1-22-1996

Measure on P: Strength of the Notion

Martin Strauss
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/cs_techreports
Part of the Theory and Algorithms Commons

Recommended Citation
http://lib.dr.iastate.edu/cs_techreports/169

This Article is brought to you for free and open access by the Computer Science at Iowa State University Digital Repository. It has been accepted for inclusion in Computer Science Technical Reports by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Measure on P: Strength of the Notion

Martin Strauss*
Department of Computer Science
Iowa State University
Ames, IA 50011-1040
mstrauss@cs.iastate.edu

January 22, 1996

Abstract

We give a notion of measure on P that overcomes some limitations of earlier formulations. In the process, we investigate the significance for resource-bounded measure of the choice of the lexicographic ordering of the words.

1 Introduction

Resource-bounded measure was introduced by Lutz in [8]. Intuitively, this theory gives a notion of big and small to sets of languages. In recent years this tool has been used with many successes to illuminate the structure of complexity classes, notably E and $E_2 = \text{EXP}$ [9].

The theory of resource-bounded measure is a parametrized tool. For many complexity classes $C$, one plugs $C$ into the general theory, and one gets out a notion of measure at $C$, in which each singleton set of a $C$ language is small, but $C$ itself is not small.

Unfortunately, Lutz’s formulation only works directly for $C \supseteq \text{E}$. Generalizing Lutz’s notion, in [1] the authors introduced a notion of measure on P, PSPACE, and other subexponential classes. This notion satisfies many nice theoretical properties, and has some applications to BPP, but provides too few measurable sets for some purposes.

In this paper we provide a new notion of measure on P and other classes between P and E. The new formulation uses, as building blocks, aspects of previous developments. In making this formulation, our primary motivation is to provide as many small sets as possible while satisfying the essential properties of measure, the “measure axioms.” As a secondary motivation, to the extent this is possible, we want the new formulation to be compatible with Lutz’s formulation at the exponential level. To these ends, we prove that the measure axioms are satisfied, and when we limit the available small sets or deviate from Lutz’s formulation we prove (or at least give strong intuition) that this is necessary.

In Section 2 we introduce notation and sketch the formulations of Lutz’s exponential measure and the earlier weak subexponential measure. In Section 3 we discuss the significance for measure of the choice of the standard lexicographic ordering of $\Sigma^*$. In Section 4 we give a formulation, stronger than that of [1], that covers more sets that “ought” to be small. This formulation is then shown to be too strong, since it violates a measure axiom. In Section 5 we provide a new, satisfactory notion of measure on P, that preserves many of the small sets covered by the notion of Section 4. Finally, in Section 6 we show that our notion of measure on PSPACE is incomparable to the notion of measure in [10].

*Research supported by NSF grant CCR-9204874 and NSF grant CCR-9157382.
2 Preliminaries

First we sketch the formulation of resource-bounded measure at \( E \) from [8], then review the weak generalization to measure at \( P \) from [1]. Finally, we give a simple counting argument that will be used in many proofs of this paper.

2.1 Measure at \( E \)

Fix an enumeration \( s_i \) of \( \Sigma^* \). For now we follow [8] in arbitrarily using the standard lexicographic order, but in subsequent sections we will discuss the significance of this choice. We identify a language \( L \) with its characteristic sequence \( \chi_L \), where by definition the \( i \)th bit of \( \chi_L \) is 1 if the \( i \)th word \( s_i \) is in \( L \). For a sequence \( \omega \), we will write \( \omega[i] \) for the \( i \)th bit of \( \omega \), and we will overload this notation in many ways: \( \omega[i..j] \) and \( \omega[s_i..s_j] \) denote the \( i \)th through \( j \)th bits, \( \omega[A] \) denotes the bits indexed by a set \( A \) of indices \( i \) or words \( s_i \), and \( \omega[s_i] \) or \( \omega[i] \) will denote \( \omega[0..(i-1)] \), i.e., the memberships of words preceding \( s_i \). If \( w_1 \) is a string and \( w_2 \) is a string or (infinite) sequence extending \( w_1 \), we write \( w_1 \sqsubseteq w_2 \). The empty string will be denoted \( \lambda = s_0 \), and we write \( \text{pos}(x) = i \) if \( x = s_i \). Note there are approximately \( 2^{2|w|} \) predecessors of \( x \), so \( |L_x| \approx 2^{2|w|} \).

Definition 1 A supermartingale is a function \( d: \Sigma^* \to \mathbb{R}_+ \), where the input to \( d \) is regarded as a prefix of a characteristic sequence of a language, such that \( d \) satisfies the following “average inequality:"

\[
 d(w) \geq \frac{d(w0) + d(w1)}{2}. \tag{1}
\]

The success set \( S_\infty[d] \) of a supermartingale \( d \) is the set of languages \( L \) such that

\[
 \limsup_{w \sqsubseteq L} d(w) = \infty.
\]

A supermartingale satisfying (1) with equality is called a martingale. (The distinction between martingales and supermartingales is not important for defining measure on classes at least as large as \( E \), but will be important for subexponential measure.)

One can regard a supermartingale as a betting strategy. For example, the supermartingale fragment defined by \( d(\lambda) = 1, d(0) = 0 \) and \( d(1) = 2 \) corresponds to a bet of all our capital that a random language contains the first word. If \( L \) is a specific language that indeed contains the first word, then we “double our money along \( L \).” This can be regarded as a “detailed verification” that the set of languages extending the characteristic string \( w = \lambda \) has measure at most \( d(\lambda)/d(w) = 1/2 \). Similarly, the set of languages on whose characteristic sequences \( d \) becomes unbounded has measure zero.

To define measure at the level of \( E = \text{DTIME}(2^{\text{linear}}) \), Lutz considers supermartingales computable in \( P \). That is, the supermartingales run in time polynomial in the length \( |w| \) of their input \( w = L_x \), where \( |w| = 2^{2|w|} \), so the supermartingales run in time \( E \) of \( x \). We will denote the collection of such supermartingales \( \Delta(E) \). There are some issues regarding the representation of real numbers, but these have largely been solved [8, 2], and in this paper we simply assume that all reasonable arithmetic is allowed. As a final bit of notation, the variable \( n \) will always stand for \( |x| = |s_{|w|}| \approx \log |w| \), so we will sometimes write “\( d \) runs in time \( 2^{-n} \)” to mean \( d(w) \) runs in time \( 2^{-|w|} \).

The objects we measure are sets of languages. We will use the term “language” and the variable \( L \) for sets of words, and reserve “set” and variables \( A, B, C \) for sets of languages.

Definition 2 A set \( A \) of languages is said to have \( \Delta(E) \)-measure zero, or be null at \( E \), if for some \( d \in \Delta(E) \) we have \( A \subseteq S_\infty[d] \).

If \( A \subseteq E \) we also say \( A \) has measure zero in \( E \).

Some of the properties of Lebesgue measure that Lutz sought to preserve are the following measure axioms:

M 1 Singleton sets are null.

M 2 The whole space (i.e., some set) is not null.

M 3 Appropriate unions of null sets are null.
M 4 A subset of a null set is null.

The following is a basic property of resource-bounded measure, but has no analog in Lebesgue measure.

M 5 For each complexity bound $f$, the union of all null sets having covers of complexity at most $f$ is null.

These require some interpretation. Definition 2 can be used to define measure within $E$, to give meaning to statements like, “almost all languages in $E$ are not SPARSE.” With this interpretation we’d want that singleton sets of $E$ languages be null (but singleton sets of languages outside $E$ not necessarily be null). Similarly, we’d want $E$ not to have measure zero (so that $E$ would not be a small subset of $E$), and this is shown by producing from any supermartingale $d$ a language $L_d \in E \setminus S^\infty[d]$. In the theory of Lebesgue measure, the union of null sets is not in general null, but a “small” union, i.e., a countable union, of null sets is null. In the resource-bounded setting we can write $E$ as a countable union of null singletons, so we can’t expect all countable unions of null sets to be null. Instead, Lutz considered the following generalization:

Definition 3 A set $A$ is a $\Delta(E)$-union of $\Delta(E)$-null sets $A_i \subseteq S^\infty[d_i]$ if a single $\Delta(E)$-machine $M(0^i, w)$ computes $d_j(w)$.

Note that $M$ runs in time polynomial in the length of $w$ and the value of $i$.

With these interpretations, Lutz has shown [8] that Definition 2 satisfies the above axioms derived from Lebesgue measure.

As hinted above, a singleton set of language $L$ may not have measure zero in $E$ if $L \not\in E$. Such languages $L$ are intuitively random, in a weak sense. More formally and generally:

Definition 4 A language $L$ is $f(n)$-random if no supermartingale running in time $f(n)$ covers $L$.

Thus there are no $E$-random languages in $E$, but for each $c$ it’s easy to find $(2^c)$-random languages in $E$. This is because, as Lutz showed, $M 5$ is satisfied so there’s a $\Delta(E)$-martingale that covers all non-$(2^n)$-random languages. Thus, for each $c$, “almost every” language in $E$ is $(2^c)$-random.

Lutz’s formulation works for other classes at least as big as $E$, notably $E_2 = 2^{\text{polynomial}}$ and the space analogs ESPACE and $E_2$SPACE, by considering supermartingales that run with the appropriate resource bounds.

2.2 Previous Measure at $P$

In this paper as well as in [1], we generalize Lutz’s work to subexponential time classes. This means that we will consider supermartingales $d(w)$ whose runtime is less than their input length $|w|$. In this setting it is especially convenient to report the runtime of the supermartingales in terms of $n = |x| = |s_{x=1}| \approx \log |w|$, instead of in terms of $|w|$. We follow [12] in using notation from the probabilistically-checkable proofs literature [4, 5].

A martingale is a function $d : \Sigma^* \rightarrow \mathbb{R}$. To define measure on $P$, say, we require that this function be computable by a polynomial time oracle machine, given $|w|$ as input and $w$ as oracle. Often, uniform sequences of martingales $d_j(w)$ are considered; these are to be computed by machines running in time polynomial in the length of input $(j, |w|)$, with oracle $w$. In practice, it is often convenient to suppress the complexity of the supermartingale altogether, when it is clear from context.

For sufficiently nice (and sufficiently large) complexity classes $C$, one plugs in the complexity class and gets out a notion of measure on that class. In defining a notion of measure on $P$, one would be tempted at first to do the same thing. Thus, one would take as null sets the sets of languages covered by polytime (in $|x|$) supermartingales.

The straightforward attempt at measure described above cannot easily be shown to satisfy axiom $M 2$; i.e., apparently too many sets are measurable. Given a supermartingale $d$ defining measure at $E$, say, Lutz shows that $d$ does not
cover all of $E$ by diagonalizing against $d$: The left-most characteristic sequence $L$ such that $d$ never increases along $L$ is clearly not covered, and $L$ is in $E$ since the straightforward decision procedure for $x \in L$ makes $2|x|$ recursive calls to an algorithm in exponential time (in $|x|$.) If one applies Lutz’s diagonalization argument against a polytime martingale, apparently the most one can say (as noted in [10]) about the resulting $L$ is that $L$ is \textit{word-decreasing self-reducible}, i.e., that there’s a polytime reduction from $L$ to $L$ that on input $x$ queries only strings $y$ that precede $x$ lexicographically. There are self-reducible languages that are hard for $E$ [6], and in general we don’t see how, with present knowledge, to prove that such a diagonal language is in $P$.

To define measure in $P$, in [1] we considered not all polytime supermartingales, but only those requiring at most polynomially many recursive calls to perform this diagonalization. The set of input bits ultimately queried by recursive calls of $d$ on original input $w$ is called the \textit{dependency set} $G_{|w|}$ of $d$, and so we require that the dependency set be printable in polynomial time. With this restriction the diagonalization can be carried out. We denote by $(P)$ the collection of machines satisfying simultaneous polynomial (in $n$, the length of input to the oracle machine) time and dependency set bounds.

With the additional restriction that

$$d(w) \geq \frac{d(w0) + d(w1)}{2}$$

be satisfied with equality, i.e., considering only martingales, one gets a notion of measure that satisfies many closure and robustness properties, but, unfortunately, this notion of measure appears to be too weak for many desired results [1]. It is shown in [1] that these martingales can equivalently be defined by martingales that look at bits of $w$ from a particular SPARSE set $S$ (having $n^c$ words of length $n$). Such a martingale cannot cover all the languages of density at most $n^{c+1}$, so no martingale covers all SPARSE.

2.3 A Counting Argument

In several proofs we will use a counting argument similar to the following form of Markov’s inequality.

\textbf{Lemma 5} Let $f$ be a real-valued function of a finite set. If the average of $f$ is at most 1, and $f \geq 1 - \epsilon$, then for all $a > 1$ at least $1 - 1/a$ of the values of $f$ are at most $1 + \epsilon a$.

\textbf{Proof.} Otherwise, if more than $1/a$ of the elements are more than $1 + \epsilon a$, then even if the other $(1 - 1/a)$ elements are all the minimum value of $1 - \epsilon$, that gives an average value of

$$(1/a)(1 + \epsilon a) + (1 - 1/a)(1 - \epsilon)$$

$$= 1 + (1/a)\epsilon a - (1 - 1/a)\epsilon$$

$$= 1 + \epsilon - \epsilon + \epsilon/a$$

$$> 1.$$  

\hfill \square

3 Ordering $\Sigma^*$ and Exponential Measure

In this section we give conditions that make a notion of measure independent of feasibly computable reorderings. The results of this section have consequences both for measure on exponential classes and for subexponential classes. The discussion here is somewhat of a digression from the development of subexponential measure, but there are some important connections and contrasts.

Resource-bounded measure is developed via supermartingales that bet on the membership in a random language $L$ of each word $x \in \Sigma^*$, given the memberships in $L$ of each $y < x$. Here $y < x$ means $y$ precedes $x$ in the lexicographic order. The lexicographic order has many nice properties, including

- It is easy to compute successors, predecessors, the $i$th word, and the position $i$ of a given word.
• When the supermartingale bets on a word $x$, it is given an input string of length between $2^i$ and $2^{i+1}$. This gives a precise, exponential relationship between the computational power of a set of supermartingales and the class on which the set of supermartingales defines a measure.

To preserve these properties, we consider only computable reorderings, according to the following definition:

**Definition 6** An E-reordering $\pi$ of $\Sigma^*$ is a permutation of $\Sigma^*$ satisfying

- $\pi$ is length preserving, i.e., $\pi_n = \pi|\Sigma^n$ is a permutation of $\Sigma^n$, and
- A single machine $M$ computes $\pi(x)$ in time exponential in $|x|$, and another machine $M^{-1}$ similarly computes $\pi^{-1}(x)$.

We also similarly define C-reorderings for other classes C.

First a warmup result, a baseline to be improved:

**Theorem 7** Measure at ESPACE is invariant under ESPACE-reorderings.

**Proof.** We are given a supermartingale $d'$ and a permutation $\pi = \bigcup \pi_n$; we wish to construct an equivalent supermartingale working in lexicographic order.

Our supermartingale $d(w)$ should return the conditional expectation $E(d'|w) = E(d'(W)|W \supseteq w)$. That is, for $2^{n-1} < |w| \leq 2^n$,

$$d(w) = \sum_{|W| \geq w} 2^{-2^n} d'(W),$$

where $w$ and $W$ are interpreted as characteristic strings, each in their own order, and $w \subseteq W$ means $W$ extends $w$. Exponential space is sufficient to do this computation.

A supermartingale need not exactly compute a conditional probability. Next we show how to convert to lexicographic order in classes smaller than exponential space, by just approximating the conditional expectation of $d'$. To perform the approximation we first use pseudorandom generators (see [7]), and afterwards we instead the use approximate counting of [13, 14].

**Definition 8** A pseudorandom generator of security $h(n)$ is a family $G_n$ of functions from $\Sigma^n$ to $\Sigma^m$ (sensible for $m \leq h(n)$), such that for any (nonuniform) circuit $C$ of size at most $h(n)$,

$$|\Pr(C(y) = 1) - \Pr(C(G(x)) = 1)| < 1/h(n),$$

where $y$ is chosen uniformly from $\Sigma^m$ and $x$ from $\Sigma^n$.

Generators of security $2^n \epsilon$ are widely thought to exist (a function derived from integer factoring is thought to be an example). Generators of security $2^n \epsilon$ are more controversial.

**Theorem 9** If, for some $\epsilon > 0$, there exist pseudorandom number generators of security $2^n \epsilon$, then measure on $E_2$ is invariant under $E_2$-reorderings. If generators of security $2^n \epsilon$ exist, then measure on $E$ is invariant under $E$-reorderings.

**Proof.** We give the proof for $E$; the proof for $E_2$ is similar (and somewhat easier). The idea is to use the generator to approximate $E(d'|w)$. But first we massage $d'$ into a more convenient form.

The supermartingale $d'$ may double $2^n$ times on $\Sigma^n$, and this rate of growth is too high. We replace $d'$ with $\min(2^{\epsilon n}d'(0^n), d'(w))$, a supermartingale that grows by at most the factor $2\epsilon/n$ on $\Sigma^n$ and therefore $d' \leq 2^n$ on $\Sigma^{\Sigma n}$ (keep the name $d'$ for the new supermartingale). It is clear that the new supermartingale covers the same set as the old. We consider only the progress of $d'$ on $\Sigma^n$. Assume $d'$ has $\varepsilon 1$ with which to bet on $\Sigma^n$. (It will become clear from the analysis below that these assumptions are justified.)

Now we show how approximate $E(d'|w)$. Partition the range $[0, 2^n]$ of reals into subintervals of length $2^{-2n}$, so there are $2^{3n}$ subintervals in all. For each subinterval $I$, we will use the generator to approximate how many extensions $W$ of $w$ satisfy $d'(W) \in I$. Let $c \geq 1$ be such that $d'$ runs in time $2^{cn}$. For each seed of size $6cn/c$,
blow up the seed to (at most) $2^n$ bits, and use the pseudorandom bits to fill in the bits of the characteristic sequence on $\Sigma^n$ missing from $w$. For each $I$ and $W$, let $C$ be a circuit that takes input $W$ and outputs 1 if $d'(W) \in I$; then $C$ has size approximately $2^{2n}$, which is within the security of the generator, $2^{c^{(6n/c)}} = 2^{6n}$. Thus for the event $d'(W) \in I$, the observed pseudoprobability differs from the actual probability by at most the reciprocal security of the generator, $2^{-6n} \leq 2^{-6n}$.

Approximate $E(d'|w)$ by

$$E(d'|w) = \sum_{I} \min(I) \Pr_{W \subseteq w}(d'(W) \in I),$$

where $\Pr$ is the pseudoprobability as observed by sampling with the generator. There are two sources of error in this approximation, sampling error and quantization error, and we show both are small.

First, the pseudoprobability may not be the real probability. We have:

$$\left| \sum_{I} \min(I) \Pr(d'(W) \in I) - \sum_{I} \min(I) \hat{\Pr}(d'(W) \in I) \right|$$

is at most $\max_{I}(\min(I)(\Pr - \hat{\Pr}))$ times the number of $I$'s. There are $2^{3n}$ intervals, $\max_I(\min(I)) \leq 2^n$ and $\Pr - \hat{\Pr} \leq 2^{-6n}$, so the sampling error is at most $2^{-2n}$.

Also, we assume that if $d'(W) \in I$ then $d'(W) = \min(I)$. In fact, for each $W$ and for $I$ containing $d'(W)$ we have $|d'(W) - \min(I)| \leq |I|$, so the expected values differ by at most $|I| = 2^{-2n}$. Thus

$$\left| E(d'|w) - \sum_{I} \min(I) \Pr(d'(W) \in I) \right| \leq 2^{-2n},$$

and considering the error from both sources,

$$\left| E(d'|w) - \sum_{I} \min(I) \hat{\Pr}(d'(W) \in I) \right| \leq 2^{1-2n}.$$

Compared with $E(d'|w)$, the approximation $\hat{E}(d'|w)$ may make $O(2^{-2n})$ error on each word $s_i$ of length $n \approx \log |w|$. This amounts to $O(2^{-n})$ error on all $\Sigma^n$, and error $O(1)$ on all $\Sigma^*$. The supermartingale returns

$$d(w) = \hat{E}(d'|w) + O(1) - \sum_{i=0}^{n} 2^{-2i}.$$

It is easily checked that since $E(d'|w)$ is a supermartingale and $\hat{E}(d'|w)$ is close to $E(d'|w)$, our slack insures that $\hat{E}(d'|w)$ is a supermartingale (i.e., satisfies the average inequality).

Next we show how to use alternation instead of randomness to approximate $E(d'|w)$.

**Theorem 10** If $P = \text{NP}$, then measure on $E$ is invariant under $E$- reorderings.

**Proof.** If $P = \text{NP}$ then $P = \Delta_3^P$ so approximate counting can be done in $P$.

For $d'$ bounded by $2^n$ and subintervals $I$ as above, we make use of the approximate counting of [13, 14] to estimate the number $g(w)$ of $W \subseteq w$ such that $d'(W) \in I$. Note this is a $\#P$ function of $w$. By [13, 14], there's a function $f$ computable in $\Delta_3^P$ of $w$ such that

$$\left| 1 - \frac{f(w)}{g(w)} \right| < 2^{-6n}$$

(note that $2^{-6n}$ is at most a polynomial in $|w|$). Then $2^{1-w-2n} f(w)$ approximates to within a factor $(1 \pm 2^{-6n})$ the probability $2^{1-w-2n} g(w)$ that $d'(W) \in I$ for a random $W$, and since a probability is at most 1, we can approximate the probability within a difference of $2^{-6n}$. The rest of the proof is the same as above.

Thus the issue of order disappears if pseudorandom generators exist, which is likely, at least for generators of strength $2^n$. In any case, we can’t hope to present a set $A$ and ordering $\pi$ so that $A$ is null under lexicographic ordering but not null under $\pi$. If one views $P = \text{NP}$ as “no hard functions” while one views the existence of pseudorandom generators as the “existence of hard functions” (in a different sense), then the two hypotheses are almost opposite, and one might hope
to prove unconditionally that measure at E or E₂ is order-invariant.

Most of the results of this section concerning E and ESPACE scale down to the subexponential measure of [2] formulated via martingales. First consider measure on PSPACE. Recall that a martingale $d'$ working in some ordering $\preceq$ bets on each word $x$ given the memberships of some predetermined $|x|^{O(1)}$ predecessors of $x$ in $\preceq$. A martingale $d$ can, in polynomial space, try all settings of the missing bits, in a manner similar to Theorem 7 above. Similarly, if there exist generators of security $2^m$, then given an $(n^c)$-time martingale $d'$, working in $\preceq$, a martingale $d$ working in lexicographic order can try all seeds of length $(2c/\epsilon)\log n$, blowing up each seed to length $n^c$. We show in Section 6 that this argument does not scale down to the measure on PSPACE of [16].

4 A Stronger Notion

In this Section we describe an attempt at a stronger notion of measure on $P$. First we show that the new notion covers more intuitively null sets than the previous notion. Next, we show that to exploit fully the new strength one needs to change the definition of supermartingale slightly compared with existing formulations. We conclude this section by showing that in fact, too many sets are covered, as we go on to show that this notion fails to satisfy the union axiom, $M₃$. This will be rectified in later sections.

In this section we are considering measure on $P$, so the supermartingales $d(w)$ have $n^{O(1)}$ time and dependency set bound, where $n = |s_w| \approx \log |w|$.

4.1 Strength

To make more sets measurable, another notion of measure was attempted in [3], by removing the equality requirement from (1). We now recall from [3] an example of a set that is covered by a supermartingale but not by a martingale.

**Theorem 11** The set of languages with density less than $\epsilon < 1/2$ is covered by a dependency-set bounded supermartingale.

**Proof.**

Partition $\Sigma^*$ into consecutive regions $R_0, R_1, \ldots$ as follows. $R_0 = \{ w : |w| < 16 \}$, and for $n \geq 16$ the $2^n$ strings of length $n$ are divided up into $2^{n-2\log n}$ regions of $2^{2\log n} > n^2$ lexicographically consecutive words. Thus $R_j$ consists of words of length $n$, for some $n > \log j$.

Let $X_j$ denote the languages with density less than $\epsilon$ on the $j^{th}$ region. By the Chernoff inequality, for some $c$ that depends on $\epsilon$,

$$\mu(X_j) \leq e^{-cn^2} \leq 2^{-3n} \leq \frac{1}{j^{2}}.$$  

We next construct a supermartingale $d_j$ that climbs from 1 to $j^3$ on $X_j$, works in time polynomial in $n$, and with dependency set $R_j$ of size $n^2$. Given input $w$, let $a$ be the number of 1's in $w[R_j]$. We can set $d_j(w)$ to the exact value

$$\sum_{i \leq cn^2-a} \binom{n^2 - |w[R_j]|}{i}$$

of $\Pr(X_j|C_w)$, since this takes time polynomial in $n = \log |w|$ and has the desired dependency set.

Classically, the argument would proceed by summing the martingales $d_j/j^2$. Intuitively, instead we will do the following:

- Make sure (inductively) that we have $1/(j - 1)$ capital available before starting to bet on $R_j$.
- Bet on $R_j$ using strategy $d_j/j^2$, risking just $1/j^2$ of our capital but winning $j^3 - 1/j^2 = j$ for infinitely many $j$'s.
- Before starting to bet on $R_{j+1}$, “throw away” the wager and potential winnings of $d_j$, and assume that we have only $1/(j - 1) - 1/j^2 \geq 1/j$, enough to continue inductively.

Continuing in this way our winnings will be unbounded, yet we will be able to keep dependency sets small.
Define $d(w)$ as follows. Determine $j$ such that $s_{[w]} \in R_j$ (note $|w|/\log^2 |w| \leq j \leq |w|$). Put

$$d(w) = \frac{1}{j} + \frac{d_j(w)}{j^2}.$$ 

Finally, let $L$ be a language of density less than $\epsilon$. Then the density of $L$ is less than $\epsilon$ on $R_j$ for infinitely-many $j$’s, and for such $j$ $d$ climbs to $d_j/j^2 = j$ along $L$.

4.2 Ordering $\Sigma^*$ and Subexponential Measure

This subsection presents a weakness of the convention that supermartingales return their current capital. Wolfgang Merkle [11] observed that this can be solved by having the supermartingales return the amount of the next bet rather than the current capital.

Above we mentioned that the choice of lexicographic is a minor issue for measure at $E$. However this choice affects measure at $E$ it also affects measure at $P$, and we now show that the choice of enumeration is actually a bigger issue for measure at $P$. The proof of Theorem 11 relies on the fact that each $\epsilon$-SPARSE language is $\epsilon$-SPARSE on infinitely many of the $R_j$’s, and the property of being $\epsilon$-SPARSE on $R_j$ is a property of the contiguous block of words in $R_j$. One can formulate similar properties of non-contiguous words, and one would expect to cover similarly the set of languages meeting infinitely many of these conditions. We show now that this cannot, in general, be done by supermartingales outputting their current capital.

Theorem 12 There’s a set $A$ and an ordering $\prec$ such that $A$ is covered by under $\prec$ but $A$ is not covered in lexicographic ordering.

Proof. Partition $\Sigma^n$ into blocks of $(3/4)n$ words. Suppose there are $3N$ blocks. Then label the blocks in the following order:

$$R_1^n S_1^n R_2^n S_2^n \ldots R_N^n S_N^n T_1^n T_2^n \ldots T_N^n.$$ 

That is, each $R_j^n$, $S_j^n$, and $T_j^n$ has length $3n/4$. All the $R$’s and $S$’s alternate, then come all the $T$’s.

Let

$$A = \{L : \forall n \exists j L[R_j^n] = L[S_j^n] = L[T_j^n]\}.$$ 

Then $A$ is covered in some ordering: In the ordering in which $R_j^n$, $S_j^n$, and $T_j^n$ are contiguous, a supermartingale risks $2^{-5n/4}$ of its capital that the $(3/2)n$ words in $S_j^n$ and $T_j^n$ match $R_j^n$. Next if successful the supermartingale wins $2^{-5n/4} \cdot 2^{3n/2} = 2^{n/4} = \infty$, and if unsuccessful the supermartingale loses only $2^{-5n/4}$ on each of the $\leq 2^n$ regions corresponding to $n$, a total of $2^{-n/4}$, and since $\sum 2^{-n/4} < \infty$ the supermartingale never runs out of money. After betting on some triple $R, S, T$, the supermartingale throws away its wager, so the dependency set for this procedure has size $|R| + |S| + |T| = 9n/4$.

On the other hand, $A$ is not covered in lexicographic order.

First, for each $n$, we want to construct a large set $J_n$ of $j$’s such that for any $j' < j \in J_n$, the first point of $S_j^n$ has no dependency on $S_{j'}^n$.

Start by collecting the largest $j$ into $J_n$. Now, cancel all the $n^{O(1)}$-many $j$’s such that the first point of $S_j^n$ has a dependency on $S_{j'}^n$. Repeat the procedure: collect the largest remaining uncanceled $j'$, then cancel the $j$’s such that $S_j^n$ contains a point in the dependency set of the first point of the most recently collected $S_{j'}^n$. By repeating the procedure until quiescence, we will collect $J_n$ of size $2^n/n^{O(1)}$.

We first give an overview of the proof:

Given a supermartingale $d$, we wish to construct an uncovered language $L \in A$. For most $j$’s, $d$ bets just a little on $S_j^n$ since there are too many $j$’s to risk a lot on each. Construct $L$ so that most $S$’s match their corresponding $R$’s, and $d$ wins, but very little, on each particular $j$. The lack of dependencies among $j$’s insures that $d$ can’t gather its winnings from all $j$’s (the total that an unrestricted supermartingale can accumulate from all the $j$’s would be large). Thus $d$ has to start betting on the $T$’s with a bounded amount of capital. But the $T$’s are too short, and thus present too few betting opportunities, for $d$ to earn significant capital.

Now quantitatively:
Suppose a supermartingale has unit capital at the start of $\Sigma^n$. Consider the words of $\Sigma^n$ in turn. Set $L[R^n_j]$ to defeat $d$, and if $j \notin J_n$, then set $L[S^n_j]$ to defeat $d$. Suppose $j \in J_n$. If some setting makes $d$ rise by at least $2^{-n/8}$ then this winning mass must be distributed over $2^{5n/8}$ other losing settings, and thus some setting $\sigma$ makes $d$ fall by at least $2^{-7n/8}$. Put $L[S^n_j] = \sigma$, and note that this action can only be taken $2^{7n/8}$ times while $|J_n| = 2^{n-O(\log n)}$. Thus for $2^{n-O(\log n)} - 2^{7n/8} \in 2^{n-o(n)}$ of the $j \in J_n$, every setting of $S^n_j$ makes $d$ fall, or rise by at most $2^{-n/8}$. For such $j$, put $L[S^n_j] = L[R^n_j]$.

Let $m$ be the earliest point following $S^n_j$ such that $m$ has no dependency on $S^n_j$. Then by the average law and definition of dependency set, the capital that $d$ has at $m$ is at most 1, the capital $d$ had at the start of $S^n_j$. Let $j'$ be the next element in $J_n$ after $j$, so $m$ occurs at the start of $S^n_j$, or earlier. The bits between $S^n_j$ and $m$ are set to defeat the supermartingale, so $d$'s capital is at most $1 + 2^{-n/8}$ there, and the bits between $m$ and $S^n_j$ are also set to defeat the supermartingale so the capital is at most 1 there. We conclude that the capital at the start of $S^n_j$ is $1$.

Continue setting the $R$'s and $S$'s this way, and let $J'_n = \{ j \in J_n : L[R^n_j] = L[S^n_j] \}$. Thus $|J'_n| \in 2^{n-o(n)}$. Note that $d$'s capital is bounded by $1 + 2^{-n/8}$ on the $R$'s and $S$'s.

Now the $T^n_j$'s. One can show directly that $\{ L : \exists j \in J'_n, L[T^n_j] = L[R^n_j] \}$ has Lebesgue measure $1 - o(1)$, so no supermartingale having bounded capital at the start of the $T$'s succeeds on $A$: we give a supermartingale-style proof. For $j \notin J'_n$, set $L[T^n_j]$ to defeat $d$. For $j \notin J'_n$, if some setting of $L[T^n_j]$ makes $d$ rise by at least $2^{-n/8}$, then some other setting $L[T^n_j] = \sigma$ makes $d$ drop by at least $2^{-7n/8}$. Set $L[T^n_j] = \sigma$, and note this action only occurs $2^{7n/8}$ times. Finally, assume that $j \in J'_n$ and no setting of $L[T^n_j]$ makes $d$ rise by more than $2^{-n/8}$. Put $L[T^n_j] = L[R^n_j]$, and set the rest of the $T$'s to defeat $d$ (we don't need more than one match). This way the language we've constructed meets the condition at $n$ for membership in $A$, but $d$'s capital is bounded by $1 + 2^{1-n/8}$ on $\Sigma^n$. Since $\sum 2^{-n/8} < \infty$, we have constructed a language $L \in A$ such that $d$ remains bounded on the $L$.

The difficulty in covering this set in lexicographic order derives from the fact that if a supermartingale wins some capital at word $x$, and wants to bet the winnings on a later word $z$, then for each $y$ between $x$ and $z$ the supermartingale needs to know how it fared at $x$, in order not to violate the average inequality (1). For example, consider how an unbounded supermartingale would bet on the language $L \in A$ constructed above. A supermartingale wants to bet approximately $2^{-5n/4}$ on $R^n_j$, $S^n_j$, and $T^n_j$, winning $2^{3n/4}2^{-5n/4} = 2^{-n/2}$ on each $R^n_j$ and $S^n_j$. The total amount an unbounded supermartingale would have at the start of the $T$'s would then be $|J'_n|2^{-n/2}$. The unbounded supermartingale would divide this among $|J'_n|T$'s, betting $2^{-n/2}$ and winning $2^{-n/2}2^{3n/4} = 2^{3n/4}$ on each.

In [11], Merkle suggests that the supermartingales return the signed amount bet (i.e., if $w'$ denotes $w$ with the last bit flipped, $b(w) = \frac{d(w) - d(w')}{2}$) instead of their current capital $d(w)$. The sort of betting strategy for an unbounded supermartingale described above is also allowed for a polynomially-bounded supermartingale under the convention that $b(w)$, and not $d(w)$, is output. While $b(w)$ is easily computable from $d(w)$, and with the same dependency set, computing $d$ from $b$ may take time $|w|$, and thus we gain by the new convention — in the example above, to compute the value of $d$ at the start of the $T$'s, a supermartingale would have to know the number of $j$'s for which it bet correctly that $L[R^n_j] = L[S^n_j]$, while in the new convention the supermartingale only needs to check one $R-S$ pair to whether to bet $2^{-n/2}$ or 0. Also, note that “if some setting makes $d$ rise by $\epsilon$” in the above proof becomes “if $b$ bets $\epsilon$ on some setting” in the new formulation, so the new formulation is perhaps more intuitive. In Section 5, we will give a formulation in which supermartingales can bet on the words in any (feasible) ordering they prefer. In particular, the supermartingales can order the words so that desired blocks of words become contiguous, and thus such supermartingales have a mecha-
nism for betting on discontinuous words. On the other hand, there may be other other agenda in choosing the ordering of words (see Section 3), so we let supermartingales bet on discontinuous words via the mechanism of returning an amount bet rather than current capital.

The ordering used in the above proof is highly scrambling; it was constructed to break up blocks of contiguous words. By contrast, switching from lexicographic to reverse lexicographic order preserves the contiguity of most blocks of words. Switching from lexicographic ordering to an ordering that is not highly scrambling seems to have a similar effect at subexponential measure as at exponential measure: Not an issue for space, and not an issue for time if pseudorandom number generators exist. Thus we can not hope to show, for example, that reverse lexicographic measure differs from lexicographic. Having the supermartingales return the amount bet rather than current capital resolves the issue of betting on contiguous blocks, but not the issues of ordering shared by subexponential and exponential measures.

4.3 Unions

In Subsection 4.1 we saw that more sets are covered by supermartingales than by martingales. One would therefore hope to define a notion of measure via supermartingales. Unfortunately, coverage by supermartingales does not satisfy the Union Axiom M 3 (despite claims in [3]). Here we present two covered sets A and B whose union is not covered. But first we need to examine the structure of dependency sets and betting strategies in more detail.

**Definition 13** A betting strategy \( b(w) \) is a real-valued function such that \( b(w) = -b(w') \) for \( w \neq \lambda \), and for all \( w \), \( \sum_{z \in w} b(z) \geq 0 \).

Starting with a supermartingale \( d(w) \), as above one can define a betting strategy \( b(w) = \frac{d(w') - d(w)}{2} \), where \( w' \) is \( w \) with the last bit flipped, and by convention \( b(\lambda) = d(\lambda) \). Intuitively, \( b(w) \) is the amount bet on the last bit of \( w \). Note that \( b(w) = -b(w') \), and if \( j \notin G_j \) then \( d(w) = d(w') \), so \( b(w) = b(w') = 0 \). Also

\[
b(w) = \frac{d(w) - d(w')}{2} = \frac{d(w) - d(w) + d(w')}{2} \geq \frac{d(w) - d(w_\infty)}{2},
\]

so \( \sum_{z \in w} b(z) \geq d(w) \). On the other hand, one can easily see that starting with any function \( b(w) \) such that \( b(w) = -b(w') \) and \( \sum_{z \in w} b(z) \geq 0 \), one can put \( d(w) = \sum_{z \in w} b(z) \), and get a well-formed supermartingale. Starting with \( d \), converting to \( b \) and then back to \( d \) gives a martingale that is greater than or equal to \( d \), with equality iff the supermartingale \( d \) was originally a martingale.

The collection of dependency sets \( \{G_k\} \) has a natural directed acyclic graph structure (also to be called \( G \)) where the set of nodes is the numbers \( j \) such that \( j \in G_j \). Put an arc from \( i \) to \( j \) when \( i \neq j \) and \( i \in G_j \); this records a dependency. We will also include the node \(-\infty\), and put an arc from \(-\infty\) to each other node (this represents that the supermartingale knows its starting capital and knows some fixed apportionment of the starting capital to each node). A dependency chain in \( G \) is a path in \( G \). Let \( I \subseteq G \) be the set of initial bets, i.e., nodes with just one path (a single edge) from \(-\infty\).

For any supermartingale \( d \), we have \( d(w) \leq 2^{\|d(\lambda)\|} \), since the supermartingale can double at most \( |w| \) times. If a supermartingale has bounded dependency sets, then strictly fewer than \( |w| \) doublings may actually occur. The number of doublings is bounded by the length of the longest dependency chain.

**Lemma 14** Let \( d \) be a supermartingale, \( k \) a number, and \( G_{\{0..k\}} \) be the dependency graph for \( d \) on \( \{0..k\} \). Let \( h \) denote the length of the longest dependency chain of \( G_{\{0..k\}} \). Suppose \( d \) is bounded from below by \( (1 - \varepsilon)d(\lambda) \) on all \( w \) of length at most \( k \). Then \( d \) is bounded from above by \( (1 + 2^h)d(\lambda) \) on \( w \) of length at most \( k \).

**Proof.** If \( d \) is bounded below by \( (1 - \varepsilon)d(\lambda) \), then \( d' = d - (1 - \varepsilon)d(\lambda) \) is nonnegative, and
hence a legitimate supermartingale. Note $d'(\lambda) = cd(\lambda)$. We will show that $d'$'s capital is bounded by $2^h d'(\lambda) = c2^h d(\lambda)$, and thus the capital of $d(w) = (1 - c)d(\lambda) + d'(w)$ is bounded by

$$(1 - c)d(\lambda) + c2^h d(\lambda) \leq (1 + c2^h)d(\lambda).$$

We will proceed by induction on $h$. For $h = 1$, we are considering some $w$ with $j \in I$ for all $j \leq |w|$. Thus $|\tilde{b}(w)|$ depends on at most $j$. For the losingest $w$,

$$b'(\lambda) - \sum_{\lambda \neq \varepsilon \in \mathcal{L}_w} b'(z) = \sum_{\lambda \neq \varepsilon \in \mathcal{L}_w} b'(z')$$

$$= d'(\overline{w})$$

$$\geq 0,$$

where $z'$ is $z$ with the last bit flipped and $\overline{w}$ is $w$ with all bits flipped. Therefore

$$\sum_{z \in \mathcal{L}_w} b'(z) = b'(\lambda) + \sum_{\lambda \neq \varepsilon \in \mathcal{L}_w} b'(z)$$

$$\leq 2b'(\lambda)$$

Now consider $h > 1$. We can win at most $b'(\lambda)$ on the bets in $I$, and the longest chain not counting edges from $-\infty$ into $I$ has length $h - 1$. We will assume we doubled our money betting on $I$, and by induction we double at most $h - 1$ more times.

Define a function $b$ as follows. Put $b(\lambda) = 2b'(\lambda)$, for $|w| \in I$ put $b(w) = 0$, and for $|w| \notin I$ put $b(w) = b'(w)$. The function $b$ has an associated acyclic graph structure formed by consolidating $-\infty$ and $I$ in which the longest chain has length at most $h - 1$. Then, as above, $b'(\lambda) \geq \sum_{\lambda \in \mathcal{L}_w} b'(w)$ so for all $w$

$$\sum_{z \in \mathcal{L}_w} b'(z) = 2b'(\lambda) + \sum_{z \in \mathcal{L}_w} b'(z)$$

$$\geq b'(\lambda) + \sum_{\lambda \neq \varepsilon \in \mathcal{L}_w} b'(z)$$

$$= \sum_{\lambda \neq \varepsilon \in \mathcal{L}_w} b'(z)$$

$$= d'(\lambda)$$

$$\geq 0,$$

and so $b$ is a legitimate betting function. Therefore, by induction,

$$\sum_{w \in \mathcal{L}_w} b(z) \leq 2^{h - 1} b'(\lambda)$$

$$\leq 2^h b'(\lambda).$$

We conclude

$$d'(w) \leq \sum_{z \in \mathcal{L}_w} b(z)$$

$$\leq 2^h b'(\lambda)$$

$$= 2^h d'(\lambda),$$

and so

$$d(w) = (1 - c)d(\lambda) + d'(w)$$

$$\leq d(\lambda) + 2^h d'(\lambda)$$

$$\leq (1 + c2^h)d(\lambda).$$

\

We now give the main theorem of 4.3.

**Theorem 15** There are sets $A$ and $B$, each covered by a dependency-set bounded supermartingale, such that $A \cup B$ is not covered by any such supermartingale.

**Proof.** For each $n$, partition the words of $\Sigma^n$ into consecutive blocks of $.6n$ consecutive words of length $n$. Let $R_j$ denote the $j$th block (so $2^n/n \leq j \leq 2^n$, and $R_j$ has about $.6 \log j$ words). Define

$$A = \{L : \exists^\infty j \; I[R_{2j}] = I[R_{2j+1}] \in 0^*\},$$

and

$$B = \{L : \exists^\infty j \; I[R_{2j-1}] = I[R_{2j}] \in 0^*\}.$$

Thus $A$ is the set of languages $L$ such that for infinitely many even-indexed blocks $R_{2j}, R_{2j+1}$ have no words.

By an argument similar to the proof of Theorem 11, one sees that each of the sets $A$ and $B$ is covered by a supermartingale. We show that $A \cup B$ is not covered.
The idea is that if a supermartingale $d$ bets an appropriately small amount on block $2j$ or $2j+1$ and still wins, then it must make a lot of dependencies. The dependencies for $A$ and $B$ overlap, so the transitive closure becomes too big.

Note that by the pigeonhole principle the set $A \cup B$ can be written as

$$\{ L : \exists j \exists i < 2 L[R_{2j+1}] = L[R_{2j+1+i}] \in 0^* \},$$

and we will find this characterization convenient.

Inductively suppose we have defined our language $L$ through words of length less than $n$. We will show how to extend the language through $\Sigma^n$, making at least one pair of empty blocks, and only increasing $d$ by $2^{-\Omega(n)}$. Since $\Sigma_{n=0}^{\infty} 2^{-\Omega(n)} < \infty$, we conclude that $d$ remains bounded on $L$.

Set $j = 1$, and start in Phase 1.

Phase 1: If more than half of the $x \in R_{j+1}$ have a dependency set that includes more than half of $R_j$, then extend $L$ through $R_j$ along the path of decreasing $d$ and remain in phase 1; otherwise go to phase 2. One can show, by induction on the number $l$ of such $j$’s in an uninterrupted run, that half the words in such an $R_{j+1}$ have dependency set of size at least $l$. Thus we can remain in phase 1 without interruptions for a number $l < n^{O(1)}$ of $j$’s, and we are in phase 1 at most $2^n(1 - 1/n^{O(1)})$ of the time. Call the other $2^n/n^{O(1)}$ $j$’s dependency breaks.

Phase 2: We are considering a $j$ such that at most half of the words in $R_{j+1}$ have in their dependency set more than half of the $x \in R_j$. If some setting of $R_j$ and $R_{j+1}$ makes $d$ drop by $2^{-95n}$, extend $L$ by that setting, and go to phase 1. Note that this can only happen $2^{95n}$ times before $d$ runs out of money, but that leaves $2^n n^{-o(n)}$ dependency breaks (in particular, at least one) in which no setting makes $d$ drop by $2^{-95n}$ (phase 3).

Phase 3: At most half of the $x \in R_{j+1}$ have a dependency set that includes more than half of the $x \in R_j$, and no setting of $R_j$ and $R_{j+1}$ makes $d$ drop by as much as $2^{-95n}$. In particular, no setting of any prefix of $R_j$ and $R_{j+1}$ makes $d$ drop that much, since one can diagonalize against $d$ for the rest of the words. Note that no point in $R_j$ or $R_{j+1}$ has dependency chain of length greater than $\frac{2}{3}(|R_j| + |R_{j+1}|) \approx .9n$, since otherwise more than half of the $x \in R_{j+1}$ would have dependency set including more than half of $R_j$. It follows from Lemma 14 that no setting of any prefix $L_1$ makes $d$ rise by more than $2^{-0.5n}$. Extend $L$ through $R_j$ and $R_{j+1}$ by all zeros, then extend $L$ through the rest of $\Sigma^n$ to diagonalize against $d$.

Thus, on $\Sigma^n$, the value of $d$ rises by at most $2^{-0.5n}$. Since $\sum_n 2^{-0.5n} < \infty$, the supermartingale remains bounded.

Above we took two supermartingales whose dependency graphs have a lot of overlap. We should note that if a pair or sequence of supermartingales have compatible dependency graphs (e.g., if all the dependency sets are the same), then the usual union theorem holds (e.g., the sum of two such supermartingales is again a supermartingale with the same small dependency graph). It will sometimes be useful to construct supermartingales with compatible dependency graphs, just so that we can add them.

The fact that the union of two intuitively small sets is not small is a serious flaw. Yet, one can make a weak case for considering this notion anyway, by analogy with the choice of lexicographic ordering. In formulating resource-bounded measure, one has to fix some enumeration of $\Sigma^*$, and the lexicographic enumeration has been chosen. On the other hand, the reverse lexicographic order (for each $n$, list the words of length $n$ backwards) is an equally suitable enumeration. For measure on $E$, it is consistent with what we know that there’s a set $A$ that’s covered by a supermartingale in lexicographic order, and $B$ that’s covered in reverse-lexicographic order, yet $A \cup B$ (which has Lebesgue measure zero) is not covered by a supermartingale in any order. By fixing a standard enumeration ahead of time and considering only supermartingales in that enumeration, one gets a notion of measure closed under unions. (See Section 3 for more discussion about why the choice of enumeration is not really a problem for measure at $E$).

It is far less natural, but nevertheless one could decide ahead of time which are the allowable dependency sets, and consider only such super-
martingales. The resulting notion of measure would be awkward, but it would include more null sets than the martingale notion of [1], and in fact include some sets not measurable by the notion of measure to be presented in Section 5. A particular drawback is this awkward notion would not satisfy M 5 in a reasonable way, i.e., the resulting notion of resource-bounded pseudorandomness would be unnatural.

Another option is to restrict what is meant by “appropriate unions.” If a single machine \( M(j, w) \) computes \( d_j(w) \), and the dependency sets for \( d_j \) obey some uniformity conditions in \( j \) making the dependency sets compatible, then there is a single martingale covering \( \bigcup_j S^\infty[d_j] \). As we saw above, this notion of union does not include all finite unions.

In the next section we present a notion of measure that balances these requirements.

## 5 Quotient Formulation

To remedy the flaws with the supermartingale measure discussed previously, in this section we propose a new notion of measure. We first give the relevant definitions and show that the measure axioms are satisfied. Next we show that desired properties of resource-bounded measure are, to a large degree, preserved: We look at resource-bounded pseudorandomness and the “density of a random language.” The notion of measure presented now balances a need to satisfy measure axioms with a desire to have as many null sets as possible.

### 5.1 Basic Definitions and Properties

**Definition 16** The quotient of a language \( L \) by a word \( y \) is the language

\[
L/y = \{ x \mid xy \in L \}.
\]

The direct product of a sequence \( \{L_i\} \) of languages is the language

\[
\bigotimes L_i = \{ x10^{i-1} \mid x \in L_i \}.
\]

We say a set \( A \) is closed under quotients if \( L \in A \) implies for all \( x, L/x \in A \). The quotient by \( 10^{i-1} \) and direct product are (essentially) inverse operations: \( \bigotimes_j L_j/10^{i-1} = L_i \), and \( \bigotimes_1 (L/10^{i-1}) = L \setminus 0^* \). We will often consider quotients by \( 10^{i-1} \), so, for convenience, for an integer \( i \) we write \( L/i \) for \( L/10^{i-1} \), and \( L/i = L \) if \( i = 0 \). Quotients can be composed: \( (L/x)/y = L/yx \).

For every set \( A \) of languages, therefore, one can define the interior \( \overline{A} \) of \( A \) by \( \{ L \in A \mid \forall x \ L/x \in A \} \); then \( \overline{A} \) is the largest subset of \( A \) closed under quotients. One can also take the closure \( \bar{A} \) of the set of all quotients in \( A \), this is the smallest superset of \( A \) closed under quotients. Note that the characteristic sequence \( \chi_{(L/y)} \) is a subsequence of the characteristic sequence \( \chi_L \), formed by taking the bits indexed by an arithmetic progression of difference \( 2^i \).

Finally, note that \( L/\bar{y} \) can be reduced to \( L \) by extremely weak reductions, and if a machine \( M(0^i, y) \in \mathbb{P} \) decides \( y \in L_i \) then \( \bigotimes L_i \in \mathbb{P} \).

**Definition 17** A subbasic null set is a set closed under quotients that is covered by a dependency-set size bounded supermartingale.

A basic null set is the enumerated union of subbasic null sets.

A null set is a subset of a basic null set.

Each basic null set has its own supermartingale, with its individual dependency set and enumeration of \( \Sigma^* \).

By an “enumerated union” of subbasic null sets \( \{A_i\} \) we mean a single machine \( M(i, w) \), that runs in time polynomial in \( |i| \) and \( \log |w| \), such that \( M(i, \cdot) \) computes a supermartingale \( d_i \), that covers the subbasic null set \( A_i \) (together with \( M \) implicitly come machines to compute the dependency set and enumeration of \( \Sigma^* \)). It will be convenient also to allow \( M(i, \cdot) \) to be any function at all, provided \( M(i, w) = 0 \) for sufficiently long \( w \) and, correspondingly, \( A_i = \emptyset \). A cover of a null set in this formulation is this machine \( M \).

This formulation has preceded in classical mathematics. Forming a basic-null set from a subbasic null set has precedent in the theory of Baire category, where a meager set is defined to
be the countable union of nowhere-dense sets, and taking subsets of basic null sets has precedent in the completion of a measure: a Lebesgue-null set is defined to be a subset of a Borel-null set. The advantage is that satisfaction of $\mathcal{M}$ is trivial.

First we show the analog of the Baire category theorem. The closure under quotients was contrived solely to make this go through:

**Theorem 18** $P$ is not a null set.

**Proof.** Suppose $P \subseteq \bigcup A_i$, for an enumeration of subbasic null sets $A_i$. Since this is an enumerated union, there is a number $c$ and a sequence $d_i$ of supermartingales such that $d_i$ covers $A_i$ and all the supermartingales have bound $n^c$. For each supermartingale $d_i$, find a language $L_i \in \text{DTIME}(n^{2c+1})$ not covered by $d_i$, and put $L = \bigotimes L_i$. Then, again using the uniformity in the union, $L$ is seen to be in $P$, but $L/i = L_i \notin A_i$ by construction, so for all $i$ we have $L \notin A_i$ since $A_i$ is closed under quotients. ■

**Theorem 19** For each $c$, the set $\text{DTIME}(n^c)$ is null.

**Proof.** Let $M(i, x)$ be a P-time machine that is universal for $\text{DTIME}(n^c)$. For each $i$, there is a supermartingale $d_i$ that succeeds on $L(M(i, \cdot))$ such that $d_i$ has dependency set in $0^*$. Finally, we can add the supermartingales $2^{-i}d_i$, getting a dependency set in $0^*$.

In particular,

**Corollary 20** For each $L \in P$, the set $\{L\}$ is null.

Thanks to Steve Fenner for suggesting that we show that a union of supermartingales fails to cover $P$, and not that a union is equivalent to some other single supermartingale.

### 5.2 Resource-Bounded Randomness

Next we look at $\mathcal{M} 5$, the reason for allowing non-martingales into the enumeration.

A useful property of measure at $E$ is the resulting property of pseudorandomness. A language $L$ is $E$-random (in Lutz’s notation, p-random) if no $\Delta(E)$-martingale covers $\{L\}$. The current formulation of measure at $E$ was chosen so that a supermartingale $d$ computable in just more than $2^n$ time covers all languages covered by any $(2^n)$-bounded supermartingale $d_i$. To do this, $d$ enumerates all functions $f_i$ computable in $2^n$ time, checks the first $|w|$ functions to see if they are legitimate supermartingales, and sums the legitimate ones.

The difficulty in extending this to subexponential time measure lies in detecting when a function is a legitimate supermartingale (i.e., satisfies the average inequality). Our solution here is to allow non-legitimate supermartingales into the enumeration, but these supermartingales must not have an infinite limsup on any language (so they each cover the empty set).

With this convention, we have

**Theorem 21** For each $c$, almost every language is $(n^c)$-random.

**Proof.** We construct a cover of the non-$(n^c)$-random languages. Enumerate all $(n^c)$ functions (together with $(n^c)$-bounded dependency sets and enumerations of $\Sigma^*$). On input $(i, w)$ check $M(i, \cdot)$ on the first $\log |w|$ prefixes of $w$, and if the average inequality is violated anywhere output zero, otherwise output $M(i, w)$. ■

### 5.3 Density

Now we consider the density of a “random language,” and compare the supermartingale, martingale, and quotient formulations of measure. With high probability, a (Lebesgue-) random language has $(1/2 - o(1))2^n$ words of length $n$, and this property is captured by supermartingales in Theorem 11. At the other extreme, no martingale (satisfying an exact average law) covers even the set of all polynomially sparse languages[1]. The
The quotient formulation is much better than the martingale formulation in this regard, but not quite so good as the supermartingales.

**Theorem 22** The set $A$ of languages having $o(2^n)$ words of length $n$ has measure zero at $\mathbb{P}$ by the quotient formulation.

That is, $L \in A$ if there exists $f_L(n) \in o(2^n)$ such that $L$ has density $f(n)$. There need not be a single $f$ for all $A$.

**Proof.** The set $A$ is the interior of the set in Theorem 11.

This theorem shows that the quotient formulation is midway between the martingales and supermartingales in regard to density. In several contexts, the interior of a set satisfying a condition “infinitely often” is a set satisfying a similar condition “with density $1 - o(1)$.” A similar observation applies to the sets in Theorem 15.

Earlier we contrasted martingales, which obey the average law with equality, with supermartingales, that satisfy an average inequality. One can show that a set is covered by a martingale if a supermartingale has an infinite limit (and not just infinite limsup) on every language in the set. The quotient formulation falls somewhere between these notions of limit: Note that $\chi_{L/x}$ is a subsequence of $\chi_L$ formed by the bits of the latter appearing in positions of some arithmetic sequence of difference $2^{|x|}$. By requiring a supermartingale $d$ to become unbounded not only on $L$ but also on every quotient of $L$, we are requiring some sort of growth on each dyadic arithmetic progression, no matter how sparse.

### 5.4 Immunity

A property of measure at $\mathbb{E}$ is that almost every language $L$ is $\mathbb{P}$-bi-immune, i.e., neither $L$ nor $L^c$ contains an infinite language in $\mathbb{P}$ (in fact, for each $c$, almost every language is $\text{DTIME}(2^{cn})$-bi-immune)[10]. In particular, for every easy-to-compute infinite set $A$, for almost every language $L \in \mathbb{E}$, we have $L \cap A \notin \text{DTIME}(2^m)$ (the idea is that a supermartingale can double its money at each “easy instance:” each word in $A$). A similar scaled property holds for the weak martingale notion of measure in [1], but as we now observe, this is not the case for the supermartingale measure of this paper. In the following, $L'$ has an infinite set of easy instances in an intuitive sense, but is not measure zero.

**Proposition 23** Let $L$ be $\mathbb{E}$-pseudorandom. Then for $L' = L \cup 0^*$, $\{L'\}$ is not, $(\mathbb{P})$-null.

**Proof.** The hypothesis says that no $\mathbb{E}$-supermartingale covers $L$.

Note that the set $L'$ has lots of intuitively easy instances, all of $0^*$. Suppose $\{L'\}$ is null. Then $\{L'\} \subseteq \bigcup A_i$, i.e., $L'$ is in some $A_i$, where $A_i$ is closed under quotients and covered by a supermartingale. It’s easy to see that $L'/""$ is $\mathbb{E}$-pseudorandom, so no supermartingale covers $L'/""$, or hence $\{L'\}$, or hence $A_i$.

Instead of a result about immunity, we have the following weakened version, replacing “infinite” with “density $1 - o(1)$.” Call a set $L$ weakly $f(n)$-immune if $L$ contains no set in $\text{DTIME}(f(n))$ having $2^n(1 - o(n))$ words of length $n$. If $L$ and $L^c$ are both immune, then $L$ is weakly-bi-immune.

**Theorem 24** For every $c$, almost every set in $\mathbb{P}$ is weakly-$n^c$-bi-immune.

**Proof.** We show immunity. Enumerate $\text{DTIME}(n^c)$, and for now fix a machine for a set $B \in \text{DTIME}(n^c)$; we will construct a supermartingale $d_B$ and afterwards take a union over all $B$. Our goal is to construct a supermartingale $d_B$ such that if $B \subseteq L$ and $B$ has density $1 - o(1)$, then for all $y \in d$ succeeds on $L/y$. That is, if $B$ turns out to have sufficient density, then $d_B$ covers the set of languages that are not immune by virtue of $B$.

The algorithm $d_B$ bets on words of length $n$ as follows: Partition $\Sigma^n$ into blocks $R_j$ of $2n^2$ words. The supermartingale $d_B$ allocates $1/j^2$ capital to bet on $R_j$, and then allocates a $(2^{-|b|})$-fraction of this $1/j^2$ to each $y$ of length at most $n$. With the allocated capital, for each word $x \in R_j \cap B/y$, $d_B$ bets that $x$ is in. The supermartingale then throws away these winnings before betting.
on \(R_{j+1}\), so the dependency sets are contained in the disjoint \(R_j\)'s.

Suppose \(B \subseteq L\) has density \(1 - o(1)\). We need to show that for each \(y\) (including \(y = \lambda\)), \(d_B\) succeeds on \(L/y\). Since \(B \subseteq L\), it follows that \((B/y) \subseteq L/y\). By the above construction, \(d_B\) bets that \((B/y) \subseteq L/y\), which is true, so \(d_B\) never loses the money allocated to \(y\) in betting on \(L/y\). Since \(B\) has density \(1 - o(1)\), it follows by the pigeonhole principle that \((B/y)\) has density \(1 - 2^{\log o(1)} = 1 - o(1)\). Thus, for each \(y\), for \(1 - o(1)\) of the \(j\)'s, \(|R_j \cap B/y| \geq |R_j|/2 \geq n^2\). Thus \(d_B\) gets to make \(n^2\) bets on \(L/y\) before throwing away its winnings, on at least one of the \(R_j\)'s.

For each \(j\), \(d_B\) allocates \(1/j^2 \approx 2^{-2n}\) to \(R_j\), and only \(2^{-n}\) of that to the longest \(y\). But it multiplies this \(2^{-3n}\) capital by nearly \(2^{n^2}\) when successful, so the supermartingale becomes unbounded on each \(L/y\) when \(L \supseteq B\).

\[\square\]

6 Space

In this section we note that our definitions hold for space bounds instead of time bounds. We then compare our PSPACE measure, denoted \(\Phi\) (PSPACE), to that of [10], which we denote by \(\Phi\) (PSPACE). A set is \(\Phi\) (PSPACE)-null if it is covered by a supermartingale that works in poly-log space, reads its input once from left to right, and is given the allowable workspace (but not the input length).

Definition 25 Let ODD denote the set of languages \(L\) such that for each \(n\), \(L\) has an odd number of words of length \(n\).

Note that ODD has Lebesgue measure zero.

Theorem 26 The set ODD has \(\Phi\) (PSPACE)-measure zero.

Proof. Immediate; also see [10].

Theorem 27 The set ODD \(\cap\) PSPACE does not have \(\Phi\) (PSPACE)-measure zero.

Proof. To show that no enumeration covers ODD, let \(d_i\) be an enumeration of supermartingales. For each \(i\), form a language \(L_{i+1}\) that defeats \(d_i\) (leaving \(L_1\) unspecified). Put \(L = \bigotimes L_i\), and now set \(L_1\) so that \(L\) is in ODD.

In the above proof, ODDness defeated closure under quotients, i.e., we constructed an ODD language \(L\) such that for each supermartingale \(d\), some quotient of \(L\) is uncovered by \(d\). It is also possible to construct a single language \(L\) that is itself both ODD and uncovered.

We now present a set measurable in our measure on PSPACE but not in the sense of [10].

Partition \(\Sigma^n\) into blocks of \(n^2\) words. Label the first half of the blocks \(R^n_j, j = 1, \ldots, 2^n/2n^2\), and label the second half of the blocks \(S^n_j\):

\[R_1R_2 \ldots R_{2n^{2n}} S_1S_2 \ldots S_{2n^{2n}}.\]

Definition 28 Let MATCH be the set of sequences \(\omega\) such that for infinitely many pairs \((n, j)\), \(\omega[R^n_j] = \omega[S^n_j]\). Let MATCH denote the interior of MATCH.

Note that MATCH has Lebesgue measure zero, which is shown by using the Borel-Cantelli lemma: MATCH is the limsup of \(n\)-sections having measure

\[1 - \left(1 - \frac{1}{2^{n^2}}\right)^{2^n \log n - 1} \approx 1 - e^{-2^n - n^2}.\]

Since \(2^n - n^2\) is small we have \(e^{-2^n - n^2} \approx 1 - 2^n - n^2\), and so \((1 - e^{-2^n - n^2}) \approx 2^{-n^2}\) is exponentially small. Since \(\sum (1 - e^{-2^n - n^2}) < \infty\), we can apply the Borel-Cantelli lemma and conclude that MATCH has measure zero.

Theorem 29 The set MATCH has \(\Phi\) (PSPACE)-measure zero.

Proof. Similar to Theorem 11 above.

Theorem 30 The set MATCH \(\cap\) PSPACE does not have \(\Phi\) (PSPACE)-measure zero.
Proof. We first show that MATCH is not covered (the bulk of the proof), then we consider MATCH ∩ PSPACE, and finally MATCH ∩ PSPACE.

First, notation:

We will concentrate on one configuration at a time for the bulk of the proof. The variable \( y, |y| = n^2 \). The variable \( w, |w| = 2^n-1 \), denotes a setting of all the \( R^n \). A configuration will mean a configuration of the machine after reading through the \( R^n \) and before reading the \( S^n \).

We are given a supermartingale \( d \) computed by a \( \log^k(n) \)-space-bounded online Turing machine, which we may assume works in the limit (see [10]): \( \omega \) will be covered by \( d \) if \( \lim_{n \to \infty} \omega[1..n] \) exists and is infinite. Therefore, for a counterexample it suffices to construct a sequence \( \omega \) with \( \{d(\omega[0..2^n]) : n \in \mathbb{N}\} \) bounded. We will let \( d \) increase by a factor of \( 1 + O(1/n^2) \) at the \( n \)th stage, and, since \( \prod (1 + c/n^2) \leq \infty \), \( d \) will be bounded.

Now, an overview. There may be a configuration reached by only one string \( w \) of length \( 2^n-1 \). From that configuration, a supermartingale is prepared to make many successful bets on the second phase. So we will begin by excluding from consideration for our language the configurations reached by too few \( w \)'s. This leaves at least one configuration, since some configuration is reached by the average number of \( w \)'s. Also, the remaining configurations are reached by most of the \( w \)'s, so the average of \( d \) over the remaining configurations is not too large. Thus there is some configuration \( C \) reached by a large number of \( w \)'s and with \( d(C) \) not rising much. (These comments will be made quantitative below.)

If \( C \) is reached by many \( w \)'s, then for many \( j \) there are many settings \( y \) of \( \omega[R^n] \) that are consistent with \( C \) ("\( \omega[R^n] = y \) allows \( C \)).

For at most half of the \( j \)'s can there be a setting \( y \) of \( \omega[S^n] \), in which \( d \) falls significantly. For these \( j \)'s we give up on a match but make sure \( d \) decreases. But there are other \( j \)'s such that no \( y \) makes \( d \) drop much, and so by Markov's inequality most \( y \)'s make \( d \) drop or rise by very little. Combining this with the last paragraph, we've found a \( j \) and a \( y \) such that \( \omega[R^n] = y \) allows \( C \) and \( \omega[S^n] = y \) makes \( d \) not rise much. We've found our match; fill in the other bits according to the path of decreasing \( d \).

Now more formally and quantitatively:

Let \( D \) be the value of \( d(\omega[0..2^n-1]) \). (That is, \( D \) is the value of \( d \) after treating the previous \( n \).

Consider only the \( C \)'s that are reached by at least \( 1/4 \cdot 2^n \) of their fair share of \( w \)'s (e.g., if there are \( 2^n \) \( C \)'s, only the \( C \)'s reached by at least \( 2^{n-1} / (2^{n} n^2) \) of the \( 2^{n-1} \) \( w \)'s). Note this leaves at least one \( C \). Also, it leaves at least \( (1 - 1/n^2) \) of the \( w \)'s, so the average, over remaining \( w \)'s, of \( d(w) \) is at most \((1 + 2/n^2) D \). Fix one of the remaining \( C \)'s with \( d(C) \leq (1 + 2/n^2) D \).

Let \( W = \{w : C \text{ is reached by } w\} \); note that \( |W| \geq 2^{2^n-1} / (2^{n} n^2) \).

Initialize \( j \) to 0, and initialize \( S \) to the empty string. (In general, as \( j \) changes, \( S \) will contain the bits in positions \( S^0 \cup \cdots \cup S^{n-1} \).) We will talk about \( d(CS) \), and mean \( d(zS) \) where \( z \) is any string that takes the machine computing \( d \) to configuration \( C \).

For half of the \( 2^{n-2log n - 1} \) \( j \)'s, at least the fraction \( 3/4 \) of the \( 2^n \) possible \( y \)'s at \( j \) allow \( C \). Otherwise, if \( a \leq 2^{n-2log n - 2} \) of the \( j \)'s have this property, then

\[
|W| \leq \left( \frac{3}{4} \right)^{2^{n-2log n - 1}} (2^n)^a = \left( \frac{3}{4} \right)^{2^{n-2log n - 1}} \frac{1}{2^{2^{n-2log n - 1}}} \leq \left( \frac{3}{4} \right)^{2^{n-2log n - 2}} \frac{1}{2^{2^{n-1}}} \leq \frac{1}{2^{2^{n-1}}},
\]

a contradiction. Let \( A \) be the set of \( j \)'s such that \( 3/4 \) of the \( y \)'s at \( j \) allow \( C \).

Consider the \( j \)'s in increasing order. If \( j \notin A \) then reset \( S = S_y \), for \( y \) along the path of decreasing \( d \). If \( j \in A \), then if there's a setting \( y \) of \( S^n \) with \( d(CSy) < d(CS) - 2^{n-2log n - 3} D \), reset \( S = S_y \). After considering \( 2^{n-2log n - 3} \) of the \( j \)'s in \( A \) (i.e., less than half the \( j \)'s in \( A \)), either the value of \( d(CS) \) has decreased to zero or we've found a \( j \) such that for all \( y \) we have \( d(CSy) \geq d(CS) - 2^{n-2log n - 3} D \). Then, by
Markov, at least 3/4 of the $y$'s make $d(CSy) < d(CS) + 2^{5+2\log n-2} D$. Since $j \in A$, 3/4 of the $y$'s at $j$ allow $C$, so 1/2 of the $y$'s satisfy both

- $\omega[R^n_j] = y$ allows $C$
- $d(CSy) < d(CS) + 2^{5+2\log n-2} D$.

Fix one of these $y$'s, extend $S - Sy$, find a setting of the unspecified $R^n_j$'s witnessing that $y$ at $j$ allows $C$, and finally find a setting of the unspecified $S^n_i$'s along the path of decreasing $d$.

We have $d(\omega[0..2^{n+1} - 1]) \leq d(C) + 2^{5+2\log n-2} D \leq (1 + O(1)/n^2) D$, which was our goal.

Next we modify the above proof to produce a language in PSPACE.

Instead of finding $C$ with $d(C) < (1 + 2/n^2) D$ reached by $1 - 1/n^2$ of the $w$'s, find $C$ with $d(C) < (1 + 2/n^2) D$ such that for half of the $j$'s as least 3/4 of the $y$'s at $j$ allow $C$ (such a $C$ exists by the previous argument). This can be done by cycling through all $j$'s (counting as we go), for each $j$ cycling through the $y$'s, and for each $(j, y)$ using a Savitch divide-and-conquer technique to determine if $y$ at $j$ allows $C$. Later, as we consider the $j$'s in turn, instead of maintaining $S$, maintain only the configuration of $S$. The rest of the proof is similar.

Finally we consider MATCH. Fix a $\Phi(PSPACE)$-martingale $d$. For each fixed string $x$, an argument like the above shows how to construct, for each $x$ and for large enough $n$, a language slice $L \cap \Sigma^n$ such that $d$ rises by at most $O(1/n^2)$ on $L$ and $L/x$ has a match. (This match will be among words of length $n - |x|$, and the block structure of bits matched in $L/x$ will be different than those in $L$.)

Let $\{x_n\}$ be a sequence in which each word appears infinitely often, and $|x_n| \leq n$. For each $n$ in turn, construct $L \cap \Sigma^n$ so that $d$ barely rises on $L$ and $L/x_n$ has a match. We conclude that $d$ remains bounded on $L$, and since for each $x$ the language $L/x$ has infinitely many matches, we conclude $L/x \in MATCH$.

Note that under some ordering of $\Sigma^*$, namely the ordering in which $S^n_j$ immediately follows $R^n_j$, the set MATCH is $\Phi(PSPACE)$-null. Thus $\Phi(PSPACE)$ measure is not robust under PSPACE-reorderings.

### 7 Conclusions

We presented a notion of measure on $P$ that, compared with previous notions, satisfies more of the intuitive properties of measure. Still, some of these properties have to be weakened numerically. For example, we show that a random language in $P$ has density at least $\Omega(1)$, whereas a Lebesgue-random language has density near 1/2. It would be interesting to know if the stronger density or immunity property can be achieved by some other notion of measure that preserves all essential properties of measure.

Alternatively, if one can write $P$ as a subset of an enumerated union of sets $A_i$, then no reasonable notion of measure makes all the $A_i$'s null. Can one write $P$ as the union of “intuitively null” sets?

Resource-bounded genericity on $P$ may escape some of the problems of resource-bounded measure on $P$. This needs to be investigated.

### Acknowledgments

Section 3 is joint work with D. Sivakumar, and Section 6 is joint work with Eric Allender (the latter essentially appeared in [3]). I thank these researchers for permitting the inclusion of this work. I thank Allender, Wolfgang Merkle, Ken Regan, and Sivakumar for helpful discussions about other parts of the work.

### References


