Computability Theory, Algorithmic Randomness and Turing's Anticipation

Rod Downey[⋆]

School of Mathematics, Statistics and Operations Research,
Victoria University of Wellington,
PO Box 600, Wellington,
New Zealand.
rod.downey@vuw.ac.nz

Abstract. This article looks at the applications of Turing's Legacy in computation, particularly to the theory of algorithmic randomness, where classical mathematical concepts such as measure could be made computational. It also traces Turing's *anticipation* of this theory in an early manuscript.

1 Introduction

Beginning with the work of Church, Kleene, Post and particularly Turing, especially in the magic year of 1936, we know what computation means. Turing's theory has substantially developed under the names of *recursion theory* and *computability theory.* Turing's work can be seen as perhaps the high point in the confluence of ideas in 1936. This paper, and Turing's 1939 paper [141] (based on his PhD Thesis of the same name), laid solid foundations to the pure theory of computation, now called computability or recursion theory. This article gives a brief history of some of the main lines of investigation in computability theory, a major part of Turing's Legacy.

Computability theory and its tools for classifying computational tasks have seen applications in many areas such as analysis, algebra, logic, computer science and the like. Such applications will be discussed in articles in this volume. The theory even has applications into what is thought of as proof theory in what is called reverse mathematics. Reverse mathematics attempts to claibrate the logical strength of theorems of mathematics according to calibrations of comprehension axioms in second order mathematics. Generally speaking most separations, that is, proofs that a theorem is true in one system but not another, are performed in normal "ω" models rather than nonstandard ones. Hence, egnerally

[⋆] Research supported by the Marsden Fund of New Zealand. Some of the work in this paper was done whilst the author was a visiting fellow at the Isaac Newton Institute, Cambridge, UK, as part of the Alan Turing "Semantics and Syntax" programme, in 2012. Some of this work was presented at CiE 2012 in Becher [7] and Downey [42]. Many thanks to Veronica Becher, Carl Jockusch, Paul Schupp, Ted Slaman and Richard Shore for numerous corrections.

such proofs are results in computability theory which yield metamathematical proof theortical corollaries. Discussing reverse mathematics would take us a bit far afield, so we chose not to include this development in the present volume. In the present article, we we will look at the pure theory of computation.

As we later see, computability theory turned out to be the needed mathematical basis for the formalization of the old concept of *randomness* of individual objects. The theory of what we call today *algorithmic randomness* was anticipated by Turing in a manuscript that remained unpublished until its inclusion in the Collected Works [143]. This article reviews the development of the theory of algorithmic randomness as part of Turing's Legacy.

Mathematics has developed many tools to utilize randomness in the development of algorithms and in combinatorial (and other) techniques. For instance, these include Markov Chain Monte Carlo and the Metropolis algorithms, methods central to modern science, the probabilistic method is central to combinatorics. Quantum physics suggests to us that randomness is essential to our understanding of the universe. Computer science uses randomness in cryptography, fast algorithms and proof techniques.

But the key question we need to ask is "What is randomness?". There are some in the physics community that suggest that the universe can generate "true randomness?" which seems a philosophical notion, and this article is *not* concerned with this notion. Here we will be interested in what is called *algorithmic randomness*, which is not a philosophical notion, but a collection of precise mathematical notions.

The underlying idea in this area is that randomness should equate to some kind of inability to describe/predict/compress the random object using algorithmic means. We will use Turing's clarification of the notion of an algorithm to make this precise. For example, if I was presented with a very long string bit by bit, if it was random, then there would seem no way I should be able to predict, algorithmically, what the $n+1$-st bit would be even knowing the first n bits.

The reader should note that this approach abandons the notion of "absolute randomness" since randomness depends on the algorithmic strength of the (algorithmic) predictor. The more powerful the algorithmic device, the fewer strings or reals will be random. The last decade has seen some quite dramatic advances in our understanding of algorithmic randomness. In particular, we have seen significant clarification as to the mathematical relationship between algorithmic computational power of infinite random sources and level algorithmic randomness. Much of this material has been reported in the short surveys Downey [41, 42], Nies [95] and long surveys [40, 47] and long monographs Downey and Hirschfeldt [46] and Nies [94]. Also the book edited by Hector Zenil [152] has a lot of discussion of randomness of varying levels of technicality, many aimed at a general audience.

To give a definition of algorithmic randomness and to understand questions like: "When is one real more random than another? What can be said about the algorithmic power of a random real?" we need a theory of computation. Fortunately this is readily at hand. We know what computation means. The theory has substantially developed, under the names of recursion theory and computability theory. As mentioned earlier, in this book, there are articles on the confluence of ideas in 1936, and the development of the theory at its roots. There are also articles on generalized computation complexity theory and applications of computability theory to algebra and model theory, complexity theory and also to analysis. However, there is none about the pure classical computability theory, underlying such applications and extensions. Thus this article will begin with a brief history of some of the main lines of investigation in this part of Turing's Legacy.

Having done this, we will return to applying the theory to understanding algorithmic randomness.

To our knowledge, whilst he did have the notion of a pseudo-random number generator, Turing himself thought that randomness was a physical phenomenon, and certainly recognized the noncomputable nature of generating random strings. For example, from Turing [142], we have the following quote.

> " An interesting variant on the idea of a digital computer is a "digital computer with a random element." These have instructions involving the throwing of a die or some equivalent electronic process; one such instruction might for instance be, "Throw the die and put the-resulting number into store 1000." Sometimes such a machine is described as having free will (though I would not use this phrase myself)."

John von Neumann (e.g. [147]) also recognized the noncomputable nature of generating randomness.

> "Any one who considers arithmetical methods of producing random digits is, of course, in a state of sin."

Arguably this idea well predated any notion of computation, but the germ of this can be seen in the following quotation of Joseph Bertrand [11] in 1889.

> "How dare we speak of the laws of chance?
> Is not chance the antithesis of all law?"

There has been a developing body of work seeking to understand not just the theory of randomness but how it arises in mathematics; and in physics, biology and the like.

For example, we have also seen an initiative (whose roots go back to work of Demuth [38]) towards using these ideas in the understanding of almost everywhere behaviour and differentiation in analysis (such as Brattka, Miller, Nies [21]). Also halting probabilities are natural and turn up in places apparently far removed from such considerations. For instance, as we later discuss, they turned up naturally in the study of subshifts of finite type (Hochman and Meyerovitch [70], Simpson [124, 126]), fractals (Braverman and Yampolsky [22, 23]). We also know that algorithmic randomness gives insight into Ergodic theory such as Avigad [5], Bienvenu et al. [15] and Franklin et al. [56].

2 Classical computability theory

There are already long books devoted to classical computability theory such as Soare [136], Odifreddi [110, 111], Rogers [117], Griffor [64], and the subject is still being developed. In this section we aim at giving a "once over lightly" with an overview of what we believe are some highlights. As discussed by Sieg, Nerode and Soare in this volume, as well as extensively analysed in Davis [37] and Herken [69], we have seen how Turing's work has led to the birth of computation, and indeed, the digital computer. What about the pure theory of computation after Turing?

The work of Turing [141] led to the notion of relative computation. We imagine a machine M with an oracle (read only memory) A which can be consulted during the computation. This give rise to the the fundamental operator called the *jump* operator: A' is the halting problem with oracle A. Then $\emptyset'$ is the usual halting problem, and $(\emptyset')' = \emptyset^{(2)}$ would be the halting problem given an oracle for the halting problem.

The use of oracles also gives us a basic calibration of the complexity of sets (languages) $A \leq_T B$ means that (the characteristic function of) A can be computed from a machine with oracle access to B. This pre-ordering $\leq_T$ is called *Turing reducibility* and the equivalence classes are called (Turing) degrees.

The jump operator is monotone in the sense that if $X \leq_T Y$ then $X' \leq_T Y'$. Due to the work of Kleene and Post [76], as we discuss below, we know that it is not one to one on the degrees. For example, there are sets $X \not\equiv_T \emptyset$ with $X' \equiv_T \emptyset'$. We call such sets *low*, since we think of them as having low information content because the jump operator cannot distinguish them from having no oracle at all. The spine of the degrees is provided by the jump operator: Start with $\emptyset$ and give it degree $\mathbf{0}$. Form the halting problem $\emptyset'$ and its degree $\mathbf{0}'$. Then $\emptyset^{(2)}$ has degree $\mathbf{0}^{(2)}$. Iterate the construction and obtain any finite ordinal jump. Using what are called effective ordinal notations we can extend this to the transfinite: $\mathbf{0}^{\omega}$ is the effective join of $\mathbf{0}^{(n)}$ for all finite n and then ever upwards. Namely, $\mathbf{0}^{(\omega+1)}$ would be the degree of jump of some representative of $\mathbf{0}^{\omega}$. To work with larger ordinals, what is done is to represent ordinals for α via "notations" which are partial computable functions specifying sequences of smaller ordinals converging to α in the case that α is a limit, and the predecessor of α if α is a successor. In some sense this is the very least one would imagine needed for giving computable representations of ordinals. Remarkably, it is enough, in that for such representations, any two representations for the same ordinal allow us to define $\mathbf{0}^{(\alpha)}$ up to Turing degree, a result of Spector.

Returning to our story, it is certainly the case that Turing's original paper [140] is written very clearly. The same *cannot* be said about much of the early work in the theory of computation, particularly that of Kleene and Church. Most of it was couched in terms of lambda calculus or recursive functions, and it all seemed forbiddingly formal.

A great paper, following the early era of the 30's, was due to Emil Post [113], who returned to Turing's clear informal style. Whilst Turing [141] did define the notion of an oracle computation, it is only in Post's article that the notion of Turing reducibility was defined, and Post focused attention on recursively (=computably) enumerable sets. Post also demonstrated the connection between arithmetical definability and the hierarchies of computability theory, establishing that the n-th jump of the empty set was Σ_n^0 complete, etc. That is, he showed that if A is Σ_n^0 then A is many-one reducible to $\emptyset^{(n)}$, where $X \leq_m B$ means that there is a computable function f with $n \in X$ iff $f(n) \in B$. Many-one reducibility was one of the many refinements of Turing reducibility noted by Post.

Post also suggested the study of the ordering structures generated by Turing reducibilities and by many other refinements of these reducibility. Myhill [109] showed that if X and Y are two versions of the halting problem (for different universal machines) then $X \equiv_m Y$. Post also noted other reducibilities such as truth table reducibility and variants such as bounded truth table, weak truth table, etc. These reducibilities are commonly found in algebra. Truth table reducibility can me thought of as a reduction procedure which must be total for all oracles. It is extensively used in algorithmic randomness as it allows for translations of effective measures. The article of Homer and Selman in this volume discuss how miniaturizations of these ideas gave rise to computational complexity theory. The time bounded version of m-reducibility is used extensively in complexity theory where it is called Karp reducibility.

We concentrate now in describing the work done on Turing reducibilities. The work on other reducibilities is also vast.

2.1 The global degrees

Until mid 1950, it was consistent with all known facts that the ordering of the degrees was a linear ordering of length 2^ω with the countable predecessor property consisting of only iterated jumps of $\emptyset$. However, Kleene and Post [76] showed that this was not the case by exhibiting a pair of degrees $\mathbf{a}, \mathbf{b}$ $(\leq \mathbf{0}')$ which were incomparable. (i.e. $\mathbf{a} \not\leq \mathbf{b}$ and $\mathbf{a} \not\geq \mathbf{b}$, which is written as $\mathbf{a}|_T\mathbf{b}$.) The method of proof introduced by Kleene and Post is a kind of primitive Cohen forcing. Thus, the degrees are a nontrivial upper semi-lattice with join induced by $A \oplus B = \{2n \mid n \in A\} \cup \{2n+1 \mid n \in B\}$. Soon after, Spector [127] proved that there was a minimal degree $\mathbf{a}$; that is $\mathbf{a} > \mathbf{0}$ and for all $\mathbf{c}$, it is not the case that $\mathbf{a} > \mathbf{c} > \mathbf{0}$. This paper implicitly uses another forcing technique which uses perfect closed sets as its conditions. In the same paper Spector proved an "exact pair" theorem showing that all countable ideals could be named by pairs of degrees as the elements below both, and the proof of this influential result introduces forcing with infinite conditions. This exact pair method allows us to show that the degrees are not a lattice.

Each of these two papers had very significant impact on the field. People showed that the degrees were very complicated indeed. The Kleene-Post method enabled the proof of the existence of low sets. This was extended by Friedberg [58] who showed that the range of the jump operator is as big as it can be: the Friedberg Jump Theorem says that if $\mathbf{a} \geq \emptyset'$ there is a degree $\mathbf{c}$ with $\mathbf{c} \cup \mathbf{0}' = \mathbf{c}' = \mathbf{a}$. If $\mathbf{c}' = \mathbf{a}$, we say that $\mathbf{a}$ inverts to $\mathbf{c}$. Friedberg observed a similar result for degrees $\mathbf{d} > \mathbf{0}^{(n)}$. The set C of degree $\mathbf{c}$ that the proof constructs is called 1-generic, meaning that it is Cohen generic for 1 quantifier arithmetic. The inversion to 1-generic degrees is not the only kind. Cooper [31] demonstrated that every degree above $\mathbf{0}'$ can be inverted to a minimal degree. This result uses a combination of the kind of coding used by Firedberg and Spector's methods. These Friedberg-Kleene-Post methods can also be pushed to the transfinite, as proven by Macintyre [98], so that given any $X > \emptyset^{(\alpha)}$, there is a set Y with $Y^{(\alpha)} \oplus \emptyset^{(\alpha)} \equiv_T Y^{(\alpha)} \equiv_T X$, for $\alpha < \omega_1^{CK}$ (the computable ordinals). Applications of such n-generic sets occur in many places in computability theory and its applications in, for instance, effective algebra, and randomness.

Spector's Theorem on minimal degrees was extended by many authors including Lachlan, Lerman, Lachlan-Lebeuf proving results on initial segments showing that these can be, respectively, all countable distributive lattices, countable lattices, and countable upper-semilattices, (see Lerman [88]) culminating in theorems like every partial ordering of size $\aleph_1$ with the countable predecessor property is an initial segment of the Turing degrees (Abraham-Shore [1]). Later it was shown that questions about further extensions often have to do with set theory (Groszek-Slaman [65]). There are still many questions open here. These results imply that the theory of the degrees is undecidable. There has been work towards understanding the quantifier level where undecidabilty occurs. The Kleene-Post theorem and Spector's exact pair theorem also had numerous extensions, heading towards definability results in the degrees, as well as combinations to extensions of embeddings, embeddings with jumps etc. Some noteworthy results here include Slaman-Woodin's proof of the definability from parameters of countable relations in the degrees, this leading to the (parameter-free) definability of the jump operator in the partial ordering of the degrees by Shore and Slaman [131] (Also see Slaman [133]). Still open here is the longstanding question of Rogers: are the Turing degrees rigid?

Related here are results on relativization. Early on it was noted that most results relativized in the sense that if they were true then relativizing everything kept them true. For example, there are sets $A, B < \emptyset'$ with $A|_T B$. This result relativized in that the proof shows that for any oracle X, there are sets $X <_T A^X|_T B^X < X'$. One question was whether "everything relativizes" and, as a consequence, the cones of degrees above each degree would all be isomorphic, or perhaps elementary equivalent? The answer turned out to be no. Beginning with work of Feiner [52] who demonstrated that there were nonisomorphic cones if you had the jump operator, and culminating with work of Shore [128] who showed non-isomorphism in the language of partial ordering, and Shore [129] who demonstrated non-elementary equivalence in the same language.

2.2 Post's Problem and the priority method

Post observed that much of the work of undecidability proofs was in coding halting sets. He called sets A which were domains of partial computable functions *recursively enumerable* and now they are either known by this name or by the name *computably enumerable*, as suggested by Soare, since it captures the intentional meaning, and their degrees similarly. Post asked a very interesting question: Does there exist a computable enumerable degree $\mathbf{a}$ with $\mathbf{0} < \mathbf{a} < \mathbf{0}'$?

This problem became known as Post's Problem, and its solution was highly influential. Post's problem was solved by two students, Friedberg [57] and Muchnik [107]. The method took the Kleene-Post method and added backtracking to give rise to a method known as the *priority method.*

Here is a brief description of the method applied in the setting of an old unpublished result of Tennenbaum. We construct a computable ordering of type $\omega + \omega^*$ with no infinite computable ascending or descending subsequences. We will build the ordering by adding two points at a time. We think of the points in the ω-part as blue and the ones in the ω^* part as red. Thus, if there were nothing happening, we would start with a blue and a red point $x_0 y_0$. At the next stage, we would add a red and a blue point to get $x_0 x_1 y_1 y_0$, etc. Now we must meet certain *requirements*, namely R_e that W_e, the e-th computably enumerable set, is not an infinite ascending sequence and B_e that W_e is not an infinite descending sequence. Lets consider R_e. This is saying that W_e if infinite is not all red. The way to force this to happen would be as follows. Suppose that at some stage we see some point x_n occur in W_e at stage s in its enumeration. Then if we changed the colours at this stage so that x_n was put into the blue section, we would be done since W_e would not be all red. That is, if we had at stage s, $x_0 \dots x_m y_m \dots y_0$, we could *recolour* so that at the next stage we would have $x_0 \dots x_{n-1} x_{s+1} y_{s+1} x_n \dots x_s y_s \dots y_0$, that is *moving* the place we build the sequences to between x_{n-1} and x_n. That is, R_e seeks to make red things blue, and in the same spirit, B_e seeks to make blue things red. Furthermore, we need to make sure that from some point on all elements have as stable colour so that the order type is $\omega + \omega^*$. To to this we give each requirement some kind of *priority.* Say $R_0 < B_0 < R_1 < B_1 \dots$. This means that R_0 has the highest priority and is allowed to make red elements blue, and if it does this, that action is not allowed to be undone. B_0 is allowed to make blue elements red, and this action cannot be undone by any other requirement *except* R_0. If it is undone by R_0 then the next element it makes red (which R_0 does not care about, as it has a satisfying element) will not be made blue by anyone. Finally, to make the order type $\omega + \omega^*$, we also ask that R_e and B_e only are allowed to move elements x_i, y_j for $i, j > e$.

The finite injury method is a mainstay of the area. It has applications in descriptive set theory, complexity theory and most other parts of computable mathematics. One longstanding question here is Sacks' questions about a *degree invariant* solution to Post's Problem. Is there a computably enumerable operator W such that for all X, $X <_T W^X <_T X'$, *and* for all $X \equiv_T Y$, $W^X \equiv_T W^Y$? Lachlan [83] showed that the answer is no if an index for the reductions witnessing $W^X \equiv_T W^Y$ can be read off from indices for the reductions witnessing $X \equiv_T Y$, and Downey and Shore [49] showed that the solution W, if there is one, needs to be reasonably constrained, low$_2$ or high. Martin has conjectured a very strong negative answer which says more or less that the only degree invariant operators on the degrees are jumps and their iterates. Slaman and Steel [134] have the strongest results here, showing, for instance, that there is no order preserving solution.

Powerful generalizations of the finite injury method came from constructions where each requirement could act infinitely often, but subsequent requirements could guess the activity and take it into account. This gave rise to *infinite injury* methods. There is no fixed method and these arguments have many classifications according to "how complex" they are. One method of classification was suggested by Leo Harrington. He said that priority arguments should be classified according to how many iterations of the jump are needed to produce an oracle which could compute how the requirements are satisfied in the construction. Finite injury arguments typically require one jump, and the easiest infinite injury arguments require 2 jumps. However, there are arguments requiring arbitrary numbers of jumps in both the pure theory and in applications such as computable model theory. A significant technical obstacle for such arguments is simply to find a way to coherently present the argument.

The early incarnations of the infinite injury method enabled the proof that the computably enumerable degrees are dense as a partial ordering (Sacks [118]), and that the diamond lattice is embeddable preserving 0 in the computably enumerable degrees. (Lachlan [81], Yates [151]). Sacks also used the method to prove the c.e. jump theorem, namely that if $X \geq \emptyset'$ is c.e. relative to $\emptyset'$ then there is a c.e. set Y with $Y' \equiv_T X$. In the c.e. case of jump theorems it is clearly necessary that "targets" be c.e. relative to $\emptyset'$. Again these results were pushed a long way. All (necessarily countable) distributive and some, but not all, finite non-distributive lattices are embeddable into the computably enumerable degrees (See, for instance, Lachlan [82], Lerman [85], Lempp-Lerman [86], and Lachlan-Soare [84]). Also, many lattices can be embedded densely, such as all distributive lattices (Slaman [132]), and some nondistributive lattices (Ambos-Spies, Hirschfeldt and Shore [2]) but not all embeddable lattices (Downey [39], Weinstein [150]).

We cannot expect that these embeddings can also jump invert, but as partial orderings, we can do embeddings with any reasonable expectation about the jumps consistent with the ordering relationships (Shore [130], building on earlier work of, for example, Robinson [116]). For example, if $\mathbf{a} < \mathbf{b}$ are computably enumerable degrees with jumps $\mathbf{c}$ and $\mathbf{d}$ respectively, and $\mathbf{e} < \mathbf{f}$ are degrees computably enumerable in $\mathbf{0}'$ with $\mathbf{c} \leq \mathbf{e} \leq \mathbf{f} \leq \mathbf{d}$, then there exist c.e. degrees $\mathbf{g}, \mathbf{h}$ with $\mathbf{a} < \mathbf{g}, \mathbf{h} < \mathbf{b}$, and (for example) $\mathbf{g}|\mathbf{h}$ and $\mathbf{g}' = \mathbf{e}$ and $\mathbf{h}' = \mathbf{f}$. Also many results were proven about the structure of the computably enumerable degrees, they are not a lattice, they have an undecidable first order theory, they have algebraic decompositions, etc. There are themes relating definability to enumerations. For example, any arithmetically definable class of computably enumerable degrees closed under double jump is definable in the c.e. degrees. (Nies, Shore and Slaman [96]) There were also a number of results relating the lattice of computably enumerable sets and the upper semi-lattice of computably enumerable degrees. For example, it was shown that maximal sets (that is coatoms in the quotient structure of the computably enumerable sets modulo finite sets-a notion from Post's paper) always have high degrees (meaning $A' \equiv_T \emptyset''$) and include all high degrees; and form an orbit in the automorphism group of the lattice of computably enumerable sets. (Martin [100] and Soare [135], respectively.) Harrington and Soare [68] used the infinite injury method and great ingenuity to show that there is a definable property of the lattice of computably enumerable sets which solves Post's problem. Cholak and Harrington [29] proved a nice definability result for double jump classes in the c.e. degrees. Namely, suppose that $\mathcal{C} = \{\mathbf{a} : \mathbf{a}$ is the Turing degree of a Σ_3 set greater than $\mathbf{0}''\}$. Let $\mathcal{D} \subseteq \mathcal{C}$ such that $\mathcal{D}$ is upward closed. Then there is an non-elementary (Ł$_{\omega_1,\omega}$) $\mathcal{L}(A)$ property $\varphi_{\mathcal{D}}(A)$ such that $D'' \in \mathcal{D}$ iff there is an A where $A \equiv D$ and $\varphi_{\mathcal{D}}(A)$. Double jumps are necessary since Rachel Epstein [51] recently showed that there is a c.e. degree $\mathbf{a}$ which is non-low and each of its members are automorphic to low sets. Cholak, Downey and Harrington [28] recently showed that determining orbits in the automorphism group of the lattice of computably enumerable sets is Σ^1_1 complete.

We refer the reader to Soare [136] for a somewhat dated but well-written account of results up to 1987.

More complex versions of the infinite injury method allowed for very complex results in involving n-th jumps, partial orderings and embeddings such as Lerman-Lempp [87], and things about arithmetical definability such as Harrington [67] (See Odifreddi [111] for this). These methods have been applied by Ash and Knight in effective algebra [3], and model theory (e.g. Marker [99]). The infinite injury method has also been applied in complexity theory such as Downey, Flum, Grohe and Weyer [43]. Modern computability theory could not exist without the infinite injury method.

2.3 Approximation techniques and Π^0_1 classes

A recurrent theme in computability theory is to use computable approximations to complex objects. This can have several forms.

For instance, the Limit Lemma of Shoenfield says that $A \leq_T \emptyset'$ iff there is a computable function $f(\cdot,\cdot)$ such that $\lim_s f(x,s)$ exists for all x and $x \in A$ iff $\lim_s f(x,s) = 1$. That is, A is computable from the halting problem iff A has a computable approximation which changes its mind only finitely often on each argument.

Another important example of approximation comes in the form of what are called the hyperimmune-free or computably dominated degrees. Such a degree $\mathbf{a}$ can be noncomputable but is defined to have the property that for all $f \leq_T \mathbf{a}$, there is a computable g with $f(x) \leq g(x)$ for all x. That is, we can compute a finite number of instances $\{0, \ldots, g(x)\}$ as the possible values of $f(x)$. The non-computably dominated degrees sort of resemble the ones below $\mathbf{0}'$ to some extent, and the class has deep connections with algorithmic randomness. If A has non-computably dominated degree then there is a $f \leq_T A$, which "escapes" any computable function, meaning that if we run a construction with oracle f, then for any computable g, we know that there exist infinitely many n with $f(n) > g(n)$. Thus we run some kind of construction and construct g to measure when we need to perform some action. Then we will argue that $g(n)$ gives the relevant information that $f(n)$ encodes.

If something fails to be approximable then this fact of "escaping" can often be used in constructions. One illustration is the fact that high degrees compute functions that dominate every computable function. This allows us to show that, for instance, every high computably enumerable degree bounds a minimal pair of computably enumerable degrees. (Cooper [32].) Another such example concerns the non-low$_2$ (i.e. $\mathbf{a}'' > \mathbf{0}''$) degrees. Following work of Martin we know that A is non-low$_2$ iff A computes a function f which infinitely often escapes any $\emptyset'$ computable function. (That is, for all $g \leq_T \emptyset'$, $\exists^\infty n(f(n) > g(n))$). This fact enables one to show, for instance, that any finite lattice can be embedded below such degrees preserving 0 and 1, and below and such degree we can find a 1-generic. With some more delicate techniques, such lon-low$_2$ methods can be adapted to the c.e. degrees, allowing lattice embeddings below such degrees, for instance. (for example, Downey and Shore [50]) Work here is ongoing with new precise characterizations of what kinds of approximations allow us to embed, for example, the 5 element modular nondiatributive lattice to be embedded below it, giving new definability results. (Downey-Greenberg [45])

Another recurrent approximation technique is the use of what are called Π^0_1 classes. (Computably bounded) Π^0_1 classes can be thought of as the collections of paths through an infinite computable binary tree. They occur often in effective mathematics. For example, if you give me a computable commutative ring then the ideals can be represented as a Π^0_1 class. The collection of complete extensions of Peano Arithmetic form a Π^0_1 class.

Many *basis* results can be proven for these useful classes. These assert that (perhaps under certain conditions) every nonempty Π^0_1 class has a member of a certain type. The classic result is the *Low Basis Theorem* of Jockusch and Soare [72] which asserts that every Π^0_1 class has a member of low degree (i.e. $A' \equiv_T \emptyset'$) and the Hyperimmune-free Basis Theorem which says it has one of computably dominated degree, and the basis theorem from the same paper that asserts that for every special Π^0_1 class $\mathcal{P}$ (i.e. with no computable members), and every $S \geq_T \emptyset'$, there is a a member $P \in \mathcal{P}$ with $P' \equiv_T S$.

The theory of Π^0_1 classes and algorithmic randomness interact very strongly. For example, the collection of Martin-Löf random reals (for a fixed constant c of randomness deficiency-as defined in the next section) forms a Π^0_1 class with no computable members, and which has positive measure. The basis theorem for special classes above therefore proves that there are random reals of low Turing degree and ones of every possible jump. Thus, tools from topology and from computability theory are simultaneously applicable. Π^0_1 classes and things like reverse mathematics are also intertwined, since Π^0_1 classes correspond to what is called *Weak König's Lemma*. For more we refer the reader to [24].

3 Basics of Algorithmic Randomness

3.1 Notation

We will refer to members of $\{0,1\}^* = 2^{<\omega}$ as *strings*, and infinite binary sequences (members of 2^ω, Cantor space) as *reals.* 2^ω is endowed with the tree topology, which has as basic clopen sets

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

where $\sigma \in 2^{<\omega}$. The *uniform* or *Lebesgue measure* on 2^ω is induced by giving each basic open set $[\sigma]$ measure $\mu([\sigma]) := 2^{-|\sigma|}$. This is simply the restatement that the uniform distribution has all strings of length n equally likely of probability 2^{-n}.

We identify an element X of 2^ω with the set $\{n : X(n) = 1\}$. The space 2^ω is measure-theoretically identical (via the usual mapping taking $[0]$ to $[0, \frac{1][2}{)}$ and $[1]$ to $[\frac{1}{2}, 1)$.) with the real interval $[0, 1)$, although the two are not homeomorphic as topological spaces, so we can also think of elements of 2^ω as elements of $[0, 1]$. We will let $X \upharpoonright n$ denote the first n bits of X.

3.2 von Mises

The theory of randomness of an individual sequence actually pre-dates the foundation of probability theory; and, arguably, one of the reasons for the latter was the unsatisfactory nature of the former until the 60's. The pioneer was Richard von Mises [146]. He said a random real should certainly obey the frequency laws like the law of large numbers for any reasonable sampling of the bits. Thus

$$\lim_{n\to\infty} \frac{|\{m \mid m < n \wedge X(m) = 1\}|}{n} = \frac{1}{2}.$$

This property is called *normality* and was studied by Borel and others. In fact, any random real clearly should be what is called *absolutely normal*, meaning it is normal to any basis (more on this later, when we return to Turing).

Inter alia, we mention that it is here that Turing later enters the picture. His interest was absolute normality and some of his ideas will anticipate those of the theory of algorithmic randomness as developed by Martin-Löf, Kolmogorov, Levin and others. We will return to this development in Section 6.

von Mises' idea was to consider any possible *selection* of a subsequence (i.e. of positions of the given real to sample) and ask that this selection be normal: Let $f : \omega \to \omega$ be an increasing injection, a selection function. Then a random X should satisfy the following.

$$\lim_{n\to\infty} \frac{|\{m \mid m \leq n \wedge X(f(m)) = 1\}|}{n} = \frac{1}{2}.$$

von Mises had no canonical choice for "acceptable selection rules". For example, if we take any real with infinitely many 1's, and make the selection the collection of places where the real is 1, then plainly the real fails to be random relative to that choice according to this criteria. Clearly that selection fails to realize the spirit of von Mises idea. What selection functions should be acceptable? Wald [148, 149] showed that for any *countable* collection of selection functions, there is a sequence that is random in the sense of von Mises. The problem is that von Mises work predated the work in the 30's of Church, Kleene, Post and Turing, culminating in the classic paper of Turing [140], clarifying the notion of computable function. Church [33] proposed restricting f to computable increasing functions. This incarnation of von Mises' idea gives rise to notions now called *computable stochasticity*, and, of we use partial computable selections, *partial computable stochasticity.*

Unfortunately, von Mises' approach, even with Church's reformulation using computability theory, was fatally injured (or at least seriously hurt) by the work of Ville [145]. In the following, $S(\alpha, n)$ is the number of 1's in the first n bits of α and similarly S_f for the selected places.

Theorem 1 (Ville's Theorem [145]). *Let E be any countable collection of selection functions. Then there is a sequence $\alpha = \alpha_0\alpha_1\dots$ such that the following hold.*

1. $\lim_n \frac{S(\alpha,n)}{n} = \frac{1}{2}$.
2. *For every $f \in E$ that selects infinitely many bits of α, we have* $\lim_n \frac{S_f(\alpha,n)}{n} = \frac{1}{2}$.
3. *For all n, we have* $\frac{S(\alpha,n)}{n} \leq \frac{1}{2}$.

The killer is item 3 which says that there are *never* situations with more 1's than 0's in the first n bits of α. Suppose you were betting on the outcomes of a sequence of coin tosses of a biased coin, where there are always fewer tails then heads. Certainly you could figure out a betting strategy to make a lot of money. This is the import of item 3.

Ville suggested adding a further statistical law, the law of the iterated logarithm, to von Mises' definition. However, we might well ask "How we can be sure that adding this law would be enough?". Why should we expect there not to be a further result like Ville's (which there is, see [46]) exhibiting a sequence that satisfies both the law of large numbers and the law of the iterated logarithm, yet clearly fails to have some other basic property that we would naturally associate with randomness?

We could add more and more statistical laws to our collection of desiderata for random sequences, but there is no reason to believe we would ever be done, and we certainly do not want a definition of randomness that changes with time, if we can avoid it.

3.3 Martin-Löf

One solution to this quandary was provided by the work of Per Martin-Löf [101], and as we later see somewhat anticipated by Turing. Martin-Löf's fundamental idea in [101] was to define an abstract notion of a performable statistical test for randomness, and require that a random sequence pass *all* such tests. He did so by effectivizing the notion of a set of measure 0. The way to think about Martin-Löf's definition below is that as we effectively shrink the measure of the open sets we regard as "tests", we are specifying reals satisfying them more and more.

In the below, a Σ^0_1 class is a computably enumerable collection $\{[\sigma] \mid \sigma \in W\}$ for some computably enumerable (c.e.) set W of strings. Alternatively think of this as a c.e. set of intervals in the interval $[0, 1]$.

Definition 1 (Martin-Löf [101]).

1. *A* Martin-Löf test *is a sequence* $\{U_n\}_{n\in\omega}$ *of uniformly* Σ^0_1 *classes such that* $\mu(U_n) \leq 2^{-n}$ *for all* n.
2. *A class* $C \subset 2^\omega$ *is* Martin-Löf null *if there is a Martin-Löf test* $\{U_n\}_{n\in\omega}$ *such that* $C \subseteq \bigcap_n U_n$.
3. *A set* $A \in 2^\omega$ *is* Martin-Löf random *if* $\{A\}$ *is not Martin-Löf null.*

For example, think of the test that every second bit of the real is 1. It is okay for a random real to have this for a long time but at some stage it must abandon having every second bit 1. Thus we could specify this test by $U_1 = \{[01]\}$, $U_2 = \{[0001], [0101]\}$, etc. Even at this point we would like to make the reader aware of the calibrations of randomness possible. This test consists only of nested sequences of clopen sets. Thus any randomness notion defined by:

"X is random iff it passes all Martin-Löf tests but restricted to tests where each level is specified by a computable function giving a canonical index for a clopen set"

would be enough to pass this test and any "similar" tests. This "clopen" notion has a name and is called *Kurtz* or *weak* randomness. It is equivalent to saying X is in every Σ^0_1 class of measure 1.

3.4 Three approaches to algorithmic randomness

The modern viewpoint has three main paradigms for defining algorithmic randomness. Martin-Löf's above is called the *measure-theoretical paradigm.*

We briefly discuss the two other main paradigms in algorithmic randomness as they are crucial to our story. The first is the *computational paradigm* : Random sequences are those whose initial segments are all hard to describe, or, equivalently, hard to compress.

We think of Turing machines U with input τ giving a string σ. We regard τ as a description of σ and the shortest such is regarded as the intrinsic information in σ. Kolmogorov [77] defined *plain U-Kolmogorov complexity* $C_U(\sigma)$ of σ as the *length* of the shortest τ with $U(\tau) = \sigma$. Turing machines can be enumerated $U_0, U_1, \ldots$ and hence we can remove the machine dependence by defining a new (universal) machine

$$U(0^e 1\tau) = U_e(\tau),$$

so that we can define for this machine M, $C(\sigma) = C_M(\sigma)$ and for all e, $C(\sigma) \leq C_{U_e}(\sigma) + e + 1$. We will use the notation $\leq^+$ to dispense with explicit mention of absolute additive constants in inequalities. For example, this inequlaity would be written as $C(\sigma) \leq^+ C_{U_e}(\sigma)$.

A simple counting argument due to Kolmogorov [77] shows that as $C(\sigma) \leq^+ |\sigma|$ (using the identity machine), there must be strings of length n with $C(\sigma) \geq n$. We call such strings C-*random*. The intuition here is that the only way to describe σ would be to hardwire σ into the program. σ is *incompressible* and, in particular, has *no* regularities to allow for compression.

We would like to define a real, an infinite sequence, to be random by saying for all n, $C(\alpha \upharpoonright n) \geq^+ n$. Unfortunately, there are no such random reals due to a phenomenon called complexity oscillations, which (in a quantitative way) say that in very long strings σ there must be segments with $C(\sigma \upharpoonright n) < n$. This oscillation really due to the fact that on input τ, we don't just get the *bits* of τ as information but the *length* of τ as well.

Specifically, imagine a sufficiently long string α. Now each initial segment σ of α has some shortest programme say σ^*. Now this program can be interpreted as a number $n^* = \sigma^*$. Consider τ the next segment of α after σ (i.e. $\sigma\tau \preceq \alpha$) *with τ having length n^*.* Then the program that has input a string ρ and does the following. First it looks at its length and interprets this as a string. Taking that strings ν it computes $U(\nu)$ and if this halts outputs $U(\nu)\rho$.

Now assume we run this algorithm on τ. Then it computes $U(\sigma^*) = \sigma$ first, and then outputs $\sigma\tau$. This shows that $C(\sigma\tau) =^+ |\tau| =^+ C(\sigma)$. For long enough σ this is a compression. The key here is that we are using the *length* as well as the *bits* of τ. Thus we are losing the intentional meaning that the bits of τ are processed by U to produce σ. To get around this first Levin [89, 90] and later Chaitin [25] suggested using *prefix-free machines* to capture this intentional meaning that the *bits* of the input encode the information of the output.

One way is to use *prefix-free complexity* via machines whose domains are prefix-free sets of strings. That is, prefix free machines work like telephone numbers. If $U(\tau)\downarrow$ (i.e. halts) then for all $\hat{\tau} \neq \tau$ comparable with τ, $U(\hat{\tau})\uparrow$.

Already we see a theme that there is not one but perhaps *many* notions of computational compressibility of relevance to understanding randomness. In the case of prefix-free complexity, in some sense we know we are on the correct track, due to the following theorem which can be interpreted as saying (for discrete spaces) that Occam's razor and Bayes' Theorem give the same result (in that the shortest description is essentially the probability that the string is output).

Theorem 2 (Coding Theorem-Levin [89, 90], Chaitin [25]). *For all σ, $K(\sigma) =^+ -\log(Q(\sigma))$ where $Q(\sigma)$ is $\mu(\{\tau \mid U(\tau) = \sigma\})$, and of course, logs here are base 2.*

Using this notion, and noticing that the universal machine above would be prefix-free if all the U_e were prefix free, we can define the prefix-free Kolmogorov complexity $K(\sigma)$.

Definition 2 (Levin [90], Chaitin [25]). *A set A is* 1-random *if* $K(A \upharpoonright n) \geq^{+} n$.

Schnorr proved that we are on the right track here:

Theorem 3 (Schnorr). *A real A is Martin-Löf random iff it is 1-random.*

It is not difficult to show that almost all reals are random, but Schnorr's Theorem give no explicit example. The oft-quoted example of a 1-random real is Chaitin's *halting probability* (for a universal prefix-free machine U):

$$\Omega = \sum_{\{\sigma | U(\sigma)\downarrow\}} 2^{-|\sigma|},$$

the measure of the domain of U (which is well-defined as the domain of U is a prefix free set of strings).

An easy proof of this fact is reminiscent of the fact that the halting problem is undecidable. We can use the Recursion Theorem to build part of the universal prefix-free machine U, via a prefix-free machine M with (known) coding constant e in U. Imagine we are monitoring $\Omega_s = \sum_{\{\sigma|U(\sigma)\downarrow\}}[s]$. Suppose that we see some $\sigma \preceq \Omega_s \upharpoonright s$ we see $K_s(\sigma) < |\sigma| - e - 1$. (That is, this segment does not look random.) This means that some ν of length $K_s(\sigma)$ has been enumerated into the domain of U describing σ. Then what we do is enumerate the same ν into the M_{s+1} describing σ also. Then, it follows that $\Omega_{s+1} \geq \Omega_s + 2^{-|\nu|}$, and in particular $\sigma \not\preceq \Omega_{s+1}$. The fact that Ω has a prefix-free domain means that M does too as we are simply recycling what U does.

The summary is "if the opponent says here's a short description of an initial segment of Ω_s, we act to show that it is not an initial segment after all."

It would seem that the definition of Ω is thoroughly machine dependent, but in the same spirit as Myhill's Theorem (showing that there is only one halting problem up to m-degree), we can define a reducibility on halting probabilities we call *Solovay reducibility.* $X \leq_S Y$ means there is a constant c and partial computable f such that for all rationals $q < Y$, $f(q) \downarrow < X$ and $c(Y - q) > (X - f(q))$. To wit, a good approximation for Y yields one for X. There is essentially one Ω: The approximation $\Omega = \lim_s \Omega_s$ is monotone from below, and Ω is what is called a *left c.e.-real.* Every left c.e. real is a halting probability in the same way that each c.e. set is a the domain of a Turing machine. Clearly $\leq_S$ is well-defined on left c.e. reals. The culmination of a series of papers is the Kučera-Slaman Theorem which states that there is really only one left-c.e. random real.

Theorem 4 (Kučera-Slaman Theorem [79]). *A left c.e. real α is* 1*-random iff for all left c.e.-reals β, $\beta \leq_S \alpha$.*

3.5 Martingales

The final randomness paradigm is the one based on prediction. The *unpredictability paradigm* is that we should not be able to predict the next bit of a random sequence even if we know all preceding bits, in the same way that a coin toss is unpredictable even given the results of previous coin tosses.

Definition 3 (Levy [92]). *A function $d : 2^{<\omega} \to \mathbb{R}^{\geq 0}$ is a* martingale[1] *if for all σ,*

$$d(\sigma) = \frac{d(\sigma 0) + d(\sigma 1)}{2}.$$

d is a supermartingale *if for all σ,*

$$d(\sigma) \geq \frac{d(\sigma 0) + d(\sigma 1)}{2}.$$

A (super)martingale d succeeds *on a set A if $\limsup_n d(A \upharpoonright n) = \infty$. The collection of all sets on which d succeeds is called the* success set *of d, and is denoted by $S[d]$.*

The idea is that a martingale $d(\sigma)$ represents the capital that we have after betting on the bits of σ while following a particular betting strategy ($d(\lambda)$ being our starting capital). The *martingale condition* $d(\sigma) = \frac{d(\sigma 0)+d(\sigma 1)}{2}$ is a fairness condition, ensuring that the expected value of our capital after a bet is equal to our capital before the bet. Ville [145] proved that the success sets of (super)martingales correspond precisely to the sets of measure 0.

Now again we will need a notion of effective betting strategy. We will say that the martingale is computable if d is a computable function (with range $\mathbb{Q}$, without loss of generality), and we will say that d is c.e. iff d is given by an effective approximation $d(\sigma) = \lim_s d_s(\sigma)$ where $d_{s+1}(\sigma) \geq d_s(\sigma)$. This means that we are allowed to bet more as we become more confident of the fact that σ is the more likely outcome in the betting, as time goes on. The following result was anticipated in Ray Solomonoff's approach to randomness as discussed in e.g. Downey-Hirschfeldt [46].

Theorem 5 (Schnorr [119, 120]). *A set is 1-random iff no c.e. (super)martingale succeeds on it.*

Schnorr argued that Theorem 5 showed that perhaps the notion of Martin-Löf randomness was not really capturing the notion of *effective randomness* as it was intrinsically *computably enumerable.* Schnorr argued that it seems strange that to define randomness we use c.e. martingales and not computable ones. Based on this possible defect, Schnorr defined two other notions of randomness, *computable randomness* (where the martingales are all computable) and Schnorr randomness (where we use the Martin-Löf definition but insist that $\mu(U_k) = 2^{-k}$ rather than $\leq 2^{-k}$ so, in particular, we know precisely the $[\sigma]$ in U_k uniformly in k and $[\sigma]$) meaning in each case that the randomness notion is a computable rather than a computably enumerable one. We know that Martin-Löf randomness

[1] A more complex notion of martingales (which are called martingale processes) is used in probability theory. We refer the reader to [46], where it is discussed how computable martingale processes can be used to characterize 1-random reals.

implies computable randomness which implies Schnorr randomness, and none of these implications are reversible.

It seemed that Ville's Theorem was a fatal blow to von Mises' program. However, there seems to be a possible resurrection. Can we define 1-randomness using computable martingales somehow? The answer is "possibly" if we allow *non-monotonicity*. The idea is to use a computable *but non-monotonic* notion of randomness, where we have a betting strategy which bets on bits one at a time, but instead of being increasing, we can bet in some arbitrary order, and need not bet on all bits. The order is determined by what has happened so far. This gives a notion called *Kolmogorov-Loveland* (or non-monotonic) randomness and the following question has been open for quite a while.

Question 1 (Muchnik, Semenov, and Uspensky [108]). Is every non-monotonically random sequence 1-random?

4 Developments

The theory of algorithmic randomness has been widely developed. First many variations of the notions of a random real or string have been introduced. We have already seen three, Kurtz, Schnorr and computable randomness. Each of these notions has its own applications and gives its own insight into the level of randomness needed for measuring the randomness of some process.

4.1 Randomness is the same as differentiability

There has been quite a bit of recent work relating "almost everywhere" behaviour in analysis to levels of randomness. This is a program going back to the work of Oswald Demuth, a constructivist from Prague.

Here we will be working in computable analysis, a subject going back to Turing's original paper [140]. This area will be discussed in detail in the article by Avigad and Brattka in this volume. Briefly, if we are doing computable analysis, then we need some representation of the individual objects we will be analysing. For example a computable real is essnetially a computablely convergent Cauchy sequence. That is, we have $\{q_n \mid n \in \omega\}$ with limit α and for all j, we can effectively compute $g(j)$ such that $|\alpha - q_{g(j)}| < 2^{-j}$. In a function space, we would have a set of effectively described functions, such as polynomials with rational coefficients, effectively converging to the function in the sense of the appropriate norm. Then a typical definition of a computable function on such a space is that if I can approximate x in the input to within 2^{-j}, then I can compute a similar approximation to $f(x)$. When formalized this is implicit in Turing's orginal paper, and now usually called "type 2" computability. An alternative and weaker notion of computable function is that it takes computable reals to computable reals. This is referrred to as Markov computability. The reader is referred to the Avigad-Brattka paper for more details and the history of the development of the area.

Using such a setting, as an example of the relationship between randomness and differntiability, recall that the Denjoy upper and lower derivatives for a function f are defined as follows.

$$\overline{D}f(x) = \limsup_{h \to 0} \frac{f(x) - f(x+h)}{h} \text{ and } \underline{D}f(x) = \liminf_{h \to 0} \frac{f(x) - f(x+h)}{h}.$$

The Denjoy derivative exists iff both of the above quantities exist and are finite. The idea in this is that slopes like those in the definitions can be considered to be martingales. Using this for one direction, various notions of randomness can be characterized by (i) varying the strength of the notion of computable real valued function (e.g. Markov computable, type 2 computable etc) (ii) varying the theorem.

For an illustration, we have the following.

Theorem 6 (Brattka, Miller and Nies [21]). *z is computably random iff every computable (in the type two sense) increasing function $f[0,1] \to \mathbb{R}$ is Denjoy differentiable at z.*

There are similar results relating 1-randomness of z to the differentiability of functions of bounded variation at z. There is still a lot of activity here, and classes like Lipschitz functions and many other classical almost everywhere behaviour in analysis are found to correlate to various notions of randomness. The paper [21] is an excellent introduction to this material.

Other almost everywhere classical behaviour comes from ergodic theory. There is a great deal of current work exploring the relationship between ergodic theory and algorithmic randomness. The simplest example is an old theorem of Kučera which says that if $\mathcal{C}$ is a Π^0_1 class of measure 1, then for any 1-random X there must be a $Y \in \mathcal{C}$ with the "tail" of X in Y. By this we mean that there is some n which that for all $m > n$, $X(n) = Y(n)$. This is related to ergodic theory as can be seen by an analysis of the the Poincaré Ergodic Theorem. To wit, let (X, μ) be a probability space, and $T : X \to X$ measure preserving so that for measurable $A \subseteq X$, $\mu(T^{-1}A) = \mu(A)$. Such a map is called T-*invariant* if $T^{-1}A = A$ except on a measure 0 set. Finally the map is *ergodic* if every T-invariant subset is either null or co-null. The shift operator on Cantor space is the mapping $T(a_0a_1\ldots) = a_1a_2\ldots$ is an ergodic action with the Bernoulli product measure. The "tail" map above can be thought of as a statement of a statement about the shift operator.

A classic theorem of Poincaré is that if T is ergodic on (X, μ), then for all $E \subseteq X$ of positive measure and *almost all* $x \in X$, $T^n(x) \in E$ for infinitely many n. For a set E of measurable subsets of X, we call an x a *Poincaré point* if $T^n(x) \in Q$ for all $Q \in E$ of positive measure. Restating the theorem Kučera [78] we see the following: X is 1-random iff X is a Poincaré point for the shift operator with respect to the collection of effectively closed subsets of 2^ω.

Bienvenu et al. proved the following extension of this result.

Theorem 7 (Bienvenu, et al. [15]). *Let T be computable ergodic on a computable probability space (X, μ). Then $x \in X$ is 1-random iff x is a Poincaré point for all effectively closed subsets of X.*

Again there is a lot of ongoing work here. For instance, one exciting development has seen the applications of algorithmic randomness to *symbolic dynamics* a sub-area of ergodic theory, with well-known applications in additive number theory and analysis. A d-dimensional *shift* of finite type is a collection of colourings of $\mathbb{Z}^d$ defined by local rules and a shift action (basically saying certain colourings are illegal). Its (Shannon) *entropy* is the asymptotic growth in the number of legal colourings. More formally, consider $G = (\mathbb{N}^d, +)$ or $(\mathbb{Z}^d, +)$, and A a finite set of symbols. We give A the discrete topology and A^G the product topology. The *shift action* of G on A^G is

$$(S^g x)(h) = x(h + g), \text{ for } g, h \in G \wedge x \in A^G.$$

A *subshift* is $X \subseteq A^G$ such that $x \in X$ implies $S^g x \in X$ (i.e. shift invariant). *Symbolic Dynamics* studies subshifts usually of "finite type." The following is a recent theorem showing that Ω occurs naturally in this setting.

Theorem 8 (Hochman and Meyerovitch, [70]). *The values of entropies of subshifts of finite type over $\mathbb{Z}^d$ for $d \geq 2$ are exactly the complements of halting probabilities.*

In this area, Jan Reimann [114] gave a new and simpler proof of a classical theorem called Frostman's Lemma. An even more notable example is due to Simpson [126]. Simpson studies topological entropy for subshifts X and the relationship with Hausdorff dimension.

Here we pause to mention that, in the same way that we can suggest that an individual sequence can be thought to be random, the theory of effective Hausdorff dimension allows us to give an individual sequence effective *dimension.*

After effectivizing the the whole theory of Hausdorff using effective versions of "weighted" inner and outer "measures," it turns out that there are simple characterizations of these notions in terms of Kolmogorov complexity.

Mayordomo [102] proves that effective Hausdorff dimension of X is equal to $\liminf_{n \to \infty} \frac{K(X \upharpoonright n)}{n}$. Athreya, Hitchcock, Lutz, and Mayordomo [4] proved that the effective packing dimension is $\limsup_{n \to \infty} \frac{K(X \upharpoonright n)}{n}$ (C can replace K in both cases).

Again, there has been a long line of development seeking to understand algorithmic dimension. An easy way to make something of effective Hausdorff dimension $\frac{1}{2}$ is to take a 1-random real and "thin it out' by inserting a 0 in every second position. A longstanding question was whether in some sense this was necessary: could randomness be extracted from any a real of nonzero effective Hausdorff dimension? Miller [105] showed that the answer is no. A strong negative answer to this question could also be obtained by constructing a real of minimal Turing degree and of effective Hausdorff dimension 1, but this remains an open question. For packing dimension, either a Turing degree has only elements of effective packing dimension 0, or the sup of the packing dimensions of the members is 1 (Fortnow, Hitchcock, Aduri, Vinochandran, and Wang [53]). However, Conidis [30] showed that the degree did not need to have a real of effective packing domension 1. In the case of effective Hausdorff dimension, Zimand [153] proved that domension 1 can be extracted from two independent sources of nonzero Hausdorff entropy.

Looking at one use of these notions, we return to Simpson's work. If $X \subset A^G$ use the standard metric $\rho(x, y) = 2^{-|F_n|}$ where n is as large as possible with $x \upharpoonright F_n = u \upharpoonright F_n$ and $F_n = \{-n, \dots, n\}^d$. In discussions with co-workers, Simpson proved that the classical dimension equals the entropy (generalizing a difficult result of Furstenburg 1967) using effective methods, which were much simpler.

Theorem 9 (Simpson [126]). *If X is a subshift (closed and shift invariant), then the effective Hausdorff dimension of X is equal to the classical Hausdorff dimension of X is equal to the entropy, moreover there are calculable relationships between the effective and classical quantities. (See Simpson's home page for his recent talks and more precise details.)*

There are many other investigations looking into other Kolmogorov complexities, resource bounded versions such as polynomial time randomness, and the like, and randomness in other spaces than Cantor space. We will finish with a short section exploring the work of the last decade which seeks to understand how computability and randomness relate.

5 Computability and Randomness

Interactions of measure, randomness and computability go back to the early years of the study of degrees of unsolvability. The classical paper was de Leeuw et al. [35] where, amongst other things, it is proven that a set X is computably enumerable from a set of oracles of positive measure iff X is computably enumerable. As a consequence, we get a result later rediscovered by Sacks that if a real X is computable from a collection of sources of positive measure, then X must be computable. Nevertheless, another classical result is the following saying that 1-random sources can have computational power.

Theorem 10 (Kućera [78], Gács [61]). *For every set X, there is a 1-random Y such that $X \leq_{wtt} Y$, where $\leq_{wtt}$ is Turing reducibility with use bounded by a computable function.*

Theorem 10 argues that 1-random reals are not random enough to correlate to the thesis that random reals should have no computational power. This intuition was clarified by Stephan who proved the following[2].

[2] Interpreted by Hirschfeldt as saying that there are two methods of passing a stupidity test. One is the be the genuine article. The other is to be like Ω and be so smart that you know what a stupid person would say.

Theorem 11 (Stephan [137]). *Suppose a random real is powerful enough to compute a $\{0,1\}$-valued function f such that for all n, $f(n) \neq \varphi_n(n)$ (i.e. of PA degree). Then $\emptyset' \leq_T X$, so that it is a "false random."*

There is a lot of material on Chaitin's Omega suggesting that it is the "number of knowledge" and this has something to do with randomness. The result above more or less says that if you are a knowledgeable random then you are essentially code such an Omega. A remarkable theorem here is the following, demonstrating a deep relationship between PA degrees and random degrees (i.e. degrees containing randoms).

Theorem 12 (Barmpalias, Lewis, and Ng [6]). *Every PA degree is the join of two* 1*-random degrees.*

We can strengthen the idea of randomness by giving the computational devices more compression power via Turing's notion of an oracle. Then if $\emptyset^{(n)}$ denotes the n-th iterate of the halting problem, we say that X is $n+1$-random iff $K^{\emptyset^{(n)}}(X \upharpoonright n) \geq^+ n$ for all n. A pretty result proven by Miller and Yu [106] is that if $X \leq_T Y$ are both 1-random and X is n-random, so is Y.

We can similarly do this with other notions of randomness with a little care. For notions like Schnorr randomness we need stronger reducibilities reflecting the "totality" of the tests.

There are also many results concerning the relationships between the randomness notions and Turing (and other) degrees. For example, it can be shown that X is weakly 2-random (i.e. in every Σ^0_2 class of measure 1) iff X is 1-random and its degree forms a minimal pair with $\emptyset'$ (Downey, Nies, Weber, and Yu [48] plus Hirschfeldt and Miller (in [48]) for the hard direction). Hence no (weakly) 2-random real can bound a PA degree.

It is a surprising fact that for all n, n-randomness can be defined purely in terms of K with no oracle. This follows by the next result.

Theorem 13 (Bienvenu, Muchnik, Shen, and Vereschagin [14]). $K^{\emptyset'}(\sigma) = \limsup_m K(\sigma \mid m) \pm O(1)$.

Hence A is 2-random iff for all n, $\limsup_m K(A \upharpoonright n \mid m) \geq^+ n$. For a small number of n, we know of "natural" definitions of n-randomness. For instance, we have seen that it is impossible for a real to have $C(X \upharpoonright n) \geq^+ n$ for *all* n, but Martin-Löf showed in his original paper [101] that there are reals X with $C(X \upharpoonright n) \geq^+ n$ for *infinitely many* n, and that these are all 1-random. Joe Miller [103] and later Nies, Stephan and Terwijn [97] showed that such randoms are precisely the 2-randoms, and later Miller [104] showed that the 2-randoms are exactly those that achieve maximal prefix-free complexity (which is $n + K(n)$) infinitely often. Also Becher and Gregorieff [10] have a kind of index set characterizations of higher notions of randomness. I know of no other natural definitions, such as for the 3-randoms. There has been a huge amount of work concerning the interplay between things like PA degrees and weakenings of the notion of fixed point free functions (that is, functions with $f(n) \neq \varphi_n(n)$ for all

n). For example, you can show that this ability corresponds to traceing, and the speed of growth of the initial segment complexity of a real. As an illustration, A is *h-complex* if $C(A \upharpoonright n) \geq h(n)$ for all n. A is *autocomplex* if there is an A-computable order h such that A is h-complex, where an order is a nondecreasing unbounded function with $f(0) \geq 1$..

Theorem 14 (Kjos-Hanssen, Merkle, and Stephan [73]). *A set is autocomplex iff it is of DNC degree.*

Another illustration of the interplay of notions of randomness and Turing degrees is the following theorem.

Theorem 15 (Nies, Stephan, and Terwijn [97]). *If a nonhigh set (i.e. $A' \not\geq_T \emptyset^{(2)}$) is Schnorr random then it is 1-random.*

On the other hand, it is possible to show that within the high degrees the separations between computable, Schnorr, and Martin-Löf randomness all occuri ([97]). In the hyperimmune-free degrees, weak randomness coincides with all of these as well as weak 2-randomness. So the degree can have great effect on what a notion of randomness means.

One long sequence of results concerns lowness and randomness. For any reasonable property P we say that X is *low for P* if $P^X = P$. For example, being low for the Turing jump means that $X' \equiv_T \emptyset'$. A set A is low for 1-randomness iff A does not make any 1-randoms nonrandom. That is, if Y is 1-random then Y^A is $A-1$-random. Normally we would expect an oracle A would enable us to compress some intital segment of Y for some Y allowing us to derandomize it. You can also have a notion of lowness for tests, meaning that every (effective nullset)^A can be covered by an effective nullset. In all cases the lowness notion for randomness and for tests have turned out to coincide with a single recent exception of "difference randomness" found by Diamondstone and Franklin (paper in preparation).

Now it is not altogether clear that noncomputable sets low for 1-randomness should exist. But they do and form a remarkable class called the K-trivials which had earlier and independently been defined purely in terms of their initial segment complexity. That is, the reals low for Martin-Löf randomness coincide with the class of reals A such that for all n, $K(A \upharpoonright n) \leq^+ K(n)$. (In fact Bienvenu and Downey [12] showed that it is enough to put a Solovay function[3] in place of $K(n)$.) Many properties of this class have been shown. The coincidence of these two concepts lowness and triviality is one of the jewels of the area. It was Andre Nies who proved the deep result that A is K-trivial iff A is low for Martin-Löf randomness iff A is useless as a compressor, meaning that for all σ, $K^A(\sigma) =^+ K(\sigma)$. (Nies [93]). A good account of this material can be found in Nies [94, 95], but things are constantly changing, with perhaps seventeen

[3] That is a computable function f with $f(m) \geq K(m)$ for all m and $f(n) =^+ K(n)$ infinitely many n. See also Beinvenu and Merkle [13] and Hölzl, Kräling, and Merkle [71].

characterizations of this class at present. We also refer to [46] for the situation up to mid-2010.

Other randomness notions give quite different lowness notions. For example, X is low for C, meaning $C^X =^+ C$, iff X is computable (essentially Chaitin [26]), and similarly Y is low for computable randomness iff Y is computable (Nies [93]). reals which are low for C nor any low for computable randomness. On the other hand, lowness for Schnorr and Kurtz randomness give interesting subclasses of the hyperimmune-free degrees characterized by notions of being computably dominated, and fixed point free functions in the case of Kurtz. (This is a detailed story with many references, begining with the beautiful paper of Terwijn and Zambella [139], seee Downey and Hirschfeldt [46]) Work here is still ongoing and many results have been proven, but the pattern remains very opaque. Even for a fixed real like Ω (i.e. when does Ω^X remain random?) results are quite interesting. In the case of Ω, X is low for Ω and X is computable from the halting problem, then X is K-trivial, but there are *random* reals low for Ω. In fact, X is 1-random and low for Ω, iff X is 2-random. (Result of Joe Miller, see [46].)

These classes again relate to various refinements of the jump and to "traceing" which means giving an effective collection of *possibilities* for (partial) functions computable from the degree at hand. Again this idea has taken on a life of its own, and such notions have been used to solve questions from classical computability theory. For instance, Downey and Greenberg [44] used "strong jump traceability" to solve a longstanding question of Jockusch and Shore on pseudo-jump operators and cone avoidance. Strongly jump traceable reals have their own techniques and theory and form a fascinating class, see e.g. [27].

We should also mention the the deep results of Reimann and Slaman who were looking at the question (first discussed by Levin):

"Given $X \not\equiv_T \emptyset$, is there a measure relative to which X is random?"

Clearly we can trivially answer Levin's question: every real is, we can concentrate a measure on a real. But clearly what is asked is for the situation where we are not allowed to do this concentration. If we allow atoms, then the answer is still that that every noncomputable real can be made random. On the other hand, if we ask that there are no atoms in the measure, the situation is very different. We get a nonempty class of *never continuously n-random* reals. For each n this class is countable, but the proof of this requires magical things like big fragments of Borel determinacy, *provably.* This metamathematical aspect of the answer seems strange in that the definitions of Martin-Löf randomness only needs a couple of quantifiers and hence we would expect a low level answer. But no. So algorithmic randomness not only interacts strongly with computability theory but also even with set theory. Reimann and Slaman's results use techniques involving models of ZFC and "master codes". The reader should look at Reimann and Slaman [115].

6 Turing

What has this got to do with Turing? What was the anticipation we alluded to? Certainly, the very notion of algorithmic randomness needs the notion of algorithm and arguably there is this weak connection: Turing clarified the notion of algorithm. However, there is something rather more remarkable than that.

We return to the notion of (absolute) normality. Recall that X is Borel normal to base n if we represent X in base n, then for all $0 \leq i \leq n-1$,

$$\lim_s \frac{|\{X(k) = i \mid k \leq s\}|}{s} = \frac{1}{n}.$$

This notion was defined by Emil Borel in 1909. We have seen that variations of the notion of normality were the basis of von Mises approach to defining randomness.

Interestingly, normality *precisely defines* an algorithmic randomness notion. To give a machine (randomness) definition of normality, we change the computational device to that of a finite automaton. A real number is normal to a base b if, and only if, no finite state gambler can make infinite winnings when betting on its base b expansion, as we see more explicitly in Theorem 18 below. (See [121, 36, 18].)

Borel demonstrated that almost every real is absolutely normal, but asked the questions "How can we construct an *explicit* absolutely normal number?" and "Can a real be normal relative to one base and not another?"

Normality is a longstanding area of research in number theory. It is also yet another area of number theoretical research where the questions rapidly outrun our ability to prove theorems. For example, it is unknown if familiar reals like e and π and the like are normal to *any* base.

Returning to Borel's questions, how should we interpret "explicit construction of a normal number"? With the machinery of computability theory developed by Turing, Church, Kleene and others, we have at least one interpretation. From the material of the previous sections, it is obvious that Ω is normal. However, in some sense, this is cheating since it is not a computable, but a c.e. object so is hardly an *explicit construction.*

In an unpublished manuscript, Turing attacked the question of an explicit construction of an absolutely normal number by interpreting "explicit" to mean *computable.* His manuscript entitled *"A note on normal numbers"*, presumably written in 1938, presents the best answer to date to Borel's first question: an algorithm that produces absolutely normal numbers. This early proof of existence of computable normal numbers remained largely unknown because Turing's manuscript was only published in 1997 in his Collected Works, edited by J. L. Britton [143]. The editorial notes say that the proof given by Turing is inadequate and speculate that the theorem could be false. In [8] Becher, Figueira and Picci reconstructed and completed Turing's manuscript, trying to preserve his ideas as accurately as possible and correcting minor errors.

As Becher [7] remarks, the very first examples of normal numbers were independently given by Henri Lebesgue and Waclaw Sierpiński[4] in 1917 [80, 123]. They also lead to computable instances by giving a computable reformulation of the original constructions [9]. Together with Turing's algorithm these are the only known constructions of computable normal numbers. It is pretty clear that Turing was unaware of the limiting constructions given in [80, 123].

What does Turing's construction do? His paper says the following:

> Although it is known that almost all numbers are [absolutely] normal no example of [an absolutely] normal number has ever been given. I propose to show how [absolutely] normal numbers may be constructed and to prove that almost all numbers are [absolutely] normal constructively.

I won't reproduce Turing's construction, save to say that it makes an ingenious extension of the law of large numbers to blocks, and basically makes a low complexity Martin-Löf type test. The details can be found in Becher [7], for instance.

What Turing actually does is something very modern. He develops an effective version of measure theory (sound familiar?) and demonstrates that the reals which are *not* absolutely normal have *computable measure 0*. Therefore, there must be a computable real which is absolutely normal. Here is what Jack Lutz said of this in a lecture at the conference *Computability, Complexity and Randomness, 2012* at Cambridge:

> Placing computability constraints on a nonconstructive theory like Lebesgue measure seems a priori to weaken the theory, but it may strengthen the theory for some purposes This vision is crucial for present-day investigations of
> - individual random sequences,
> - dimensions of individual sequences,
> - measure and category in complexity classes, etc.

7 From a modern perspective

How have investigations into normality played out? Using polynomial martingales and hence a polynomial notion of randomness, we have the following.

Theorem 16 (Strauss [138]). *There are absolutely normal numbers computable in exponential time.*

Using a more delicate construction, Elvira Mayordomo brought the complexity of an *explicit* computable absolutely normal real down as follows.

Theorem 17 (Mayordomo). *We can construct an absolutely normal number in time $O(n \log n)$.*

[4] Both published their works in the same journal issue, but Lebesgue's dates back to 1909, immediately after Borel's question.

As mentioned earlier, this is all related to the theory of *finite state compressors* and the corresponding notion of dimension.

Definition 4 (Schnorr and Stimm [121]).

1. *A finite state gambler is a quadruple* $G = (Q, \delta, q_0, B)$ *where* (Q, δ, q_0) *is a finite state automaton, and* $B : Q \to \Delta_{\mathbb{Q}}(\Sigma)$ *is a betting function, where* $\Delta_{\mathbb{Q}}(\Sigma)$ *is the collection of rational-valued probability measures on* Σ*.*
2. *A* martingale *is a function* $d_G : \Sigma^* \to [0, \infty)$ *with* $d_G(\lambda) = 1$ *(Here* λ *is the empty string), and again the fairness condition:*

$$d_G(wa) = |\Sigma| d_G(w) B(\delta(w))(a).$$

3. *for* $s \in [0, \infty)$ *the* s-gale *of* G *is* $d_G^{(s)}(w) = 2^{(s-1)|w|} d_G(w)$*.*

As usual, we say that d succeeds on X if $\limsup_{n\to\infty} d(X \upharpoonright n) = \infty$ and that d succeeds strongly if $\liminf_{n\to\infty} d(X \upharpoonright n) = \infty$. Then the *finite state dimension* of a real X is

$$\dim_F S(X) = \inf\{s \in [0, \infty) \mid \exists \text{ finite state } G \text{ s.t. } d_G^{(s)} \text{ succeeds on } X\}.$$

By a theorem of Dai, Lathrop, Lutz and Moyordomo [36], this quantity equals the infimum over all finite state compressors F of

$$\liminf_{n\to\infty} \frac{C_F(X \upharpoonright n)}{n \log \Sigma},$$

aligning with the definition met before for effective Hausdorff dimension, with a similar formula holding for effective string dimension below. There is a similar definition for strong dimension and strong success.

The theorem is the following.

Theorem 18 (Schnorr and Stimm [121]). *X is normal base b iff the base b finite state dimension of X is 1.*

There is a very active program concerned with the analysis of finite state dimensions. Many modern text compressors such as ZIP are examples of finite state compressors so this theory seems quite pertinent to applications. We refer the reader to Dai et al. [36] and to Lutz's home page for much more on this topic, and its relationship to things like DNA self-assembly.

References

1. U. Abraham and R. Shore. Initial segments of the Turing degrees of size $\aleph_1$. *Israel Journal of Mathematics*, Vol. 55 (1986), 1-51.
2. K. Ambos-Spies, D. Hirschfeldt, and R. Shore Undecidability and 1-types in intervals of the c.e. degrees. *Annals of Pure and Applied Logic*, Vol. 106 (2000), 1-48.

3. Chris Ash and Julia Knight. *Computable Structures and the Hyperarithmetical Hierarchy.* Elsevier, 2000.
4. K. Athreya, J. Hitchcock, J. Lutz, and E. Mayordomo. Effective strong dimension in algorithmic information and computational complexity. *SIAM Jour. Comput.*, 37 (2007), 671–705.
5. J. Avigad. The metamathematics of ergodic theory. *Annals of Pure and Applied Logic,* 157 (2009), 64-76.
6. G. Barmpalias, A. Lewis, and K. M. Ng. The importance of Π^0_1 classes in effective randomness. JSL, 75(1) (2010), 387–400.
7. V. Becher. Turing's normal numbers: towards randomness. In S.B. Cooper, A.Dawar, B. Löwe (eds.), CiE 2012, Lecture Notes in Computer Science 7318: 35-45 Springer, Heidelberg, 2012.
8. V. Becher, S. Figueira, R. Picchi. Turing's unpublished algorithm for normal numbers. *Theoretical Computer Science* 377 (2007), 126–138.
9. V. Becher and S. Figueira. An example of a computable absolutely normal number. *Theoretical Computer Science* 270 (2002), 947–958.
10. V.Becher, S.Grigorieff. From index sets to randomness in $\emptyset^n$, Random reals and possibly infinite computations. *Journal of Symbolic Logic,* 74:1 (2009), 124–156.
11. J. Bertrand, *Calcul des Probabilités*, 1889.
12. L. Bienvenu and R. Downey. Kolmogorov complexity and Solovay functions. in STACS 2009, 147–158.
13. L. Bienvenu and W. Merkle. Reconciling data compression and Kolmogorov complexity. In ICALP 2007, Lecture Notes in Computer Science 4596. Springer, 2007.
14. L. Bienvenu, An. A. Muchnik, A. Shen, and N. Vereshchagin. *Limit complexities revisited*, in STACS 2008.
15. L. Bienvenu, A. Day, M. Hoyrup, I. Mezhirov, and A. Shen. Ergodic-type characterizations of algorithmic randomness. To appear in *Information and Computation.*
16. Émil Borel. Les probabilités dénombrables et leurs applications arithmétiques. *Rendiconti del Circolo Matematico di Palermo* 27 (1909), 247–271.
17. Émil Borel. *Leçons sur la thèorie des fonctions.* Gauthier Villars, 2nd ed. 1914.
18. C. Bourke, J. Hitchcock, N. Vinodchandran, Entropy rates and finite-state dimension. *Theoretical Computer Science* 349(3) (2005), 392–406.
19. Yann Bugeaud, Nombres de Liouville et nombres normaux, *Comptes Rendus de l'Académie des Sciences de Paris* 335 (2002), 117–120.
20. Yann Bugeaud, *Distribution Modulo One and Diophantine Approximation*, Cambridge University Press. 2012.
21. V. Brattka, J. Miller, and A. Nies, *Randomness and differentiability*, to appear.
22. M. Braverman and M. Yampolsky, Non-Computable Julia Sets. *Journ. Amer. Math. Soc.*, Vol. 19(3), 2006
23. M. Braverman and M. Yampolsky, *Computability of Julia Sets*, Springer-Verlag, 2008.
24. Douglas Cenzer and Carl Jockusch. Π^0_1 classes - Structure and applications. In *Contemporary Mathematics* Vol. 257 (2000), 39-59.
25. G. Chaitin. A theory of program size formally identical to information theory. *Journal of the ACM,* Vol. 22 (1975), 329–340.
26. G. Chaitin. Information-theoretical characterizations of recursive infinite strings. *Theoretical Computer Science*, Vol. 2 (1976), 45–48.

27. P. Cholak, R. Downey, and N. Greenberg. Strong-jump traceablilty. I. The computably enumerable case. *Advances in Mathematics,* Vol. 217 (2008) 2045–2074.
28. Peter Cholak, Rod Downey and Leo Harrington. On the orbits of computably enumerable sets. *Journal of the American Mathematical Society* Vol. 21, No. 4 (2008), 1105-1135.
29. Peter Cholak and Leo Harrington. On the definability of the double jump in the computably enumerable sets. *J. Math. Log.,* Vol. 2(2) (2002), 261-296,
30. C. Conidis. A real of strictly positive effective packing dimension that does not compute a real of effective packing dimension one. *Journal of Symbolic Logic,* Vol. 77(2) (2012), 447–474.
31. Barry Cooper. Minimal degrees and the jump operator. *Journal of Symbolic Logic*, Vol. 38 (1973), 249-271.
32. Barry Cooper. Minimal pairs and high recursively enumerable degrees. *JṠymbolic Logic* Vol. 39 (1974) 655-660.
33. Alonzo Church. On the concept of a random sequence. *Bulletin of the American Mathematical Society,* Vol. 46 (1940), 130–135.
34. R. Cilibrasi, P.M.B. Vitanyi, and R. de Wolf. Algorithmic clustering of music based on string compression. *Computer Music J.,* Vol. 28 (2004), 49-67.
35. K. de Leeuw, E. F. Moore, C. E. Shannon, and N. Shapiro. Computability by probabilistic machines. In C. E. Shannon and J. McCarthy, editors, *Automata studies*, number 34 in Annals of Mathematics Studies, pages 183–212. Princeton University Press, Princeton, N. J., 1956.
36. L. Dai, J. Lutz, E. Mayordomo. Finite-state dimension. *Theoretical Computer Science* 310 (2004), 1–33.
37. Martin Davis. *Computability and Unsolvability.* Dover, 1985.
38. O. Demuth. The differentiability of constructive functions of weakly bounded variation on pseude-numbers. *Comment. Math. Univ. Carolina*, Vol. 16 (1975), 583-599.
39. R. Downey. Lattice nonembeddings and initial segments of the recursively enumerable degrees. *Annals Pure and Appl. Logic*, Vol. 49 (1990), 97-119.
40. R. Downey. Five Lectures on Algorithmic Randomness. In *Computational Prospects of Infinity, Part I: Tutorials* (Ed. C. Chong, Q. Feng, T. A. Slaman, W. H. Woodin and Y. Yang) Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore Vol 14, World Scientific, Singapore, 2008, 3-82.
41. R. Downey. Algorithmic randomness and computability. In *Proceedings of the 2006 International Congress of Mathematicians,* Vol 2, *European Mathematical Society*, (2006), 1-26.
42. Rod Downey. Randomness, Computation and Mathematics. In S.B. Cooper, A.Dawar, B. Löwe (eds.), CiE 2012, Lecture Notes in Computer Science 7318: Springer, Heidelberg, 2012.
43. R. Downey, J. Flum, M. Grohe and M. Weyer. Bounded fixed-parameter tractability and reducibility, *Annals of Pure and Applied Logic* Vol. 148 (2007), 1-19.
44. R. Downey and N. Greenberg, Pseudo-jump operators and SJTHard sets. to appear *Advances in Mathematics.*
45. R. Downey and N. Greenberg, *A Transfinite Hierarchy of Computably Enumerable Degrees, Unifying Classes, and Natural Definability.* monograph in preparation.

46. R. Downey and D. Hirschfeldt, *Algorithmic Randomness and Complexity*, Springer-Verlag, 2010.
47. Downey, R., D. Hirschfeldt, A. Nies, and S. Terwijn. Calibrating randomness. *Bulletin Symbolic Logic* Vol. 12 (2006), 411-491.
48. R. Downey, A. Nies, R. Weber, and L. Yu. Lowness and Π^0_2 nullsets. *The Journal of Symbolic Logic*, Vol. 71 (2006), 1044–1052.
49. R. Downey and R. Shore. Thre is no degree invariant half-jump. *Proc. AMS*, Vol. 125 (1997), 3033-3037.
50. R. Downey and R. Shore. Lattice embeddings below a non-low$_2$ recursively enumerable degree. *Israel Journal of Mathematics*, Vol. 94 (1996), 221-246.
51. Rachel Epstein. The nonlow computably enumerable degrees are not invariant in E. *Trans. Amer. Math. Soc.*, to appear.
52. Lawrence Feiner. The strong homogeneity conjecture. *J. Symb. Logic,* Vol. 35 (1970), 373–377.
53. L. Fortnow, J. Hitchcock, P. Aduri, V. Vinochandran, and F. Wang. Extracting Kolmogorov complexity with applications to dimension zero-one laws. In *Proceedings of the 33rd International Colloquium on Automata, Languages, and Programming* (*ICALP 2006*), Lecture Notes in Computer Science 4051, pages 335–345. Springer, 2006.
54. W. Fouche. The descriptive complexity of Brownian motion. *Advances in Mathematics*, Vol. 155 (2000), 317–343.
55. W. Fouche. Dynamics of a generic Brownian motion: Recursive aspects. *Theoretical Computer Science*, Vol. 394 (2008), 175-186.
56. J. Franklin, N. Greenberg, J. Miller and Keng Meng Ng, Martin-Loef random points satisfy Birkhoff's ergodic theorem for effectively closed sets. to appear, *Proc. Amer. Math. Soc.*
57. R. Friedberg Two recursively enumerable sets of incomparible degrees of unsolvability. *Proc. Nat. Acad. Sci. U.S.A.*, Vol. 43 (1957), 236–238.
58. R. Friedberg. A criterion for completeness of degrees of unsolvability. *Journal of Symbolic Logic,* Vol. 22 (1957), 159-160.
59. H. Fuchs and C. Schnorr. Monte Carlo methods and patternless sequences. In *Operations Research Verfahren*, Vol XXV, Symp. Heidelberg, 1977, 443-450.
60. P. Gács. On the relation between descriptional complexity and algorithmic probability. *Theoretical Computer Science,* Vol. 22 (1983), 71–93.
61. P. Gács. Every set is reducible to a random one. *Information and Control,*, Vol 70, (1986), 186–192.
62. P. Gács, M. Hoyrup and C. Rojas. Randomness on computable probability spaces, a dynamical point of view. to appear, *Theory of Computing Systems.*
63. S. Gregorieff and M. Ferbus, *Is Randomness native to Computer Science? Ten years after* in [152], (2011) 243-263.
64. E Griffor. *Handbook of Computability Theory.* Elsevier, 1999.
65. Marcia J. Groszek and Theodore A. Slaman. Independence results on the global structure of the Turing degrees. *Trans. Amer. Math. Soc.,* Vol. 277(2) (1983), 579-588.
66. Hardy,G.H., Wright,E.M. 1979. *An Introduction to the Theory of Numbers.* Oxford University Press. First edition in 1938.
67. Leo Harrington. MacLachlin's Conjecture. Handwritten Notes 1970's.
68. Leo Harrington and Robert Soare. Post's Program and incomplete recursively enumerable sets. *Proc. Natl. Acad. of Sci. USA*, Vol. 88 (1991), 10242-10246.
69. Rolf Herken. *The Universal Turing Machine: A Half-Century Survey.* Springer-Verlag, 1995.

70. M. Hochman and T. Meyerovitch. A characterization of the entropies of multidimensional shifts of finite type. *Annals of Mathematics*, Vol. 171 (2010), No. 3, 2011-2038

71. R. Hölzl, T. Kräling, and W. Merkle. *Time bounded Kolmogorov complexity and Solovay functions*, in MFCS 2009, volume 5734 of *Lecture Notes in Computer Science*, pages 392–402. Springer, 2009.

72. Carl Jockusch and Robert Soare. Degrees of members of Π^0_1 classes. *Pacific J. Math.*, Vol. 40 (1972), 605-616.

73. B. Kjos-Hanssen, W. Merkle, and F. Stephan. *Kolmogorov complexity and the recursion theorem*, in STACS 2006, LNCS 3884, 149–161. Springer.

74. B. Kjos-Hanssen and A. Nerode, *Effective dimension of points visited by Brownian motion* Theoretical Computer Science 410 (2009), no. 4-5, 347-354.

75. B. Kjos-Hanssen and T. Szabados, *Kolmogorov complexity and strong approximation of Brownian motion,* Proc. Amer. Math. Soc. 139 (2011) no. 9, 3307-3316.

76. Stephen Kleene and Emil Post. The upper semi-lattice of degrees of recursive unsolvability. *Annals of Mathematics* Vol. 59 (3)(1954), 379–407

77. A. N. Kolmogorov, *Three approaches to the quantitative definition of information*, Problems of Information Transmission, 1 (1965), 1–7.

78. A. Kučera. *Measure, Π^0_1 classes, and complete extensions of PA*, In *Recursion Theory Week*, volume 1141 of *Lecture Notes in Mathematics*, pages 245–259, Oberwolfach, 1984, 1985. Springer, Berlin.

79. A. Kučera and T. Slaman, *Randomness and recursive enumerability*, SIAM J. on Comp., 31 (2001), 199–211.

80. Henri Lebesgue. Sur certaines démonstrations d'existence. *Bulletin de la Société Mathématique de France* 45 (1917), 132–144.

81. A. Lachlan. Lower bounds for pairs of recursively enumerable degrees. *Proc. Lond. Math. Soc.*, Vol. 16, (1966), 537-569.

82. A. H. Lachlan. Embedding nondistributive lattices in the recursively enumerable degrees. In *Conference in Mathematical Logic—London '70 (Proc. Conf., Bedford Coll., London, 1970)*, pages 149–177. Lecture Notes in Math., Vol. 255. Springer, Berlin, 1972.

83. A. Lachlan. Uniform enumeration operators. *J. Symb. Logic*, Vol. 40 (1975), 401-409.

84. A. H. Lachlan and R. Soare. Not every finite lattice is embeddable in the recursively enumerable degrees. *Advances in Mathematics* , Vol. 37 (1980), 7482.

85. Manuel Lerman. The embedding problem for the recursively enumerable degrees. In *Recursion theory (Ithaca, N.Y., 1982)*, volume 42 of *Proc. Sympos. Pure Math.*, pages 13–20. Amer. Math. Soc., Providence, RI, 1985.

86. Steffen Lempp and Manuel Lerman. A finite lattice without critical triple that cannot be embedded into the enumerable Turing degrees. *Ann. Pure Appl. Logic*, 87(2):167–185, 1997. Logic Colloquium '95 Haifa.

87. Steffen Lempp and Manuel Lerman. The decidability of the existential theory of the poset of the recursively enumerable degrees with jump relations. *Advances in Mathematics,* 1996.

88. M. Lerman. *Degrees of Unsolvability: Local and Global Theory.* Springer-Verlag, 1983.

89. L. Levin. *Some theorems on the algorithmic approach to probability theory and information theory*, Dissertation in Mathematics Moscow University, 1971.

90. L. Levin. Laws of information conservation (non-growth) and aspects of the foundation of probability theory. *Problems of Information Transmission,* Vol. 10 (1974), 206–210.
91. Levin, M.B. 1979. On absolutely normal numbers. English translation in *Moscow University Mathematics Bulletin* 34:32–39.
92. P. Lévy. *Théorie de l'Addition des Variables Aléatoires.* Gauthier-Villars, 1937.
93. A. Nies. Lowness properties and randomness. *Advances in Mathematics*, Vol. 197, 1 (2005), 274-305..
94. A. Nies, *Computability and Randomness,* Oxford University Press, 2009.
95. A. Nies. Interactions of computability and randomness. In *Proceedings of the International Congress of Mathematicians,* (S. Ragunathan, ed.) 30-57 (2010).
96. A. Nies, R. Shore and T. Slaman. Interpretability and definability in the Recursively Enumerable Degrees. *Proc. London Math. Soc.* Vol. 77 (1998), 241-291.
97. A. Nies, F. Stephan, and S. A. Terwijn. Randomness, relativization, and Turing degrees. *Journal of Symb. Logic.*, Vol. 70(2) (2005), 515–535.
98. Angus Macintyre. Transfinite iterations of Friedberg's Completeness Criterion. *Journal of Symbolic Logic*, Vol. 38 (1977), 1-10.
99. David Marker. Degrees of models of true arithmetic. In *Proceedings Herbrand Symposio,* (J. Stern, ed.), North-Holland, Amsterdam, (1982), 233-242.
100. D. A. Martin. Classes of recursively enumerable sets and degrees of unsolvability. *Z. Math. Logik Grundlag. Math.*, Vol. 12 (1966), 295–310.
101. P. Martin-Löf. The definition of random sequences. *Information and Control,* Vol. 9 (1966) 602–619.
102. E. Mayordomo. A Kolmogorov complexity characterization of constructive Hausdorff dimension. *Infor. Proc.Lett.*, Vol. 84 (2002), 1–3.
103. J. Miller. Kolmogorov random reals are 2-random. *The Journal of Symbolic Logic*, Vol. 69(3) (2004), 907–913.
104. J. Miller. The K-degrees, low for K-degrees, and weakly low for K sets. *Notre Dame Journal of Formal Logic*, Vol. 50(4) (2010), 381–391.
105. J. Miller. Extracting information is hard: a Turing degree of non-integral effective Hausdorff dimension. *Advances in Mathematics.* Vol. 226(1) (2011), 373384.
106. J. S. Miller and L. Yu. On initial segment complexity and degrees of randomness. *Transactions of the American Mathematical Society*, 360(6) (2008), 3193–3210.
107. A. Muchnik. On the unsolvability of the problem of reducibility in the theory of algorithms. *Dokl. Akad. Nauk. S.S.S.R.*, Vol. 106 (1956) 194–197.
108. An. A. Muchnik, A. Semenov, and V. Uspensky, Mathematical metaphysics of randomness. *Theor. Comp. Sci.* Vol. 207(2) (1998), 263–317.
109. J. Myhill. Creative sets. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, Vol. 1 (1955), 97-108.
110. P. Odifreddi. Classical Recursion Theory, Vol. 1 North-Holland, 1989.
111. P. Odifreddi. Classical Recursion Theory, Vol. 2 North-Holland, 1999.
112. M. Poul-El and I. Richards, *Computability in Analysis and Physics*, Springer-Verlag, 1989.
113. Emil Post. Recursively enumerable sets of positive integers and their decision problems. *Bulletin of the American Mathematical Society* Vol. 50 (5) (1944), 284–316.
114. J. Reimann. Effectively closed classes of measures and randomness. *Annals of Pure and Applied Logic,* Vol. 156(1) (2008), 170–182.
115. J. Reimann and T. Slaman, *Randomness for continuous measures*, to appear. (draft available from Reimann's web site.)

116. Robert Robinson. Jump restricted interpolation in the recursively enumerable degrees. *Annals of Math.*, Vol. 93 (3) (1971), 586-596.
117. Hartley Rogers, Jr. *Theory of Recursive Functions and Effective Computabilty.* McGraw-Hill, 1967.
118. Gerald Sacks. The recursively enumerable degrees are dense. *Annals of Mathematics*, Vol. 80 (1964), 300–312.
119. C. P. Schnorr. A unified approach to the definition of a random sequence. *Mathematical Systems Theory,* Vol. 5 (1971), 246–258.
120. C. P. Schnorr, *Zufälligkeit und Wahrscheinlichkeit. Eine algorithmische Begründung der Wahrscheinlichkeitstheorie*, volume 218 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin–New York, 1971.
121. C. P. Schnorr and H. Stimm. Endliche Automaten und Zufallsfolgen. *Acta Informatica* 1 (1972), 345–359.
122. W. M. Schmidt. On normal numbers. *Pacific Journal of Math.* 10 (1960), 61–672.
123. Waclaw Sierpiński. Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'un tel nombre. *Bulletin de la Société Mathématique de France* 45(1917), 127–132.
124. S. Simpson. Medvedev Degrees of 2-Dimensional Subshifts of Finite Type. to appear, *Ergodic Theory and Dynamical Systems.*
125. S. Simpson. Mass Problems Associated with Effectively Closed Sets. to appear *Tohoku Mathematical Journal.*
126. S. Simpson, *Symbolic Dynamics: Entropy = Dimension = Complexity* (2011) to appear.
127. Clifford Spector. On the degrees of recursive unsolvability. *Annals of Mathematics,* Vol. 64 (1956), 581-592.
128. Richard Shore. The homogeneity conjecture. *Proceedings of the National Academy of Sciences USA,* Vol. 76 (1979), 4218-4219.
129. Richard Shore. On homogeneity and definability in the first order theory of the Turing degrees. *Journal of Symbolic Logic*, Vol. 47 (1982), 8-16.
130. Richard Shore. A non-inversion theorem for the jump operator. *Annals of Pure and Applied Logic*, Vol. 40 (1988), 277-303.
131. Richard Shore and Theodore Slaman. Defining the Turing jump. *Mathematical Research Letters*, Vol. 6 (1999), 711722.
132. Theodore A. Slaman. The density of infima in the recursively enumerable degrees. *Ann. Pure Appl. Logic,* Vol. 52(1-2) (1991), 155-179.
133. Theodore A. Slaman. Global properties of the Turing degrees and the Turing jump. In *Computational prospects of infinity. Part I. Tutorials,* volume 14 of Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap., pages 83-101. World Sci. Publ., Hackensack, NJ, 2008.
134. Theodore Slaman and John Steel. Definable functions on degrees. In *Cabal Seminar 81-85*, Volume 1333 of Lecture Notes in Math., pages 37-55. Springer, Berlin, 1988.
135. Robert Soare Automorphisms of the lattice of recursively enumerable sets I: maximal sets. *Ann. of Math.*, Vol. 100 (1974), 80-120
136. Robert Soare. *Recursively Enumerable Sets and Degrees.* Springer-Verlag, 1987.
137. F. Stephan. Martin-Löf random sets and PA-complete sets. In *Logic Colloquium '02*, volume 27 of *Lecture Notes in Logic*, 342–348. Association for Symbolic Logic, 2006.
138. Martin Strauss. Normal numbers and sources for BPP. *Theoretical Computer Science* 178 (1997), 155-169.

139. S. Terwijn and D. Zambella. Algorithmic randomness and lowness. *The Journal of Symbolic Logic*, Vol. 66 (2001), 1199–1205.

140. A. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society,* Vol. 42 (1936), 230–265, 1936. Correction in *Proceedings of the London Mathematical Society,* Vol. 43 (1937), 544–546.

141. A. Turing. Systems of logic based on ordinals. *Proc. Lond. Math. Soc.*, (2) Vol. 45 (1939), 161-228.

142. A. Turing. Computing machinery and intelligence. *Mind,* Vol. 59 (1950), 433-460.

143. A. Turing. A note on normal numbers. In J.L.Britton, editor *Collected Works of A.M. Turing: Pure Mathematics.* North Holland, Amsterdam, 1992, 117–119, with notes of the editor in 263–265.

144. P. Vitanyi. Information distance in multiples. *IEEE Trans. Inform. Theory,* Vol. 57:4(2011), 2451-2456.

145. J. Ville, *Étude Critique de la Notion de Collectif*, Gauthier-Villars, 1939.

146. R. von Mises. Grundlagen der Wahrscheinlichkeitsrechnung. *Math. Z.*, Vol. 5 (1919), 52–99.

147. J. von Neumann. Various techniques used in connection with random digits, in in *Monte Carlo Method.* In (A.S. Householder, G.E. Forsythe, and H.H. Germond, editors), National Bureau of Standards Applied Mathematics Series, vol. 12 1951: 36-38.

148. A. Wald. Sur le notion de collectif dans la calcul des probabilitiés. *Comptes Rendes des Seances de l'Académie des Sciences,* Vol. 202 (1936), 1080–1083.

149. A. Wald. Die Weiderspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung. *Ergebnisse eines mathematischen Kolloquiums,* 8 (1937), 38–72.

150. B. Weinstein. *On embeddings of the 1-3-1 lattice into the recursively enumerable degrees.* PhD thesis, University of California, Berkeley, 1988.

151. C. Yates. A minimal pair of recursively enumerable degrees. *J. Symb. Logic*, (1966), 159-168.

152. H. Zenil, *Randomness Through Computation: Some Answers, More Questions,* (Hector Zenil editor), World Scientific, Singapore, 2011.

153. M. Zimand. Two sources are better than one for increasing the Kolmogorov complexity of infinite sequences. In E. Hirsch, A. Razborov, A. Semenov, and A. Slissenko, editors, *Computer Science—Theory and Applications*, Lecture Notes in Computer Science 5010, pages 326–338. Springer, Berlin, 2008.