Optimal Proof Systems and Sparse Sets*

Harry Buhrman**[1], Steve Fenner***[2], Lance Fortnow†[3], and Dieter van Melkebeek‡[4]

[1] CWI
[2] University of South Carolina
[3] University of Chicago
[4] University of Chicago and DIMACS

Abstract. We exhibit a relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ has no complete sets. This gives the first relativized world where no optimal proof systems exist.

We also examine under what reductions $\mathbf{NP} \cap \mathbf{SPARSE}$ can have complete sets. We show a close connection between these issues and reductions from sparse to tally sets. We also consider the question as to whether the $\mathbf{NP} \cap \mathbf{SPARSE}$ languages have a computable enumeration.

1 Introduction

Computer scientists study lower bounds in proof complexity with the ultimate hope of actual complexity class separation. Cook and Reckhow [CR79] formalize this approach. They create a general notion of a proof system and show that polynomial-size proof systems exist if and only if $\mathbf{NP} = \mathbf{coNP}$.

Cook and Reckhow also ask about the possibility of whether optimal proof systems exist. Informally an optimal proof system would have proofs which are no more than polynomially longer than any other proof system.

An optimal proof system would play a role similar to $\mathbf{NP}$-complete sets. There exists a polynomial-time algorithm for Satisfiability if and only if $\mathbf{P} = \mathbf{NP}$. Likewise, if we have an optimal proof system, then this system would have polynomial-size proofs if and only if $\mathbf{NP} = \mathbf{coNP}$.

---

* Several proofs have been omitted to conserve space. A full version can be found at http://www.neci.nj.nec.com/homepages/fortnow/papers.

** CWI, INS4, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands. Email: buhrman@cwi.nl.

*** Supported in part by NSF grants CCR-9501794 and CCR-9996310. Address: Department of Computer Science, The University of South Carolina, Columbia, SC 29208. Email: fenner@cs.sc.edu.

† Supported in part by NSF grant CCR-9732922. Current Address: NEC Research, 4 Independence Way, Princeton, NJ 08540. Email: fortnow@research.nj.nec.com.

‡ Supported in part by NSF grant CCR-9732922. Current Address: DIMACS Center, Rutgers University, 96 Frelinghuysen Road, Piscataway, NJ 08854. Email: dieter@dimacs.rutgers.edu.

The existence of optimal proof systems remained an interesting open question. No one could exhibit such a system except under various unrealistic assumptions [KP89, MT98]. Nor has anyone exhibited a relativized world where optimal proof systems do not exist.

We construct such a world by building the first oracle relative to which $\mathbf{NP} \cap \mathbf{SPARSE}$ does not have complete sets. Messner and Torán [MT98] give a relativizable proof that if an optimal proof system exists than $\mathbf{NP} \cap \mathbf{SPARSE}$ does have complete sets.

We also consider whether $\mathbf{NP} \cap \mathbf{SPARSE}$-complete sets exist under other more general reductions than the standard many-one reductions. We show several results such as:

- There exists a relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ has no disjunctive-truth-table complete sets.
- There exists a relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ has no complete sets under truth-table reductions using $o(n/\log n)$ queries.
- For any positive constant c, there exists an oracle relative to which the class $\mathbf{NP} \cap \mathbf{SPARSE}$ has no complete sets under truth-table reductions using $o(n/\log n)$ queries and $c \cdot \log n$ bits of advice.
- Under a reasonable assumption for all values of $k > 0$, $\mathbf{NP} \cap \mathbf{SPARSE}$ has a complete set under conjunctive truth-table reductions that ask $\frac{n}{k \log n}$ queries and use $O(\log n)$ bits of advice.

The techniques used for relativized results on $\mathbf{NP} \cap \mathbf{SPARSE}$-complete sets also apply to the question of reducing sparse sets to tally sets. We show several results along these lines as well.

- Every sparse set S is reducible to some tally set T under a 2-round truth-table reduction asking $O(n)$ queries.
- Let c be any positive constant. There exists a sparse set S that does not reduce to any tally set T under truth-table reductions using $o(n/\log n)$ queries even with $c \cdot \log n$ bits of advice.
- Under a reasonable assumption for every sparse set S and every positive constant k, there exists a tally set T and a ctt-reduction from S to T that asks $\frac{n}{k \log n}$ queries and $O(\log n)$ bits of advice. We can also have a 2-round truth-table reduction using $\frac{n}{k \log n}$ queries and no advice.

We use the "reasonable assumptions" to derandomize some of our constructions building on techniques of Klivans and Van Melkebeek [KvM99]. The assumption we need is that there exists a set in $\mathbf{DTIME}[2^{O(n)}]$ that requires circuits of size $2^{\Omega(n)}$ even when the circuits have access to an oracle for $\mathbf{SAT}$. Under this assumption we get tight bounds as described above.

We also examine how $\mathbf{NP} \cap \mathbf{SPARSE}$ compares with other promise classes such as $\mathbf{UP}$ and $\mathbf{BPP}$ in particular looking at whether $\mathbf{NP} \cap \mathbf{SPARSE}$ has a uniform enumeration.

The proofs in our paper heavily use techniques from Kolmogorov complexity. We recommend the book of Li and Vitányi [LV97] for an excellent treatment of this subject.

1.1 Reductions and Relativizations

We measure the relative power of sets using reductions. In this paper all reductions will be computed by polynomial-time machines.

We say a set A reduces to a set B if there exists a polynomial-time computable function f such that for all strings x, x is in A if and only if $f(x)$ is in B. We also call this an m-reduction, "m" for many-one.

For more general reductions we need to use oracle machines. The set A Turing-reduces to B if there is a polynomial-time oracle Turing machine M such that $M^B(x)$ accepts exactly when x is in A. A tt-reduction (truth-table) requires that all queries be made before any answers are received.

A 2-round tt-reduction allows a second set of queries to be made after the answers from the first set of queries is known. This can be generalized to k-round tt-reductions but we will not need $k > 2$ in this paper.

We can think of a (one-round) tt-reduction R as consisting of two polynomial-time computable functions: One that creates a list of queries to make and an evaluator that takes the input and the value of B on those queries and either accepts or rejects. We use the notation $Q_R(x)$ to denote the set of queries made by reduction R on input x. For a set of inputs X, we let $Q_R(X) = \cup_{x \in X} Q_R(x)$.

A dtt-reduction (disjunctive-truth-table) means that $M^B(x)$ accepts if any of the queries it makes are in B. A ctt-reduction (conjunctive-truth-table) means that $M^B(x)$ accepts if all of the queries it makes are in B. A $q(n)$-tt reduction is a tt-reduction that makes at most $q(n)$ queries. A btt-reduction (bounded-truth-table) is a k-tt reduction for some fixed k.

We say a language L is r-hard for a class $\mathcal{C}$ if every language in $\mathcal{C}$ r-reduces to L. If L also sits in $\mathcal{C}$ then we say L is r-complete for $\mathcal{C}$.

All the results mentioned and cited in this paper relativize, that is they hold if all machines involved can access the same oracle. If we show that a statement holds in a relativized world that means that proving the negation would require radically different techniques. Please see the survey by Fortnow [For94] for a further discussion on relativization.

1.2 Optimal Proof Systems

A proof system is simply a polynomial-time function whose range is the set of tautological formulae, i.e., formulae that remain true for all assignments. Cook and Reckhow [CR79] developed this concept to give a general proof system that generalizes proof systems such as resolution and Frege proofs. They also give an alternate characterization of the **NP** versus **coNP** question:

Theorem 1 (Cook-Reckhow). **NP** $=$ **coNP** *if and only if there exists a proof system f and a polynomial p such that for all tautologies ϕ, there is a y, $|y| \leq p(|\phi|)$ and $f(y) = \phi$.*

Cook and Reckhow [CR79] also defined optimal and p-optimal proof systems.

Definition 1. *A proof system g is* optimal *if for all proof systems f, there is a polynomial p such that for all x, there is a y such that $|y| \leq p(|x|)$ and $g(y) = f(x)$. A proof system g is* p-optimal *if y can be computed in polynomial time from x.*

Messner and Torán [MT98] building on work of Krajíček and Pudlák [KP89] show that if **NEE** = **coNEE** then optimal proof systems exist and if **NEE** = **EE** then p-optimal proof systems exist. Here **EE**, double exponential time, is equal to $\mathbf{DTIME}[2^{O(2^n)}]$. The class **NEE** is the nondeterministic version of **EE**.

Messner and Torán [MT98] show consequences of the existence of optimal proof systems.

Theorem 2 (Messner-Torán).

- *If p-optimal proof systems exist then* **UP** *has complete sets.*
- *If optimal proof systems exist then* $\mathbf{NP} \cap \mathbf{SPARSE}$ *has complete sets.*

Hartmanis and Hemachandra [HH84] give a relativized world where **UP** does not have complete sets. Since all of the results mentioned here relativize, Messner and Torán get the following corollary.

Corollary 1 (Messner-Torán). *There exists an oracle relative to which p-optimal proof systems do not exist.*

However Messner and Torán leave open the question as to whether a relativized world exists where there are no optimal proof systems. Combining our relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ has no complete sets with Theorem 2 answers this question in the positive.

1.3 Reducing SPARSE to TALLY

A tally set is any subset of 1^*. Given a set S, the census function $c_S(n)$ is the number of strings of length n in S. A set S is sparse if the census function is bounded by a polynomial.

In some sense both sparse sets and tally sets contain the same amount of information but in sparse sets the information may be harder to find. Determining for which kind of reductions **SPARSE** can reduce to **TALLY** is an exciting research area.

Book and Ko [BK88] show that every sparse set tt-reduces to some tally set but there is some sparse set that does not btt-reduce to any tally set.

Ko [Ko89] shows that there is a sparse set that does not dtt-reduce to any tally set. He left open the conjunctive case.

Buhrman, Hemaspaandra and Longpré [BHL95] give the surprising result that every sparse set ctt-reduces to some tally set. Later Saluja [Sal93] proves the same result using slightly different techniques.

Schöning [Sch93] uses these ideas to show that **SPARSE** many-one reduces to **TALLY** with randomized reductions. In particular he shows that for every sparse set S and polynomial p there is a tally set T and a probabilistic polynomial-time computable f such that

– If x is in S then $f(x)$ is always in T.
– If x is not in S then $\Pr[f(x) \in T] \leq 1/p(|x|)$.

We say that S *co-rp-reduces* to T. Schöning notes that his reduction only requires $O(\log n)$ random bits.

1.4 Complete sets for NP ∩ SPARSE

Hartmanis and Yesha [HY84] first considered the question as to whether the class $\mathbf{NP} \cap \mathbf{SPARSE}$ has complete sets. They show that there exists a tally set T that is Turing-complete for $\mathbf{NP} \cap \mathbf{SPARSE}$. They also give a relativized world where there is no tally set that is m-complete for $\mathbf{NP} \cap \mathbf{SPARSE}$.

We should note that $\mathbf{NP} \cap \mathbf{TALLY}$ has m-complete sets. Let M_i be an enumeration of polynomial-time nondeterministic machines and consider

$$\{1^{\langle i,n,k\rangle} \mid M_i(1^n) \text{ accepts in } k \text{ steps}\}. \tag{1}$$

Also there exists a set in $\mathbf{D}_p \cap \mathbf{SPARSE}$ that is m-hard for $\mathbf{NP} \cap \mathbf{SPARSE}$. The class $\mathbf{D}_p$ contains the sets that can be written as the difference of two $\mathbf{NP}$ sets. For the $\mathbf{NP} \cap \mathbf{SPARSE}$-hard language we need to consider the difference $A - B$ where:

$$A = \{\langle x, 1^i, 1^k\rangle \mid M_i(x) \text{ accepts in } k \text{ steps}\}$$
$$B = \{\langle x, 1^i, 1^k\rangle \mid M_i \text{ accepts more than } k \text{ strings of length } |x| \text{ in } k \text{ steps}\}$$

As a simple corollary we get that if $\mathbf{NP} = \mathbf{coNP}$ then $\mathbf{NP} \cap \mathbf{SPARSE}$ has complete sets. However the results mentioned in Section 1.2 imply that one only needs the assumption of $\mathbf{NEE} = \mathbf{coNEE}$.

Schöning [Sch93] notes that from his work mentioned in Section 1.3 if the sparse set S is in $\mathbf{NP}$ then the corresponding tally set T is also in $\mathbf{NP}$. Since $\mathbf{NP} \cap \mathbf{TALLY}$ has complete sets we get that $\mathbf{NP} \cap \mathbf{SPARSE}$ has a complete set under co-rp-reductions. The same argument applied to Buhrman-Hemaspaandra-Longpré shows that $\mathbf{NP} \cap \mathbf{SPARSE}$ has complete sets under ctt-reductions.

2 NP ∩ SPARSE-Complete Sets

In this section, we establish our main result.

Theorem 3. *There exists a relativized world where* $\mathbf{NP} \cap \mathbf{SPARSE}$ *has no complete sets under many-one reductions.*

Proof. Let M_i be a standard enumeration of nondeterministic polynomial-time Turing machines and f_i be an enumeration of polynomial-time reductions where M_i and f_i use at most time n^i.

Let $t(m)$ be the tower function, i.e., $t(0) = 1$ and $t(m+1) = 2^{t(m)}$.

We will build an oracle A. For each i we will let

$$L_i(A) = \{x \mid \text{There is some } y,\ |y| = 2|x| \text{ and } \langle i, x, y\rangle \in A\}. \tag{2}$$

The idea of the proof is that for each i and j, we will guarantee that either $L(M_i^A)$ has more than n^j elements at some input length n or $L_i(A)$ is sparse and f_j^A does not reduce $L_i(A)$ to $L(M_i^A)$.

We start with the oracle A empty and build it up in stages. At each stage $m = \langle i, j\rangle$ we will add strings of the form $\langle i, x, y\rangle$ to A where $|x| = n = t(m)$ and $|y| = 2n$. For each stage m we will do one of the following:

1. Put more than r^j strings into $L(M_i^A)$ for some length r, or
2. Make $L_i(A) \cap \Sigma^n$ have exactly one string and for some x in Σ^n, have

$$x \in L_i(A) \Leftrightarrow f_j^A(x) \notin L(M_i^A). \tag{3}$$

By the usual tower arguments we can focus only on the strings in A of length n: Smaller strings can all be queried in polynomial-time; larger strings are too long to be queried.

Pick a string z of length $2n2^n$ that is Kolmogorov random conditioned on the construction of A so far. Read off 2^n strings y_x of length $2n$ for each x in Σ^n. Consider $B = \{\langle i, x, y_x\rangle \mid x \in \Sigma^n\}$.

If $L(M_i^B)$ has more than r^j strings of any length r then we can fulfill the requirement for this stage by letting $A = B$. So let us assume this is not the case.

Note that $f_j^B(x)$ for x of length n cannot query any string y_w in B or we would have a shorter description of z by describing y_w by x and the index of the query made by $f_j^B(x)$. Our final oracle will be a subset of B so we can just use $f_j^\emptyset$ as the reduction.

Suppose $f_j^\emptyset(x) = f_j^\emptyset(w)$ for some x and w of length n. We just let A contain the single string $\langle i, x, y_x\rangle$ and $f_j^\emptyset$ cannot be a reduction. Let us now assume that there is no such x and w.

So by counting there must be some $x \in \Sigma^n$ such that $f_j^\emptyset(x) \notin L(M_i^B)$. Let $v = f_j^\emptyset(x)$. We are not done yet since $L_i(B)$ has too many strings.

Now let A again consist of the single string $\langle i, x, y_x\rangle$. If we still have $v \notin L(M_i^A)$ then we have now fulfilled the requirement.

Otherwise it must be the case that $M_i^A(v)$ accepts but $M_i^B(v)$ rejects. Thus every accepting path (and in particular the lexicographically least) of $M_i^A(v)$ must query some string in $B - A$. Since we can describe v by x this allows us a short description of some y_w given y_x for $w \neq x$ which gives us a shorter description of z, so this case cannot happen. □

Corollary 2. *There exists a relativized world where optimal proof systems do not exist.*

Proof. Messner and Torán [MT98] give a relativizable proof that if optimal proof systems exist then **NP** $\cap$ **SPARSE** has complete sets. □

3 More Powerful Reductions

In the previous section, we constructed a relativized world where the class **NP** $\cap$ **SPARSE** has no complete sets under m-reductions. We now strengthen that construction to more powerful reductions. Using the same techniques as well as other ones, we will also obtain new results on the reducibility of **SPARSE** to **TALLY**.

3.1 Relativized Worlds

We start by extending Theorem 3 to dtt-reductions.

Theorem 4. *There exists a relativized world where* **NP** $\cap$ **SPARSE** *has no dtt-complete sets.*

The proof of Theorem 4 works for any subexponential density bound. In particular, it yields a relativized world where the class of **NP** sets with no more than $2^{n^{o(1)}}$ strings of any length n has no dtt-complete sets.

We can handle polynomial-time tt-reductions with *arbitrary* evaluators provided the number of queries remains in $o(n/\log n)$.

Theorem 5. *There exists a relativized world where* **NP** $\cap$ **SPARSE** *has no complete sets under* $o(n/\log n)$*-tt-reductions.*

For sets of subexponential density the proof of Theorem 5 yields a relativized world where the class of **NP** sets containing no more than $2^{n^{o(1)}}$ strings of any length n, has no complete sets under tt-reductions of which the number of queries is at most n^α for some $\alpha < 1$.

On the positive side, recall from Section 1.4 that **NP** $\cap$ **SPARSE** has complete sets under ctt-reductions as well as under co-rp-reductions.

3.2 SPARSE to TALLY

The techniques used in the proofs of Theorems 3, 4, and 5 also allow us to construct a sparse set S that does not reduce to any tally set under the type of reductions considered. As mentioned in Section 1.3, such sets were already known for m-reductions and for dtt-reductions. For $o(n/\log n)$-tt-reductions we provide the first construction.

Theorem 6. *There exists a sparse set S that does not* $o(n/\log n)$*-tt-reduce to any tally set.*

On the other side, $O(n)$ queries suffice to reduce any sparse set to a tally set. Previously, it was known that **SPARSE** ctt- and co-rp-reduces to **TALLY** (see Section 1.3). We give the first deterministic reduction for which the degree of the polynomial bounding the number of queries does not depend on the density of the sparse set.

Theorem 7. *Every sparse set S is reducible to some tally set T under a 2-round tt-reduction asking $O(n)$ queries.*

Proof. Schöning [Sch93] shows that for any constant $k > 0$ there exists a tally set T_1 and a polynomial-time reduction R such that for any string x of any length n

$$x \in S \Rightarrow \Pr[R(x,\rho) \in T_1] = 1$$
$$x \notin S \Rightarrow \Pr[R(x,\rho) \in T_1] < \frac{1}{n^k}, \tag{4}$$

where the probabilities are uniform over strings ρ of length $O(\log n)$.

By picking $\frac{n}{k \log n}$ independent samples ρ_i, we have for any $x \in \Sigma^n$:

$$x \in S \Rightarrow \Pr[(\forall\, i) R(x,\rho_i) \in T_1] = 1$$
$$x \notin S \Rightarrow \Pr[(\forall\, i) R(x,\rho_i) \in T_1] < (\frac{1}{n^k})^{\frac{n}{k \log n}} = \frac{1}{2^n}.$$

Therefore, there exists a sequence $\tilde{\rho}_i$, $i = 1, \ldots, \frac{n}{k \log n}$, such that

$$\forall\, x \in \Sigma^n : x \in S \Leftrightarrow (\forall\, i) R(x, \tilde{\rho}_i) \in T_1. \tag{5}$$

Since each $\tilde{\rho}_i$ is of length $O(\log n)$, we can encode them in a tally set T_2 from which we can recover them using $O(\frac{n}{k \log n} \cdot \log n)$ nonadaptive queries. This way, we obtain a 2-round tt-reduction from S to $T_1 \oplus T_2$ using $O(n)$ queries: The first round determines the $\tilde{\rho}_i$'s, and the second round applies (5). Since $T_1 \oplus T_2$ m-reduces to a tally set T, we are done. □

In Section 4.1, we will show that under a reasonable hypothesis we can reduce the number of queries in Theorem 7 from $O(n)$ to $\frac{n}{k \log n}$ for any constant $k > 0$. See Corollary 3.

We do not know whether the **NP** ∩ **SPARSE** equivalent of Theorem 7 holds: Does **NP** ∩ **SPARSE** have a complete set under reductions asking $O(n)$ queries? See Section 6 for a discussion.

4 Reductions With Advice — Tight Results

Our results in Section 3 pointed out a difference in the power of reductions making $o(n/\log n)$ queries and reductions making $O(n)$ queries. In this section we close the remaining gap between $o(n/\log n)$ and $O(n)$ by considering reductions that take some advice. The approach works for both the **NP** ∩ **SPARSE** setting and the **SPARSE**-to-**TALLY** setting.

4.1 SPARSE to TALLY

We first observe that Theorem 6 also holds when we allow the reduction $O(\log n)$ bits of advice.

Theorem 8. *Let c be any positive constant. There exists a sparse set S that does not reduce to any tally set T under $o(n/\log n)$-tt-reductions that take $c \cdot \log n$ bits of advice.*

Theorem 8 is essentially optimal under a reasonable assumption as the next result shows.

Theorem 9. *Suppose there exists a set in* $\mathbf{DTIME}[2^{O(n)}]$ *that requires circuits of size $2^{\Omega(n)}$ even when the circuits have access to an oracle for* **SAT**. *Then for all relativized worlds, every sparse set S and every positive constant k, there exists a tally set T and a ctt-reduction from S to T that asks $\frac{n}{k\log n}$ queries and $O(\log n)$ bits of advice.*

Proof. Let S be a sparse set. The construction in the proof of Theorem 7 can be seen as a ctt-reduction of S to the tally set T_1 that makes $\frac{n}{k\log n}$ queries and gets $O(n)$ bits as advice, namely the sequence of $\frac{n}{k\log n}$ $\tilde{\rho}_i$'s, each of length $\ell(n) \in O(\log n)$.

We will now show how the hypothesis of Theorem 9 allows us to reduce the required advice from $O(n)$ to $O(\log n)$ bits.

The requirement the $\tilde{\rho}_i$'s have to fulfill is condition (5). By a slight change in the parameters of the proof of Theorem 7 (namely, by replacing k by $2k$ in (4)), we can guarantee that most sequences $\tilde{\rho}_i$ actually satisfy (5). Since the implication from left to right in (5) holds for any choice of $\tilde{\rho}_i$'s, we really only have to check

$$\forall x \in \Sigma^n : x \notin S \Rightarrow (\exists\, i) R(x, \tilde{\rho}_i) \notin T_1. \tag{6}$$

Without loss of generality, we can assume that $Q_R(\Sigma^n) \cap T_1 = Q_R(S \cap \Sigma^n) \cap T_1$, where $Q_R(X) = \{R(x,\rho) \mid x \in X \text{ and } |\rho| = \ell(|x|)\}$. Therefore, we can replace (6) by the condition

$$\forall x \in \Sigma^n : x \notin S \Rightarrow (\exists\, i) R(x, \tilde{\rho}_i) \notin Q_R(S \cap \Sigma^n). \tag{7}$$

Since S is sparse, this condition on the $\tilde{\rho}_i$'s can be checked by a polynomial-size family of circuits with access to an oracle for **SAT**: The circuit has a enumeration of the elements of $S \cap \Sigma^n$ built in, and once a polynomial-time enumeration of $S \cap \Sigma^n$ is available, (7) becomes a **coNP** predicate.

Under the hypothesis of Theorem 9, Klivans and Van Melkebeek [KvM99, Theorem 4.2] construct a polynomial-time computable function f that maps strings of $O(\log n)$ bits to sequences $\tilde{\rho}_i$ such that most of the inputs map to sequences satisfying (7). An explicit input to f for which this holds, suffices as advice for our reduction from S to $T = T_1$. □

Since we can encode the advice in a tally set and recover it from the tally set using $O(\log n)$ queries, we obtain the following in the terminology of Theorem 7.

Corollary 3. *Under the same hypothesis as in Theorem 9, for any constant $k > 0$ every sparse set S is reducible to some tally set T under a 2-round tt-reduction asking $\frac{n}{k\log n}$ queries.*

4.2 Relativized Worlds

Our tight results about the reducibility of **SPARSE** to **TALLY** carry over to the $\mathbf{NP} \cap \mathbf{SPARSE}$ setting.

Theorem 10. *For any constant $c > 0$, there exists a relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ has no complete sets under $o(n/\log n)$-tt reductions that take $c \cdot \log n$ bits of advice.*

We also note that Theorem 4 can take up to $n - \omega(\log n)$ bits of advice.

Theorem 11. *There exists a relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ has no complete sets under dtt-reductions that take $n - \omega(\log n)$ bits of advice.*

On the positive side, we obtain:

Theorem 12. *Suppose there exists a set in $\mathbf{DTIME}[2^{O(n)}]$ that requires circuits of size $2^{\Omega(n)}$ even when the circuits have access to an oracle for **SAT**. Then for all relativized worlds and all values of $k > 0$, $\mathbf{NP} \cap \mathbf{SPARSE}$ has a complete set under ctt-reductions that ask $\frac{n}{k \log n}$ queries and $O(\log n)$ bits of advice.*

5 NP ∩ SPARSE and Other Promise Classes

Informally, a promise class has a restriction on the set of allowable machines beyond the usual time and space bounds. For example, **UP** consists of languages accepted by **NP**-machines with at most one accepting path. Other common promise classes included $\mathbf{NP} \cap \mathbf{coNP}$, **BPP** (randomized polynomial time), **BQP** (quantum polynomial time) and $\mathbf{NP} \cap \mathbf{SPARSE}$.

Nonpromise classes have easy complete sets, for example:

$$\{\langle i, x, 1^j \rangle \mid M_i(x) \text{ accepts in at most } j \text{ steps}\} \tag{8}$$

is complete for **NP** if M_i are nondeterministic machines, but no such analogue works for **UP**.

We say that **UP** has a uniform enumeration if there exists a computable function ϕ such that for each i and input x, $M_{\phi(i)}(x)$ uses time at most $|x|^i$ and has at most one accepting path on every input and $\mathbf{UP} = \cup_i L(M_{\phi(i)})$. Uniform enumerations for the other promise classes are similarly defined.

It turns out that for most promise classes, having a complete set and a uniform enumeration are equivalent. Hartmanis and Hemachandra [HH84] show this for **UP** and their proof easily generalizes to the other classes. We include a proof here for completeness.

Theorem 13 (Hartmanis-Hemachandra). *The classes **UP**, $\mathbf{NP} \cap \mathbf{coNP}$, **BPP** and **BQP** have complete sets under many-one reductions if and only if they have uniform enumerations.*

For $\mathbf{NP} \cap \mathbf{SPARSE}$ neither direction of the proof goes through. In fact despite Theorem 3, $\mathbf{NP} \cap \mathbf{SPARSE}$ has a uniform enumeration (in all relativized worlds).

Theorem 14. *The class* $\mathbf{NP} \cap \mathbf{SPARSE}$ *has a uniform enumeration.*

In some sense Theorem 14 is a cheat. In the uniform enumeration, all the sets are sparse but we cannot be sure of the census function at a given input length. To examine this case we extend the definition of uniform enumeration.

Definition 2. *We say* $\mathbf{NP} \cap \mathbf{SPARSE}$ *has a* uniform enumeration with size bounds *if there exists a computable function* ϕ *such that* $\mathbf{NP} \cap \mathbf{SPARSE} = \cup_i L(M_{\phi(i)})$*, and for all* i *and* n*,* $M_{\phi(i)}$ *accepts at most* n^i *strings of length* n *using at most* n^i *time.*

Hemaspaandra, Jain and Vereshchagin [HJV93] defined a similar extension for the class **FewP**.

We can use Definition 2 to prove a result similar to Theorem 13 for the class $\mathbf{NP} \cap \mathbf{SPARSE}$.

Theorem 15. $\mathbf{NP} \cap \mathbf{SPARSE}$ *has complete sets under invertible reductions if and only if* $\mathbf{NP} \cap \mathbf{SPARSE}$ *has a uniform enumeration with size bounds.*

The promise class $\mathbf{NP} \cap \mathbf{SPARSE}$ differs from the other classes in another interesting way. Consider the question as to whether there exists a language accepted by a nondeterministic machine using time n^3 which has at most one accepting path on each input that is not accepted by any such machine using time n^2. This remains a murky open question for **UP** and the other usual promise classes.

For $\mathbf{NP} \cap \mathbf{SPARSE}$ the situation is quite different as shown by Seiferas, Fischer and Meyer [SFM78] and Žàk [Žàk83].

Theorem 16 (Seiferas-Fischer-Meyer,Žàk). *Let the functions* t_1 *and* t_2 *be time-constructible such that* $t_1(n+1) = o(t_2(n))$*. There exists a tally set accepted by a nondeterministic machine in time* $t_2(n)$ *but not in time* $O(t_1(n))$*.*

6 Open Problems

Several interesting questions remain including the following.

- Theorem 7 which shows that every sparse set reduces to a tally set using $O(n)$ queries does not seem to give a corresponding result for $\mathbf{NP} \cap \mathbf{SPARSE}$-complete sets. Is there a relativized world where $\mathbf{NP} \cap \mathbf{SPARSE}$ does not have complete sets under Turing reductions using $O(n)$ queries? If we can construct the $\tilde{\rho}_i$'s in the proof of Theorem 7 in polynomial time using access to a set in $\mathbf{NP} \cap \mathbf{coNP}$, the answer is yes. However, the best we know is to construct them in polynomial time with oracle access to $\mathbf{NP}^{\mathbf{NP}}$.
- Can we reduce or eliminate the assumption needed for Theorem 9, Corollary 3, and Theorem 12? If we knew how to construct the $\tilde{\rho}_i$'s from the proof of Theorem 9 in polynomial time with $O(\log n)$ bits of advice, we could drop the assumption.
- Does $\mathbf{NP} \cap \mathbf{SPARSE}$ having m-complete sets imply $\mathbf{NP} \cap \mathbf{SPARSE}$ has a uniform enumeration with size bounds? Can we construct in a relativized world a complete set for $\mathbf{NP} \cap \mathbf{SPARSE}$ that is not complete under invertible reductions?

References

[BHL95] H. Buhrman, E. Hemaspaandra, and L. Longpré. SPARSE reduces conjunctively to TALLY. *SIAM Journal on Computing*, 24(3):673–681, 1995.

[BK88] R. Book and K. Ko. On sets truth-table reducible to sparse sets. *SIAM Journal on Computing*, 17(5):903–919, 1988.

[CR79] S. Cook and R. Reckhow. The relative efficiency of propositional proof systems. *Journal of Symbolic Logic*, 44:36–50, 1979.

[For94] L. Fortnow. The role of relativization in complexity theory. *Bulletin of the European Association for Theoretical Computer Science*, 52:229–244, February 1994.

[HH84] J. Hartmanis and L. Hemachandra. Complexity classes without machines: On complete languages for UP. *Theoretical Computer Science*, 34:17–32, 1984.

[HJV93] L. Hemaspaandra, S. Jain, and N. Vereshchagin. Banishing robust turing completeness. *International Journal of Foundations of Computer Science*, 4(3):245–265, 1993.

[HY84] J. Hartmanis and Y. Yesha. Computation times of NP sets of different densities. *Theoretical Computer Science*, 34(1-2):17–32, November 1984.

[Ko89] K. Ko. Distinguishing conjunctive and disjunctive reducibilities by sparse sets. *Information and Computation*, 81(1):62–87, 1989.

[KP89] J. Krajíček and P. Pudlák. Propositional proof systems, the consistency of first order theories and the complexity of computations. *Journal of Symbolic Logic*, 54:1063–1079, 1989.

[KvM99] A. Klivans and D. van Melkebeek. Graph nonisomorhism has subexponential size proofs unless the polynomial-time hierarchy collapses. In *Proceedings of the 31st ACM Symposium on the Theory of Computing*, pages 659–667. ACM, New York, 1999.

[LV97] M. Li and P. Vitányi. *An Introduction to Kolmogorov Complexity and Its Applications*. Graduate Texts in Computer Science. Springer, New York, second edition, 1997.

[MT98] J. Messner and J. Torán. Optimal proof systems for propositional logic and complete sets. In *Proceedings of the 15th Symposium on Theoretical Aspects of Computer Science*, volume 1373 of *Lecture Notes in Computer Science*, pages 477–487. Springer, 1998.

[Sal93] S. Saluja. Relativized limitations of left set technique and closure classes of sparse sets. In *Proceedings of the 8th IEEE Structure in Complexity Theory Conference*, pages 215–223. IEEE, New York, 1993.

[Sch93] U. Schöning. On random reductions from sparse sets to tally sets. *Information Processing Letters*, 46(5):239–241, July 1993.

[SFM78] J. Seiferas, M. Fischer, and A. Meyer. Separating nondeterministic time complexity classes. *Journal of the ACM*, 25(1):146–167, 1978.

[Žàk83] S. Žàk. A Turing machine time hierarchy. *Theoretical Computer Science*, 26(3):327–333, 1983.