JOURNAL OF COMPUTER AND SYSTEM SCIENCES 53, 261–266 (1996)
ARTICLE NO. 0067

Random Strings Make Hard Instances*

Harry Buhrman[†]

Centrum voor Wiskunde en Informatica, P.O. Box 94070, 1090 GB Amsterdam, The Netherlands

and

Pekka Orponen[§]

Department of Computer Science, University of Helsinki, P.O. Box 26, FIN-00014 Helsinki, Finland

Received September 5, 1994

We establish the truth of the "instance complexity conjecture" in the case of DEXT-complete sets over polynomial time computations, and r.e. complete sets over recursive computations. Specifically, we obtain for every DEXT-complete set A an exponentially dense subset C and a constant k such that for every nondecreasing polynomial $t(n) = \Omega(n^k)$, $\mathrm{ic}^t(x: A) \geqslant K^t(x) - c$ holds for some constant c and all $x \in C$, where ic^t and K^t are the t-bounded instance complexity and Kolmogorov complexity measures, respectively. For any r.e. complete set A we obtain an infinite set $C \subseteq \bar{A}$ such that $\mathrm{ic}(x: A) \geqslant K(x) - c$ holds for some constant c and all $x \in C$, where ic and K denote the time-unbounded versions of instance and Kolmogorov complexities, respectively. The proofs are based on the observation that *Kolmogorov random strings* are individually hard to recognize by bounded computations.

1. INTRODUCTION

The notion of "instance complexity" was introduced in [6] to quantify the complexity of solving individual instances of decision problems. The basic idea here is to measure the complexity of each individual problem instance by the size of the simplest "special case" algorithm applicable to it. An instance is then "inherently hard" if even the simplest applicable algorithm essentially requires table look-up on that instance.

In [6, 11] it was conjectured that any problem not decidable in a given time bound will have infinitely many such inherently hard instances with respect to that time bound. In the present paper, we establish this result for DEXT-complete problems over polynomial time computations, and for r.e. complete problems over recursive computations. The basic observation underlying the proofs is that *random* strings are guaranteed to have no distinguishable features, and hence to be inherently hard to recognize.

To make these ideas more precise, let A be a set of binary strings to be recognized, and let t be some time bound function. Consider Turing machines that run in time t, and on each input x output either 1 ("yes"), 0 ("no") or $\perp$ ("don't know"). Say that a machine M *decides* string x if $M(x) \neq \perp$. Machine M is *consistent* with a set A if on each input x that M decides, $M(x) = 1$ if and only if $x \in A$. The *t-bounded instance complexity* of a string x with respect to A is then defined as

$$\mathrm{ic}^t(x: A) = \min\{|M|: M \text{ is a } t\text{-time bounded machine consistent with } A \text{ and deciding } x\}.$$

(We are being somewhat sloppy here, as the notion of "the size of Turning machine M" is encoding-dependent. The proper definition, in terms of programs to a fixed universal machine, is given below in Section 2.)

A table look-up argument shows that the t-bounded instance complexity of any string x is upper bounded (roughly) by its *t-bounded Kolmogorov complexity*,

$$K^t(x) = \min\{|M|: M(\lambda) = x \text{ in time } t(|x|)\}.$$

The "instance complexity conjecture" proposed in [6, 11] states that for any set $A \notin \mathrm{DTIME}(t)$ this upper bound is reached infinitely often, i.e., for some constant c there are infinitely many strings x such that $\mathrm{ic}^t(x: A) \geqslant K^t(x) - c$. (For $A \in \mathrm{DTIME}(t)$, the instance complexity is constant-bounded.) Some partial results supporting the conjecture were obtained in [6, 11, 12] for NP- and DEXT-hard sets.

* A preliminary version of this paper was presented at the Ninth Annual Conference on Structure in Complexity Theory (Amsterdam, 28.6–1.7.1994).

† E-mail address: buhrman@cwi.nl. Partially supported by the Dutch Foundation for Scientific Research (NWO) through NFI project ALADDIN under contract NF 62-376 and a TALENT stipend.

§ E-mail address: orponen@cs.helsinki.fi. Part of this work was done while this author was visiting the Centrum voor Wiskunde en Informatica, Amsterdam.

0022-0000/96 $18.00

571/53/2-8

Specifically, it was shown that for every set A that is DEXT-hard with respect to honest $\leqslant^{p}_{1-tt}$-reductions it is the case that (i) for any polynomial t there exist another polynomial t' and a constant c such that $\mathrm{ic}^{t'}(x: A) \geqslant K^{t'}(x) - c$ holds for infinitely many x; and (ii) there is an exponentially dense set C such that for every polynomial t and some constant c, $\mathrm{ic}^{t'}(x: A) \geqslant K^{\exp'}(x) - 2\log K^{\exp'}(x) - c$ holds for all $x \in C$, with $\exp'(n) = cn\,2^{2n} + c$. (A simpler proof for result (i) was recently given by Fortnow and Kummer in a paper [4] which includes also many other interesting results on resource-bounded instance complexity.)

In this paper, we prove a strong version of the conjecture in the case of many-one complete sets for DEXT over polynomial time computations: we show that every DEXT-complete set A has an exponentially dense subset C such that for some constant k and every nondecreasing polynomial $t(n) = \Omega(n^k)$, the lower bound $\mathrm{ic}^{t'}(x: A) \geqslant K^{t'}(x) - c$ holds for some constant c and all $x \in C$. Besides being a considerable improvement to both of the results (i) and (ii) above, the proof of this theorem is astonishingly simple, as compared to the complicated diagonalizations required earlier. The fundamental observation underlying the proof is that all the 2^{2n}-*bounded Kolmogorov random strings* are hard to recognize as such in polynomial time, i.e., given any polynomial t there is a constant c such that the inequality $\mathrm{ic}^{t'}(x: R^{\exp}) \geqslant K^{t'}(x) - c$ holds for all x in $R^{\exp} = \{x: K^{2^{2n}}(x) \geqslant |x|\}$. The main theorem follows from this by a simple reducibility argument.

In [11, 12] it was also conjectured that for any r.e., nonrecursive set A there is a constant c such that $\mathrm{ic}(x: A) \geqslant K(x) - c$ holds for infinitely many x, where ic and K (or ic^{∞} and K^{∞}) are the time-unbounded versions of instance and Kolmogorov complexity, respectively. An analogous argument as in the time-bounded case, this time based on considering the set of *recursive random strings* $R = \{x: K(x) \geqslant |x|\}$, proves this conjecture in the case of the r.e. complete sets.

The connection between instance complexity and (pseudo-)randomness was studied earlier by Ko in [5], but in that case in the context of pseudorandom *sequences*, not individual Kolmogorov random strings.

Very recently, Kummer.[8] has shown that the instance complexity conjecture is *not* valid for the r.e. incomplete sets, by constructing an r.e., nonrecursive set A for which $\mathrm{ic}(x: A) \leqslant \log K(x) + c$ holds for some constant c and all x. On the other hand, Fortnow and Kummer [4] have proved that the time-bounded version of the conjecture holds, with small "slack factors" in the time bounds, for all sets not in DEXT.

2. PRELIMINARIES

We shall consider decision problems for languages over the alphabet $\Sigma = \{0, 1\}$. The length of a string $x \in \Sigma^*$ is denoted $|x|$; λ denotes the empty string. Given strings x, y, we represent the pair $\langle x, y\rangle$ as the string $\bar{x}10y$, where $\bar{x}$ denotes x with each of its characters doubled. Note that for all x and y, $|\langle x, y\rangle| = 2|x| + |y| + 2$.

Complexity classes of languages are defined in the standard manner [1]; we shall study specifically the class $\mathrm{DEXT} = \bigcup_{c \geqslant 0} \mathrm{DTIME}(2^{cn})$.

The completeness notion we use is the one induced by many-to-one reductions.

An *interpreter* is a deterministic Turing machine M with two input tapes (a "program" tape and a "real input" tape) and an arbitrary number of work tapes, one of which is a designated output tape. The input and output tape alphabets of M are $\Sigma \cup \{\text{blank}\}$. M accepts its input if at the end of a computation, the output tape contains the string "1", rejects if the output tape contains a "0", and is *undecided* if the computation does not halt or if at its end the output tape contains something else—we denote both of these outcomes generically as "$\perp$". The partial mapping from $\Sigma^* \times \Sigma^*$ to Σ^* computed by M is denoted $M(p, x)$, and the time requirement of M on input (p, x) is denoted $\mathrm{time}_M(p, x)$. Given any function t on the natural numbers, an (M, t)-*program* is a string p such that $\mathrm{time}_M(p, x) \leqslant t(|x|)$ for all strings x.

For a set of strings A, $A(x)$ denotes the characteristic function of A, i.e., $A(x) = 1$ if $x \in A$ and $A(x) = 0$ if $x \notin A$. For $b \in \{0, 1\}$, we denote $M(p, x) \simeq b$ (read $M(p, x)$ is *consistent with* b) if $M(p, x) = b$ or $M(p, x) = \perp$. An (M, t)-program p *is consistent with* A if $M(p, x) \simeq A(x)$ for all x. Program p *decides* string x if $M(p, x) \neq \perp$.

DEFINITION 2.1. Let t be a function on the natural numbers. The *t-bounded instance complexity* of a string x with respect to set A using interpreter M is defined as

$$\mathrm{ic}^t_M(x: A) = \min\{|p|: p \text{ is an } (M, t)\text{-program that is consistent with } A \text{ and decides } x\}.$$

If no (M, t)-program consistent with A decides x, $\mathrm{ic}^t_M(x: A)$ is taken to be infinite.

DEFINITION 2.2. Let t be a function on the natural numbers. The *t-bounded Kolmogorov complexity* of string x using interpreter M is defined as

$$K^t_M(x) = \min\{|p|: M(p, \lambda) = x \text{ and } \mathrm{time}_M(p, \lambda) \leqslant t(|x|)\}.$$

If no M-program produces x in time $t(|x|)$, $K^t_M(x)$ is taken to be infinite.

As is well known [7, 12][9, p. 91], such notions can be defined robustly by means of a universal interpreter.

THEOREM 2.1 (Invariance). *There exists an interpreter U such that corresponding to any other interpreter M there*

is a constant c, such that for all sets A, time bounds t and strings x,

$$\mathrm{ic}^{t'}_{U}(x: A) \leqslant \mathrm{ic}^{t}_{M}(x: A) + c,$$
$$K^{t'}_{U}(x) \leqslant K^{t}_{M}(x) + c,$$

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

This invariance enables us to define the (absolute) t-bounded instance complexity of x with respect to A as $\mathrm{ic}^{t}(x: A) = \mathrm{ic}^{t}_{U}(x: A)$, and the (absolute) t-bounded Kolmogorov complexity of x as $K^{t}(x) = K^{t}_{U}(x)$. We also call a (U, t)-program p simply a t-program, and denote $\mathrm{time}_{p}(x) = \mathrm{time}_{U}(p, x)$.

We shall also refer to the time-unbounded versions of instance and Kolmogorov complexity. Let us say that a program p is *total* if $U(p, y)$ halts on all y, and define:

$$\mathrm{ic}(x: A) = \min\{|p|: p \text{ is a total program that is consistent with } A \text{ and decides } x\},$$
$$K(x) = \min\{|p|: U(p, \lambda) = x\}.$$

The Kolmogorov complexity of a string is easily seen to be an upper bound on its instance complexity with respect to any set [6, 12].

PROPOSITION 2.2. *For any time constructible function t, there exists a constant c such that for any set A and string x,*

$$\mathrm{ic}^{t'}(x: A) \leqslant K^{t}(x) + c,$$

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

Proof. Given a time constructible t, consider an interpreter M that works as follows: on input $(\langle b, p\rangle, y)$, where $b \in \Sigma$, $p \in \Sigma^*$, $y \in \Sigma^*$, simulates $U(p, \lambda)$ for $t(|y|)$ steps. If $U(p, \lambda)$ halts in this time with output y, M outputs b and halts, otherwise M halts with output $\perp$. Clearly there is a constant d such that for any b, p, and y, M halts in time bounded by $t(|y|) + d$. Let then A be any set, and x a string. Let $b = A(x)$, and let p be a minimal length t-program for producing x. Then $\langle b, p\rangle$ is an $(M, t + d)$-program for A deciding x, and so

$$\mathrm{ic}^{t+d}_{M}(x: A) \leqslant |\langle b, p\rangle| \leqslant |p| + 4 = K^{t}(x) + 4.$$

By invariance (Theorem 2.1), then, there is a constant c, independent of A and x, such that

$$\mathrm{ic}^{t'}(x: A) \leqslant K^{t}(x) + c,$$

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

The analogous result naturally holds for the time-unbounded versions of the measures.

3. RANDOM STRINGS ARE HARD TO RECOGNIZE

In this section we establish our main lemma showing that all exponential time Kolmogorov random strings are hard to recognize in polynomial time.

DEFINITION 3.1. Let T be a function on the natural numbers. The set of *T-bounded Kolmogorov random strings* is defined as

$$R^{T} = \{x \in \Sigma^*: K^{T}(x) \geqslant |x|\}.$$

In the following, we specifically consider the set $R^{\exp}$, where $\exp(n) = 2^{2^n}$.

By a simple counting argument [7][9, p. 96], it is easy to show that each of the sets R^{T} contains at least one string of each length. In fact, if one considers more generally the sets

$$R^{T}_{r} = \{x \in \Sigma^*: K^{T}(x) \geqslant |x| - r\},$$

then for each $r \geqslant 0$ the fraction of strings of each length *not* in R^{T}_{r} is smaller than 2^{-r}.

Another simple observation, to be used later, is that for every time constructible function T, the set R^{T} is in the class $\mathrm{DTIME}(2^{n}T(n))$. In particular, $R^{\exp} \in \mathrm{DEXT}$.

LEMMA 3.1. *Let t be a nondecreasing polynomial. Then there is a constant d such that for every $x \in R^{\exp}$,*

$$\mathrm{ic}^{t}(x: R^{\exp}) \geqslant |x| - d.$$

Proof. We prove the result by establishing the following strong "immunity" property of the set $R^{\exp}$ (cf. [10, p. 265] and Definition 3.2 below).

Claim. Let t be a nondecreasing polynomial. Then any t-program p consistent with $R^{\exp}$ accepts only strings x such that $|x| \leqslant |p| + d$, for some constant d independent of p.

Observe that the claim implies the statement of the lemma: let x be any string in $R^{\exp}$, and let p be a minimal size t-program consistent with $R^{\exp}$ and deciding x. Then

$$\mathrm{ic}^{t}(x: R^{\exp}) = |p| \geqslant |x| - d.$$

To prove the claim, consider an interpreter M that on input $(\langle d, p\rangle, \lambda)$ attempts to find and output the lexicographically first string x of length greater than $|p| + d$ that program p accepts (i.e., for which $U(p, x) = 1$). (If there are no strings matching the description, then M need not halt.) Clearly M can be implemented so that if p is some t-program and $M(\langle d, p\rangle, \lambda) = x$, then

$$K^{t'}_{M}(x) \leqslant |\langle d, p\rangle| = 2|d| + |p| + 2,$$

where $t'(n) = 2^n t(n)$. By invariance, there is a constant c' such that for any x for which $K^{t'}_M(x)$ is defined,

$$K^{t''}(x) \leqslant K^{t'}_M(x) + c',$$

where $t''(n) = c' t'(n) \log t'(n) + c' = O(n\, 2^n t(n))$.

Choose then a constant d so large that both $d - 2|d| - 2 \geqslant c'$ and $2^{2d} \geqslant t''(d)$. Assume, contrary to the claim, that some t-program p that is consistent with $R^{\exp}$ accepts a string x such that $|x| > |p| + d$. Then for the least such x,

$$\begin{aligned} K^{2^{2n}}(x) &\leqslant K^{t''}(x) \\ &\leqslant K^{t'}_M(x) + c' \\ &\leqslant 2|d| + |p| + 2 + c' \\ &< 2|d| + (|x| - d) + 2 + c' \\ &\leqslant |x|. \end{aligned}$$

But this contradicts the assumption that p is consistent with $R^{\exp}$ and hence accepts only 2^{2n}-bounded Kolmogorov random strings. ∎

THEOREM 3.2. *For every nondecreasing polynomial $t(n) = \Omega(n \log n)$ there is a constant c such that for every $x \in R^{\exp}$,*

$$\mathrm{ic}^t(x: R^{\exp}) \geqslant K^t(x) - c.$$

Proof. Let $t(n) = \Omega(n \log n)$ be some nondecreasing polynomial time bound. Let c' be a constant (cf. [9, p. 92]) such that for any x,

$$K^t(x) \leqslant |x| + c',$$

and set $c = c' + d$, where d is chosen as in Lemma 3.1. Then for any $x \in R^{\exp}$,

$$\begin{aligned} \mathrm{ic}^t(x: R^{\exp} &\geqslant |x| - d \\ &\geqslant K^T(x) - c' - d \\ &= K^t(x) - c. \end{aligned}$$ ∎

We point out two simple modifications of the proofs. First, the results can easily be extended to the set $R^{\exp}_r$, for any constant $r \geqslant 0$. by choosing the constant d in the proof of Lemma 3.1 so large that $d - 2|d| - 2 \geqslant c' + r$. Secondly, the analogous results hold also in the time-unbounded case. Define

$$R = \{x \in \Sigma^*: K(x) \geqslant |x|\}.$$

Concerning the recursion-theoretic complexity of the set R it is known that R is T-complete but not m-complete in the class of co-r.e. sets [10, pp. 264–265]. The same proof as in Lemma 3.1, but with the time bounds removed, establishes the following:

THEOREM 3.3. *There is a constant c such that for every $x \in R$.*

$$\mathrm{ic}(x: R) \geqslant K(x) - c.$$ ∎

A weaker, but more general version of this result can also be proved by the following recursion-theoretic argument.

DEFINITION 3.2. A set A is *strongly effectively immune* (cf. [10, p. 263]), if A is infinite and there exists a recursive function g such that for any total program p consistent with A, $g(p) \geqslant \max\{|x|: U(p, x) = 1\}$.

In particular, the proof of Lemma 3.1 establishes (when the time bounds are removed) that the set R is strongly effectively immune via the function $g(p) = |p| + d$.

THEOREM 3.4. *Let A be a strongly effectively immune set. Then there exist a constant c and infinitely many strings $x \in A$ such that:*

$$\mathrm{ic}(x: A) \geqslant K(x) - c.$$

Proof. Let x_0 be some string in A, and let p_{x_0} be a program that witnesses the instance complexity of x_0 (i.e., p_{x_0} is a total program consistent with A, $U(p_{x_0}, x_0) = 1$, and p_{x_0} is of minimal length).

Since A is strongly effectively immune, the program p_{x_0} accepts only finitely many strings. Let x_1 be the maximal string (in the lexicographic ordering) that p_{x_0} accepts. Next consider similarly the program p_{x_1} to find the maximal string x_2 it accepts, and so forth. Repeat this process until for some k, $|p_{x_k}| = |p_{x_{k+1}}|$. Note that $|p_{x_{i+1}}| < |p_{x_i}|$ for all $i < k$, so the process must terminate.

We claim now that the inequality

$$\mathrm{ic}(x_{k+1}: A) \geqslant K(x_{k+1}) - c$$

holds, for some constant c independent of x_{k+1}. By construction, x_{k+1} is the maximal string accepted by program p_{x_k}. Moreover, it can actually be computed from p_{x_k} using the recursive length bound $|x_{k+1}| \leqslant g(p_{x_k})$, and the recursive test "$U(p_{x_k}, \cdot) = 1$". Thus, for some constant c independent of x_{k+1},

$$\begin{aligned} K(x_{k+1}) &\leqslant |p_{x_k}| + c \\ &= |p_{x_{k+1}}| + c \\ &= \mathrm{ic}(x_{k+1}: A) + c, \end{aligned}$$

establishing the claim.

In order to locate a new string in A satisfying the condition of the theorem, we restart the above procedure from some string $x_0' \in A$ that follows x_{k+1} in the lexicographic ordering. Since A is infinite, the procedure can be repeated infinitely often. ∎

4. HARD INSTANCES FOR COMPLETE SETS

The following lemma, quoted from [6, 11, 12], establishes that instance complexity cannot decrease by more than a constant in $a \leqslant_m^p$-reduction, i.e., "hard" instances cannot be reduced to "easy" ones. This property enables us to translate the hardness results of Theorems 3.2 and 3.3 upwards in the reducibility ordering. (In [11, 12] the lemma was actually formulated in terms of $\leqslant_{1-tt}^p$-reductions; however, we do not need the stronger version here.)

LEMMA 4.1. *Let f be a $\leqslant_m^p$-reduction from a set A to a set B. Then for some constant c and any polynomial t there is a polynomial t' such that for all x,*

$$\mathrm{ic}^{t'}(x: A) \leqslant \mathrm{ic}^{t}(f(x): B) + c.$$

Proof. Let M be an interpreter that on input (q, x) first computes the value $f(x)$, and then simulates the computation of interpreter U on input $(q, f(x))$. Assume that the reduction f can be computed in time bounded by a nondecreasing polynomial r. Let t be any polynomial and x any string; we may clearly assume that t is nondecreasing. Now if q is any t-program that is consistent with B and decides $f(x)$, then, viewed through the interpreter M, q is also an (M, t'')-program consistent with A deciding x where $t''(n) = r(n) + t(r(n))$. Hence $\mathrm{ic}_M^{t''}(x: A) \leqslant \mathrm{ic}^t(f(x): B)$ for all x. But by invariance, there is a constant c, independent of t and t'', such that for all x, $\mathrm{ic}^{t'}(x: A) \leqslant \mathrm{ic}^{t''}(x: A) + c$, where $t'(n) = ct''(n) \log t''(n) + c$. ∎

The analogous result again holds in the recursion theoretic setting.

LEMMA 4.2. *Let f be a $\leqslant_m$-reduction from a set A to a set B. Then there is a constant c such that for all x,*

$$\mathrm{ic}(x: A) \leqslant \mathrm{ic}(f(x): B) + c. \quad ∎$$

Let $\Sigma^{(n)}$ denote the set of strings of length at most n. A set of strings C is *exponentially dense* if there is a constant $\varepsilon > 0$ such that $|C \cap \Sigma^{(n)}| \geqslant 2^{n^\varepsilon}$ holds for all but finitely many n.

THEOREM 4.3. *Every DEXT-complete set A contains an exponentially dense subset of strings $C \subseteq A$ such that for some constant k and every nondecreasing polynomial $t(n) = \Omega(n^k)$,*

$$\mathrm{ic}^t(x: A) \geqslant K^t(x) - c$$

holds for some constant c and all $x \in C$.

Proof. Let A be any DEXT-complete set, and let f be a $\leqslant_m^p$-reduction from the exponentially dense set $R_1^{\exp}$ to A. Because all DEXT-complete sets are related by one-to-one length-increasing reductions [3, 13], we may assume that also the reduction $f(z)$ is one-to-one and length-increasing. (If necessary, we may perform the reduction initially via some linearly paddable DEXT-complete set to ensure that these properties hold [2, p. 123].) By simple counting [2, p. 138], based on the properties of f and the fact that $|R_1^{\exp} \cap \Sigma^{(n)}| \geqslant 2^{\varepsilon n}$ for all n, we know that the set $C = f(R_1^{\exp}) \subseteq A$ is exponentially dense.

Let us then verify that the inequality of the theorem holds for all x of the form $x = f(z)$ for $z \in R_1^{\exp}$. Let $k = r + 1$ be a constant such that the reduction $f(z)$ can be computed by some interpreter M in time $n^r + r$, where $n = |f(z)|$. Then for any $z \in \Sigma^*$,

$$K_M^{n^r + r}(f(z)) \leqslant |z|,$$

and by invariance, there is for any polynomial $t(n) = \Omega(n^{r+1}) = \Omega(n^k)$ a constant c_1 such that for all $z \in \Sigma^*$,

$$K^t(f(z)) \leqslant |z| + c_1.$$

By Lemma 4.1, there exist a nondecreasing polynomial t'' and a constant c_2 such that for all strings $z \in \Sigma^*$,

$$\mathrm{ic}^{t''}(z: R_1^{\exp}) \leqslant \mathrm{ic}^t(f(z): A) + c_2$$

On the other hand, by Lemma 3.1, there is a constant d such that for all $z \in R_1^{\exp}$,

$$|z| \leqslant \mathrm{ic}^{t''}(z: R_1^{\exp}) + d.$$

Combining the inequalities and choosing $c = c_1 + c_2 + d$ shows that for all $z \in R_1^{\exp}$,

$$K^t(f(z)) \leqslant \mathrm{ic}^t(f(z): A) + c,$$

i.e. the desired result. ∎

We note that the density of the set C guarantees that, for some $\varepsilon > 0$, most of the strings $x \in C$ are of Kolmogorov complexity at least $K(x) \geqslant |x|^\varepsilon$. In summary, one could thus say that every DEXT-complete set contains a dense subset of hard instances whose absolute complexity is at least a polynomial fraction of the maximum possible.

Again, a result analogous to Theorem 4.3 holds for all r.e. complete sets A, although in this case we get no bound on the density of the set of hard instances. Also, as the class of r.e. sets is not closed under complement, the co-r.e. set R gets in this case translated into a set of hard instances in the *complement* of the complete set A.

THEOREM 4.4. *For every r.e. complete set A there exists a constant c such that for infinitely many $x \in \bar{A}$:*

$$\mathrm{ic}(x: A) \geqslant K(x) - c.$$

Proof. Similar to the proof of Theorem 4.3, using the well-known fact (e.g. [10, p. 321]) that all r.e. complete sets are in fact complete with respect to one-to-one reductions. ∎

5. CONCLUSION AND OPEN PROBLEMS

We have proved strong versions of the "instance complexity conjecture" of [6, 11] in the case of DEXT-complete and r.e. complete sets. Specifically, in the former case we have shown that for every DEXT-complete set A, there exists a constant k and an exponentially dense subset C such that for every nondecreasing polynomial $t(n) = \Omega(n^k)$, $\mathrm{ic}'(x: A) \geqslant K'(x) - c$ holds for some constant c and all $x \in C$. For r.e. complete sets A we have proved the analogous result, but without the density and time bounds; also in this case the "hard instances" x are located in the complement of A. (However, Kummer [8] has subsequently shown that hard instances also exist in A itself.)

The proofs of these results use in a fundamental way the observation that random strings by definition have no distinguishing features, and hence are individually hard to recognize. It will be interesting to investigate whether some analogue of this idea can be extended to prove the instance complexity conjecture in this strong form also in the case of NP-complete sets, under the appropriate assumptions. (A slightly weaker version of the conjecture for NP-complete sets was recently settled by Fortnow and Kummer [4] using diagonalization.) Furthermore it will be interesting to extend the techniques to work for sets that are immune or bi-immune for DTIME(2^{2^n}) instead of effectively immune.

REFERENCES

1. J. L. Balcázar, J. Díaz, and J. Gabarró, "Structural Complexity," Vol. I, Springer-Verlag, Berlin, 1988.
2. J. L. Balcázar, J. Díaz, and J. Gabarró, "Structural Complexity," Vol. II, Springer-Verlag, Berlin, 1990.
3. L. Berman, "Polynomial Reducibilities and Complete Sets," Ph.D. thesis, Cornell University, 1977.
4. L. Fortnow, and M. Kummer, Resource-bounded instance complexity, *in* "Proc. of the 12th Symposium on Theoretical Aspects of Computer Science (Munich, March 1995)," Lecture Notes in Computer Science, Vol. 900, pp. 557–608, Springer-Verlag, Berlin, 1995.
5. K. Ko, A note on the instance complexity of pseudorandom sets, *in* "Proc. of the 7th Annual Conference on Structure in complexity Theory (Boston, MA, June 1992)," pp. 327–337, IEEE Press, New York, NY, 1992.
6. K. Ko, P. Orponen, U. Schöndig, and O. Watanabe, What is a hard instance of a computational problem?, *in* "Proc. of the Conference on Structure in Complexity Theory (Berkeley, CA, June 1986)," Lecture Notes in Computer Science, Vol. 223, pp. 197–217, Springer-Verlag, Berlin, 1986.
7. A. N. Kolmogorov, Three approaches to the quantitative definition of information, *Problems Info. Transmission* **1** (1965), 1–7.
8. M. Kummer, The instance complexity conjecture, *in* "Proc. of the 10th Annual Conference on Structure in Complexity Theory (Minneapolis, MN, June 1995)," pp. 111–124, IEEE Press, New York, NY, 1995.
9. M. Li, and P. M. B. Vitányi, "An Introduction to Kolmogorov Complexity and Its Applications," Springer-Verlag, New York, NY, 1993.
10. P. Odifreddi, "Classical Recursion Theory," Elsevier, Amsterdam, 1989.
11. P. Orponen, On the instance complexity of NP-hard problems, *in* "Proc of the 5th Annual Conference on Structure in Complexity Theory (Barcelona, June 1990)," pp. 20–27, IEEE Press, New York, NY, 1990.
12. P. Orponen, K. Ko, U. Schöning, and O. Watanabe, Instance complexity, *J. Assoc. Comput. Mach.* **41** (1994), 96–121.
13. O. Watanabe, On one-one polynomial time equivalence relations, *Theoret. Comput. Sci.* **38** (1985), 157–165.