On reductions to sets that avoid EXPSPACE *

V. Arvind[†] J. Köbler[‡¶] M. Mundhenk[§¶]

Abstract

A set B is called *EXPSPACE-avoiding*, if every subset of B in EXPSPACE is sparse. Sparse sets and sets of high information density (called $\mathcal{HIGH}$ sets in [5]) are shown to be EXPSPACE-avoiding. Investigating the complexity of sets A in EXPSPACE that honestly reduce to EXPSPACE-avoiding sets, we show that if the reducibility used has a property, called *range-constructibility*, then A must also reduce to a sparse set under the same reducibility.

Keywords: Computational Complexity, Reducibilities, Sparse Sets.

1 Introduction

The study of reductions to low information content sets has received much attention in structural complexity theory research in recent years. There is a series of results showing that complexity classes containing intractable problems cannot be reduced to sets of low information content unless there is an unlikely collapse of complexity classes. The class of sparse sets [8, 10, 11] is an example of a well-studied class of low information content sets. A research trend is to identify different classes of low information content sets, and to study the consequences of the existence of hard sets of low information content for intractable complexity classes under different kinds of reducibilities.

Recently, Book and Lutz [5] introduced and studied sets whose characteristic sequences are of very high space-bounded Kolmogorov complexity (they call the class of such sets $\mathcal{HIGH}$). The existence of $\mathcal{HIGH}$ sets follows from the fact that $\mathcal{RAND}$, the class of algorithmically random languages, is of measure 1 and is a subclass of $\mathcal{HIGH}$ [5]. In [5] it is shown that every set in ESPACE that is (polynomial time) bounded truth-table reducible to a set in $\mathcal{HIGH}$ is actually bounded

*A preliminary version was presented at MFCS'93.

†Department of Computer Science, Institute of Mathematical Sciences, C.I.T. Campus, Madras 600113, India. Work done at Universität Ulm. Supported in part by an Alexander von Humboldt research fellowship.

‡Abt. Theoretische Informatik, Universität Ulm, D-89069 Ulm, Germany.

§FB IV - Informatik, Universität Trier, D-54286 Trier, Germany.

¶Work supported in part by the DAAD through Acciones Integradas 1992, 313-AI-e-es/zk

truth-table reducible to some sparse set (it is shown in [5] that this even holds for $O(\log n)$-truth-table reducibilities). The reason for considering the class ESPACE is that most intractable complexity classes of interest like NP, PSPACE etc. are contained in ESPACE, and $\mathcal{HIGH}$ itself is defined using exponential space-bounded Kolmogorov complexity. Consequently, if an NP-complete set, say SAT, bounded truth-table reduces to a set in $\mathcal{HIGH}$, then SAT bounded truth-table reduces to a sparse set, and by [11] it follows that P=NP. This and similar consequences for other complexity classes (e.g. PSPACE and PP) are derived in [5].

In this paper, we address the following question: are there further polynomial time reducibilities α such that every set in ESPACE that is α reducible to a set in $\mathcal{HIGH}$ is actually α reducible to some sparse set?

The answer we provide to the above question is based on the following observation which is easy to prove (see Theorem 3.2): for any $A \in \mathcal{HIGH}$ it holds that every subset of A in EXPSPACE is sparse. This property is captured by a class of sets (larger than $\mathcal{HIGH}$) which we call EXPSPACE-*avoiding*. A set A is EXPSPACE-avoiding (abbreviated as $\mathcal{EA}$) if every subset of A in EXPSPACE is sparse. Theorem 3.2 rephrased states that every set in $\mathcal{HIGH}$ is in $\mathcal{EA}$. Next, we identify a simple property for reducibilities called *range-constructibility*. Informally speaking, whenever A honestly α reduces to B via a range-constructible reducibility, then A α reduces to a subset C of B such that C can be constructed by an ESPACE machine that uses A as oracle. We show that several natural reducibilities are range-constructible, for example the many-one, conjunctive, and Hausdorff reducibilities.

In Theorem 4.2 we show that every set in EXPSPACE which *honestly* reduces to a set in $\mathcal{EA}$ via a range-constructible reducibility in fact reduces to a sparse set. From known collapse results for reductions to sparse sets under the considered range-constructible reducibilities, it follows that the existence of hard sets in $\mathcal{EA}$ under honest range-constructible reductions for complexity classes like UP, NP, PP, or PSPACE implies unlikely collapses of complexity classes. It turns out that these collapse consequences hold for range-constructible reductions to sets in $\mathcal{HIGH}$ without the honesty assumption.

2 Preliminaries

Let $\Sigma = \{0, 1\}$ be the standard alphabet, and let $A \subseteq \Sigma^*$ be a set. The length of a string x is denoted by $|x|$. $A^{=n}$ ($A^{\leq n}$) denotes the set of all strings in A of length n (up to length n, respectively). χ_A denotes the characteristic function of A. $\chi^A_{\leq n}$ denotes the characteristic sequence of A for all strings up to length n, i.e., $|\chi^A_{\leq n}| = 2^{n+1} - 1$, and the i-th bit of $\chi^A_{\leq n}$ equals $\chi_A(s_i)$ where s_i is the i-th string in Σ^* in lexicographic order. The cardinality of A is denoted by $|A|$. The census function of a set A is $census_A(1^n) = |A^{\leq n}|$. A set S is called sparse if its census function is bounded above by a polynomial. A set T is called a tally set if

$T \subseteq 0^*$. We use $\mathcal{TALLY}$ and $\mathcal{SPARSE}$ to denote the classes of tally and sparse sets, respectively. ESPACE denotes the complexity class DSPACE($2^{O(n)}$), whereas EXPSPACE = DSPACE($2^{n^{O(1)}}$).

The deterministic and non-deterministic reducibilities discussed in this paper are the standard polynomial-time reducibilities defined in [9] and the Hausdorff reducibility introduced in [12].

Definition 2.1 *1. [12] A is Hausdorff reducible to B (in symbols: $A \leq_{hd}^{p} B$), if there exists a polynomial-time computable function f mapping every string x to a sequence of queries, such that for all $x \in \Sigma^*$, if $f(x) = \langle y_1, \ldots, y_k \rangle$ then*

- *$y_{i+1} \in B$ implies $y_i \in B$ for all $i = 1, \ldots, k-1$, and*
- *$x \in A \Leftrightarrow \max(\{j \mid 1 \leq j \leq k : y_j \in B\} \cup \{0\})$ is odd.*

In this context, the i-th query y_i, $1 \leq i \leq k(x)$, computed by $f(x)$ is also denoted by $f(x,i)$. We call f a bounded Hausdorff reduction ($A \leq_{bhd}^{p} B$) if the number $k(x)$ of queries produced by f on x is bounded by a constant for all x.

2. [9] A set A is co-np many-one reducible to a set B (denoted $A \leq_{m}^{co\text{-}np} B$) if there exists a polynomial-time nondeterministic Turing transducer M such that for every $x \in \Sigma^*$, $x \in A$ if and only if all outputs of M on input x are members of B.

3. (cf. [1]) A set A is co-rp many-one reducible to a set B (denoted $A \leq_{m}^{co\text{-}rp} B$) if there exist a polynomial-time computable function f and a polynomial q such that for all $x \in \Sigma^*$,

$$x \in A \Rightarrow \mathrm{Prob}_{w \in \Sigma^{q(|x|)}}[f(x,w) \in B] = 1, \text{ and}$$
$$x \notin A \Rightarrow \mathrm{Prob}_{w \in \Sigma^{q(|x|)}}[f(x,w) \notin B] \geq 1/q(|x|).$$

Here, the string w is chosen uniformly at random from the set $\Sigma^{q(|x|)}$.

Note that for all sets A and B, $A \leq_{c}^{p} B$ implies $A \leq_{m}^{co\text{-}rp} B$, and $A \leq_{m}^{co\text{-}rp} B$ implies $A \leq_{m}^{co\text{-}np} B$. A reduction f (of some type α) from a set A to a set B is called *honest* if there is a polynomial p such that for every x, and for every query y generated by $f(x)$, it holds that $|x| \leq p(|y|)$.[1] For any reducibility type

[1]Note that for a reduction f of any reducibility type α discussed in this paper it is meaningful to talk of queries generated by f on input x.

3 EXPSPACE-avoiding sets

Book and Lutz [5] measured the information content of a set by the Kolmogorov complexity of its characteristic sequences. Let M be a Turing machine, z be a string and let d, s be natural numbers. We say that $z \in KS_M[d, s]$, if M on some input of length at most d outputs z using space at most s. In other words, $KS_M[d, s]$ is the set of strings whose *s-space-bounded Kolmogorov complexity* relative to M is bounded by d. Well known simulation-techniques (see [4]) show that there is a Universal Turing machine U such that for every machine M there is a constant c such that for all d, s: $KS_M[d, s] \subseteq KS_U[d + c, cs + c]$. Henceforth, we fix U as such a Universal Turing machine and omit the subscript. Note that there is a constant c such that for every set A and for all n, the characteristic sequence $\chi^A_{\leq n}$ of $A^{\leq n}$ is in $KS[2^{n+1} + c, 2^{cn}]$. A set A is said to have maximum information content [5], if for every c and almost every n

$$\chi^A_{\leq n} \notin KS[2^{n+1} - 2n, 2^{cn}] \quad .$$

As in [5], $\mathcal{HIGH}$ is used to denote the class of sets of maximum information content. Intuitively, there is a large gap between the information content of sets in EXPSPACE and sets in $\mathcal{HIGH}$. We capture this property in the following definition.

Definition 3.1 *A set B is* EXPSPACE-avoiding *(in short: B is in $\mathcal{EA}$) if every subset of B in* EXPSPACE *is sparse.*

This can be seen as a weaker form of EXPSPACE-immunity (cf. [6]) since only non-sparse subsets in EXPSPACE are forbidden in EXPSPACE-avoiding sets. We show next that only sparse subsets of a $\mathcal{HIGH}$ set can be in EXPSPACE and therefore $\mathcal{HIGH} \subseteq \mathcal{EA}$.[2] Since sparse sets are in $\mathcal{EA}$ and not in $\mathcal{HIGH}$, the next theorem implies that $\mathcal{HIGH}$ is a proper subset of $\mathcal{EA}$.

Theorem 3.2 *Every set A in $\mathcal{HIGH}$ is in $\mathcal{EA}$. Moreover, even the padded version $pad(A) = \{x10^n \mid x \in A, n \geq 0\}$ of A is in $\mathcal{EA}$.*

Proof. Assume that A is in $\mathcal{HIGH}$ but $pad(A)$ is not in $\mathcal{EA}$. Then there exists a nonsparse subset B of $pad(A)$ which is in EXPSPACE. Since $pad(A)$ is paddable, it follows that there exists a nonsparse subset C of $pad(A)$ which is in ESPACE. For every n, let A_n be the subset $\{x \mid \exists k \leq n - |x| - 1, x10^k \in C\}$ of $A^{\leq n}$, and let v_n be the length $2^{n+1} - 1 - |A_n|$ substring of $\chi^A_{\leq n}$ obtained by deleting all 1's of $\chi^A_{\leq n}$ which correspond to some $x \in A_n$. Since C is nonsparse, it follows that for all polynomials q there exist infinitely many n such that $|A_n| > q(n)$. Then it is clear that the following algorithm outputs $\chi^A_{\leq n}$ and can be implemented in space $2^{O(n)}$. (ε denotes the empty string.)

[2] This and subsequent results concerning EXPSPACE in the paper can be easily extended to EXPSPACE/poly.

input $v_n 10^n$
$i := 0$;
for $x := \varepsilon$ **to** 1^n (in lexicographic order) **do**
  if $x \in A_n$ **then output** 1
  else $i := i + 1$; **output** the i-th bit of v_n

Since $|v_n 10^n| = 2^{n+1} - |A_n| + n$, it follows that there is a constant c such that $\chi^A_{\leq n} \in KS[2^{n+1} - q(n), 2^{cn}]$ for every polynomial q and infinitely many n, contradicting the fact that $A \in \mathcal{HIGH}$.

It can be similarly shown that if $pad(A)$ is in $\mathcal{EA}$ then also A is in $\mathcal{EA}$. ∎

4 Range-constructible reducibilities and collapse implications

We first formally define the notion of range-constructibility.

Definition 4.1 *A reducibility $\leq_\alpha$ is* range-constructible *if for all sets A, B such that $A \leq_\alpha B$, there is an* EXPSPACE *oracle transducer which on input 0^n and with oracle $A^{\leq n}$ outputs a subset B_n of B to which A correctly reduces via the given reduction for all inputs of length up to n.*

Using this definition the proof of the following theorem is straightforward.

Theorem 4.2 *Every set A in* EXPSPACE *which honestly reduces to a set B in $\mathcal{EA}$ via a range-constructible reducibility also reduces via the same type of reducibility to a sparse set.*

Proof. Let p be the polynomial witnessing the honesty of the reduction from A to B, i.e. each query y on input x fulfills $p(|y|) \geq |x|$. According to the definition of range-constructibility, let B_n be the set produced by the EXPSPACE oracle transducer on input 0^n with oracle $A^{\leq n}$. At first we define (for every n) a subset of B_n to which A correctly reduces via the given reduction for all inputs of length exactly n. Since the reduction is honest, $B'_n = \{y \mid y \in B_n \wedge p(|y|) \geq n\}$ has this property. Thus it is straightforward to see that by replacing in the reduction from A to B on input x each query y by the query $y10^{|x|}$, we obtain a reduction from A to $\hat{B} = \{y10^n \mid y \in B'_n\}$. We will show now that $\hat{B}$ is sparse. Since A is in EXPSPACE, it follows that the set $B' = \bigcup_{n \geq 0} B'_n$ also is in EXPSPACE. Since $B' \subseteq B$ and $B \in \mathcal{EA}$, B' is sparse. $\hat{B}$ is a subset of $pad(B')$, thus it finally follows that $\hat{B}$ is sparse. ∎

Next we give various examples of reducibilities that are range-constructible and apply Theorem 4.2 to these reducibilities to derive strong collapse consequences for various complexity classes, under the assumption that they honestly reduce to sets in $\mathcal{EA}$. We start by showing the range-constructibility of the $\leq_m^{co\text{-}np}$ and $\leq_m^{co\text{-}rp}$ reducibilities.

Proposition 4.3 *The co-np many-one and the co-rp many-one reducibilities are range-constructible.*

Proof. Let $A \leq_m^{co\text{-}np} B$ via a polynomial time nondeterministic Turing transducer M, that is, for every $x \in \Sigma^*$, $x \in A$ if and only if all output strings produced by M on input x are in B. Let $B_n = \{y \mid \exists x \in A^{\leq n}, M(x) \text{ outputs } y\}$. It is clear that B_n can be computed in exponential space with oracle $A^{\leq n}$. Also, B_n is a subset of B to which A reduces via M for all instances up to length n. The proof for the *co-rp* many-one reducibility proceeds analogously. ∎

Theorem 4.4 *1. The composition of the Hausdorff reducibility and the co-np many-one reducibility is range-constructible.*

2. The composition of the Hausdorff and the co-rp many-one reducibilities is also range-constructible.

Proof. We prove the first part. The second part's proof is analogous. Let $A \leq_{hd}^{p} B$ via a polynomial time computable function f, and let B *co-np* many-one reduce to C via a polynomial time nondeterministic Turing transducer M. Let $k(x)$ be the number of queries in the list $f(x)$, and let $M(y)$ denote the set of all outputs produced by M on input y. Define C_n as the set computed by the following algorithm on input 0^n with oracle $A^{\leq n}$.

> **input** 0^n
> $C' := \emptyset$
> **repeat**
>   **for** all $x \in \Sigma^{\leq n}$ **do**
> $l := \max(\{0\} \cup \{j \mid 1 \leq j \leq k(x) \text{ and } M(f(x,j)) \subseteq C'\})$
> $C' := C' \cup \bigcup\{M(f(x,j)) \mid 1 \leq j \leq l\}$
>   **if** $x \in A \Leftrightarrow l$ is even **then** $C' := C' \cup M(f(x,l+1))$
> **until** C' remained unchanged during the last execution of the loop
> **output** C'

At first we show that C_n is a subset of C for all $n \geq 0$. We proceed by induction on the number of iterations of the repeat-loop. Assume that $C' \subseteq C$ after the $(m-1)$-st iteration, and that C' is extended by some set $M(f(x,j))$ during the m-th iteration. There are two cases. In the first case there exists an index $l > j$ such that $M(f(x,l)) \subseteq C'$, and since $C' \subseteq C$, this implies $M(f(x,j)) \subseteq C$ by the monotonicity of the Hausdorff reduction. In the second case it holds that $j = l+1$ where $l = \max\{j \mid 0 \leq j \leq k(x) \text{ and for all } i = 1, \ldots, j : M(f(x,i)) \subseteq C'\}$ is even if and only if x is in A. It is clear that $l < k(x)$ since $C' \subseteq C$. By way of a contradiction, assume $M(f(l+1,x)) \not\subseteq C$. Because $C' \subseteq C$, this contradicts the fact that A reduces to C via f and M since it would imply that $x \in A \Leftrightarrow \max\{j \mid 0 \leq j \leq k(|x|) \text{ and for all } i = 1, \ldots, j : M(f(x,i)) \subseteq C\}$ is even.

Let p be a polynomial bounding the length of the queries produced by the composition of f and M. Observe that $C_n \subseteq \Sigma^{\leq p(n)}$, implying that the algorithm terminates. Moreover, since $|C_n| = 2^{O(n)}$ the algorithm stops after $2^{O(n)}$ steps. Finally, it is clear from the definition of the algorithm that A reduces to C_n via f and M for all instances $x \in \Sigma^{\leq n}$. ■

It follows that the many-one ($\leq_m^p$), conjunctive ($\leq_c^p$), (bounded) Hausdorff, and the composition of the (bounded) Hausdorff and conjunctive reducibilities are also range-constructible, because they are special cases of the reducibilities considered in Theorem 4.4. As a consequence of Theorem 4.2 and the above range-constructibility results we obtain the following theorem.

Theorem 4.5 *Let A be in* EXPSPACE *and let α be one of the following reducibility types:*

> *conjunctive, co-rp many-one, co-np many one, or the composition of the (bounded) Hausdorff reducibility with one of these reducibilities.*

Then, the following conditions are equivalent:

i) A is honestly α reducible to some set in $\mathcal{EA}$

ii) $A \in R_\alpha(\mathcal{SPARSE})$,

iii) $A \in R_\alpha(\mathcal{HIGH})$.

Proof. The implication $i) \to ii)$ directly follows from Corollary 4.2 and from the range-constructibility of the considered reducibilities. To show $ii) \to iii)$, observe that $R_\alpha(\mathcal{SPARSE}) = R_\alpha(\mathcal{TALLY})$ follows from $\mathcal{SPARSE} \subseteq R_c^p(\mathcal{TALLY})$ [7] and the fact that $R_\alpha(R_c^p(\mathcal{C})) = R_\alpha(\mathcal{C})$ for all the reducibilities α considered here. Since for every tally set T there exists a set B in $\mathcal{HIGH}$ such that $T = 0^* \cap B$, it follows that $R_\alpha(\mathcal{SPARSE}) \subseteq R_\alpha(\mathcal{HIGH})$. Finally, consider $A \in$ EXPSPACE which α reduces to a set $B \in \mathcal{HIGH}$. Then A also honestly α reduces to the set $pad(B)$ which is in $\mathcal{EA}$ by Theorem 3.2. This shows $iii) \to i)$. ■

We now have the corollary of collapse consequences.

Corollary 4.6

1. *Let $\mathcal{C}$ be any complexity class from $\{\mathrm{UP}, \mathrm{NP}, \mathrm{C_=P}, \mathrm{PP}\}$. If $\mathcal{C} \subseteq R_{bhd}^p(R_c^p(\mathcal{HIGH}))$, then $\mathcal{C} = \mathrm{P}$.*

2. *Let $\mathcal{C}$ be any of the complexity classes from $\{\mathrm{UP}, \mathrm{NP}, \mathrm{C_=P}, \mathrm{PP}\}$. If $\mathcal{C} \subseteq R_d^p(\mathcal{HIGH})$, then $\mathcal{C} = \mathrm{P}$.*

3. *For $\mathcal{C} \in \{\mathrm{NP}, \mathrm{PSPACE}\}$, if $\mathcal{C} \subseteq R_{hd}^p(R_c^p(\mathcal{HIGH}))$, then $\mathcal{C}$ is low for Δ_2^p.*

4. *If NP is contained in $R_{bhd}^p(R_m^{co\text{-}rp}(\mathcal{HIGH}))$, then $\mathrm{NP} = \mathrm{RP}$.*

Proof. The first part is a direct consequence of Theorem 4.5 and the result that $\mathcal{C} \subseteq R^p_{bhd}(R^p_c(\mathcal{SPARSE}))$ implies $\mathcal{C} = \mathrm{P}$ for any of the above complexity classes [2]. The second part follows from the first since $R^p_d(\mathcal{HIGH}) \subseteq R^p_{bhd}(R^p_c(\mathcal{HIGH}))$. The third holds since the existence of a sparse hard set for NP (or PSPACE) with respect to the composed Hausdorff and conjunctive reducibility implies the collapse of the polynomial time hierarchy to Δ^p_2 (respectively, PSPACE $= \Delta^p_2$) [3]. The fourth part follows from the result that NP is not contained in $R^p_{bhd}(R^{co\text{-}rp}_m(\mathcal{SPARSE}))$ unless NP = RP (cf. [2]). ■

It is easy to see from Theorem 4.5, that the above corollary also holds for honest reductions to $\mathcal{EA}$ sets for the considered reducibilities.

Acknowledgments. We thank Montse Hermo and Elvira Mayordomo for helpful comments.

References

[1] L. Adleman and K. Manders. Reducibility, randomness, and intractability. *Proc. 9th ACM Symp. on Theory of Computing* (1977) 151-163.

[2] V. Arvind, J. Köbler, and M. Mundhenk. On bounded truth-table, conjunctive, and randomized reductions to sparse sets. In *Proc. 12th FST& TCS,* Lecture Notes in Computer Science 652 (1992) 140–151.

[3] V. Arvind, J. Köbler, and M. Mundhenk. Upper bounds on the complexity of sparse and tally descriptions. *Mathematical Systems Theory*, to appear.

[4] J.L. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity I/II.* EATCS Monographs on Theoretical Computer Science (Springer Verlag 1988/1990).

[5] R. Book and J. Lutz. On languages with very high space-bounded Kolmogorov complexity. *SIAM Journal on Computing* 22(2) (1993) 395-402.

[6] R. Book and U.Schöning. Immunity, relativizations, and nondeterminism. *SIAM Journal on Computing* 13 (1984) 329-337.

[7] H. Buhrman, L. Longpré, and E. Spaan. SPARSE reduces conjunctively to TALLY. In *Proc. 8th Structure in Complexity Theory Conference* (1993) 208-214.

[8] R. Karp and R. Lipton. Some connections between nonuniform and uniform complexity classes. *Proc. 12th ACM Symposium on Theory of Computing,* (1980) 302-309.

[9] R. Ladner, N. Lynch, and A. Selman. A comparison of polynomial time reducibilities. *Theoretical Computer Science*, 1(2) (1975) 103-124.

[10] S. Mahaney. Sparse complete sets for NP: solution of a conjecture of Berman and Hartmanis. *J. of Computer and System Sciences*, 25(2) (1982) 130-143.

[11] M. Ogiwara and O. Watanabe. On polynomial-time bounded truth-table reducibility of NP sets to sparse sets. *SIAM J. on Computing*, 20(3) (1991) 471-483.

[12] K.W. Wagner. More complicated questions about maxima and minima, and some closures of NP. *Theoretical Computer Science*, 51 (1987) 53-80.