Computability 1 (2012) 39–43
DOI 10.3233/COM-2012-006
IOS Press

Limitations of Efficient Reducibility to the Kolmogorov Random Strings

John M. Hitchcock*
Department of Computer Science, University of Wyoming
1000 E. University Ave. Laramie, WY 82071, USA

Abstract. We show the following results for polynomial-time reducibility to R_C, the set of Kolmogorov random strings.

1. If P $\neq$ NP, then SAT does not dtt-reduce to R_C.
2. If PH does not collapse, then SAT does not n^{α}--reduce to R_C for any $\alpha < 1$.
3. If PH does not collapse, then SAT does not n^{α}-T-reduce to R_C for any $\alpha < \frac{1}{2}$.
4. There is a problem in E that does not dtt-reduce to R_C.
5. There is a problem in E that does not n^{α}--reduce to R_C, for any $\alpha < 1$.
6. There is a problem in E that does not n^{α}-T-reduce to R_C, for any $\alpha < \frac{1}{2}$.

These results hold for both the plain and prefix-free variants of Kolmogorov complexity and are also independent of the choice of the universal machine.

Keywords: Kolmogorov random strings, polynominal-time reducibility, Turing reduction, universal machine

1. Introduction

Because the Kolmogorov complexity function $C(x)$ is noncomputable, the set

$$R_C = \{x \mid C(x) > |x|\}$$

of Kolmogorov random strings is undecidable. In fact, R_C has no infinite computably enumerable subset. From this and the fact that the complement $\overline{R_C}$ is computably enumerable, Arslanov's completeness criterion implies that R_C is hard for the c.e. sets under Turing reductions. Kummer [7] showed a stronger result: $\overline{H} \leq_{\text{dtt}} R_C$, where $\overline{H}$ is the complement of the halting problem and $\leq_{\text{dtt}}$ denotes a disjunctive truth-table reduction. Neither of these reductions from the halting problem to R_C is efficient. This raises the question [1]: what can be efficiently reduced to R_C?

Recall that the Kolmogorov complexity [9] of a binary string x is the length of a shortest program that prints x on a universal Turing machine U:

$$C_U(x) = \min\{|p| \mid U(p) \text{ prints } x\}.$$

For the most part, the theory of Kolmogorov complexity does not depend on the choice of the universal machine U: for any two universal machines U and V, C_U and C_V are within an additive constant of each other. As usual, we fix a universal machine U and omit it from the notation, writing $C(x)$ instead of $C_U(x)$. There are, however, situations when the choice of universal machine matters and then we will be explicit with the subscript. We use the notation $\mathrm{P}_\tau(A)$ to denote the class of problems that reduce to A by $\leq_\tau^{\mathrm{p}}$-reductions.

Kummer's result [7] implies there is a computable time bound $t(n)$ such that for every decidable A, $A \leq_{\text{dtt}}^{t(n)} R_C$. Kummer's proof is nonconstructive and does not yield any information about the function $t(n)$. In fact, Allender et al. [1] show that some uncertainty about the time bound $t(n)$ is inevitable. They show that the $t(n)$ in Kummer's theorem

*This research was supported in part by NSF grants 0652601 and 0917417. Part of this research was done while the author was on sabbatical at CWI, supported by an NWO visiting scholar grant. A preliminary version of this paper appeared in the proceedings of the 6th Conference on Computability in Europe (2010).

may be arbitrarily large, depending on the choice of the universal machine U. Formally, for every computable time bound $t(n)$, there exists a universal machine U and a decidable set A such that A does not $\leq_{\mathrm{dtt}}^{t(n)}$-reduce to R_{C_U}. On the other hand, independent of U, there exist decidable sets with arbitrarily high time complexity that reduce to R_{C_U} via a polynomial-time dtt-reduction: for every computable time bound $t(n)$ and every universal machine U, there is a set $A \in \mathrm{DEC} - \mathrm{DTIME}(t(n))$ such that $A \leq_{\mathrm{dtt}}^{\mathrm{p}} R_{C_U}$. While this result shows $\mathrm{P}_{\mathrm{dtt}}(R_C)$ contains sets of high time complexity, the set A in this theorem is constructed via padding, which makes A very sparse. Thus while A has high time complexity, A is very simple in other terms. We show that this simplicity is inherent: any such A is highly predictable in the sense of polynomial-time dimension. From this it follows that R_C is not hard for E under $\leq_{\mathrm{dtt}}^{\mathrm{p}}$-reductions. This holds for every universal machine, i.e. $\mathrm{E} \not\subseteq \mathrm{P}_{\mathrm{dtt}}(R_{C_U})$ for every U. We also show that R_C is not polynomial-time dtt-hard for NP unless P = NP. Both of these results follow from showing that if a decidable set $\leq_{\mathrm{dtt}}^{\mathrm{p}}$-reduces to R_C, then the set $\leq_{\mathrm{dtt}}^{\mathrm{p}}$-reduces to a tally set. These results complement the result of Allender et al. [1] that

$$\mathrm{P} = \mathrm{DEC} \cap \bigcap_{U} \mathrm{P}_{\mathrm{dtt}}(R_{C_U}),$$

where the intersection is over all universal machines. While the class $\mathrm{DEC} \cap \mathrm{P}_{\mathrm{dtt}}(R_{C_U})$ contains arbitrarily complex sets, it is intuitively “close” to P for every U, in that it has small dimension and cannot contain NP unless P = NP.

Allender et al. [2] showed that R_C is hard for PSPACE under polynomial-time Turing reductions: $\mathrm{PSPACE} \subseteq \mathrm{P}_{\mathrm{T}}(R_C)$. Buhrman et al. [3] showed that R_C is hard for BPP under polynomial-time truth-table reductions: $\mathrm{BPP} \subseteq \mathrm{P}_{\mathrm{tt}}(R_C)$. We consider bounded query Turing and truth-table reductions. Based on the Winnow algorithm [10] and polynomial-time dimension [6], we show that R_C is not $\leq_{n^\alpha\text{-tt}}^{\mathrm{p}}$-hard for E, for any $\alpha < 1$. This is an improvement of a result in [1] which obtained the same consequence for EE. Also, we use the techniques of [4, 5] to show that R_C is not $\leq_{n^\alpha\text{-tt}}^{\mathrm{p}}$-hard for NP unless $\mathrm{NP} \subseteq \mathrm{coNP/poly}$ and the polynomial-time hierarchy collapses by Yap's theorem [13]. Finally, we obtain the same consequences for $\leq_{n^\alpha\text{-T}}^{\mathrm{p}}$-reductions, for all $\alpha < \frac{1}{2}$.

2. Preliminaries

We use standard notions of polynomial-time reducibilities [8]. We also need the following two notions of reducibility.

Definition 2.1. Let $\mathcal{B} = (B_n \mid n \geq 0)$ be a family of subsets of Σ^*. We say that A NP-*reduces to* $\mathcal{B}$ if there is an NPMV function N such that for all n, for all $x \in \Sigma^n$, $x \in A$ iff at least one output of $N(x)$ is in B_n.

Definition 2.2. Let $\mathcal{B} = (B_n \mid n \geq 0)$ be a family of subsets of Σ^*. We say that A *disjunctively reduces to* $\mathcal{B}$ *in* $t(n)$ *time* if there is an algorithm M such that for all n, for all $x \in \Sigma^n$, $M(x)$ outputs a list of strings in $t(n)$ time and $x \in A$ iff at least one output of $M(x)$ is in B_n.

The following lemma is from [4], based on a technique of [5]. An AND-function (of order 1) for a set A is a polynomial-time computable function g such that for all strings $x_1, x_2, \ldots, x_n$, $|g(x_1, \ldots, x_n)| = O\big(\sum_{i=1}^{n} |x_i|\big)$ and $g(x_1, x_2, \ldots, x_n) \in A$ iff $x_i \in A$ for all i.

Lemma 2.3. *Let A have an AND-function and let $\alpha < 1$. Let $\mathcal{B} = (B_n \mid n \geq 0)$ be a family of sets with $|B_n| \leq 2^{n^\alpha}$ for sufficiently large n. If A NP-reduces to $\mathcal{B}$, then $A \in \mathrm{NP/poly}$.*

The p-dimension [11] of a complexity class is a real number in $[0, 1]$. The p-dimension of P is 0 and the p-dimension of E is 1. For this paper, we do not need the full details of p-dimension; all we require is the fact that a p-dimension 0 class cannot contain E and the following lemma. The proof of this lemma relies on the Winnow online learning algorithm [10] and is straightforward to prove using the approach of [6].

Lemma 2.4. *Let $\alpha < 1$ and let $c \geq 1$. Let X be the class of all A for which there exists a family $\mathcal{B} = (B_n \mid n \geq 0)$ with $|B_n| \leq 2^{n^\alpha}$ such that A disjunctively reduces to $\mathcal{B}$ in 2^{cn} time. Then X has* p*-dimension 0. In particular, X does not contain* E.

3. Disjunctive Reductions

Theorem 3.1. *If A is decidable and $A \leq^{\mathrm{p}}_{\mathrm{dtt}} R_C$, then $A \leq^{\mathrm{p}}_{\mathrm{dtt}} B$ for some $B \in$ TALLY.*

Proof. We use the proof technique from [1] that A is decidable and $A \leq^{\mathrm{p}}_{\mathrm{mtt}} R_C$ (monotone truth-table) implies $A \in \mathrm{P/poly}$, observing that we can encode in a tally set to obtain the stronger result.

Suppose A is decidable and $A \leq^{\mathrm{p}}_{\mathrm{dtt}} R_C$ via a reduction computable in time n^d. Let the queries on input x be denoted by $Q(x)$. For some constant c, we claim only the queries of length at most $l(n) = c \log n$ "matter."

For any x, we have $x \in A$ iff $Q(x) \cap R_C \neq \emptyset$. Define $Q'(x) = Q(x) \cap \Sigma^{\leq l(n)}$, where $n = |x|$. We claim that for each $x \in A$, there is some $q \in Q'(x)$ such that for all y with $|y| = |x|$, $q \in Q'(y)$ implies $y \in A$.

Suppose the claim is false. Then given n, we can find the first string x of length n such that $x \in A$ and each query $q \in Q'(x)$ belongs to $Q'(y)$ for some $y \notin A$. This implies that $Q'(x) \cap R_C = \emptyset$. Since $x \in A$, it follows that $Q(x) - Q'(x)$ contains a string $r \in R_C$. This string r has $C(r) > l(n)$ because $r \notin Q'(x)$. We can describe r by describing n and the index of r in $Q(x)$. Since $|Q(x)| \leq n^d$, this takes at most $(d+3) \log n$ bits, a contradiction if we choose $c = d + 4$.

Let $\{w_1, \ldots, w_N\}$ be an enumeration of $\Sigma^{\leq l(n)}$. Let I_n be the collection of all i where for all y of length n, $w_i \in Q(y)$ implies $y \in A$. Our desired tally set is $\{0^{\langle n,i \rangle} \mid n \geq 0 \text{ and } i \in I_n\}$, where $\langle \cdot, \cdot \rangle$ is a pairing function on the natural numbers. □

Corollary 3.2. *If* $\mathrm{P} \neq \mathrm{NP}$, *then* $\mathrm{NP} \not\subseteq \mathrm{P_{dtt}}(R_C)$.

Proof. Suppose that $\mathrm{NP} \subseteq \mathrm{P_{dtt}}(R_C)$. By Theorem 3.1, $\mathrm{SAT} \leq^{\mathrm{p}}_{\mathrm{dtt}} B$ for a tally set B. Then $\overline{\mathrm{SAT}} \leq^{\mathrm{p}}_{\mathrm{ctt}} \overline{B} \cap 0^*$. Ukkonen [12] showed that $\mathrm{P} = \mathrm{NP}$ if coNP has a sparse $\leq^{\mathrm{p}}_{\mathrm{ctt}}$-hard set. □

Corollary 3.3. *The class* $\mathrm{P_{dtt}}(R_C) \cap \mathrm{DEC}$ *has* p*-dimension 0.*

Proof. Theorem 3.1 implies $\mathrm{P_{dtt}}(R_C) \cap \mathrm{DEC} \subseteq \mathrm{P_{dtt}}(\mathrm{TALLY}) \subseteq \mathrm{P_{dtt}}(\mathrm{SPARSE})$. This last class was shown to have p-dimension 0 in [6]. □

Corollary 3.4. $\mathrm{E} \not\subseteq \mathrm{P_{dtt}}(R_C)$.

Proof. This follows from Corollary 3.3 because E has p-dimension 1. □

4. Truth-Table Reductions

Theorem 4.1. *Let $\alpha < 1$.*

1. *If A is decidable, A has an AND-function, and $A \leq^{\mathrm{p}}_{n^\alpha\text{-tt}} R_C$, then $A \in \mathrm{NP/poly}$.*
2. *The class* $\mathrm{P}_{n^\alpha-\mathrm{tt}}(R_C) \cap \mathrm{DEC}$ *has* p*-dimension 0.*

Proof. The main idea of the proof is from [1]. We expound the argument here and show how to apply Lemmas 2.3 and 2.4.

Let A be decidable such that $A \leq^{\mathrm{p}}_{n^\alpha\text{-tt}} R_C$. Write $Q(x)$ for the truth-table reduction's queries on input x and $Z_x \subseteq \Sigma^{n^\alpha}$ for the query answer sequences that cause the reduction to accept x. That is, if $Q(x) = \{q_1, \ldots, q_{n^\alpha}\}$ in lexicographic order, then $x \in A$ if and only if $R_C[q_1] \cdots R_C[q_{n^\alpha}] \in Z_x$.

Let $l(n) = n^\epsilon$, where $0 < \epsilon < 1 - \alpha$. We claim that the truth-table reduction is still correct if we only use the queries of length at most $l(n)$. Formally, let $Q'(x) = Q(x) \cap \Sigma^{\leq l(n)}$ and let Z'_x be the restriction of Z_x with bits corresponding to strings in $Q(x) - Q'(x)$ removed.

Call two strings x and y of the same length *equivalent* if $Q'(x) = Q'(y)$. We claim that for each $x \in A$, there is some $z_x \in Z'_x$ such that for all y equivalent to x, $z_x \in Z'_y$ iff $y \in A$.

Suppose the claim is false. We can find the least $x \in A$ such that for all $z \in Z'_x$, there is some y_z equivalent to x such that $z \in Z'_y$ iff $y_z \notin A$. Let v be the correct answer sequence for $Q'(x) \cap R_C$ and let r be the number of 1's in v;

that is, $r = |Q'(x) \cap R_C|$. Given x and r, we can enumerate $\overline{R_C}$ to compute $Q'(x) \cap R_C$ and obtain v. Then we can compute y_v such that query answers v are incorrect for y_v. This means that $Q(y_v) - Q'(y_v)$ must contain a string in R_C with length $> l(n)$. However, we can describe this string by describing n, r, and its index in $Q(y_v)$, which takes $O(\log n)$ bits, a contradiction.

We define a family of sets $\mathcal{B} = (B_n \mid n \geq 0)$ as follows. For each equivalence class $[x]$ with queries $Q'(x) = \{w_1, \ldots, w_{n^\alpha}\}$ and $z_x \in Z'_x$ the answer sequence that is correct for all strings in the equivalence class, we put the tuple $\langle w_1, \ldots, w_{n^\alpha}, z_x\rangle$ in B_n. Note that $|B_n| < 2^{n^\gamma}$ where $\alpha + \epsilon < \gamma < 1$. By the claim, A NP-reduces to $\mathcal{B}$. It follows from Lemma 2.3 that $A \in \mathrm{NP/poly}$ if A has an AND-function.

We also have that A is disjunctively reducible in 2^n time to $\mathcal{B}$. Therefore Lemma 2.4 applies to show $\mathrm{P}_{n^\alpha-\mathrm{tt}}(R_C) \cap$ DEC has p-dimension 0. □

Corollary 4.2. *If* $\mathrm{NP} \subseteq \mathrm{P}_{n^\alpha-\mathrm{tt}}(R_C)$ *for some* $\alpha < 1$, *then* $\mathrm{NP} \subseteq \mathrm{coNP/poly}$.

Proof. This follows from Theorem 4.1 because the hypothesis implies $\overline{\mathrm{SAT}} \leq^{\mathrm{p}}_{n^\alpha\text{-tt}} R_C$ and $\overline{\mathrm{SAT}}$ has an AND-function. □

Corollary 4.3. *If the polynomial-time hierarchy does not collapse, then* $\mathrm{NP} \not\subseteq \mathrm{P}_{n^\alpha-\mathrm{tt}}(R_C)$ *for any* $\alpha < 1$.

Proof. This is immediate from Corollary 4.2 and Yap's theorem [13] that $\mathrm{NP} \subseteq \mathrm{coNP/poly}$ implies the polynomial-time hierarchy collapses to its third level. □

Corollary 4.4. *For any* $\alpha < 1$, $\mathrm{E} \not\subseteq \mathrm{P}_{n^\alpha-\mathrm{tt}}(R_C)$.

Proof. This follows from Theorem 4.1 because E has p-dimension 1. □

5. Turing Reductions

Theorem 5.1. *Let* $\alpha < \frac{1}{2}$.

1. *If* A *is decidable,* A *has an AND-function, and* $A \leq^{\mathrm{p}}_{n^\alpha\text{-T}} R_C$, *then* $A \in \mathrm{NP/poly}$.
2. *The class* $\mathrm{P}_{n^\alpha-\mathrm{T}}(R_C) \cap \mathrm{DEC}$ *has* p*-dimension 0.*

Proof. Let $\alpha < \beta < \frac{1}{2}$. Suppose that $A \in \mathrm{DEC}$ and $A \leq^{\mathrm{p}}_{n^\alpha\text{-T}} R_C$ via M. Let M' be the Turing machine that simulates M and whenever M makes a query of length at least n^β, M' makes no query and proceeds as if the answer to the query were no. We use the following concepts:

- An *advice* is a tuple $(z, w_1, \ldots, w_{n^\alpha})$ such that $z \in \Sigma^{n^\alpha}$ and each $w_i \in \Sigma^{<n^\beta}$.
- A string y is *accepted with advice* $(z, w_1, \ldots, w_{n^\alpha})$ if $M'(y)$ queries $w_1, \ldots, w_{n^\alpha}$ and accepts y when M' is given $z[1], \ldots, z[n^\alpha]$ as the query answers.
- An advice $(z, w_1, \ldots, w_{n^\alpha})$ is *safe* if for all $y \in \Sigma^n$, y is accepted with advice $(z, w_1, \ldots, w_{n^\alpha})$ implies $y \in A$.

We claim that for all $x \in A_{=n}$, there is a safe advice $(z, \vec{w})$ such that x is accepted with advice $(z, \vec{w})$.

Suppose the claim is false. Then we can find the least $x \in A_{=n}$ that does not have a safe advice. We can specify the correct answer sequence $z \in \Sigma^{n^\alpha}$ for $M(x)$ when querying oracle R_C. With this correct answer sequence z, M must query some string in R_C that is not in $\Sigma^{<n^\beta}$. Therefore we can describe a string r with $C(r) \geq n^\beta$ by describing n, z, and the index of r in $M(x)$'s query set on query answer sequence z. Thus $C(r) \leq n^\alpha + O(\log n)$, which is a contradiction since $\alpha < \beta$.

We define a family of sets $\mathcal{B}$ by putting into B_n all advices $(z, w_1, \ldots, w_{n^\alpha})$ that are safe. Let $1 > \gamma > \alpha + \beta$. The total number of possible advices is at most $2^{n^\alpha} \cdot (2^{n^\beta})^{n^\alpha} < 2^{n^\gamma}$, so $|B_n| < 2^{n^\gamma}$. We have that A NP-reduces to $\mathcal{B}$ and A disjunctively reduces in 2^n time to $\mathcal{B}$, so the theorem follows from Lemmas 2.3 and 2.4. □

Corollary 5.2. *If* $\mathrm{NP} \subseteq \mathrm{P}_{n^\alpha-\mathrm{T}}(R_C)$ *for some* $\alpha < \frac{1}{2}$, *then* $\mathrm{NP} \subseteq \mathrm{coNP/poly}$.

Corollary 5.3. *If the polynomial-time hierarchy does not collapse, then* NP $\not\subseteq \mathrm{P}_{n^\alpha-\mathrm{T}}(R_C)$ *for any* $\alpha < \frac{1}{2}$.

Corollary 5.4. *For any* $\alpha < \frac{1}{2}$, E $\not\subseteq \mathrm{P}_{n^\alpha-\mathrm{T}}(R_C)$.

6. Open Problems

We believe the following open problems should be tractable but appear to require techniques beyond those used in this paper.

Problem 6.1. Show that E $\not\subseteq \mathrm{P}_{n^\alpha-\mathrm{T}}(R_C)$ for $\frac{1}{2} \leq \alpha < 1$.

Problem 6.2. Show that NP $\not\subseteq \mathrm{P}_{n^\alpha-\mathrm{T}}(R_C)$ for $\frac{1}{2} \leq \alpha < 1$ under a reasonable hypothesis.

It is unknown whether even every decidable problem is polynomial-time Turing reducible to R_C. We conjecture that ESPACE $\not\subseteq \mathrm{P}_\mathrm{T}(R_C)$ and that this can be proved using resource-bounded dimension or measure:

Problem 6.3. Show that $\mathrm{P}_\mathrm{T}(R_C) \cap$ DEC has pspace-dimension 0.

Lastly, we know SAT $\leq_{\mathrm{dtt}} R_C$ and that SAT $\leq^{\mathrm{p}}_{\mathrm{dtt}} R_C$ iff P = NP. What more can be said about the amount of time it takes to disjunctively reduce SAT to R_C?

References

[1] E. Allender, H. Buhrman, and M. Koucký. What can be efficiently reduced to the Kolmogorov-random strings? *Annals of Pure and Applied Logic*, 138:2–19, 2006.

[2] E. Allender, H. Buhrman, M. Koucký, D. van Melkebeek, and D. Ronneburger. Power from random strings. *SIAM Journal on Computing*, 35:1467–1493, 2006.

[3] H. Buhrman, L. Fortnow, M. Koucký, and B. Loff. Derandomizing from random strings. In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity*, pages 58–63. IEEE Computer Society, 2010.

[4] H. Buhrman and J. M. Hitchcock. NP-hard sets are exponentially dense unless NP $\subseteq$ coNP/poly. In *Proceedings of the 23rd Annual IEEE Conference on Computational Complexity*, pages 1–7. IEEE Computer Society, 2008.

[5] L. Fortnow and R. Santhanam. Infeasibility of instance compression and succinct PCPs for NP. *Journal of Computer and System Sciences*, 77(1):96–106, 2011.

[6] J. M. Hitchcock. Online learning and resource-bounded dimension: Winnow yields new lower bounds for hard sets. *SIAM Journal on Computing*, 36(6):1696–1708, 2007.

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

[8] R. E. Ladner, N. A. Lynch, and A. L. Selman. A comparison of polynomial-time reducibilities. *Theoretical Computer Science*, 1(2):103–123, 1975.

[9] M. Li and P. M. B. Vitányi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, 3rd edition, 2008.

[10] N. Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2(4):285–318, 1988.

[11] J. H. Lutz. Dimension in complexity classes. *SIAM Journal on Computing*, 32(5):1236–1259, 2003.

[12] E. Ukkonen. Two results on polynomial time truth-table reductions to sparse sets. *SIAM Journal on Computing*, 12(3), 1983.

[13] C. K. Yap. Some consequences of non-uniform conditions on uniform classes. *Theoretical Computer Science*, 26:287–300, 1983.