examples that give a flavor of the correspondence between instances of a.e. typicality and notions of algorithmic randomness. In a standard course on real analysis, one encounters the following theorems:

(1) For every non-decreasing real-valued function $f : [0,1] \to \mathbb{R}$, f is differentiable almost everywhere.
(2) For every real-valued function $f : [0,1] \to \mathbb{R}$ of bounded variation, f is differentiable almost everywhere.[33]

There are several observations to make here. First, the statements of these two theorems have precisely the same form, the only difference being that they involve quantification over different collections of real-valued functions. Second, note that (2) is strictly stronger than (1), in the sense that

[31]I should emphasize that these glosses here are not, strictly speaking, formal, insofar as the notions of "typicality" or "random point" are not precisely defined in the contexts in which one finds these glosses.

[32]A function $f : [0,1] \to \mathbb{R}$ 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 $g : \omega \to \omega$ (called a *modulus function*) such that for every $x, y \in [0,1]$ and every $n \in \omega$,

$$|x - y| \leq 2^{-g(n)} \Rightarrow |f(x) - f(y)| \leq 2^{-n}.$$

[33]A function $f : [0,1] \to \mathbb{R}$ 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.

(i) having established (2), we are justified in concluding that a non-decreasing real-valued function is differentiable almost everywhere (since every non-decreasing function is of bounded variation), but

(ii) having established (1), we are not justified in concluding that a real-valued function of bounded variation is differentiable almost everywhere (at least not without first establishing that every real-valued function of bounded variation is the difference of two non-decreasing real-valued functions).

Third, the function quantifier in each of theorems ranges over sets of size $2^{\mathfrak{c}}$, the size of the power set of the continuum. Lastly, the properties "being a point of differentiability of some non-decreasing real-valued function on $[0, 1]$" and "being a point of non-differentiability of some non-decreasing real-valued function [0,1]" are satisfied by every point in [0,1] (and similarly for the analogous properties for functions of bounded variation). To wit, for each $x \in [0, 1]$, one can find a non-decreasing functions f and g such that f is differentiable at x but g is not.

With these observations in mind, let us now consider the following modified versions of (1) and (2) above, given by restricting each class of functions to the computable members of that class. I will refer to these statements as *effectively restricted versions* of (1) and (2).

(1^*) For every *computable* non-decreasing real-valued function $f : [0, 1] \to \mathbb{R}$, f is differentiable almost everywhere.

(2^*) For every *computable* real-valued function $f : [0, 1] \to \mathbb{R}$ of bounded variation, f is differentiable almost everywhere.

Although it is still the case that we can immediately establish (1^*) from (2^*), there are some significant differences between these results and their unrestricted counterparts.

First, in both (1^*) and (2^*), the function quantifier only ranges over countably many functions, as there are only countably many computable functions $g : \omega \to \omega$ that can be used to define the modulus of a computable real-valued function. Thus, each of (1^*) and (2^*) yields a countable sequence of sets of Lebesgue measure zero, where each set corresponds to the points of non-differentiability of some computable real-valued function of the relevant kind.

This leads to the second significant difference between the restricted and unrestricted versions of the two theorems, namely that the properties "being a point of differentiability of some computable non-decreasing real-valued function" and "being a point of non-differentiability of some computable non-decreasing real-valued function" are *not* satisfied by every point in [0,1]. Since there are only countably many computable real-valued functions, the collection of points x such that *some* computable non-decreasing real-valued function f is not differentiable at x is the union of countably many sets of measure zero, which is thus also a set of measure zero. It follows that the collection of points x such that *every* computable non-decreasing real-valued function f is differentiable at x is a set of measure one. Thus, whereas the unrestricted results (1) and (2) do not single out any collection of points as a collection of "typical" points, the restricted results (1^*) and (2^*) yield a set of measure one many "typical" points.

As stated above, statements of the form "Property $\mathcal{P}$ holds almost everywhere" can be glossed as "The randomly chosen point has property $\mathcal{P}$." Significantly, such a parallel persists with effectively restricted versions of such statements. That is, in a number of cases, when we effectively restrict statements that hold almost everywhere as described above, the resulting collection of a.e. typical points picks out a collection of algorithmically random points. Thus with respect to (1^*) and (2^*) we have:

($\dagger_1$) Every computable non-decreasing real-valued function $f : [0,1] \to \mathbb{R}$ is differentiable at $x \in [0,1]$ if and only if x is computably random.[34]

($\dagger_2$) Every computable real-valued function $f : [0,1] \to \mathbb{R}$ of bounded variation is differentiable at $x \in [0,1]$ if and only if x is Martin-Löf random.[35]

As we observed above, since every non-decreasing real-valued function is of bounded variation, for $x \in [0,1]$ to be a point of differentiability for every function of bounded variation is more demanding than to be a point of differentiability for every non-decreasing function. This is reflected in the fact that $\mathsf{MLR} \subsetneq \mathsf{CR}$.

As another example of this phenomenon, if we consider the points of differentiability of computable real-valued functions that are differentiable almost everywhere (a class of functions that includes every computable function of bounded variation), we have the following:

($\dagger_3$) Every computable almost everywhere differentiable real-valued function $f : [0,1] \to \mathbb{R}$ is differentiable at $x \in [0,1]$ if and only if x is weakly 2-random.[36]

Again, by considering a larger class of functions, the resulting notion of a.e. typicality is more restricted than the previous two examples, which is reflected in the fact that $\mathsf{W2R} \subsetneq \mathsf{MLR} \subsetneq \mathsf{CR}$.

Not every notion of a.e. typicality is given in terms of computable real-valued functions. For instance, consider the Lebesgue differentiation theorem, which states that for every L_1-integrable function[37] $f : [0,1] \to \mathbb{R}$, for almost every $z \in [0,1]$ the following holds:

$$(*) \qquad \lim_{r\to\infty} \frac{1}{\lambda(B_r(z))} \int_{B_r(z)} f(x)dx = f(z).^{38}$$

For any z satisfying $(*)$, we say that z is a *Lebesgue point* of f.

Now, one way to effectively restrict this theorem is to consider functions that are not only computable, but computably approximable as elements of the Hilbert space $L_1[0,1]$, which are known as L_1-*computable* functions. Restricting to this collection of functions yields the following:

($\dagger_4$) For $z \in [0,1]$, z is a Lebesgue point of every L_1-computable real-valued function $f : [0,1] \to \mathbb{R}$ if and only if z is Schnorr random.[39]

However, we can also consider the Lebesgue differentiation theorem in terms of partial computable functions $f : [0,1] \to \mathbb{R}$ that are defined almost everywhere, which are called *a.e. computable* functions. With this choice of effective restriction of the class of functions in the statement of the theorem, we have

[34] This is due to Brattka, Miller, and Nies [BMN11].

[35] This is originally due to Demuth [Dem75] but was rediscovered by Brattka, Miller, and Nies [BMN11].

[36] This is also due to Brattka, Miller, and Nies [BMN11].

[37] A function $f : [0,1] \to \mathbb{R}$ is L_1-integrable if $\int |f| d\lambda < \infty$

[38] $B_r(z)$ is the ball of radius r centered at z.

[39] This result was obtained by Pathak, Rojas, and Simpson [PRS11] and independently by Rute [Rut11].

($\dagger_5$) For $z \in [0,1]$, z is a Lebesgue point of every a.e. computable real-valued function $f : [0,1] \to \mathbb{R}$ if and only if z is Kurtz random.[40]

In sum, we have at least five notions of a.e. typicality derived from results in classical analysis, each of which corresponds to a notion of randomness introduced in §2 above.

Let us now consider the significance of this data for the status of the various randomness-theoretic theses and for the alternative account of adequacy for definitions of randomness that I am attempting to articulate.

4. Significance of the data

4.1. **The status of the randomness-theoretic theses.** In light of the data presented in the previous section, how do the three main randomness-theoretic theses fare? As a first step towards answering this question, let us consider the status of the Martin-Löf-Chaitin thesis in the face of the correspondence between notions of a.e. typicality and definitions of randomness. Given this correspondence, it is clear that the advocate of the Martin-Löf-Chaitin thesis faces a dilemma:

(i) explain why it is that the points of differentiability of computable functions of bounded variation pick out precisely the intuitively random points while the points of differentiability of non-decreasing computable functions or of almost everywhere differentiable computable functions do not, or

(ii) concede that there is nothing more intuitively random about points in the former collection than the points in the latter two.

Of course, if the advocate of the Martin-Löf-Chaitin thesis were to opt for (ii), this would amount to a rejection of the thesis. So if the advocate of the Martin-Löf-Chaitin thesis does not want to reject her preferred thesis, she must opt for (i). But I think we ought to be skeptical that the requisite explanation as described in (i) can be successfully provided.

As I see it, the difficulty in providing the explanation as described in (i) is that each of the relevant collections of points of differentiability has a reasonable claim as a collection of intuitively random points, insofar as the points are understood in terms of the informal gloss of almost everywhere results as discussed in §3.1. More specifically, earlier we saw that a statement of the form "almost every $x \in [0,1]$ has property $\mathcal{P}$" can be glossed as "the random point $x \in [0,1]$ has property $\mathcal{P}$." Thus, there is precedent for considering such points as random, at least in an informal sense.

But it would peculiar if this informal gloss were only legitimately applicable in the context of the points of differentiability of functions of bounded variation but not in the context of the points of differentiability of other kinds of functions. Thus, whatever the advocate of the Martin-Löf-Chaitin thesis takes the "intuitive conception of randomness" to be, she needs to explain why only one collection of points of differentiability captures the intuitively random points while all others fail to do so, despite there being grounds for taking these latter points to be random in the informal sense specified above. That is, if we take a conception of randomness to be one on the basis of which we attribute randomness to various objects, the advocate of the Martin-Löf-Chaitin thesis owes us an

[40] This, and similar characterizations of Kurtz randomness in terms of some other almost everywhere phenomena, is due to Miyabe [Miy13].

explanation as to why her understanding of the intuitive conception of randomness is preferable to one on the basis of which we attribute randomness to, say, the points of differentiability of non-decreasing computable functions.

The advocate of the Martin-Löf-Chaitin thesis might attempt to split the horns of this dilemma by simply holding that these differentiability results, while interesting, do not really tells us anything significant about the various definitions of randomness. However, to go this route would be to give up too much, for ($\dagger_2$) is precisely the sort of result to which the advocate of the Martin-Löf-Chaitin thesis should appeal to demonstrate the significance and value of the notion of Martin-Löf randomness.

Mutatis mutandis, a similar dilemma can also be posed to the advocate of the weak 2-randomness thesis. Furthermore, we can also pose such a dilemma for the advocate of Schnorr's thesis in terms of the Lebesgue points for L_1-computable functions and a.e. computable functions. Thus, the data presented in §3 suggest that it is much too strong a position to hold that any of the randomness thesis candidates picks out precisely the intuitively random sequences.

4.2. **The status of $Adeq_1$.** Even if we were to accept the claim that none of randomness thesis candidates picks out precisely the intuitively random sequences, might we reasonably hold out hope that some hitherto unknown definition of randomness captures the intuitive conception of randomness?

To hold such a position, one must reject the claim that the correspondence between definitions of algorithmic randomness and notions of a.e. typicality has any significance for our understanding of the concept of randomness. In particular, one must reject the informal gloss that licenses us to consider a.e. typical points as randomly chosen points.

But what grounds are there for rejecting this gloss? Moreover, what grounds are there for rejecting this gloss while keeping open the possibility that some hitherto unknown definition of randomness will satisfy $Adeq_1$? To hold such a position, the adherent of $Adeq_1$ owes us an account of the intuitive conception of randomness according to which (i) the three main randomness thesis candidates fail to capture this intuitive conception and (ii) the informal gloss should be rejected. But what could such a conception of randomness be?

In my view, instead of striving to develop such an account of the intuitive conception of randomness that severs the correspondence between definitions of algorithmic randomness and notions of a.e. typicality, our conception of randomness should be informed by this correspondence. And if our conception of randomness is informed by this correspondence, we will see that no single definition of randomness can do the work of capturing every mathematically significant collection of typical points; rather, multiple non-equivalent definitions are necessary to carry out this task.

4.3. **The merits of $Adeq_2$ as an alternative to $Adeq_1$.** Recall that $Adeq_2$ is a criterion of adequacy for definitions of randomness according to which a definition $\mathscr{D}$ is adequate if it captures some notion of a.e. typicality. In this section, I will argue that two primary merits that $Adeq_2$ has compared to $Adeq_1$ are:

(1) the use of $Adeq_2$ is informed by theorems from classical mathematics, and thus there are identifiable grounds for determining that a definition of randomness is adequate according to $Adeq_2$, unlike $Adeq_1$, and
(2) $Adeq_2$ is consistent with a research program to classify various notions of a.e. typicality that occur in classical mathematics in terms of the definitions of randomness to which they correspond, whereas $Adeq_1$ is not.

Let us consider each of these merits in turn.

First, regarding the claim that the use of $Adeq_2$ is informed by theorems from classical mathematics, this should be clear from the examples that we considered in §3. In order for a definition $\mathscr{D}$ of randomness to be judged adequate according to $Adeq_2$, there must be some almost everywhere theorem that, when suitably restricted, yields a notion of a.e. typicality that is coextensive with $\mathscr{D}$. Moreover, in the examples we have seen, as well as many others not discussed here, the almost everywhere theorems in question are not contrived but rather come directly from classical mathematics. In this respect, one looks to results from classical mathematics to ground the $Adeq_2$-adequacy of a given definition of randomness, something that cannot be claimed for $Adeq_1$.

It follows from the discussion in the previous paragraph that there are identifiable grounds for determining that a definition of randomness is adequate according to $Adeq_2$: if there is a notion of a.e. typicality that is coextensive with the definition in question, we judge it to be adequate.[41] By contrast, no such identifiable grounds are available in judging a whether a definition of randomness is adequate according to $Adeq_1$. It is simply not sufficient to appeal to the intuitive conception of randomness without providing a clear account of the conditions under which we should judge a sequence to be intuitively random or intuitively non-random. But as noted in §2.2, no such account has been offered. Furthermore, in light of the data considered in §3, it is far from clear how one might develop an account of the intuitive conception of randomness according to which at most one definition of algorithmic randomness is coextensive with it.

Second, $Adeq_2$ is consistent with a specific research program in the theory of algorithmic randomness, namely, that of classifying notions of a.e. typicality in terms of *degrees of randomness*, each of which is picked out by a different definition of algorithmic randomness, whereas $Adeq_1$ is not. Before I describe this research program, let me explain this latter claim, that the various definitions of algorithmic randomness pick out different degrees of randomness.

In §2, I appealed to the relations "weaker than" and "stronger than" in comparing two definitions of randomness. To be explicit, one definition $\mathscr{D}_1$ is *stronger* (resp. *weaker*) than another $\mathscr{D}_2$ if $\mathscr{D}_1$ counts strictly fewer (resp. more) sequences as random than $\mathscr{D}_2$ does, i.e., $\mathsf{ext}(\mathscr{D}_1) \subsetneq \mathsf{ext}(\mathscr{D}_2)$, where $\mathsf{ext}(\mathscr{D})$ is the extension of the definition $\mathscr{D}$. Additionally, for definitions $\mathscr{D}_1$ and $\mathscr{D}_2$ such that $\mathsf{ext}(\mathscr{D}_1) \subsetneq \mathsf{ext}(\mathscr{D}_2)$, given that sequences that are $\mathscr{D}_1$-random pass more tests than sequences that are merely $\mathscr{D}_2$-random, it is not unreasonable to say that $\mathscr{D}_1$-random sequences are *more random* than sequences that are $\mathscr{D}_2$-random but not $\mathscr{D}_1$-random (and thus that the $\mathscr{D}_2$-random sequences

[41] If no such notion of a.e. typicality is available, all other things being equal, we should withhold judgment concerning the adequacy of the definition in question.

that are not $\mathscr{D}_1$-random are *less random* than the $\mathscr{D}_1$-random sequences). Thus it is natural to interpret the various notions of algorithmic randomness in terms of degrees of randomness.

Now, the research program mentioned above is as follows: for each result in classical mathematics involving an almost everywhere quantifier in the sense of Lebesgue measure, one can investigate the degree of randomness necessary and sufficient for the result to hold in an effective setting by identifying a definition of algorithmic randomness that captures the relevant notion of a.e. typicality.

We have already seen, for instance, that the points of differentiability of computable functions of bounded variation form a proper subset of the points of differentiability of computable non-decreasing functions. That the former collection is a subset of the latter is hardly surprising, given that every non-decreasing function is of bounded variation. However, that the former collection is a *proper* subset of the latter is much less apparent. By identifying the definitions of randomness corresponding to each of the two classes of points of differentiability, we learn that the degree of randomness necessary and sufficient for being a point of differentiability of computable functions of bounded variation is strictly higher than the degree of randomness necessary and sufficient for being a point of differentiability of computable non-decreasing functions. Thus, such a study can be informative not just from the point of view of those interested in algorithmic randomness, but also from the point of view of those interested in the effective content of mathematics.

As mentioned in the introduction, a number of results along these lines have been established from areas such as classical analysis, probability theory, ergodic theory, and information theory. This opens up the possibility that we can compare two notions of a.e. typicality from different areas of mathematics, which would yield further insights into the effective content of the classical theorems from which the instances of a.e. typicality are derived. For example, we have seen that Martin-Löf randomness picks out the points of differentiability of computable functions of bounded variation. It has also been shown that an effective version of Birkhoff's ergodic theorem for effectively closed classes holds at precisely the Martin-Löf random points (see [BDH$^+$11] or [FGMN11]). Thus, we can conclude that the almost everywhere quantifiers in the restricted versions of each of these results are of the same strength, a result that one could hardly have anticipated prior to introducing the tools of algorithmic randomness into the analysis. Thus, we have a research program akin to reverse mathematics, which one might call *reverse randomness*.[42] Much work remains to be done in this direction.[43]

It should be clear that this research program is consistent with judging definitions to be adequate according to $Adeq_2$. However, on the approach taken by the advocate of $Adeq_1$, only one definition, up to extensional equivalence, can capture the intuitive conception of randomness. Thus, for those who prefer that we judge the adequacy of the various definitions of randomness according to $Adeq_1$, it is not clear what value there is in the classification program described above.

[42] I owe this term to Laurent Bienvenu.

[43] A similar program can be carried out for results in classical mathematics involving an almost everywhere quantifier in the sense of Baire category. For instance, Kuyper and Terwijn recently showed that the points at which everywhere differentiable real-valued functions have continuous derivative are the 1-generic points (which form a comeager set, but one of measure zero). See [KT13] for more details.

Summing up the discussion in this section, the advocate of the randomness-theoretic thesis for $\mathscr{D}$-randomness, where $\mathscr{D}$ is one of the three randomness thesis candidates discussed here, has two viable options: either (i) continue to accept the adequacy of $\mathscr{D}$-randomness according to $Adeq_1$, in spite of the evidence suggesting the untenability of this approach, or (ii) to recognize that the judging the adequacy of a definition of randomness according to $Adeq_2$ provides a more fruitful approach to the definitions of randomness, one that is consistent with the potentially informative research program described above.[44] 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 mathematically significant purposes. This, I claim, is considerable progress in understanding the concept of randomness as it occurs in classical mathematics.

5. Two Objections

In this final section, I will consider two objections one might raise against my account. The first objection is directed towards the notions of a.e. typicality that the various definitions of randomness purportedly capture. The second objection involves a worry that the various definitions of randomness, and the corresponding notions of a.e. typicality, are artificial, in which case it would follow that these definitions do not capture what mathematicians consider to be significant truths about randomness.

5.1. **The Overspill Problem.** Consider the formula $x = x$, which obviously holds for almost every sequence in 2^ω, as it holds for *every* sequence in 2^ω. Why should this not be counted as a notion of a.e. typicality? Equivalently, why should the collection of sequences satisfying $x = x$ not be counted as picking out a degree of randomness?

Specifically, the objection here is that I have not said enough about what counts as a *legitimate* notion of a.e. typicality, so that without some constraint on the notion of a.e. typicality, my account runs the risk of counting all sorts of collections sequences as corresponding to various degrees of randomness even though they clearly ought not be considered as such. Let us refer to this problem as the *overspill problem.*

As a first step towards responding to the overspill problem, let us stipulate that a notion of a.e. typicality is legitimate if it is obtained by effectively restricting some almost everywhere theorem along the lines of the examples we considered in §3. Does this solve the problem? Not quite yet, for one can argue that since any universal statement from classical mathematics can be seen as a special kind of "almost everywhere" statement, the formula $x = x$ defines a notion of a.e. typicality.

Of course, one might further argue that we should 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* object in a given domain as typical is a defective notion of typicality. But what exactly is defective about such a notion of typicality?

To answer this question, one might appeal to intuitive considerations. Consider, for instance, certain everyday, informal uses of "typical": "the typical American family," "the typical income,"

[44] Of course, there is a third option, which is to simply reject all definitions of randomness as inadequate. But clearly this is not a likely response from anyone who is already committed to the correctness of some definition of randomness.

"the typical interview questions," etc. We do not 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. In each of these cases, the predicate "x is typical" picks out a proper subcollection of the domain in question, a property of notions of typicality we might refer to as being *discriminative.* So, the argument might go, just as everyday, informal notions of typicality are discriminative, good mathematical definitions of typicality ought to be discriminative.

While there may be some merit to this argument, a stronger argument can be given if we consider how notions of Lebesgue typicality are used in classical mathematics. In considering the usage of "the Lebesgue typical point" in mathematical discourse, one will observe that when some property is asserted to hold at the Lebesgue-typical object, this just means that there is a set of objects of measure zero where this property fails to hold. Furthermore, if the property in question were to hold everywhere and not just almost everywhere, there would be no reason to state that the property holds at the Lebesgue-typical point; instead, one would simply state that the property holds of *every* object in the domain, which is clearly a stronger claim. Thus, the notion of Lebesgue-typicality as it appears in classical mathematics is discriminative. Consequently, if we further require that a legitimate notion of a.e. typicality is one that is obtained by restricting some *discriminative* notion of Lebesgue typicality from classical mathematics, then it follows that non-discriminative properties such as the one given by $x = x$ do not yield legitimate notions of a.e. typicality.

We should note that a related objection can be raised that certain notions of a.e. typicality, which are legitimate according to the criterion specified above, should not be counted as corresponding to any degree of randomness, since such collections contain sequences that are seemingly non-random. For instance, consider the notions of a.e. typicality corresponding to Kurtz randomness (one example of which was given in §3). Although such notions of a.e. typicality satisfy the criterion of legitimacy, the collection of Kurtz random sequences includes all sequences that fail to satisfy even the most basic laws of probability, such as the law of large numbers (namely, the weakly 1-generic sequences). Such sequences, one might argue, have no business being considered random.

To give a fully developed response to this particular concern would take us beyond the scope of this paper, as it requires a more nuanced account of the relationship between the different definitions of randomness (in particular, the dependence of each such definition on some underlying collection of resources), a detailed discussion as to why seemingly non-random sequences can nonetheless be counted as formally random, and an argument that this latter fact does not undermine the claim that the definitions of algorithmic randomness pick out various degrees of randomness.[45] For the present purpose, it suffices to note that this concern does not undermine the original argument given against the three randomness theses, as the adherent of each randomness thesis still has to account for the equivalence of the other two definitions of randomness with some notion of a.e. typicality, even if other notions of a.e. typicality (i.e., those corresponding to Kurtz randomness) are less well-behaved.

[45]This very objection and my response to it is the focus of the paper, "Randomness and accessible mathematical objects" (in progress).

5.2. **The Artificiality Objection.** The second objection that we will consider here is that these notions of randomness and the notions of a.e. typicality associated to them are artificially concocted and as such, they are not relevant to understanding the concept of randomness as it occurs in classical mathematics.

An objection along these lines is provided by Michiel van Lambalgen in his dissertation, "Random Sequences" [vL87]. 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. [vL87, p. 92]

While I mostly agree with van Lambalgen here,[46] this is the extent of our agreement 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.* [vL87, p. 92, emphasis added]

If there 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 is 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 us call this problem the *artificiality problem*.

We should first note that the various definitions of randomness were formulated prior to the discovery that they are equivalent to various notions of a.e. typicality; that is, the definitions of randomness were not reverse engineered to capture the notions of a.e. typicality discussed in §3. In light of this fact, it is not unreasonable to hold that the various definitions of randomness

[46]One point of disagreement: it is unclear to me that van Lambalgen's 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 is not the mere presence of multiple definitions of randomness that should be troubling to the advocate of one of the randomness-theoretic theses. Thus I take my argument to be an improvement of van Lambalgen's insofar as I account for why this multiplicity proves to be problematic for advocates of a given randomness-theoretic thesis.

are far from artificial and should be of interest to mathematicians outside of the community of computability theorists (and even outside of the community of mathematical logicians).

In response to this, one might claim that the restriction of almost everywhere results to computable instances is itself an artificial restriction. Moreover, the objection might continue, there are no notions of typicality implicit in almost everywhere theorems from classical mathematics to begin with; these so-called notions of typicality only emerge once we have introduced artificial restrictions to the theorems in question. 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 as I laid out in §3.2, in the unrestricted setting, for every $x \in [0,1]$, we can always find a non-decreasing real-valued function f such 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.[47] Thus, there is no notion of typicality associated with this classical theorem.

It is true that the different notions of randomness are not explicitly identifiable when we consider the almost everywhere results in their unrestricted forms. But to hold that these definitions are therefore artificial due to the fact that they only become relevant to the discussion once we effectively restrict the almost everywhere results in question 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.[48] 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,[49] and studying the extent to which the theorems still hold true in these restricted settings. In many cases, in such a restricted setting, additional information about the objects in question is uncovered, information which likely would not have been uncovered in the unrestricted setting.[50]

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. 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]

[47]One might further object that the restriction to computable real-valued functions here 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 restriction 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 test. 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.

[48]For a brief survey of the history of work in effective mathematics, see [Har98, pp. 5-7].

[49]For instance, "nicely definable" can be taken to mean "Δ^0_1-definable in the language of arithmetic," "arithmetically definable," "definable at level n of the arithmetical hierarchy," and so on.

[50]For examples of this phenomenon, see the many selections in [EGN$^+$98a] and [EGN$^+$98b].

This is precisely what I take the analysis of almost everywhere theorems to deliver for us: In the context of this larger project of unearthing the effective content of classical mathematical theorems, these results give us insight into the effective content of certain almost everywhere theorems from classical mathematics. In this respect, the definitions of algorithmic randomness are far from artificial.

References

[BBJ07] George S. Boolos, John P. Burgess, and Richard C. Jeffrey. *Computability and logic*. Cambridge University Press, Cambridge, fifth edition, 2007.

[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.

[BHW91] Maya Bar-Hillel and Willem A Wagenaar. The perception of randomness. *Advances in applied mathematics*, 12(4):428–454, 1991.

[BMN11] Vasco Brattka, Joseph S. Miller, and André Nies. Randomness and differentiability. *Preprint*, 2011.

[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.

[Del11] Jean-Paul Delahaye. The Martin-Löf-Chaitin Thesis: The identification by recursion theory of the mathematical notion of random sequence. In Hector Zenil, editor, *Randomness through computation*, chapter 9, pages 121–140. World Scientific Publishing Company, 2011.

[Dem75] Osvald Demuth. The differentiability of constructive functions of weakly bounded variation on pseudo numbers. *Comment. Math. Univ. Carolin*, 16(3):583–599, 1975.

[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.

[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.

[FK97] Ruma Falk and Clifford Konold. Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104(2):301, 1997.

[Göd46] Kurt Gödel. Remarks before the princeton bicentennial conference on problems in mathematics. *Collected works*, 2:144–153, 1946.

[GS82] Haim Gaifman and Marc Snir. Probabilities over rich languages, testing and randomness. *The journal of 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.

[Hoy11] Mathieu Hoyrup. The dimension of ergodic random sequences. *arXiv preprint arXiv:1107.1149*, 2011.

[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.

[Kol83] A. N. Kolmogorov. Combinatorial foundations of information theory and the calculus of probabilities. *Uspekhi Mat. Nauk*, 38(4(232)):27–36, 1983.

[KT13] Rutger Kuyper and Sebastiaan A Terwijn. Effective genericity and differentiability. *arXiv preprint arXiv:1309.3983*, 2013.

[Kur82] Stuart Kurtz. Randomness and genericity in the degrees of unsolvability. *Dissertation Abstracts International Part B: Science and Engineering[DISS. ABST. INT. PT. B-SCI. & ENG.],*, 42(9):1982, 1982.

[Miy13] Kenshi Miyabe. Characterization of kurtz randomness by a differentiation theorem. *Theory of Computing Systems*, 52(1):113–132, 2013.

[ML66] Per Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.

[MR06] Wolfgang Merkle and Jan Reimann. Selection functions that do not preserve normality. *Theory Comput. Syst.*, 39(5):685–697, 2006.

[OW08] Daniel Osherson and Scott Weinstein. Recognizing strong random reals. *The Review of Symbolic Logic*, 1(01):56–63, 2008.

[Por12] C. Porter. *Mathematical and Philosophical Perspectives on Algorithmic Randomness.* PhD thesis, University of Notre Dame, 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.

[Rut11] Jason Rute. Randomness and the Lebesgue differentation theorem. 2011.

[Rut12] Jason Rute. Algorithmic randomness, martingales, and differentiability i. *preparation*, 5:33, 2012.

[Sch71a] Claus-Peter Schnorr. A unified approach to the definition of random sequences. *Math. Systems Theory*, 5:246–258, 1971.

[Sch71b] Claus-Peter Schnorr. *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie.* Lecture Notes in Mathematics, Vol. 218. Springer-Verlag, Berlin, 1971.

[Sch77] Claus-Peter 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.

[Tur36] A. M. Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proc. London Math. Soc.*, S2-42(1):230–265, 1936.

[Vil39] Jean Ville. *Étude critique de la notion de collectif.* Gauthier-Villars, 1939.

[vL87] Michiel van Lambalgen. *Random sequences.* PhD thesis, University of Amsterdam, 1987.

[Wan99] Yongge Wang. A separation of two randomness concepts. *Information Processing Letters*, 69(3):115–118, 1999.