Instance complexities of hard and weakly hard problems

by

Vikram Shantveer Mhetre

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Computer Science

Major Professor: Jack H. Lutz

Iowa State University

Ames, Iowa

1999

Graduate College
Iowa State University

This is to certify that the Master's thesis of

Vikram Shantveer Mhetre

has met the thesis requirements of Iowa State University

Signatures have been redacted for privacy

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Jack Lutz for his patience, encouragement and enthusiasm throughout this research and the writing of this thesis. I would also like to thank my fellow graduate student, Sridhar Srinivasan for useful discussions and help during the entire course of this work.

CHAPTER 1. INTRODUCTION

A problem that is computationally intractable in the worst case may or may not be intractable in the average case. In applications such as cryptography and derandomization, where intractability is a valuable resource, worst-case intractability seldom suffices, average-case intractability often suffices, and almost-everywhere intractability is sometimes required. Implicit in these distinctions is the truism that some instances of a computational problem may be hard while others are easy.

The complexity of an individual instance of a problem cannot be measured simply in terms of the running time required to solve that instance, because any algorithm for that problem can be modified to solve that instance quickly via a look-up table. Orponen, Ko, Schöning, and Watanabe [36] used ideas from algorithmic information theory to circumvent this difficulty, thereby introducing a precise formulation of the complexities of individual instances of computational problems.

Given a decision problem $A \subseteq \{0,1\}^*$, an instance $x \in \{0,1\}^*$, and a time bound $t : \mathbb{N} \to \mathbb{N}$, Orponen, Ko, Schöning, and Watanabe [36] defined the *t-time-bounded instance complexity of* x *relative to* A, written $ic^t(x : A)$, to be the number of bits in the shortest program π such that π decides x in at most $t(|x|)$ steps and π does not decide any string incorrectly for A. (See chapters 2 and 3 for complete definitions of this and other terms used in this introduction.) Instance complexity has now been investigated and applied in a number of papers, including [36, 22, 9, 17, 24, 8, 34], and is discussed at some length in the text [25].

In this paper we investigate the instance complexities of problems that are hard or weakly-hard for exponential time under polynomial time, many-one reductions. Our most technical results establish the measure-theoretic abundance of problems for which almost all instances

have essentially maximal instance complexities. From these results we derive our main results, which are lower bounds on the instance complexities of weakly hard problems, and we separately establish upper bounds on the instance complexities of hard problems. We now discuss these results in a little more detail.

The *t-time-bounded plain Kolmogorov complexity* of a string x, written $C^t(x)$, is the number of bits in the shortest program π that describes (i.e., prints) x in at most $t(|x|)$ steps. As observed in [36], it is easy to see that, for t' modestly larger than t, $ic^{t'}(x : A)$ cannot be much larger than $C^t(x)$, since a description of x contains all but one bit of the information required for a program to correctly decide whether $x \in A$ and decline to decide all other strings. An instance x thus has *essentially maximal* t-time-bounded instance complexity if $ic^t(x : A)$ is nearly as large as $C^{t'}(x)$, where t' is modestly larger than t. Orponen, Ko, Schöning, and Watanabe [36] established the existence of a problem $A \in \mathrm{E} = \mathrm{DTIME}(2^{\mathrm{linear}})$ for which all but finitely many instances x have instance complexities that are essentially maximal in the sense that $ic^{2^n}(x : A) > C^{t'}(x) - 2\log C^{t'}(x) - c$, where c is a constant and $t'(n) = cn2^{2n} + c$. In contrast with this existence result, we prove in this paper that *almost every* language $A \in \mathrm{E}$ has the property that all but finitely many instances x have essentially maximal instance complexities in the slightly weaker (but still very strong) sense that $ic^{2^n}(x : A) > (1 - \epsilon)C^{t'}(x)$, for any fixed real $\epsilon > 0$, where $t'(n) = 2^{3n}$.

Naturally arising problems that are – or are presumed to be – intractable have usually turned out to be complete for NP or some natural complexity class containing NP. The complexities of such problems are thus of greater interest than the complexities of arbitrary problems. The instance complexities of problems that are complete (or just hard) for NP or exponential time under $\leq^{\mathrm{P}}_m$-reductions have consequently been a focus of investigation.

Regarding problems that are $\leq^{\mathrm{P}}_m$-hard for exponential time, Orponen, Ko, Schöning and Watanabe [36] have shown that every such problem H must have an exponentially dense set of instances x that are hard in the sense that for every polynomial t, $ic^t(x : H) > C^{t'}(x) - 2\log C^{t'}(x) - c$, where c is a constant and $t'(n) = cn2^{2n} + c$. Buhrman and Orponen [9] proved a related result stating that, if H is actually $\leq^{\mathrm{P}}_m$-complete for exponential time, then H has

a dense set of instances x that are hard in the sense that for every polynomial $t(n) \geq n^2$, $ic^t(x : H) > C^t(x) - c$, where c is a constant.

The main results of this paper show that this phenomenon - a dense set of instances whose complexities are essentially maximal - holds not only for $\leq^{\mathrm{P}}_m$-hard problems for exponential time, but in fact for all *weakly* $\leq^{\mathrm{P}}_m$-hard problems for exponential time (with slight technical modifications in the instance complexity bounds). This is a significant extension of the earlier work because Ambos-Spies, Terwijn and Zheng [2] have shown that almost every problem in E is weakly $\leq^{\mathrm{P}}_m$-hard, but not $\leq^{\mathrm{P}}_m$-hard, for E, and similarly for E_2.

To be precise, we prove that for every weakly $\leq^{\mathrm{P}}_m$-hard language H for E_2 and every $\epsilon > 0$ there exists $\delta > 0$ such that the set of all instances x with $ic^{2^{n^\delta}}(x : H) > (1 - \epsilon)C^{2^{4n}}(x)$ is dense. Since Juedes and Lutz [21] have shown that every language that is weakly $\leq^{\mathrm{P}}_m$-hard for E is weakly $\leq^{\mathrm{P}}_m$-hard for E_2 (but not conversely, even for languages in E), our results hold *a fortiori* for problems that are weakly $\leq^{\mathrm{P}}_m$-hard for E.

Regarding problems that are NP-complete (of which we take SAT to be the canonical example), any nontrivial lower bound on instance complexity must be derived from some unproven hypothesis (or entail a proof that P $\neq$ NP) because languages in P have bounded instance complexities [36]. Assuming P $\neq$ NP, Orponen, Ko, Schöning and Watanabe [36] showed that for every polynomial t and constant c, the set $\{x | ic^t(x : SAT) \geq c \log |x|\}$ is infinite. Assuming the hypothesis that nonuniformly secure one-way functions exist (which implies P $\neq$ NP), Ko [22] proved that this set is nonsparse. Assuming E $\neq$ NE (which also implies P$\neq$NP), Orponen, Ko, Schöning and Watanabe [36] showed that SAT has an infinite set of instances of essentially maximal complexity in the sense that for every polynomial t there exist a polynomial t', a constant c, and infinitely many x such that $ic^t(x : SAT) > C^{t'}(SAT) - c$.

The hypothesis that NP does not have p-measure 0, written $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$, has been proposed by Lutz. This hypothesis has been shown to imply reasonable answers to many complexity-theoretic questions not known to be resolvable using P $\neq$ NP or other "traditional" complexity-theoretic hypotheses. (Such results are discussed in the surveys [28, 1, 27, 10].) The $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$ hypothesis implies the hypothesis E$\neq$NE [29] and is equivalent to the assertion that NP does

not have measure 0 in E_2[2]. Here we note that, if $\mu_p(NP) \neq 0$, then SAT is weakly $\leq_m^P$-hard for E_2, whence our above-mentioned results imply that SAT has a *dense* set of instances of essentially maximal complexity. That is, if $\mu_p(NP) \neq 0$, then for every $\epsilon > 0$ there exists $\delta > 0$ such that the set of all x for which $ic^{2^{n^\delta}}(x : SAT) > (1 - \epsilon)C^{2^{4n}}(x)$ is dense.

In the course of this introduction, we have seen that almost every problem A in exponential time has both of the following properties.

1. All but finitely many instances of A have essentially maximal instance complexity (our abundance results).

2. A is weakly $\leq_m^P$-hard for exponential time [2].

Thus weakly hard problems can have essentially maximal complexity at almost every instance. In contrast, we also show that every problem H that is actually $\leq_m^P$-hard for exponential time must have a dense set of instances x that are unusually easy in the very strong sense that $ic^{2^{6n}}(x : H)$ is bounded above by a constant. Our proof of this fact is based largely on the proof by Juedes and Lutz [20] of an analogous result for complexity cores.

It should be mentioned here that most of the work in this thesis is joint work with my advisor, Dr.Jack Lutz. Theorem 4.1 is joint work with Dr. Jack Lutz and my fellow graduate student, Sridhar Srinivasan. Also, most of the results in this thesis are for the complexity class E. It should be noted that these results can be extended to obtain similar results for the complexity class E_2 as in [30].

The rest of this thesis is organized as follows. In chapter 2 we summarize our basic terminology and notation and briefly review some basic aspects of resource-bounded measure. In chapter 3 we review the definition and basic properties of instance complexity, and we note that, for a language A, the condition of having very high t-time-bounded instance complexity is strongly incomparable with the condition of being incompressible by $\leq_m^{DTIME(t)}$-reductions. (In particular this implies that our lower bound results are much stronger than the analogous lower bounds for complexity cores proven in [20].) Chapter 4 is the main chapter of this thesis. In this chapter we prove our abundance theorems, derive our lower bounds on the instance complexities of weakly hard problems, and note the consequences for the complexity of SAT if

$\mu_p(\mathrm{NP}) \neq 0$. In chapter 5 we prove that every hard problem for exponential time has a dense set of unusually easy instances. Finally in chapter 6, we summarize the results in this thesis.

CHAPTER 2. PRELIMINARIES

We write $\mathbb{N}$ for the set of natural numbers, $\mathbb{Z}$ for the set of integers, and $\mathbb{Z}^+$ for the set of positive integers. All *polynomials* here have coefficients in $\mathbb{N}$, and all logarithms are base 2.

We write $[\![\varphi]\!]$ for the Boolean value of a condition φ, i.e., $[\![\varphi]\!] =$ if φ then 1 else 0. We write $|S|$ for the cardinality of a set S and S^c for the complement of S. In addition to the quantifiers $\exists x$ and $\forall x$, we use the quantifiers $\exists^\infty x$ ("there exists infinitely many x such that ... ") and $\forall^\infty x$ ("for all but finitely many x, ... ").

All *strings* in this paper are binary strings $x \in \{0,1\}^*$. We write $|x|$ for the *length* of x, and we use the *standard enumeration* $s_0 = \lambda, s_1 = 0, s_2 = 1, s_3 = 00, \ldots$ of $\{0,1\}^*$. A string x is a *prefix* of a string y, and we write $x \sqsubseteq y$, if there is a string z such that $xz = y$.

All *languages* (equivalently, *decision problems*) here are sets $A \subseteq \{0,1\}^*$. We identify each language A with its *characteristic sequence*

$$\chi_A = [\![s_0 \in A]\!][\![s_1 \in A]\!][\![s_2 \in A]\!] \ldots .$$

Relying on this identification, we write $A[0..n-1]$ for the binary string consisting of the first n bits of χ_A.

A language $A \subseteq \{0,1\}^*$ is *sparse* if there exists a polynomial q such that $(\forall n)\ |A \cap \{0,1\}^{\leq n}| \leq q(n)$, and *exponentially dense* (or, simply, *dense*) if there exists a real number $\epsilon > 0$ such that $(\forall^\infty n)\ |A \cap \{0,1\}^{\leq n}| > 2^{n^\epsilon}$.

Our main results involve *resource-bounded measure*, which was developed by Lutz [26, 28]. We briefly review a fragment of the theory that is sufficient for the purposes of this thesis. The interested reader is referred to any of the surveys [28, 27, 1, 10] for further discussion.

Definition. A *martingale* is a function $d : \{0,1\}^* \to [0,\infty)$ with the property that, for all

$w \in \{0,1\}^*$,

$$d(w) = \frac{d(w0) + d(w1)}{2}. \qquad (*)$$

A martingale d *succeeds* on a language $A \subseteq \{0,1\}^*$ if

$$\limsup_{n\to\infty} d(A[0..n-1]) = \infty.$$

The *success set* of a martingale d is

$$S^\infty[d] = \{A | d \text{ succeeds on } A\}.$$

Intuitively, a martingale d is a betting strategy that, given a language A, starts with capital (amount of money) $d(\lambda)$ and bets on the membership or nonmembership of the successive strings $s_0, s_1, s_2, \ldots$ in A. Prior to betting on a string s_n, the strategy has capital $d(w)$, where

$$w = [\![s_0 \in A]\!] \cdots [\![s_{n-1} \in A]\!].$$

After betting on the string s_n, the strategy has capital $d(wb)$, where $b = [\![s_n \in A]\!]$. Condition $(*)$ ensures that the betting is fair. The strategy succeeds on A if its capital is unbounded as the betting progresses.

Notation. The classes $\mathrm{p}_1 = \mathrm{p}$ and p_2, both consisting of functions $f : \{0,1\}^* \to \{0,1\}^*$, are defined by

$$\begin{aligned} \mathrm{p}_1 &= \mathrm{p} = \{f \mid f \text{ is computable in polynomial time}\}, \\ \mathrm{p}_2 &= \{f \mid f \text{ is computable in } n^{(\log n)^{O(1)}} \text{ time}\}. \end{aligned}$$

These classes induce measure structure on the classes $\mathrm{E}_1 = \mathrm{E}$ and E_2, respectively.

Definition. Let $i \in \{1,2\}$. A martingale d is p_i-*computable* if there is a function $\hat{d} : \mathbb{N} \times \{0,1\}^* \to \mathbb{Q}$ such that $\hat{d} \in \mathrm{p}_i$ (with input (r,w) coded in the form $0^r 1 w$) and, for all $r \in \mathbb{N}$ and $w \in \{0,1\}^*$,

$$|\hat{d}(r,w) - d(w)| \leq 2^{-r}.$$

A p_i-*martingale* is a martingale that is p_i-computable.

Martingales were introduced by Ville [44], and resource-bounded martingales were used extensively by Schnorr [39, 40, 41, 42] in his investigations of random and pseudorandom

sequences. Lutz [26] used resource-bounded martingales to induce measure structure on E and E_2 by means of the following definitions. Let X be a set of languages and let $i \in \{1,2\}$.

1. X has p_i-*measure 0*, and we write $\mu_{p_i}(X) = 0$, if there is a p_i-martingale d such that $X \subseteq S^\infty[d]$.

2. X has p_i-*measure 1*, and we write $\mu_{p_i}(X) = 1$, if $\mu_{p_i}(X^c) = 0$.

3. X has *measure 0 in* E_i, and we write $\mu(X \mid E_i) = 0$, if $\mu_{p_i}(X \cap E_i) = 0$.

4. X has *measure 1 in* E_i, and we write $\mu(X \mid E_i) = 1$, if $\mu(X^c \mid E_i) = 0$. In this case, we say that X contains *almost every* language in E_i.

We write $\mu(X|E_i) \neq 0$ to indicate that X does *not* have measure 0 in E_i. Note that this does *not* assert that "$\mu(X|E_i)$" has some nonzero value.

The following is obvious but useful.

Fact 2.1. For every set $X \subseteq \{0,1\}^*$,

$$\begin{array}{ccccc} \mu_p(X) = 0 & \Longrightarrow & \mu_{p_2}(X) = 0 & \Longrightarrow & \Pr[A \in X] = 0 \\ \Downarrow & & \Downarrow & & \\ \mu(X \mid E) = 0 & & \mu(X \mid E_2) = 0, & & \end{array}$$

where the probability $\Pr[A \in X]$ is computed according to the random experiment in which a language $A \subseteq \{0,1\}^*$ is chosen probabilistically, using an independent toss of a fair coin to decide whether each string $x \in \{0,1\}^*$ is in A.

It is shown in [26] that these definitions endow E and E_2 with internal measure structure. This structure justifies the intuition that, if $\mu(X|E) = 0$, then $X \cap E$ is a *negligibly small* subset of E (and similarly for E_2).

In addition to using the above definitions, we will use the resource-bounded first Borel-Cantelli lemma. The statement of this lemma uses the *unitary success set*

$$S^1[d] = \{A \mid (\exists n) d(A[0..n-1]) \geq 1\}$$

of a martingale d and the following notion of effective convergence.

Definition. A series $\sum_{n=0}^{\infty} a_n$ of nonnegative real numbers a_n is p-*convergent* if there is a function $m : \mathbb{N} \to \mathbb{N}$ such that $m \in \mathrm{p}$ (with input and output coded in unary) and

$$\sum_{n=m(k)}^{\infty} a_n \leq 2^{-k}$$

for all $k \in \mathbb{N}$.

Routine calculus proves the following lemma.

Lemma 2.2. Let $\epsilon > 0$. The series $\sum_{\pi \in \{0,1\}^*} 2^{-2^{\epsilon|\pi|}}$ is p-convergent.

The following lemma gives the case p of the resource-bounded generalization of the classical first Borel-Cantelli lemma.

Lemma 2.3.(Lutz [26]) Let $Z_0, Z_1, Z_2, \ldots$ be sets of languages, and let

$$Z = \{A | (\exists^{\infty} k) A \in Z_k\}.$$

Assume that there is a function $d : \mathbb{N} \times \{0,1\}^* \to [0, \infty)$ satisfying the following four conditions, where we write $d_k(w) = d(k, w)$.

1. d is p-computable.
2. For each $k \in \mathbb{N}$, d_k is a martingale.
3. For each $k \in \mathbb{N}$, $Z_k \subseteq S^1[d_k]$.
4. The series $\sum_{k=0}^{\infty} d_k(\lambda)$ is p-convergent.

Then $\mu_{\mathrm{p}}(Z) = 0$.

Recall that a language H is $\leq_m^{\mathrm{P}}$-*hard* for a class $\mathcal{C}$ of languages if $A \leq_m^{\mathrm{P}} H$ for all $A \in \mathcal{C}$, and $\leq_m^{\mathrm{P}}$-*complete* for $\mathcal{C}$ if $H \in \mathcal{C}$ and H is $\leq_m^{\mathrm{P}}$-hard for $\mathcal{C}$. Resource-bounded measure allowed Lutz to generalize these notions as follows. (We write $\mathrm{P}_m(H) = \{A | A \leq_m^{\mathrm{P}} H\}$.)

Definition. A language $H \subseteq \{0,1\}^*$ is *weakly* $\leq_m^{\mathrm{P}}$*-hard* for E (respectively, for E_2) if $\mu(\mathrm{P}_m(H)|\mathrm{E}) \neq 0$ (respectively, $\mu(\mathrm{P}_m(H)|\mathrm{E}_2) \neq 0$). A language $H \subseteq \{0,1\}^*$ is *weakly* $\leq_m^{\mathrm{P}}$*-complete* for E (respectively, for E_2) if $H \in \mathrm{E}$ (respectively, $H \in \mathrm{E}_2$) and H is weakly $\leq_m^{\mathrm{P}}$-hard for E (respectively, for E_2).

It is clear that every $\leq_m^{\mathrm{P}}$-hard language for E is weakly $\leq_m^{\mathrm{P}}$-hard for E, and similarly for E_2.

CHAPTER 3. INSTANCE COMPLEXITY AND RELATED MEASURES

In this section we review the basic properties of instance complexity and discuss its relationships with Kolmogorov complexity, complexity cores, bi-immunity and incompressibility by many-one reductions.

Following [36], we define an *interpreter* to be a deterministic Turing machine with a read-only *program tape*, a read-only *input tape*, a write-only *output tape*, and an arbitrary number of read/write *work tapes*, all with alphabet $\{0, 1, \sqcup\}$, where $\sqcup$ is the blank symbol. Given a program $\pi \in \{0,1\}^*$ on the program tape and an input $x \in \{0,1\}^*$ on the input tape, an interpreter M may eventually halt in an *accepting* configuration, a *rejecting* configuration, an *undecided* configuration, or an *output* configuration, or it may fail to halt. If M halts in an accepting configuration, we say that π *accepts* x on M, and we write $M(\pi, x) = 1$. If M halts in a rejecting configuration, we say that π *rejects* x on M, and we write $M(\pi, x) = 0$. In either of these two cases, we say that π *decides* x on M. If M halts in an undecided configuration, or if M fails to halt, we say that π *fails to decide* x on M, and we write $M(\pi, x) = \perp$. If M halts in an output configuration with output $y \in \{0,1\}^*$ on the output tape, we write $M(\pi, x) = y$. (If y is 0 or 1, the context will always make it clear whether "$M(\pi, x) = y$" refers to a decision or an output.)

We write $time_M(\pi, x)$ for the running time of M with program π and input x. If $M(\pi, x) = \perp$, we stipulate that $time_M(\pi, x) = \infty$.

A program π is *consistent* with a language $A \subseteq \{0,1\}^*$ relative to an interpreter M if for all $x \in \{0,1\}^*$, $M(\pi, x) \in \{[\![x \in A]\!], \perp\}$, i.e., π either decides x correctly for A or else fails to decide x.

We now recall the definition of time-bounded instance complexity, which is the main topic

of this paper.

Definition.(Orponen, Ko, Schöning and Watanabe[36]) Let M be an interpreter, $t : \mathbb{N} \to \mathbb{N}$, $A \subseteq \{0,1\}^*$, and $x \in \{0,1\}^*$. The *t-time-bounded instance complexity of x with respect to A* given M is

$$ic^t_M(x : A) = min\{|\pi| \mid \pi \text{ is consistent with } A \text{ relative to } M \text{ and } time_M(\pi, x) \leq t(|x|)\},$$

where min $\phi = \infty$.

Thus $ic^t_M(x : A)$ is the minimum number of bits required for a program π to decide x correctly for A on M, subject to the constraints that π is consistent with A relative to M and $M(\pi, x)$ does not run for more than $t(|x|)$ steps.

Note. Our definition of $ic^t_M(x : A)$ differs from that in [36] in that we do not require $M(\pi, y)$ to halt within $t(|y|)$ steps – or even to halt at all – for $y \neq x$. In our complexity-theoretic setting, with time-constructible functions t, this difference is technical and minor (at most a constant number of bits and a $\log t$ factor in the time bound), and it simplifies results such as Lemma 3.3 below. In other settings, such as that of the time-unbounded instance complexity conjecture [36], the halting behavior for $y \neq x$ is a more critical issue.

We next recall the definition of plain Kolmogorov complexity.

Definition.(Solomonoff[43], Kolmogorov[23] and Chaitin[11, 12]) Let M be an interpreter, $t : \mathbb{N} \to \mathbb{N}$, and $x \in \{0,1\}^*$.

1. The *plain Kolmogorov complexity* of x relative to M is

$$C_M(x) = min\{|\pi| \mid M(\pi, \lambda) = x\}.$$

2. The *t-time bounded plain Kolmogorov complexity* of x relative to M is

$$C^t_M(x) = min\{|\pi| \mid M(\pi, \lambda) = x \text{ and } time_M(\pi, \lambda) \leq t(|x|)\}.$$

(We again stipulate that min $\phi = \infty$.)

The plain Kolmogorov complexity of x is thus the minimum number of bits required to "describe" x using the interpreter M. This information content measure and its time-bounded variant have been discussed extensively in the literature. We refer the reader to the text by Li and Vitanyi [25] for a comprehensive treatment.

The existence of optimal interpreters is of fundamental importance for both instance complexity and Kolmogorov complexity.

Definition. Let U be an interpreter.

1. U is *optimal for plain Kolmogorov complexity* if for every interpreter M there is a constant $c_M \in \mathbb{N}$ such that for all $x \in \{0,1\}^*$,

$$C_U(x) \leq C_M(x) + c_M.$$

2. U is *efficiently optimal for plain Kolmogorov complexity* if for every interpreter M there is a constant $c_M \in \mathbb{N}$ such that for all time bounds $t : \mathbb{N} \to \mathbb{N}$ and all $x \in \{0,1\}^*$,

$$C_U^{t'}(x) \leq C_M^t(x) + c_M,$$

where $t'(n) = c_M t(n) \log(t(n)) + c_M$.

3. U is *efficiently optimal for instance complexity* if for every interpreter M there is a constant $c_M \in \mathbb{N}$ such that for all time bounds $t : \mathbb{N} \to \mathbb{N}$, all $A \subseteq \{0,1\}^*$, and all $x \in \{0,1\}^*$,

$$ic_U^{t'}(x : A) \leq ic_M^t(x : A) + c_M,$$

where $t'(n) = c_M t(n) \log(t(n)) + c_M$.

The existence of optimal interpreters for plain Kolmogorov complexity was proven by Solomonoff [43], Kolmogorov [23], and Chaitin [12]. Standard techniques extend this to efficient optimality, and Orponen, Ko, Schöning, and Watanabe [36] noted that this also achieves efficient optimality for instance complexity. We thus have the following well-known theorem.

Theorem 3.1.(Optimality Theorem) There is an interpreter U that is efficiently optimal for both plain Kolmogorov complexity and instance complexity.

Following standard practice, we fix an interpreter U as in Theorem 3.1 and omit it from the notation, writing $C(x) = C_U(x)$, $C^t(x) = C_U^t(x)$, and $ic^t(x : A) = ic_U^t(x : A)$. These three quantities are then simply called the *plain Kolmogorov complexity* of x, the *t-time-bounded plain Kolmogorov complexity* of x, and the *t-time-bounded instance complexity* of x *with respect to* A, respectively.

Intuitively, the instance complexity of a string cannot be much greater than its Kolmogorov complexity, since a description of the string is all but one bit of the information needed to correctly decide that string and decline to decide all others. The following known result formalizes this intuition.

Theorem 3.2.(Orponen, Ko, Schöning, and Watanabe [36]) For every time constructible function $t : \mathbb{N} \to \mathbb{N}$ there is a constant $c \in \mathbb{N}$ such that for all $A \subseteq \{0,1\}^*$ and $x \in \{0,1\}^*$

$$ic^{t'}(x : A) \leq C^t(x) + c,$$

where $t'(n) = ct(n)\log(t(n)) + c$.

Complexity cores were first introduced by Lynch [31] and have been studied extensively over the past fifteen years [16, 14, 35, 37, 6, 18, 38, 7, 15, 19, 20].

Definition. Let $t : \mathbb{N} \to \mathbb{N}$ be a time bound and let $A, K \subseteq \{0,1\}^*$. Then K is a DTIME(t)-*complexity core* of A if for every $c \in \mathbb{N}$ and every program $\pi \in \{0,1\}^*$ that is consistent with A on U, the "fast set"

$$F = \{x |\ time_U(\pi, x) \leq ct(|x|) + c\}$$

satisfies $|F \cap K| < \infty$.

Note that every subset of a DTIME(t)-complexity core of A is a DTIME(t)-complexity core of A. Also, if $s(n) = O(t(n))$, then every DTIME(t)-complexity core of A is a DTIME(s)-complexity core of A.

Definition. Let $A, K \subseteq \{0,1\}^*$.

1. K is a *polynomial complexity core* (briefly, a P-*complexity core*) of A if for every $k \in \mathbb{N}$, K is a DTIME(n^k)-complexity core of A.

2. K is an *exponential complexity core* of A if there is a real number $\epsilon > 0$ such that K is a DTIME(2^{n^ϵ})-complexity core of A.

The following lemma, which is a straightforward extension of a result of Orponen, Ko, Schöning and Watanabe [36] (see Corollary 3.4 below), relates instance complexity to complexity cores.

Lemma 3.3. Let $t : \mathbb{N} \to \mathbb{N}$ be a time bound, and let $A, K \subseteq \{0,1\}^*$. Then K is a DTIME(t)-complexity core of A if and only if for every $c \in \mathbb{N}$ there are only finitely many $x \in K$ for which $ic^{ct+c}(x : A) \leq c$.

Proof. To prove the if part, suppose there exists $c \in \mathbb{N}$ such that there are infinitely many $x \in K$ for which $ic^{ct+c}(x : A) \leq c$. Hence there are infinitely many $x \in K$, for which there exists a program π testifying to the value of $ic^{ct+c}(x : A)$ with $|\pi| \leq c$. Also, there exist only finitely many programs π such that $|\pi| \leq c$. Therefore there is at least one program π_0 with $|\pi_0| \leq c$, which testifies to the value of $ic^{ct+c}(x : A)$ for infinitely many $x \in K$. Hence π_0 is consistent with A on U and the "fast set"

$$F = \{x |\ time_U(\pi_0, x) \leq ct(|x|) + c\}$$

satisfies $|F \cap K| = \infty$. This implies that K is not a DTIME(t)-complexity core of A.

To prove the only if part, suppose K is not a DTIME(t)-complexity core of A. Hence, there exists $c_0 \in \mathbb{N}$ and a program $\pi \in \{0,1\}^*$ that is consistent with A on U, such that the "fast set"

$$F = \{x |\ time_U(\pi, x) \leq c_0 t(|x|) + c_0\}$$

satisfies $|F \cap K| = \infty$. Let $c = max(|\pi|, c_0)$. Therefore we have, $\forall x \in F \cap K$,

$$ic^{ct+c}(x : A) \leq ic^{c_0 t + c_0}(x : A) \leq |\pi| \leq c.$$

Hence there are infinitely many $x \in K$ for which $ic^{ct+c}(x : A) \leq c$. □

Corollary 3.4. Let $A, K \subseteq \{0,1\}^*$.

1. (Orponen, Ko, Schöning and Watanabe [36]). K is a polynomial complexity core of A if and only if for every polynomial t and every constant $c \in \mathbb{N}$, there are only finitely many $x \in K$ for which $ic^t(x : A) \leq c$.

2. K is an exponential complexity core of A if and only if there is a real number $\epsilon > 0$ such that for every constant $c \in \mathbb{N}$, there are only finitely many $x \in K$ for which $ic^{2^{n^\epsilon}}(x : A) \leq c$.

Corollary 3.5. Let $A \subseteq \{0,1\}^*$ and let $t : \mathbb{N} \to \mathbb{N}$ be time constructible. Then $\{0,1\}^*$ is a DTIME(t)-complexity core of A if and only if for every $c \in \mathbb{N}$,

$$\lim_{n\to\infty} ic^{ct+c}(s_n : A) = \infty,$$

where s_0,s_1, ... is the standard enumeration of $\{0,1\}^*$.

Remark 3.6. Having $\{0,1\}^*$ as a complexity core is a very strong intractability property that Balcázar and Schöning [4] have shown to be closely related to complexity-theoretic bi-immunity. Specifically, every language that is DTIME(t)-bi-immune has $\{0,1\}^*$as a DTIME(t)-complexity core, and, almost conversely, there is a constant $c \in \mathbb{N}$ such that if $t'(n) = ct(n)\log t(n) + c$, then every language that has $\{0,1\}^*$ as a DTIME(t')-complexity core is DTIME(t)-bi-immune. In particular, a language is P-bi-immune if and only if it has $\{0,1\}^*$as a P-complexity core.

Corollary 3.7.(Orponen, Ko, Schöning and Watanabe [36]) A language $A \subseteq \{0,1\}^*$ has $\{0,1\}^*$ as a P-complexity core if and only if for every polynomial t,

$$\lim_{n\to\infty} ic^t(s_n : A) = \infty.$$

Incompressibility by many-one reductions, an idea originally introduced by Meyer [33], has played a central role in earlier work on complexity cores and instance complexity and continues this role in the present thesis.

Definition. Let $f : \{0,1\}^* \to \{0,1\}^*$.

1. The *collision set* of f is the set

$$C_f = \{x \in \{0,1\}^* | (\exists y < x) f(y) = f(x)\},$$

where "$<$" refers to the standard ordering of $\{0,1\}^*$.

2. f is *one-to-one almost everywhere* (briefly, *one-to-one a.e.*) if C_f is finite. (Note that f is one-to-one if and only if $C_f = \phi$.)

Definition. Let $A, B \subseteq \{0,1\}^*$, and let $t : \mathbb{N} \to \mathbb{N}$.

1. A $\leq_m^{\mathrm{DTIME}(t)}$*-reduction of A to B* is a function $f \in \mathrm{DTIMEF}(t)$ such that $A = f^{-1}(B)$, i.e., for all $x \in \{0,1\}^*$, $x \in A$ iff $f(x) \in B$.

2. A $\leq_m^{\mathrm{DTIME}(t)}$*-reduction of A* is a $\leq_m^{\mathrm{DTIME}(t)}$-reduction of A to $f(A)$. (Note that f is a $\leq_m^{\mathrm{DTIME}(t)}$-reduction of A if and only if there exists $B \subseteq \{0,1\}^*$ such that f is a $\leq_m^{\mathrm{DTIME}(t)}$-reduction of A to B.)

3. A is *incompressible by* $\leq_m^{\mathrm{DTIME}(t)}$*-reductions* if every $\leq_m^{\mathrm{DTIME}(t)}$-reduction of A is one-to-one a.e..

4. A is *incompressible by* $\leq_m^{\mathrm{P}}$*-reductions* if A is incompressible by $\leq_m^{\mathrm{DTIME}(t)}$-reductions for every polynomial t.

Lemma 3.8.(Juedes and Lutz [20]) Let $A \subseteq \{0,1\}^*$, and let $t : \mathbb{N} \to \mathbb{N}$ be time constructible. If A is incompressible by $\leq_m^{\mathrm{DTIME}(t)}$-reductions, then A has $\{0,1\}^*$ as a $\mathrm{DTIME}(t)$-complexity core.

Corollary 3.9.(Balcázar and Schöning [4]) Every language that is incompressible by $\leq_m^{\mathrm{P}}$-reductions has $\{0,1\}^*$ as a P-complexity core.

In light of Theorem 3.2, we consider a language $A \subseteq \{0,1\}^*$ to have *essentially maximal* t-time-bounded instance complexity if for all but finitely many $x \in \{0,1\}^*$, $ic^t(x : A)$ is nearly as large as $C^{t'}(x)$, where the growth rate of t' is only modestly greater than that of t. The relations "nearly as large as" and "only modestly greater than" here can (and will) be made

precise in a variety of ways, depending upon the particular application. The following known theorem establishes the existence of languages in E that have essentially maximal time-bounded instance complexity in a very strong sense.

Theorem 3.10.(Orponen, Ko, Schöning and Watanabe [36]) There exist a language $A \in \mathrm{E}$ and a constant $c \in \mathbb{N}$ such that for all $x \in \{0,1\}^*$,

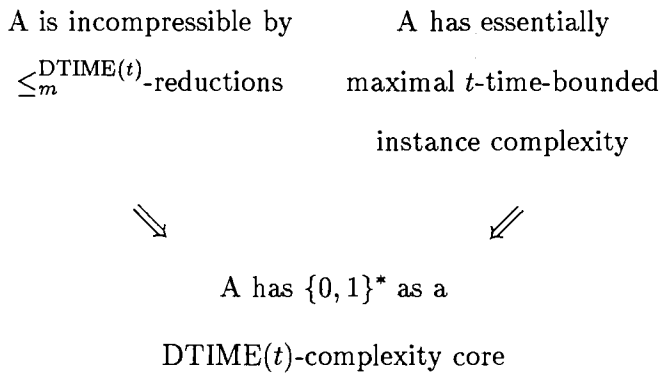
$$ic^t(x : A) \geq C^{t'}(x) - 2\log C^{t'}(x) - c,$$

where $t(n) = 2^n$ and $t'(n) = cn2^{2n} + c$.

It is well-known and easy to see that $\lim_{n\to\infty} C(s_n) = \infty$. Since $C(x) \leq C^t(x)$ for all t and x, it follows that $\lim_{n\to\infty} C^t(s_n) = \infty$ for all t. Hence, if a language $A \subseteq \{0,1\}^*$ has essentially maximal t-time-bounded instance complexity in any reasonable sense, it will satisfy the condition

$$\lim_{n\to\infty} ic^t(s_n : A) = \infty.$$

Corollary 3.5 and Lemma 3.8 thus give us the following implications for all $A \subseteq \{0,1\}^*$ and all time-constructible $t : \mathbb{N} \to \mathbb{N}$.



Due to Remark 3.6 and the implication on the left (and the failure of its converse), incompressibility by many-one reductions is sometimes called "strong bi-immunity."

We conclude this section by showing that incompressibility by $\leq_m^{\mathrm{DTIME}(t)}$-reductions and essentially maximal t-time-bounded instance complexity are incomparable conditions, (i.e.,

neither implies the other), whence the above-displayed implications are the *only* implications that hold among these three strong intractability properties.

Theorem 3.11. For all $c \in \mathbb{Z}^+$ and $\epsilon > 0$, there exist $A, B \in \mathrm{E}$ with the following properties.

1. For all but finitely many $x \in \{0,1\}^*$,

$$ic^{2^{cn}}(x : A) > (1-\epsilon)C^{2^{(c+4)n}}(x).$$

2. A is not incompressible by $\leq_m^{\mathrm{P}}$-reductions.

3. B is incompressible by $\leq_m^{\mathrm{DTIME}(2^{cn})}$-reductions.

4. For all sufficiently large $n \in \mathbb{N}$,

$$Pr_{x \in \{0,1\}^n}[ic^{n^2}(x : B) < \epsilon C(x)] > 1 - \epsilon,$$

where x is chosen according to the uniform probability measure on $\{0,1\}^n$.

Proof. Fix $c \in \mathbb{Z}^+$ and $\epsilon > 0$. By Theorem 4.1 there is a language $D \in \mathrm{E}$ such that for all but finitely many $x \in \{0,1\}^*$,

$$ic^{2^{(c+1)n}}(x : D) > (1 - \frac{\epsilon}{2})C^{2^{(c+3)n}}(x).$$

Let $A = \{bx | x \in D \text{ and } b \in \{0,1\}\}$. It is clear $A \in \mathrm{E}$, and A has property 2 because the function that deletes the first bit of every nonempty string is a $\leq_m^{\mathrm{P}}$-reduction of A to D that is not one-to-one a.e. To see that A has property 1, let $b \in \{0,1\}$ be arbitrary, and let M_1 be an interpreter such that for all π,$x \in \{0,1\}^*$, $M_1(\pi, x)$ simulates $U(\pi, bx)$. For all $x \in \{0,1\}^*$, $x \in D \Leftrightarrow bx \in A$, so if a program π testifies to the value of $ic^{2^{cn}}(bx : A)$, then $ic_{M_1}^{2^{c(n+1)}}(x : D) \leq |\pi|$, i.e., $ic_{M_1}^{2^{cn+c}}(x : D) \leq ic^{2^{cn}}(bx : A)$. It follows by an application of Theorem 3.1 that there is a constant $a \in \mathbb{N}$ such that for all sufficiently large $x \in \mathbb{N}$, $ic^{2^{(c+1)n}}(x : D) \leq ic^{2^{cn}}(bx : A) + a$. We now have that for all but finitely many $x \in \mathbb{N}$,

$$\begin{aligned} ic^{2^{cn}}(bx : A) &\geq ic^{2^{(c+1)n}}(x : D) - a \\ &> (1 - \frac{\epsilon}{2})C^{2^{(c+3)n}}(x) - a \\ &\geq (1-\epsilon)C^{2^{(c+4)n}}(bx). \end{aligned}$$

Since $b \in \{0,1\}$ is arbitrary here, it follows that A has property 1.

It is well known that there is a language in E that is incompressible by $\leq_m^{\mathrm{P}}$-reductions and contains at most one string of each length. (For example, this is Theorem 6.2 in the text [3].)It is routine to modify this construction to obtain a language $B \in \mathrm{E}$ that is incompressible by $\leq_{\mathrm{m}}^{\mathrm{DTIME}(2^{cn})}$-reductions and contains at most one string of each length. This language B clearly has property 3. To see that it has property 4, let $l = \lceil \log(\frac{2}{\epsilon}) \rceil$, so that $2^{-l} \leq \frac{\epsilon}{2}$. For each $\pi \in \{0,1\}^l$ and $n \in \mathbb{N}$, let $\pi * n = 0^{|\pi|}1\pi s_n$, where s_n is the n^{th} string in the standard enumeration of $\{0,1\}^*$. Let M_2 be an interpreter such that for all $n \in \mathbb{N}$ and $\pi, x \in \{0,1\}^*$,

$$M_2(\pi * n, x) = \begin{cases} 0 & \text{if } |x| = n \geq l \text{ and } \pi \not\sqsubseteq x \\ \perp & \text{otherwise.} \end{cases}$$

It is clear that $time_{M_2}(\pi * n, x) = O(n)$ for all $\pi \in \{0,1\}^l$, $n \geq l$, and $x \in \{0,1\}^n$. For each $n \geq l$, define $\pi_n \in \{0,1\}^l$ as follows. If $B \cap \{0,1\}^n = \{w\}$, then $\pi_n = w[0..l-1]$ is the l-bit prefix of w. Otherwise (i.e., if $B \cap \{0,1\}^n = \phi$), $\pi_n = 0^l$. Note that for all $n \geq l$, the program $\pi_n * n$ is consistent with B relative to M_2.

For each $n \geq l$, let $S_n = \{x \in \{0,1\}^n | \pi_n \not\sqsubseteq x\}$. Then for all $n \geq l$ and $x \in S_n$, the program $\pi * n$ decides x in $O(|x|)$ steps on M_2, so there is a constant $a \in \mathbb{N}$ such that

$$ic_{M_2}^{an+a}(x : B) \leq |\pi * n| \leq \log(|x|+1) + 2l + 1.$$

It follows by an application of Theorem 3.1 that there are constants $n_1 \geq l$ and $b \in \mathbb{N}$ such that for all $n \geq n_1$ and all $x \in S_n$,

$$ic^{n^2}(x : B) < \log(n+1) + 2l + b.$$

Choose $n_2 \geq n_1$ such that for all $n \geq n_2$, $\epsilon n > \log(n+1) + (2+\epsilon)l + b$.

Let $n \geq n_2$, and let $x \in \{0,1\}^n$ be chosen according to the uniform distribution on $\{0,1\}^n$.

It is well-known [25] that $Pr[C(x) < n - l] < 2^{-l}$, so

$$\begin{aligned} Pr[ic^{n^2}(x:B) < \epsilon C(x)] &\geq Pr[ic^{n^2}(x:B) < \log(n+1) + 2l + b] \\ &\quad - Pr[\epsilon C(x) \leq \log(n+1) + 2l + b] \\ &\geq Pr[x \in S_n] - Pr[\epsilon C(x) < \epsilon(n-l)] \\ &= Pr[x \in S_n] - Pr[C(x) < (n-l)] \\ &> 1 - 2^{-l} - 2^{-l} \\ &\geq 1 - \epsilon. \end{aligned}$$

Thus B has property 4. □

CHAPTER 4. HARD INSTANCES

In this chapter we prove our main results. We show that almost every instance of almost every problem in E has essentially maximal instance complexity. Using this, we show that every problem that is weakly $\leq_m^{\mathrm{P}}$-hard for E has an exponentially dense set of such maximally hard instances.

4.1 Abundance of problems having hard instances almost everywhere

In this section we prove our abundance theorem in E. In contrast with Theorem 3.10, which asserts the *existence* of a language in E with essentially maximal instance complexity, the following result says that *almost every* language in E has this property, albeit with a slightly weaker interpretation of "essentially maximal".

Theorem 4.1. For all $c \in \mathbb{Z}^+$ and $\epsilon > 0$, the set

$$X(c,\epsilon) = \{A | (\forall^\infty x) ic^{2^{cn}}(x : A) > (1-\epsilon) C^{2^{(c+2)n}}(x)\}$$

has p-measure 1, hence measure 1 in E.

Proof. Fix $c \in \mathbb{Z}^+$ and $\epsilon > 0$, assuming without loss of generality that ϵ is rational, and let $X(c,\epsilon)$ be the indicated set. For each $\pi \in \{0,1\}^*$, define the sets

$$Cons(\pi) = \{A | \pi \text{ is consistent with } A \text{ relative to } U\},$$

$$D(\pi) = \{x | time_U(\pi, x) \leq 2^{c|x|}\},$$

$$Y_\pi = \{A \in Cons(\pi) \mid |D(\pi)| = \infty\},$$

$$Z_\pi = \{A \in Cons(\pi) \mid |D(\pi)| \geq 2^{\frac{\epsilon|\pi|}{4}}\}.$$

(Note that our definition of $time_U(\pi, x)$ implies that π decides x on U for all $x \in D(\pi)$.) Let

$$Y = \{A | (\exists \pi \in \{0,1\}^*) A \in Y_\pi\},$$

$$Z = \{A | (\exists^\infty \pi \in \{0,1\}^*) A \in Z_\pi\}.$$

It clearly suffices to prove the following three claims.

Claim 1. $Y^c \cap Z^c \subseteq X(c, \epsilon)$.

Claim 2. $\mu_{\mathrm{p}}(Y) = 0$.

Claim 3. $\mu_{\mathrm{p}}(Z) = 0$.

To prove Claim 1, let $A \in Y^c \cap Z^c$. Define the sets $B = \{\pi | A \in Z_\pi\}$, $D = \bigcup_{\pi \in B} D(\pi)$. Note that each $D(\pi)$ is finite because $A \in Y^c$ and B is finite because $A \in Z^c$. Thus the set D is finite.

For each $\pi \in \{0,1\}^*$ and $k \in \mathbb{N}$, let $\pi * k = 0^{|s_k|} 1 s_k \pi$. It is routine to design an interpreter M for which there is a constant $a \in \mathbb{N}$ such that the following two conditions hold whenever $\pi \in \{0,1\}^*$ and $0 \leq k < |D(\pi)|$.

1. $M(\pi * k, \lambda)$ is the k^{th} element of $D(\pi)$ in the standard ordering of $\{0,1\}^*$.
2. $time_M(\pi * k, \lambda) \leq 2^{(c+1)n+a}$, where $n = |M(\pi * k, \lambda)|$. (This is enough time for M to simulate 2^{n+1} computations of the form $U(\pi, x)$, each for up to $2^{c|x|}$ steps, for strings $x \in \{0,1\}^{\leq n}$).

Now assume that $x \notin D$. Let π be a program testifying to the value of $ic^{2^{cn}}(x : A)$. Then $x \in D(\pi)$, so $\pi \notin B$, so $|D(\pi)| < 2^{\frac{\epsilon|\pi|}{4}}$. This implies that x is the k^{th} element of $D(\pi)$ for some $0 \leq k \leq 2^{\frac{\epsilon|\pi|}{4}} - 1$, whence $M(\pi * k, \lambda) = x$ and $time_M(\pi * k, \lambda) \leq 2^{(c+1)|x|+a}$. Letting

$t(n) = 2^{(c+1)n+a}$, it follows that

$$\begin{aligned} C_M^t(x) &\leq |\pi * k| \\ &= |\pi| + 2|s_k| + 1 \\ &\leq |\pi| + 2\log(k+1) + 1 \\ &\leq |\pi| + \frac{\epsilon|\pi|}{2} + 1 \\ &= (1 + \frac{\epsilon}{2})|\pi| + 1. \end{aligned}$$

This argument shows that, for all $x \notin D$,

$$C_M^t(x) \leq (1 + \frac{\epsilon}{2}) ic^{2^{cn}}(x : A) + 1.$$

Let c_M be the optimality constant for M, and let $t'(n) = c_M t(n) \log t(n) + c_M$. Then the set

$$D' = \{x | 2^{(c+2)|x|} < t'(|x|)\}$$

is finite. Since $A \notin Y$, the set

$$D'' = \{x | ic^{2^{cn}}(x : A) \leq \frac{2(c_M + 1)}{\epsilon}\}$$

is also finite. Hence the set $\widehat{D} = D \cup D' \cup D''$ is finite. For all $x \notin \widehat{D}$, we now have

$$\begin{aligned} C^{2^{(c+2)n}}(x) &\leq C^{t'}(x) \\ &\leq C_M^t(x) + c_M \\ &\leq (1 + \frac{\epsilon}{2}) ic^{2^{cn}}(x : A) + c_M + 1 \\ &< (1 + \epsilon) ic^{2^{cn}}(x : A), \end{aligned}$$

whence

$$\begin{aligned} ic^{2^{cn}}(x : A) &> (1 - \epsilon^2) ic^{2^{cn}}(x : A) \\ &> (1 - \epsilon) C^{2^{(c+2)n}}(x). \end{aligned}$$

This proves that $A \in X(c, \epsilon)$, completing the proof of Claim 1.

To prove Claim 2, let $d = \sum_{i=0}^{\infty} 2^{-i} d_{s_i}$, where $s_0, s_1, \ldots$ is the standard enumeration of $\{0,1\}^*$ and for each $\pi \in \{0,1\}^*$, the function $d_\pi : \{0,1\}^* \to \mathbb{Q}$ is defined by the following recursion.

1. $d_\pi(\lambda) = 1$.

2. If $w \in \{0,1\}^*$, $b \in \{0,1\}$, and π does not decide $s_{|w|}$ on U in at most $2^{c|s_{|w|}|}$ steps, then $d_\pi(wb) = d_\pi(w)$.

3. If $w \in \{0,1\}^*$, $b \in \{0,1\}$, and π decides $s_{|w|}$ on U in at most $2^{c|s_{|w|}|}$ steps, then $d_\pi(wb) = 2d_\pi(w)[\![b = U(\pi, s_{|w|})]\!]$.

It is clear that each d_π is a martingale, whence d is a martingale.

To see that $Y \subseteq S^\infty[d]$, let $A \in Y$. Then there exists $\pi = s_i \in \{0,1\}^*$ such that $A \in Y_\pi$, i.e., $A \in Cons(\pi)$ and $|D(\pi)| = \infty$. Since $A \in Cons(\pi)$ we have $d_\pi(A[0..n]) \geq d_\pi(A[0..n-1])$ for all $n \in \mathbb{N}$. Since $|D(\pi)| = \infty$, we have $d_\pi(A[0..n]) = 2d_\pi(A[0..n-1])$ for infinitely many $n \in \mathbb{N}$. It follows that $\lim_{n\to\infty} d_\pi(A[0..n-1]) = \infty$, whence $A \in S^\infty[d_\pi] = S^\infty[d_{s_i}] \subseteq S^\infty[d]$.

To see that d is p-computable define $\hat{d} : \mathbb{N} \times \{0,1\}^* \to \mathbb{Q}$ by

$$\hat{d}(r, w) = \sum_{i=0}^{r+|w|} 2^{-i} d_{s_i(w)}.$$

Then $\hat{d} \in \mathrm{p}$ and for all $r \in \mathbb{N}$ and $w \in \{0,1\}^*$,

$$\begin{aligned} |\hat{d}(r,w) - d(w)| &= \sum_{i=r+|w|+1}^{\infty} 2^{-i} d_{s_i}(w) \\ &\leq 2^{|w|} \sum_{i=r+|w|+1}^{\infty} 2^{-i} \\ &= 2^{-r}, \end{aligned}$$

so $\hat{d}$ testifies that d is p-computable. We have now shown that d is a p-martingale with $Y \subseteq S^\infty[d]$, thereby proving Claim 2.

To prove Claim 3, we use the resource-bounded first Borel-Cantelli lemma (Lemma 2.3). Specifically, define $d : \mathbb{N} \times \{0,1\}^* \to [0,\infty)$ by

$$d(k, w) = 2^{-2^{\frac{\epsilon|s_k|}{4}}} d_{s_k}(w),$$

where each d_π (i.e., each d_{s_k}) is defined exactly as in the proof of Claim 2 above. It is clear that each $d_k = 2^{-2^{\frac{\epsilon|s_k|}{4}}} d_{s_k}$ is a martingale and that d is p-computable. To see that each

$Z_{s_k} \subseteq S^1[d_k]$, fix $k \in \mathbb{N}$, let $\pi = s_k$, and let $A \in Z_\pi$. Then $A \in Cons(\pi)$, so $d_\pi(A[0..n]) \geq d_\pi(A[0..n-1])$ for all $n \in \mathbb{N}$. Also, $|D(\pi)| \geq 2^{\frac{\epsilon|\pi|}{4}}$, so there are at least $2^{\frac{\epsilon|\pi|}{4}}$ values of n for which $d_\pi(A[0..n]) = 2d_\pi(A[0..n-1])$. It follows that, for sufficiently large n,

$$d_\pi(A[0..n-1]) \geq 2^{2^{\frac{\epsilon|\pi|}{4}}},$$

whence $d_k(A[0..n-1]) \geq 1$. Thus $A \in S^1[d_k]$, completing the proof that $Z_{s_k} \subseteq S^1[d_k]$.

The series $\sum_{k=0}^{\infty} d_k(\lambda) = \sum_{\pi \in \{0,1\}^*} 2^{-2^{\frac{\epsilon|\pi|}{4}}}$ is p-convergent by Lemma 2.2, so Claim 3 now follows from Lemma 2.3. □

Before proceeding, we note that Theorem 4.1 implies the following known fact, which was proven independently by Juedes and Lutz [20] (as stated) and Mayordomo [32] (in terms of bi-immunity, which is equivalent by Remark 3.6).

Corollary 4.2.(Juedes and Lutz [20], Mayordomo [32]) Let $c \in \mathbb{Z}^+$. Almost every language in E has $\{0,1\}^*$ as a DTIME(2^{cn})-complexity core.

Proof. This follows immediately from Theorem 4.1, Corollary 3.5, and the fact that $\lim_{n\to\infty} C(s_n) = \infty$. □

4.2 Hard instances of weakly hard problems

Our next task is to use Theorem 4.1 to prove that every weakly $\leq_m^P$-hard language for exponential time has a dense set of very hard instances. For this purpose we need a few basic facts about the behavior of polynomial-time reductions in connection with time-bounded Kolmogorov complexity, time-bounded instance complexity, and density.

The *data processing inequality* of classical information theory [13] says that the entropy (Shannon information content) of a source cannot be increased by performing a deterministic computation on its output. The analogous data processing inequality for plain Kolmogorov complexity [25] says that if f is a computable function, then $C(f(x))$, which is the algorithmic information content of $f(x)$, cannot exceed $C(x)$, the algorithmic information content of x, by

more than a constant number of bits. The following lemma is a time-bounded version of this fact. It is essentially well-known, though perhaps not in precisely this form.

Lemma 4.3.(data processing inequality) For each $f \in \mathrm{PF}$, there exist a polynomial q and a constant $c \in \mathbb{N}$ such that for all $x \in \{0,1\}^*$ and all nondecreasing $t : \mathbb{N} \to \mathbb{N}$,

$$|f(x)| \geq |x| \Rightarrow C^{t''}(f(x)) \leq C^t(x) + c,$$

where $t''(n) = ct'(n)\log(t'(n)) + c$ and $t'(n) = t(n) + q(n)$.

Our next lemma is a straightforward extension of Proposition 3.5 of [36].

Lemma 4.4. For each $f \in \mathrm{PF}$ there exist a polynomial q and a constant $c \in \mathbb{N}$ such that for all $A \subseteq \{0,1\}^*$, $x \in \{0,1\}^*$, and nondecreasing $t : \mathbb{N} \to \mathbb{N}$,

$$ic^{t''}(x : f^{-1}(A)) \leq ic^t(f(x) : A) + c,$$

where $t''(n) = ct'(n)\log(t'(n)) + c$ and $t'(n) = q(n) + t(q(n))$.

The following consequence of Lemma 4.4 is especially useful here.

Corollary 4.5. For each $f \in \mathrm{PF}$ there exist $\delta > 0$ and $c \in \mathbb{N}$ such that for all but finitely many $x \in \{0,1\}^*$, for all $A \subseteq \{0,1\}^*$,

$$ic^{2^{n^\delta}}(f(x) : A) \geq ic^{2^n}(x : f^{-1}(A)) - c.$$

Proof. Let $f \in \mathrm{PF}$ and choose q and c for f as in Lemma 4.4. Let $\delta = 1/(d+1)$, where d is the degree of the polynomial q, and let $t(n) = 2^{n^\delta}$. Define t' and t'' from t as in Lemma 4.4. Then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $t''(n) < 2^n$. It follows by Lemma 4.4 that for all $x \in \{0,1\}^*$ such that $|x| \geq n_0$, for all $A \subseteq \{0,1\}^*$,

$$\begin{aligned} ic^{2^{n^\delta}}(f(x) : A) &= ic^t(f(x) : A) \\ &\geq ic^{t''}(x : f^{-1}(A)) - c \\ &\geq ic^{2^n}(x : f^{-1}(A)) - c. \end{aligned}$$

□

Juedes and Lutz [20] introduced the following useful notation. The *nonreduced image* of a language $S \subseteq \{0,1\}^*$ under a function $f : \{0,1\}^* \to \{0,1\}^*$ is the language

$$f^{\geq}(S) = \{f(x) | x \in S \text{ and } |f(x)| \geq |x|\}.$$

Lemma 4.6.(Juedes and Lutz [20]) If $f \in$ PF is one-to-one a.e. and $S \subseteq \{0,1\}^*$ is cofinite, then $f^{\geq}(S)$ is dense.

We now prove that every weakly $\leq_m^{\mathrm{P}}$-hard language for exponential time has a dense set of very hard instances. Orponen, Ko, Schöning, and Watanabe [36] have shown that every $\leq_m^{\mathrm{P}}$-hard language for exponential time has a dense set of very hard instances, and Buhrman and Orponen [9] have proven a similar result with improved time bounds and density for languages that are $\leq_m^{\mathrm{P}}$-complete for exponential time. Theorem 4.7 below can be regarded as extending this phenomenon (with some modification in the precise bounds) to all *weakly* $\leq_m^{\mathrm{P}}$-hard languages for exponential time.

Juedes and Lutz [21] have proven that every weakly $\leq_m^{\mathrm{P}}$-hard language for E is weakly $\leq_m^{\mathrm{P}}$-hard for E_2, but that the converse fails, even for languages in E. We thus state our results in terms of weakly $\leq_m^{\mathrm{P}}$-hard languages for E_2, noting that they hold *a fortiori* for languages that are weakly $\leq_m^{\mathrm{P}}$-hard for E.

Theorem 4.7. If H is weakly $\leq_m^{\mathrm{P}}$-hard for E_2, then for every $\epsilon > 0$ there exists $\delta > 0$ such that the set

$$HI^{\epsilon,\delta}(H) = \{x | ic^{2^{n^\delta}}(x : H) > (1-\epsilon)C^{2^{4n}}(x)\}$$

is dense.

Proof. Let H be weakly $\leq_m^{\mathrm{P}}$-hard for E_2, and let $\epsilon > 0$. Let $X = X(1, \frac{\epsilon}{2})$ be defined as in Theorem 4.1, and let

$$Y = \{A |\ A \text{ is incompressible by } \leq_m^{\mathrm{P}}\text{-reductions}\}.$$

By Theorem 4.1 we have $\mu_{\mathrm{p}}(X) = 1$, and Juedes and Lutz [20] proved that $\mu_{\mathrm{p}}(Y) = 1$, so we have $\mu_{\mathrm{p}}(X \cap Y) = 1$. It follows that $\mu_{\mathrm{p}_2}(X \cap Y) = 1$, whence $\mu(X \cap Y | \mathrm{E}_2) = 1$. Since H is

weakly $\leq_m^{\mathrm{P}}$-hard for E_2, we have $\mu(\mathrm{P}_m(H)|\mathrm{E}_2) \neq 0$, so it follows that there exists

$$A \in X \cap Y \cap \mathrm{P}_m(H) \cap \mathrm{E}_2.$$

Since $A \in \mathrm{P}_m(H)$, there exists $f \in \mathrm{PF}$ such that $A = f^{-1}(H)$. By Corollary 4.5, there exist $\delta > 0$ and $c_1 \in \mathbb{N}$ such that the set

$$S_1 = \{x |\ ic^{2^{n^\delta}}(f(x) : H) \geq ic^{2^n}(x : A) - c_1\}$$

is cofinite. It suffices to show that the set $HI^{\epsilon,\delta}(H)$ is dense.

Since $A \in X$, the set

$$S_2 = \{x |\ ic^{2^n}(x : A) > (1 - \frac{\epsilon}{2})C^{2^{3n}}(x)\}$$

is cofinite. By Lemma 4.3, there exist a polynomial q and a constant $c_2 \in \mathbb{N}$ such that for all $x \in \{0,1\}^*$,

$$|f(x)| \geq |x| \Rightarrow C^{2^{3n}}(x) \geq C^{t''}(f(x)) - c_2$$

where t'' is defined from q and $t(n) = 2^{3n}$ as in that lemma. Since $t''(n) = o(2^{4n})$, the set

$$S_3 = \{x \,\Big|\, 2^{4|f(x)|} \geq t''(|f(x)|)\}$$

is cofinite. Finally, since $\lim_{n\to\infty} C(s_n) = \infty$, the set

$$S_4 = \{x \,\Big|\, C^{2^{4n}}(f(x)) > 2\frac{(c_1 + c_2)}{\epsilon}\}$$

is cofinite. It follows that the set

$$S = S_1 \cap S_2 \cap S_3 \cap S_4$$

is cofinite. Since $A \in Y$, Lemma 4.6 tells us that the nonreduced image $f^{\geq}(S)$ is dense. We complete the proof by showing that $f^{\geq}(S) \subseteq HI^{\epsilon,\delta}(H)$.

Assume that $y \in f^{\geq}(S)$. Then there exists $x \in S$ such that $y = f(x)$ and $|f(x)| \geq |x|$. Since $x \in S_1 \cap S_2$, we have

$$ic^{2^{n^\delta}}(y : H) > (1 - \frac{\epsilon}{2})C^{2^{3n}}(x) - c_1.$$

Since $x \in S_3 \cap S_4$ and $|f(x)| \geq |x|$, it follows that

$$\begin{aligned} ic^{2^{n^\delta}}(y : H) &> (1 - \frac{\epsilon}{2})[C^{t''}(f(x)) - c_2] - c_1 \\ &\geq (1 - \frac{\epsilon}{2})C^{2^{4n}}(f(x)) - (c_1 + c_2) \\ &> (1 - \epsilon)C^{2^{4n}}(f(x)) \\ &= (1 - \epsilon)C^{2^{4n}}(y), \end{aligned}$$

whence $y \in HI^{\epsilon,\delta}(H)$. □

By Lemma 3.3, Theorem 4.7 implies (and by Theorem 3.11 is much stronger than) the following known result.

Corollary 4.8.(Juedes and Lutz [20]). If H is weakly $\leq_m^{\mathrm{P}}$-hard for E_2, then H has a dense exponential complexity core.

We know that, for most strings x, $C^t(x)$ and $C(x)$ are both very close to $|x|$, so the time bound on $C^t(x)$ is often of secondary significance. Thus for many purposes, the following simple consequence of Theorem 4.7 suffices.

Corollary 4.9. If H is weakly $\leq_m^{\mathrm{P}}$-hard for E or E_2, then for every $\epsilon > 0$ there exists $\delta > 0$ such that the set

$$HI_0^{\epsilon,\delta}(H) = \{x | ic^{2^{n^\delta}}(x : H) > (1 - \epsilon)C(x)\}$$

is dense.

We conclude this section with a discussion of the instance complexities of NP-complete problems. For simplicity of exposition we focus on SAT, but the entire discussion extends routinely to other NP-complete problems.

We start with three known facts. The first says that the hypothesis $\mathrm{P} \neq \mathrm{NP}$ implies a lower bound on the instance complexity of SAT.

Theorem 4.10.(Orponen, Ko, Schöning and Watanabe [36]) If $\mathrm{P} \neq \mathrm{NP}$, then for every polynomial t and constant $c \in \mathbb{N}$, the set

$$\{x | ic^t(x : SAT) > c \log |x|\}$$

is infinite.

Each of the next two facts derives a stronger conclusion than Theorem 4.10 from a stronger hypothesis.

Theorem 4.11.(Ko [22]) If nonuniformly secure one-way functions exist, then for every polynomial t and constant $c \in \mathbb{N}$, the set

$$\{x | ic^t(x : SAT) > c \log |x|\}$$

is nonsparse.

Theorem 4.12.(Orponen, Ko, Schöning and Watanabe [36]) If E $\neq$ NE, then for every polynomial t there exist a polynomial t' and a constant $c \in \mathbb{N}$ such that the set

$$\{x | ic^t(x : SAT) > C^{t'}(x) - c\}$$

is infinite.

The following theorem derives a strong lower bound on the instance complexity of SAT from the hypothesis that $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$. This hypothesis, which was proposed by Lutz, has been proven to have many reasonable consequences [28, 1, 27, 10]. The $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$ hypothesis implies E $\neq$ NE [29] and is equivalent to the assertion that NP does not have measure 0 in E_2 [2]. Its relationship to the hypothesis of Theorem 4.11 is an open question.

Theorem 4.13. If $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$, then for every $\epsilon > 0$ there exists $\delta > 0$ such that the set

$$HI_1^{\epsilon,\delta}(SAT) = \{x | ic^{2^{n^\delta}}(x : SAT) > (1 - \epsilon)C^{2^{4n}}(x)\},$$

is dense.

Proof. If $\mu_{\mathrm{p}}(\mathrm{NP}) \neq 0$, then SAT is weakly $\leq_m^{\mathrm{P}}$-complete for E_2, so this follows from Theorem 4.7. □

CHAPTER 5. EASY INSTANCES

In this brief chapter, we note that languages that are $\leq_m^{\mathrm{P}}$-hard for exponential time have instance complexities that are *unusually low* in the sense that they obey an upper bound that is violated by almost every language in exponential time. Our proof is based on the following known result.

Theorem 5.1.(Juedes and Lutz [20]) For every $\leq_m^{\mathrm{P}}$-hard language H for E, there exist $B, D \in \mathrm{DTIME}(2^{4n})$ such that D is dense and $B = H \cap D$.

The following theorem gives an upper bound on the instance complexities of hard problems for exponential time. It says that every such problem has a *dense set of* (relatively) *easy instances.*

Theorem 5.2. For every $\leq_m^{\mathrm{P}}$-hard language H for E there is a constant $c \in \mathbb{N}$ such that the set

$$EI_c(H) = \{x | ic^{2^{6n}}(x : H) \leq c\}$$

is dense.

Proof. Let H be $\leq_m^{\mathrm{P}}$-hard for E. Choose B and D for H as in Theorem 5.1. It is routine to design an interpreter M with the following properties.

1. For all π, $x \in \{0,1\}^*$,

$$M(\pi, x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \in D - B \\ \perp & \text{if } x \notin D. \end{cases}$$

2. For all $x \in D$, $time_M(\lambda, x) \leq 2^{5|x|}$.

(Note that $M(\pi, x)$ does not depend on π.)

Then the program λ is consistent with H relative to M, so for all $x \in D$ we have $ic_M^{2^{5n}}(x : H) = 0$. By the optimality of U, it follows that there is a constant $c \in \mathbb{N}$ such that for all $x \in D$, $ic^{t'}(x : H) \leq c$, where $t'(n) = ct(n)\log(t(n)) + c$ and $t(n) = 2^{5n}$. We thus have $D \cap S \subseteq EI_c(H)$, where

$$S = \{x | 2^{6|x|} \geq t'(|x|)\}.$$

Since D is dense and S is cofinite, it follows that $EI_c(H)$ is dense. □

By Theorem 4.1, almost every language in exponential time violates the upper bound given by Theorem 5.2. Thus these two results together imply the known fact [20] that the set of $\leq_m^{\mathrm{P}}$-hard languages for exponential time has p-measure 0. It should also be noted that Ambos-Spies, Terwijn and Zheng [2] have shown that almost every language in E is weakly $\leq_m^{\mathrm{P}}$-hard for E. It follows by Theorem 4.1 that almost every language in E is weakly $\leq_m^{\mathrm{P}}$-hard for E and violates the instance complexity upper bound given by Theorem 5.2. Thus Theorem 5.2 cannot be extended to the weakly $\leq_m^{\mathrm{P}}$-hard problems for E.

CHAPTER 6. CONCLUSION

In this thesis we have investigated the instance complexities of problems that are hard or weakly hard for E under polynomial time, many-one reductions. In addition, we have clarified the relationships among three notions of intractability namely DTIME(t)-complexity cores, incompressibility by $\leq_m^{\mathrm{DTIME}(t)}$ reductions and essentially maximal t-time-bounded instance complexity. The main results in this thesis are:

1. Almost every instance of *almost every* problem in E has essentially maximal instance complexity. This extends a theorem in [36], which established the *existence* of such a problem in E.

2. *Every weakly hard* problem for E has a dense set of instances with essentially maximal instance complexity. This extends a theorem in [36], which showed that every *hard* problem for E has a dense set of such maximally hard instances.

3. Every hard problem for E has a dense set of *unusually easy* instances. It follows that the set of $\leq_m^{\mathrm{P}}$-hard languages for E form a very small (p-measure 0) set, which was first shown in [20].

We hope that the results here will provide a basis for further research along these lines.

BIBLIOGRAPHY

[1] K. Ambos-Spies and E. Mayordomo. Resource-bounded measure and randomness. In A. Sorbi, editor, *Complexity, Logic and Recursion Theory*, Lecture Notes in Pure and Applied Mathematics, pages 1–47. Marcel Dekker, New York, N.Y., 1997.

[2] K. Ambos-Spies, S. A. Terwijn, and X. Zheng. Resource bounded randomness and weakly complete problems. *Theoretical Computer Science*, 172:195–207, 1997.

[3] J. L. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity II.* Springer-Verlag, Berlin, 1990.

[4] J. L. Balcázar and U. Schöning. Bi-immune sets for complexity classes. *Mathematical Systems Theory*, 18:1–10, 1985.

[5] L. Berman and J. Hartmanis. On isomorphism and density of NP and other complete sets. *SIAM Journal on Computing*, 6:305–322, 1977.

[6] R. Book and D.-Z. Du. The existence and density of generalized complexity cores. *Journal of the ACM*, 34:718–730, 1987.

[7] R. Book, D.-Z Du, and D. Russo. On polynomial and generalized complexity cores. In *Proceedings of the Third Structure in Complexity Theory Conference*, pages 236–250, 1988.

[8] H. Buhrman and E. Mayordomo. An excursion to the Kolmogorov random strings. *Journal of Computer and System Sciences*, 54:393–399, 1997.

[9] H. Buhrman and P. Orponen. Random strings make hard instances. *Journal of Computer and System Sciences*, 53:261–266, 1996.

[10] H. Buhrman and L. Torenvliet. Complete sets and structure in subrecursive classes. In *Proceedings of Logic Colloquium '96*, pages 45–78. Springer-Verlag, 1998.

[11] G. J. Chaitin. On the length of programs for computing finite binary sequences. *Journal of the Association for Computing Machinery*, 13:547–569, 1966.

[12] G. J. Chaitin. On the length of programs for computing finite binary sequences: statistical considerations. *Journal of the ACM*, 16:145–159, 1969.

[13] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley & Sons, Inc., New York, N.Y., 1991.

[14] D.-Z. Du. *Generalized complexity cores and levelability of intractable sets.* PhD thesis, University of California, Santa Barbara, 1985.

[15] D.-Z. Du and R. Book. On inefficient special cases of NP-complete problems. *Theoretical Computer Science*, 63:239–252, 1989.

[16] S. Even, A. Selman, and Y. Yacobi. Hard core theorems for complexity classes. *Journal of the ACM*, 35:205–217, 1985.

[17] L. Fortnow and M. Kummer. On resource-bounded instance complexity. *Theoretical Computer Science*, 161:123–140, 1996.

[18] D. T. Huynh. On solving hard problems by polynomial-size circuits. *Information Processing Letters*, 24:171–176, 1987.

[19] D. W. Juedes and J. H. Lutz. Kolmogorov complexity, complexity cores, and the distribution of hardness. In O. Watanabe, editor, *Kolmogorov Complexity and Computational Complexity*, pages 43–65. Springer-Verlag, Berlin, 1992.

[20] D. W. Juedes and J. H. Lutz. The complexity and distribution of hard problems. *SIAM Journal on Computing*, 24(2):279–295, 1995.

[21] D. W. Juedes and J. H. Lutz. Weak completeness in E and E_2. *Theoretical Computer Science*, 143:149–158, 1995.

[22] K. Ko. A note on the instance complexity of pseudorandom sets. In *Proceedings of the Seventh Annual Structure in Complexity Theory Conference*, pages 327–337. IEEE Comput. Soc. Press, 1992.

[23] A. N. Kolmogorov. Three approaches to the quantitative definition of 'information'. *Problems of Information Transmission*, 1:1–7, 1965.

[24] Martin Kummer. On the complexity of random strings. In *13th Annual Symposium on Theoretical Aspects of Computer Science*, pages 25–36. Springer, 1996.

[25] M. Li and P. M. B. Vitányi. *An Introduction to Kolmogorov Complexity and its Applications.* Springer-Verlag, Berlin, 1997. Second Edition.

[26] J. H. Lutz. Almost everywhere high nonuniform complexity. *Journal of Computer and System Sciences*, 44:220–258, 1992.

[27] J. H. Lutz. The quantitative structure of exponential time. In L.A. Hemaspaandra and A.L. Selman, editors, *Complexity Theory Retrospective II*, pages 225–254. Springer-Verlag, 1997.

[28] J. H. Lutz. Resource-bounded measure. In *Proceedings of the 13th IEEE Conference on Computational Complexity*, pages 236–248, New York, 1998. IEEE.

[29] J. H. Lutz and E. Mayordomo. Cook versus Karp-Levin: Separating completeness notions if NP is not small. *Theoretical Computer Science*, 164:141–163, 1996.

[30] J. H. Lutz, V. Mhetre, and S. Srinivasan. Hard instances of hard problems. Submitted.

[31] N. Lynch. On reducibility to complex or sparse sets. *Journal of the ACM*, 22:341–345, 1975.

[32] E. Mayordomo. Almost every set in exponential time is P-bi-immune. *Theoretical Computer Science*, 136(2):487–506, 1994.

[33] A. R. Meyer, 1977. Reported in [5], pages 317–318.

[34] M. Mundhenk. NP-hard sets have many hard instances. In *Mathematical foundations of computer science 1997*, pages 428–437. Springer-Verlag, 1997.

[35] P. Orponen. A classification of complexity core lattices. *Theoretical Computer Science*, 70:121–130, 1986.

[36] P. Orponen, K. Ko, U. Schöning, and O. Watanabe. Instance complexity. *Journal of the Association of Computing Machinery*, 41:96–121, 1994.

[37] P. Orponen and U. Schöning. The density and complexity of polynomial cores for intractable sets. *Information and Control*, 70:54–68, 1986.

[38] D. A. Russo and P. Orponen. On P-subset structures. *Mathematical Systems Theory*, 20:129–136, 1987.

[39] C. P. Schnorr. Klassifikation der Zufallsgesetze nach Komplexität und Ordnung. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 16:1–21, 1970.

[40] C. P. Schnorr. A unified approach to the definition of random sequences. *Mathematical Systems Theory*, 5:246–258, 1971.

[41] C. P. Schnorr. Zufälligkeit und Wahrscheinlichkeit. *Lecture Notes in Mathematics*, 218, 1971.

[42] C. P. Schnorr. Process complexity and effective random tests. *Journal of Computer and System Sciences*, 7:376–388, 1973.

[43] R. J. Solomonoff. A formal theory of inductive inference. *Information and Control*, 7:1–22, 224–254, 1964.

[44] J. Ville. *Étude Critique de la Notion de Collectif.* Gauthier–Villars, Paris, 1939.