Six Hypotheses in Search of a Theorem

Harry Buhrman* Lance Fortnow† Leen Torenvliet‡

> Sir, we are truly six special and interesting characters. Believe us. However we have gone lost.
> – "Six Characters in Search of an Author," Luigi Pirandello.

Abstract

We consider the following six hypotheses:

- $\mathbf{P} = \mathbf{NP}$.
- **SAT** is truth-table reducible to a **P**-selective set.
- **SAT** is truth-table reducible to a k-approximable set for some k.
- $\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP}[\log]}$
- **SAT** is $O(\log n)$-approximable.
- Solving **SAT** is in **P** on formulae with at most one assignment.

We discuss their importance and relationships among them.

*URL: http://www.cwi.nl/cwi/people/Harry.Buhrman.html. E-mail: buhrman@cwi.nl. Partially supported by the Dutch foundation for scientific research (NWO) by SION project 612-34-002, and by the European Union through Neuro-COLT ESPRIT Working Group Nr. 8556, and HC&M grant nr. ERB4050PL93-0516. CWI, Kruislaan 413, 1098SJ Amsterdam, The Netherlands.

†URL: http://www.cs.uchicago.edu/~fortnow. Email: fortnow@cs.uchicago.edu. Supported in part by NSF grant CCR 92-53582, the Dutch Foundation for Scientific Research (NWO) and a Fulbright Scholar award. CWI and University of Chicago, Department of Computer Science 1100 E. 58th. St. Chicago, IL 60637, USA.

‡URL: http://turing.wins.uva.nl/~leen/. E-mail: leen@wins.uva.nl. University of Amsterdam, Department of Computer Science, Plantage Muidergracht 24, 1018TV Amsterdam, The Netherlands.

1 Introduction

Complexity theorists have put considerable effort into investigating the structure and properties of sets in **NP**. This research led to various hypotheses. In this survey paper we put together, for the first time, six hypotheses that we encountered in our own research as well as in the literature. We believe that these hypotheses are important and are closely related to each other.

The first hypothesis is: "$\mathbf{P} = \mathbf{NP}$." This is the most famous and important one and does not need any further introduction.

Most sets in **NP** that arise from practice turn out to be **NP**-complete. Moreover since complete sets reflect the structure of a complexity class they receive close attention. Three of our six hypotheses concern sets that are complete or hard for **NP**.

Selman [Sel82] introduced the **P**-selective sets in analogue of recursion theory. A set is called **P**-*selective* iff there exists a polynomial time computable function that from two strings x and y selects one that (if at least one belongs to A) is in A. He investigated the possibility for **NP** to have hard sets that are **P**-selective. He showed [Sel82] that this can not be the case for many-one reductions (unless $\mathbf{P} = \mathbf{NP}$). This was later improved to $\leq_{1-tt}^{p}$ reductions by Buhrman and Torenvliet [BT96b]. The hypothesis we are interested in is: "**NP** has a truth-table hard set that is **P**-selective."

Beigel [Bei87a], looking at properties of bounded queries to sets (in **NP**), developed a generalization of **P**-selective sets later dubbed the approximable sets. A set A is k-approximable if there exists a polynomial time computable function that with k strings $x_1, \ldots, x_k$ as input, generates k bits $b_1, \ldots, b_k$ such that for at least 1 bit it is true that $b_i \neq \chi_A(x_i)$. That is from the 2^k possible settings of $x_1, \ldots, x_k$ one is excluded. Beigel, Kummer and Stephan [BKS95], Agrawal and Arvind [AA96], and Ogihara [Ogi95]

1093-0159/97 $10.00 © 1997 IEEE

showed that **NP** can not have $\leq^p_{btt}$ -hard sets that are k-approximable for some k (unless **P** = **NP**). Since **P**-selective sets are in fact 2-approximable sets this result also improves the bound for **P**-selective sets. The hypothesis related to this work is: "**NP** has a truth-table hard set that is k-approximable for some k."

Ogihara [Ogi95] working on the hypothesis that **NP** has a truth-table hard **P**-selective set, took it one step further and considered $f(n)$-approximable sets for non-constant functions $f(n)$. He showed that if **SAT** is not $a\log(n)$-approximable for $a < 1$ unless **P** = **NP**. This result subsumes the results on truth-table reductions to k-approximable sets (see Section 3). The hypothesis connected to this work is: "**SAT** is $O(\log(n))$-approximable."

The next hypothesis states that it is possible to compute **SAT** in polynomial time when we only consider formulae with at most *one* satisfying assignment. It is possible to phrase this in terms of sets as: "Unique-**SAT** $\in$ **P**" (see Section 2). Valiant and Vazirani [VV86] showed that this set problem for **SAT** is hard for **NP** under randomized reductions.

The last hypothesis deals with functions that are computable in polynomial time relative to some set in **NP**. There are essentially three different ways to define this. The most unrestricted way is that the polynomial time computable function has unrestricted access to an **NP** oracle and is called $\mathbf{FP}^{\mathbf{NP}}$. The next restriction to the oracle mechanism is that the queries have to be non-adaptive: $\mathbf{FP}^{\mathbf{NP}}_{||}$. The last and most restrictive version is that only $O(\log(n))$ queries are allowed on inputs of length n: $\mathbf{FP}^{\mathbf{NP[log]}}$. The last hypothesis can now be stated as: $\mathbf{FP}^{\mathbf{NP[log]}} = \mathbf{FP}^{\mathbf{NP}}_{||}$.

These are the main characters of our paper. We show that these hypotheses are closely related to each other and in Section 3 we show which of these hypotheses implies any of the others. Furthermore we give background information on each of them individually and we indicate which problems are still open. The main open question however is to show that any two of these six hypotheses are equivalent.

We should note that probably all of the six hypotheses are false since all of them imply that **NP** $\subseteq$ **P**/*poly* and this on its turn implies that the polynomial time hierarchy collapses to its second level [KL80].

Until recently no oracles were known that showed that any of these hypotheses are different from each other. However recent progress has been made in this direction (see Section 7).

2 Preliminaries

We assume the reader familiar with basic notions of computation and complexity theory as can be found e.g. in [HU79, BDG88, BDG90, GJ79] and many other textbooks.

Central to the six hypotheses in this paper however are the following notions, which we will highlight here by separately defining them.

Definition 2.1 *A set A is called* **P**-*selective iff there exists a polynomial time computable function f (called p-selector function) such that for any two strings x and y, $f(x,y) \in \{x,y\}$ and if x or y is in A then $f(x,y)$ is in A.*

For a set A we will identify A with its characteristic function. Hence for a string x, $A(x) \in \{0,1\}$ and $A(x) = 1$ iff $x \in A$. For two strings x and y and a **P**-selective set A, a p-selector excludes one of the four possibilities for the string $A(x)A(y)$ (either 01 or 10 is impossible). A generalization extends this exclusion to one of the possible settings for the string $A(x_1)\ldots A(x_k)$ for some function $k(n)$. For constant k, this notion was called "approximability" of sets (see Beigel et al. [BKS95]).

Definition 2.2 *A function g is called an f-approximator for a set A if for every $x_1,\ldots,x_m$ with $m \geq f(\max\{|x_1|,\ldots,|x_m|\})$,*

$$g(x_1,\ldots,x_m) \in \{0,1\}^m$$
$$\textit{and}$$
$$(A(x_1),\ldots,A(x_m)) \neq g(x_1,\ldots,x_m)$$

A set A is then called f-approximable if it has an f-approximator. A is bounded-approximable, *or $A \in$ **bAPP** if A is k-approximable for some constant k.*

The notion f-approximability was called f-membership comparability by Ogihara [Ogi95] who was the first to consider this notion for nonconstant functions. Beigel [Bei87a] uses the term "approximable" to represent **bAPP**. Sets which are not in **bAPP** Beigel calls superterse.

Amir, Beigel and Gasarch [ABG90] show that every **bAPP** language is in **P**/poly. Ogihara [Ogi95] notices that their proof generalizes.

Theorem 2.3 (Amir-Beigel-Gasarch-Ogihara) *If A is $f(n)$-approximable for any polynomial $f(n)$ then A is in* **P**/*poly.*

We use the function $\mathbf{F_{SAT}}$ which on input $\phi_1,\ldots,\phi_n$ returns a string $x \in \{0,1\}^n$, where $x_i = 1$ iff $\phi_i \in$ **SAT**. We will also need classes of functions

that are computable by queries to **SAT**. Depending on the number of queries and the type of oracle access these are defined as follows.

Definition 2.4 *A function f is in $\mathbf{FP}_{||}^{\mathbf{NP}}$ if there exists a polynomial time bounded oracle machine M that computes f with non-adaptive queries to some language in* **NP**.

Note that $\mathbf{F_{SAT}}$ is $\mathbf{FP}_{||}^{\mathbf{NP}}$ complete. A set is sparse if there exists a polynomial p such that for each length n it contains at most $p(n)$ strings. Let **SPARSE** denote the class of all sparse sets.

A truth-table reduction from A to B is disjunctive ($A \leq_{dtt}^{p} B$) if it accepts iff one of it queries is in B.

Definition 2.5 *A function f is in $\mathbf{FP}^{\mathbf{NP}[\log]}$ if there is a polynomial time bounded oracle machine that computes f using $O(\log n)$ (adaptive) queries to some language in* **NP**.

Definition 2.6 *Let Q denote a boolean predicate. we define the set* Unique-$\mathbf{SAT}_Q$ *as follows.*
For any formula x

$$\textit{Unique-}\mathbf{SAT}_Q(x) = \begin{cases} 0 & \text{if } x \notin \mathbf{SAT} \\ 1 & \text{if } x \text{ has 1 satisfying assignment} \\ Q(x) & \text{Otherwise} \end{cases}$$

If there exists a predicate Q such that Unique-$\mathbf{SAT}_Q$ is polynomial time computable then we will say "Unique-$\mathbf{SAT} \in P$*."*

The notion of bounded nondeterminism was introduced by Kintala and Fischer in [KF80].

Definition 2.7 *Let f be any function. We define $\mathbf{NP}(f(n)) = \{L \mid L \subseteq \{0,1\}^*$ and there is a constant c such that L is accepted by a polynomial time bounded Turing machine making at most $f(n)$ c-ary nondeterministic moves$\}$*

Kintala and Fischer denote $\mathbf{NP}(f(n))$ as $\mathcal{P}_{f(n)}$.

Definition 2.8 *A function $f(x)$ is $h(n)$-enumerable iff there exists a polynomial-time computable function $g(x) = \{y_1, \ldots, y_{h(n)}\}$ such that for every x, $f(x) \in g(x)$. A function $f(x)$ is poly-enumerable $f(x)$ is n^c enumerable for some c.*

There is a very useful connection between $\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP}[\log]}$ and the enumerability of $\mathbf{F_{SAT}}$ [Bei87a].

Lemma 2.9 (Beigel) $\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP}[\log]}$ *if and only if* $\mathbf{F_{SAT}}$ *is poly-enumerable.*

Proof:
($\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP}} \Rightarrow \mathbf{F_{SAT}}$ is poly-enumerable)
$\mathbf{F_{SAT}} \in \mathbf{FP}_{||}^{\mathbf{NP}}$, so by assumption it is in $\mathbf{FP}^{\mathbf{NP}[\log]}$. There are polynomially possible answers for the oracle queries of the $\mathbf{FP}^{\mathbf{NP}[\log]}$ machine. Cycling through them yields an enumeration of $\mathbf{F_{SAT}}$.
($\mathbf{F_{SAT}}$ is polynomial enumerable $\Rightarrow \mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP}}$)
On input $\phi_1, \ldots, \phi_l$ each of size at most n one can enumerate n^c vectors $b_1, \ldots b_{n^c}$ such that $b_i = \mathbf{F_{SAT}}$ for some i. Next one can use binary search to some suitable oracle in **NP** to find b_i, using $\log(n^c) + 1$ queries. □

We will need the following definition of the dimension of a family of sets, called Vapnik-Chervonenkis dimension [VC71]:

Definition 2.10 *Given a family of sets $\mathcal{F}$ the Vapnik-Chervonenkis dimension of $\mathcal{F}$ or VC-dimension is the largest number d such that there exists a set A with $\|A\| = d$ and $\|\{A \cap F \mid F \in \mathcal{F}\}\| = 2^d$. If such a d does not exist the VC-dimension of $\mathcal{F}$ is ∞.*

Sauer [Sau72] and independently Shelah [She72] proved the following lemma. Sauer notes that Paul Erdös originally posed this as a question.

Lemma 2.11 *If $\mathcal{F}$ is a family of sets with VC-dimension at most d then for any set A with $\|A\| = n$:*

$$\|\{A \cap F \mid F \in \mathcal{F}\}\| \leq \sum_{i=0}^{d} \binom{n}{i}$$

For $n \geq d \geq 1$, $\sum_{i=0}^{d} \binom{n}{i}$ is bounded by $n^d + 1$. Moreover the proof of Lemma 2.11 is constructible: Suppose we have a polynomial-time algorithm that on $S = x_1, \ldots, x_{d+1}$ computes a subset of S that is not in $\{A \cap F \mid F \in \mathcal{F}\}$. Lemma 2.11 gives us a polynomial-time algorithm to compute $\{A \cap F \mid F \in \mathcal{F}\}$ in time polynomial in n and the sizes of the elements of A.

3 Relations

In this section we will show which of the six hypotheses implies any of the others. The relations are given in Figure 1.

Theorem 3.1 $\mathbf{P} = \mathbf{NP} \Rightarrow \mathbf{SAT} \leq_{tt}^{p} \mathbf{Psel}$.

Proof: If $\mathbf{P} = \mathbf{NP}$ then **SAT** is in **P** and reduces to any set. □

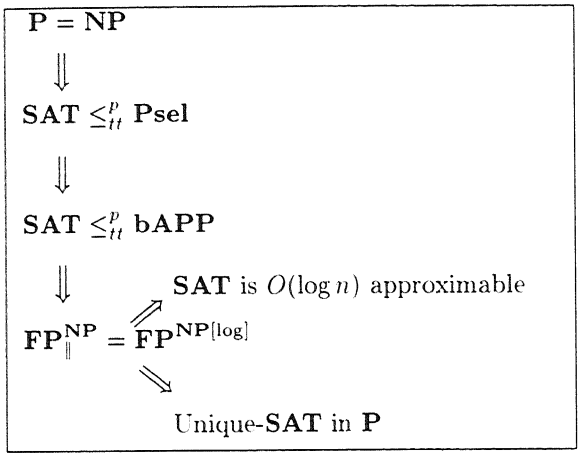

Figure 1: Relations

Theorem 3.2 $\mathbf{SAT} \leq^p_{tt} \mathbf{Psel} \Rightarrow \mathbf{SAT} \leq^p_{tt} \mathbf{bAPP}$

Proof: Note that every **P**-selective set is 2-approximable. □

Theorem 3.3 $\mathbf{SAT} \leq^p_{tt} \mathbf{bAPP} \Rightarrow \mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP[log]}}$.

We first prove the following lemma due to Beigel [Bei87a, Bei87b].

Lemma 3.4 (Beigel) *If A is k-approximable then there exists a function f which computes for any n numbers $x_1, \ldots, x_n$ a set of at most $\sum_{i=0}^{k-1} \binom{n}{i}$ vectors from $\{0,1\}^n$ which contains $\mathbf{F}^A_n(x_1, \ldots, x_n)$. Moreover f runs in time polynomial in n and the size of the largest string in $x_1, \ldots, x_n$.*

Proof: Let g be the function that k-approximates A. Define the following family of sets:

$$\mathcal{F} = \{B \mid g \text{ is a } k\text{-approximator for } B\}$$

It follows that the VC-dimension of $\mathcal{F}$ is at most $k \Leftrightarrow 1$. We then apply the constructible version of Lemma 2.11. □

We now give the proof of Theorem 3.3.

Proof: Let M witness the fact that **SAT** truth-table reduces to a k-approximable set A. let $f \in \mathbf{FP}^{\mathbf{NP}}_{||}$ via machine M_f. On input x, M_f computes the following queries $q_1, \ldots, q_l$ to **SAT**, for l some polynomial. Next reduce each of these queries to A with M, yielding a set of queries $q'_1, \ldots, q'_{l'}$, for l' a polynomial. Next we apply Lemma 3.4 to generate l'^k many different vectors, containing $\mathbf{F}^A_{l'}(q'_1, \ldots, q'_{l'})$. From these vectors one can generate l'^k many vectors containing $\mathbf{F}^{\mathbf{SAT}}_l(q_1, \ldots, q_l)$. $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP[log]}}$ follows from Lemma 2.9. □

The following theorem is implicit in [Bei88, Tod91b]

Theorem 3.5 (Beigel-Toda)
$\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP[log]}} \Rightarrow$ *Unique*-**SAT** *is in* **P**

Proof: We have to show that there is a polynomial time algorithm that tells formulae with exactly one satisfying assignment apart from ones that are unsatisfiable. Consider the function $f(\phi)$ that on input ϕ with variables $x_1, \ldots, x_k$ returns $b_1 \ldots b_k$ such that $b_i = 1$ iff there is a satisfying assignment to ϕ with $x_i = 1$. This function is in $\mathbf{FP}^{\mathbf{NP}}_{||}$ and hence, by assumption in $\mathbf{FP}^{\mathbf{NP[log]}}$. Suppose we are given a formula ϕ with exactly 1 satisfying assignment. Then f will return exactly this assignment. Since there are only polynomial many possible answers to the $\log(n)$ queries to **SAT**, one can enumerate all the possible values of f in **P**. We can check that one of the generated values is indeed a satisfying assignment to ϕ. On the other hand if ϕ was unsatisfiable we would not have generated a satisfying assignment, since none exists. □

Theorem 3.6 $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP[log]}} \Rightarrow$ **SAT** *is* $O(\log(n))$*-approximable.*

Proof: By Lemma 2.9 we have that $\mathbf{F_{SAT}}$ is m^c enumerable for some c where m is the input length of $\mathbf{F_{SAT}}$. Given any $2c\log(n)$ formulae $\phi_1, \ldots, \phi_{2c\log(n)}$ each of size at most n. The size of these $2c\log(n)$ formulae is bounded by $2c\log(n) \times n$ and thus $\mathbf{F}_{SAT}(\phi_1, \ldots, \phi_{2c\log(n)})$ is $2^{c\log(2c\log(n)\times n)} < n^{c+1}$ enumerable. Thus one of the n^{2c} vectors for $\mathbf{F_{SAT}}$ has not been enumerated. □

4 Selective and Approximable

The question whether sets that have simple structure could be hard for **NP** dates back to the Berman-Hartmanis conjecture [BH77] and subsequent work by Mahaney for sparse sets [Mah82]. Following sparse sets, the first sets of simple structure to be considered were the **P**-selective sets introduced by [Sel79].

P-selective sets, though of arbitrary complexity, are structurally simple sets. The p-selector function induces an ordering that reduces the number of possible "membership configurations" of two strings. For a **P**-selective set A and two strings x and y either $x \in A \wedge y \notin A$ or $y \in A \wedge x \notin A$ is ruled out

by the p-selector. This property makes **P**-selective sets structurally as simple as being Turing equivalent to tally sets [Sel82]. Generalizing the structural restriction: "Not all 2^n membership configurations of n strings are possible" has induced many related notions. Among the many notions that pertain to this idea are: **P**-selective sets [Sel79, HHN+95], near-testable sets [GHJY91], k-approximable sets (see below), $(a,b)_p$-recursive sets [KS91], Easily countable sets [HN93], Cheatable sets [Bei87a, BGGO93], $(a,b)_p$-verbose sets [BKS], and Membership comparable sets [Ogi95].

Because of the structural relation between **P**-selective sets and sparse sets, one might not be too surprised that hardness of **P**-selective sets for **NP** is as unlikely as hardness for **NP** of sparse sets. It is quite easy to see that **SAT** itself cannot be **P**-selective unless **P** = **NP**. Buhrman and Torenvliet [BT96b] showed that **SAT** cannot be 1-tt reducible to a **P**-selective set.

Toda [Tod91a], building upon insights provided by Ko [Ko83], proved that in the special case of the existence of only one satisfying assignment, reduction to a **P**-selective set would imply polynomial time decidability. In fact Toda's results hold for the more general k-approximable sets. In this section we cite all results for k-approximable sets. Since **P**-selective sets are k-approximable sets with $k = 2$, all these results also hold for **P**-selective sets. Similar ideas were obtained independently by Beigel [Bei88].

Theorem 4.1 (Beigel-Toda)

1. **P** = **UP** *if and only if* **UP** $\leq_{tt}^{p}$ **bAPP**.
2. *Unique*-**SAT** $\in$ **P** *if and only if Unique*-**SAT**$_Q$ $\leq_{tt}^{p}$ **bAPP** *for some* Q.
3. **P** = **NP** *if and only iff* $\Delta_2^p \leq_{tt}^{p}$ **bAPP**
4. **P** = **PSPACE** *if and only* **PSPACE** $\leq_{tt}^{p}$ **bAPP**.
5. **EXP** $\not\leq_{tt}^{p}$ **bAPP**

The Turing reduction of **bAPP** sets to sparse sets (Theorem 2.3) allows us to apply the famous Karp-Lipton theorem [KL80] showing a collapse of the polynomial-hierarchy if **SAT** is Turing-reducible to a sparse sets.

Theorem 4.2 (Karp-Lipton) *If* **SAT** $\leq_T^p$ **bAPP** *then* **PH** = Σ_2^p

or in its currently sharpest form proved in [BCG+96, KW95].

Theorem 4.3 (BCGKTKW) *If* **SAT** $\leq_T^p$ **bAPP** *then* **PH** = **ZPP**$^{\mathbf{NP}}$

Both directions of strengthening the consequence of **SAT** $\leq_T^p$ **bAPP** and weakening the reduction type r in **SAT** $\leq_r^p$ **bAPP** $\Rightarrow$ **P** = **NP** are currently the subject of active research. Of course in the present context the latter type is the more interesting. In 1994 a major breakthrough was achieved by three independent sets of authors: Beigel, Kummer and Stephan[BKS95], Agrawal and Arvind[AA96] and Ogihara[Ogi95].

Theorem 4.4 (AABKOS) *If* **SAT** $\leq_{btt}^{p}$ **bAPP** *then* **P** = **NP**

Or in its currently strongest form

Theorem 4.5 (AABKOS) *If* **SAT** $\leq_{n^\alpha - tt}^{p}$ *to some k-approximable set for some* $\alpha < \frac{1}{k-1}$ *then* **P** = **NP**.

For **P**-selective sets $k = 2$ and hence $\alpha < 1$ follows.

To understand this result we first show a relationship between reducing to **bAPP** and $r \log n$-approximability.

Theorem 4.6 *If* **SAT** $\leq_{n^\alpha - tt}^{p}$ *to some k-approximable set for some* $\alpha < \frac{1}{k-1}$ *then* **SAT** *is $r \log n$ approximable for some* $r < 1$.

Proof: Note that in Lemma 3.4 the number of vectors is actually bounded by $k \times n^{k-1}$. Hence if we have $r \log n$ formulae $\phi_1, \ldots, \phi_{r \log n}$ we can reduce these to a k-approximable set A via a reduction that produces n^α queries for a total of $(r \log n) n^\alpha < r \times n^\beta$ where $\beta < \frac{1}{k-1}$. Applying Lemma 3.4 gives $(r \times n^\beta)^{k-1}$ vectors including the characteristic vector of these formulae. Hence if $1 > r > \frac{\beta}{k-1}$ we can exclude at least one possibility, which means that **SAT** is $r \log n$-approximable. □

We can then apply the following result from [AA96, BKS95, Ogi95].

Theorem 4.7 (AABKOS) *If* **SAT** *is $r \log n$-approximable for some* $r < 1$ *then* **P** = **NP**.

To give a flavor of the proof we prove the following weaker result.

Theorem 4.8 *If* **SAT** *is 2-approximable, then* **P** = **NP**.

Proof: Given a formula ϕ, apply the standard self-reduction to produce four formulae ϕ_1, ϕ_2, ϕ_3, ϕ_4 with the property that ϕ is satisfiable iff at least one

of these formulae is satisfiable. Now let f be a 2-approximator and let $f(\phi_1 \vee \phi_2, \phi_1 \vee \phi_3) = (b_1, b_2)$. If $b_1 = b_2 = 0$ then ϕ is satisfiable and we're done. If (b_1, b_2) is $(1,0)$ then ϕ_2 can not be the only satisfiable formula. If $(b_1, b_2) = (0,1)$ then ϕ_3 can not be the only satisfiable formula. Finally, if $(b_1, b_2) = (1,1)$ then ϕ_1 is not satisfiable.

In all cases one formula in the self-reduction can be discarded and the corresponding branch in the self-reduction tree ends. Hence the self-reduction can be expanded always keeping only four formulae in the game. When all remaining self-reduction branches are extended to their full length, satisfiability of ϕ can be decided trivially. □

A polynomial (even fixed) number of queries in Theorem 4.5 is not yet in sight, nor does the proof technique seem to be extendible to obtain such a result. On the other hand there is no known oracle where $\mathbf{P} \neq \mathbf{NP}$ and $\mathbf{SAT} \leq^p_{tt} \mathbf{Psel}$.

The notion of **P**-selectivity has been extended to other types of selector functions ([HHN+95]) for these (mostly nondeterministic) selector types similar results are known. These are however outside the scope of this paper.

The value $r < 1$ seems to be a real bottleneck of the technique (see [Ogi95] for a discussion) used for the proof, but on the other hand no oracle is known where $\mathbf{P} \neq \mathbf{NP}$ and **SAT** is $O(\log n)$-approximable.

5 $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$

At first glance one might think that $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ since this is true for the language classes: $\mathbf{P}^{\mathbf{NP}}_{||} = \mathbf{P}^{\mathbf{NP}[\log]}$ [BH91, Wag90]. Indeed this result yields that $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ when only functions are considered that compute $\log(n)$ output bits (i.e. functions from $\{0,1\}^n$ to $\{0,1\}^{O(\log(n))}$). However $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ implies Unique-**SAT** in **P** and this implies that the polynomial hierarchy collapses (see Section 6). For overview papers on functions classes and related problems see [JT95, JT97, Sel96].

In Lemma 2.9 we saw that $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ is equivalent to $\mathbf{F_{SAT}}$ being polynomial enumerable. We can use these ideas to get equivalences of $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ to many other hypotheses.

Theorem 5.1 *The following are equivalent:*

- $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$
- $\mathbf{FP}^{\mathbf{NP}}_{||} \subseteq \mathbf{FP}^{X[\log]}$ *for some oracle* X. *[Bei88]*
- $\mathbf{F_{SAT}}$ *is polynomial enumerable.*
- *Every* **NPSV** *function is polynomial enumerable.*

where **NPSV** is the class of single-valued nondeterministic functions (see [Sel96]).

Some progress has been made on showing the equivalence with $\mathbf{P} = \mathbf{NP}$. Jenner and Toran [JT95] showed that $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ implies that **SAT** can be computed in less than 2^n time. They also showed that languages recognized by nondeterministic polynomial time machines that make $\log^k(n)$ nondeterministic moves are in **P**.

Theorem 5.2 (Jenner-Toran) *If* $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ *then*

1. $\mathbf{NP} \subseteq \mathbf{DTIME}(2^{n^{O(1/\log\log(n))}})$.
2. $\mathbf{NP}(\log^k(n)) \subseteq \mathbf{P}$.

Buhrman and Fortnow showed that the $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ question can be phrased as a question on resource bounded Kolmogorov complexity [BF97].

Theorem 5.3 (Buhrman-Fortnow) *The following are equivalent:*

1. $\mathbf{CND}^{poly}(x \mid y) \leq \mathbf{C}^{poly}(x \mid y) + O(\log(|x|))$.
2. $\mathbf{CND}^{poly}(x \mid y) \leq \mathbf{CD}^{poly}(x \mid y) + O(\log(|x|))$.
3. $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$.

The connection with Kolmogorov complexity enables one to use Theorem 5.2 to prove:

Theorem 5.4 (Buhrman-Fortnow) *If* $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]}$ *then the class of languages accepted by nondeterministic polynomial time machines that have at most* $2^{\log^k(n)}$ *accepting paths on inputs of length* n *is included in* **P**.

On the other hand it follows from [Ogi95] that

Theorem 5.5 *If* $\mathbf{FP}^{\mathbf{NP}}_{\beta \log n} \subseteq \mathbf{FP}^{\mathbf{NP}[\alpha \log n]}$ *for some* $1 > \beta > \alpha$ *then* $\mathbf{P} = \mathbf{NP}$.

All the above results have not established the equivalence with $\mathbf{P} = \mathbf{NP}$. We note here that in order to obtain an equivalence it is sufficient to prove that $\mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]} \Rightarrow \mathbf{P}^{\mathbf{NP}} = \mathbf{P}^{\mathbf{NP}}_{||}$ by the following theorem.

Theorem 5.6

$$\mathbf{P}^{\mathbf{NP}} = \mathbf{P}^{\mathbf{NP}}_{||} \text{ and } \mathbf{FP}^{\mathbf{NP}}_{||} = \mathbf{FP}^{\mathbf{NP}[\log]} \Longrightarrow \mathbf{P} = \mathbf{NP}$$

Proof: If $\mathbf{P}^{\mathbf{NP}} = \mathbf{P}_{||}^{\mathbf{NP}}$ then the leftmost satisfying assignment can be computed in $\mathbf{FP}_{||}^{\mathbf{NP}}$ and hence via assumption in $\mathbf{FP}^{\mathbf{NP[log]}}$. We can then cycle through all the possible oracle queries as in the proof of Lemma 2.9 and find the assignment in **FP**. □

Corollary 5.7 *If* NP = coNP *and* $\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP[log]}}$ *then* P = NP.

This argument shows that it is actually sufficient to prove that $\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP[log]}}$ implies that *some* satisfying assignment can be found in $\mathbf{FP}_{||}^{\mathbf{NP}}$. See [WT93, BT96a] for this question. Watanabe and Toda show that relative to a random oracle it is the case that some satisfying assignment can be found in $\mathbf{FP}_{||}^{\mathbf{NP}}$. However relative to a random oracle all of the six hypotheses fail (see Section 7).

6 Unique-SAT is in P

The hypotheses "Unique-**SAT** is in **P**" is a promise problem. It states the existence of a polynomial time algorithm that, under the promise that a formula has either no or a single satisfying assignment, decides whether this formula is satisfiable. Valiant and Vazirani showed that **SAT** is randomly reducible to Unique-$\mathbf{SAT}_Q$ for any predicate Q.

Theorem 6.1 (Valiant-Vazirani) *There is a polynomial time randomized procedure that given a formula ϕ of length n produces a list of n^c formulae $\phi_1, \ldots, \phi_{n^c}$ with the property that:*

- $\phi \in \mathbf{SAT}$ *then with probability* $(1 - 2^{-n})$ *there is an i such that ϕ_i has exactly 1 satisfying assignment.*
- $\phi \notin \mathbf{SAT}$ *then for all i, $\phi_i \notin \mathbf{SAT}$.*

Theorem 6.1 is the key to show that Unique-**SAT** in **P** implies that NP = **R**.

Theorem 6.2 (Valiant-Vazirani) *If Unique-*SAT *is in* P *then* NP = R *and the polynomial hierarchy collapses.*

Proof: Given a formula ϕ, use Theorem 6.1 to randomly produce a list of n^c formulae. Next for each of these formula ϕ_i, use that Unique-**SAT** is in **P** algorithm to try generate a satisfying assignment for ϕ_i. This can be done using the selfreducibility of **SAT** (See [BDG88] for details). If a satisfying assignment has been found accept ϕ and if for every i no assignment was found reject. The fact that the polynomial hierarchy collapses follows since it is known that $\mathbf{R} \in \mathbf{P}/poly$ and $\mathbf{NP} \in \mathbf{P}/poly$ implies that the polynomial hierarchy collapses [KL80]. □

Another consequence of Unique-$\mathbf{SAT} \in \mathbf{P}$ is that **FewP**, the class of languages that are accepted by nondeterministic polynomial time Turing machines that have at most a polynomial number of accepting paths, is in **P**. This was essentially proved in Toda's paper [Tod91a].

Theorem 6.3 (Toda) *If Unique-*$\mathbf{SAT} \in \mathbf{P}$ *then* $\mathbf{FewP} = \mathbf{P}$

Fortnow and Kummer [FK96] showed that the assumption that Unique-**SAT** is in **P** is linked to resource bounded Kolmogorov complexity:

Theorem 6.4 (Fortnow-Kummer) *Unique-*SAT *is in* P *if and only if*

$$\mathbf{CD}^{poly}(x \mid y) \leq \mathbf{C}^{poly}(x \mid y) + O(\log |x|))$$

We mentioned before that all the six hypotheses imply that $\mathbf{NP} \in \mathbf{P}/poly$. This is equivalent to $\mathbf{SAT} \leq_{tt}^{p} \mathbf{SPARSE}$. Ogihara and Watanabe [OW91] showed that if $\mathbf{SAT} \leq_{btt}^{p} \mathbf{SPARSE}$ then $\mathbf{P} = \mathbf{NP}$. With a slightly weaker hypothesis Cai, Naik, and Sivakumar [CNS96] proved the following:

Theorem 6.5 (Cai-Naik-Sivakumar) *If* $\mathbf{SAT} \leq_{dtt}^{p} \mathbf{SPARSE}$ *then Unique-*$\mathbf{SAT} \in \mathbf{P}$.

7 Relativization

To understand the difficulty of proving results about the six hypotheses, it is useful to turn to the theory of relativization. All of the results in this paper relativize, i.e., hold if every machine has access to the same oracle. See Fortnow [For94] for a discussion of the importance and limitations of relativization results.

In order to relativize some of the questions related to the six hypotheses we need a relativized version of **SAT** developed by Goldsmith and Joseph [GJ93]. Relativized $\mathbf{SAT}^A$ has several extra predicates $A_0, A_1, \ldots$ such that $A_m(x_1, \ldots, x_m)$ has the property that

$$x_1 \ldots x_m \in A \Leftrightarrow A_m(x_1, \ldots, x_m)$$

For every oracle A, $\mathbf{SAT}^A$ has the following properties:

1. $\mathbf{SAT}^A$ is $\mathbf{NP}^A$ complete.

2. Whether ϕ is in $\mathbf{SAT}^A$ depends only on strings in A of length less than $|\phi|$.

Baker, Gill and Solovay [BGS75] in their seminal paper on relativization give an oracle A such that $\mathbf{P}^A = \mathbf{NP}^A$. Relative to this oracle all of the six hypotheses are true.

All of the six hypotheses imply that **NP** has polynomial-size circuits (Theorems 2.3 and 6.2) and thus $\Pi_2^p = \Sigma_2^p$. Baker and Selman [BS79] give a relativized world where $\Pi_2^p \neq \Sigma_2^p$ and thus all of the six hypotheses are false. The six hypotheses also fail relative to generic and random oracles.

Creating relativized worlds where some of the six hypotheses are true while others fail appears considerably more difficult. Recently Beigel, Buhrman, and Fortnow [BBF97] have made some progress in this direction.

Theorem 7.1 (Beigel-Buhrman-Fortnow) *There exists an oracle A such that*

$$\mathbf{P}^A = \oplus\mathbf{P}^A \neq \mathbf{NP}^A = \mathbf{EXP}^A$$

One can use $\oplus\mathbf{P}$ to solve Unique-**SAT** questions. Toda [Tod91b] uses this fact in his celebrated proof that $\mathbf{PH} \subseteq \mathbf{P}^{\#\mathbf{P}}$. Combined with Corollary 5.7, this gives us the following conclusion.

Corollary 7.2 (Beigel-Buhrman-Fortnow) *There exists a relativized world where Unique-***SAT** *is in* **P** *and* $\mathbf{FP}_{||}^{\mathbf{NP}} \neq \mathbf{FP}^{\mathbf{NP}[\log]}$.

Other relativized separations of the six hypotheses remain important open problems.

We can get some more relativized separations if we weaken some of the hypotheses.

Theorem 7.3 *Let $f(n) = \omega(\log n)$. There exists a relativized world where* **SAT** *is $f(n)$-approximable but* $\mathbf{P} \neq \mathbf{NP}$.

The proof uses ideas from Homer and Longpré [HL94].

Proof: First start with an oracle that makes $\mathbf{P} = \mathbf{PSPACE}$. We build a new oracle on top of this one.

Define the language $L(A) =$

$$\{1^n \mid \text{There exists a string } x \text{ of length } n \text{ in } A.\}$$

For all A we have $L(A) \in \mathbf{NP}^A$.

We diagonalize **P** from **NP** in the same way as Baker, Gill and Solovay [BGS75]. However we will guarantee that we put at most one string in at every length and the string we put in will be among the first $2^{f(n)} \Leftrightarrow 2$ strings of length n. Since $2^{f(n)} \Leftrightarrow 2$ is greater than every polynomial, we are able to use the Baker, Gill and Solovay diagonalization technique.

Suppose we are given $f(n)$ formulae. Note there are only $2^{f(n)} \Leftrightarrow 1$ possibilities for the oracle strings of length n (the oracle could be empty). Using our $\mathbf{P} = \mathbf{PSPACE}$ base oracle we can compute some possible setting of the $f(n)$ formulae that cannot occur. □

Generalizing these techniques we get additional relativized worlds.

Theorem 7.4 *Let $f(n) = \omega(\log n)$. There exists a relativized world where*

1. $\mathbf{P} \neq \mathbf{NP}$.

2. **SAT** *is $f(n)$-Turing reducible to a* **P***-selective set and thus a k approximable set for $k \geq 2$.*

3. $\mathbf{FP}_{||}^{\mathbf{NP}} \subseteq \mathbf{FP}^{\mathbf{NP}[f(n)]}$.

8 Open Problems

In this section we summarize the open problems. For most of these problems it is not even known whether there are relativized worlds where they fail, so relativized results are welcome too.

The main open problems are the following.

1. Show that any two of the six hypotheses are equivalent to each other.

2. Show that **SAT** is $O(\log n)$-approximable implies Unique-**SAT** is in **P** or vice versa.

3. Show that if $\mathbf{FP}_{\omega(\log(n))-tt}^{\mathbf{NP}} \subseteq \mathbf{FP}^{\mathbf{NP}[\log]}$ then $\mathbf{P} = \mathbf{NP}$. This is the stronger version of the hypothesis $\mathbf{FP}_{||}^{\mathbf{NP}} = \mathbf{FP}^{\mathbf{NP}[\log]}$.

4. Show that if there is a Σ_2^p -complete set that is $O(\log n)$-approximable then $\mathbf{P} = \mathbf{NP}$. (Similar for **PSPACE**).

5. (related to Section 6) Is $\Sigma_2^p = \mathbf{UP}^{\mathbf{NP}}$.

6. Show that if $\mathbf{SAT} \leq_{dtt}^p \mathbf{SPARSE}$ then $\mathbf{P} = \mathbf{NP}$.

Acknowledgments

We like to thank Richard Beigel for proof reading an earlier version of our paper. We thank Dieter van Melkebeek for helpful discussions and Sophie Laplante for useful legwork.

References

[AA96] M. Agrawal and V. Arvind. Quasi-linear truth-table reductions to p-selective sets. *Theoretical Computer Science*, 158:361–370, 1996.

[ABG90] A. Amir, R. Beigel, and W. I. Gasarch. Some connections between bounded query classes and nonuniform complexity. In *Proceedings of the 5th Annual Conference on Structure in Complexity Theory*, pages 232–243, 1990.

[BBF97] R. Beigel, H. Buhrman, and L. Fortnow, 1997. Manuscript.

[BCG+96] N. H. Bshouty, R. Cleve, R. Gavaldà, S. Kannan, and Ch. Tamon. Oracles and queries that are sufficient for exact learning. *J. Computer and System Sciences*, 52(3):421–433, June 1996.

[BDG88] J. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity I.* Springer-Verlag, 1988.

[BDG90] J. Balcázar, J. Díaz, and J. Gabarró. *Structural Complexity II.* Springer-Verlag, 1990.

[Bei87a] R. Beigel. *Query-limited reducibilities.* PhD thesis, Stanford University, 1987.

[Bei87b] R. Beigel. A structural theorem that depends quantitatively on the complexity of SAT. In *Proceedings of the 2nd conference on Structure in Complexity Theory*, pages 28–32. IEEE Computer Science Press, 1987.

[Bei88] R. Beigel. NP-hard sets are P-superterse unless R = NP. Technical Report TR 88-4, Johns Hopkins University, 1988.

[BF97] H. Buhrman and L. Fortnow. Resource bounded kolmogorov complexity revisited. In Reischuk and Morvan, editors, *14th Annual Symposium on Theoretcal Computer Science*, volume 1200 of *Lecture Notes in Computer Scienc*, pages 105–116. Springer, 1997.

[BGGO93] R. Beigel, W. Gasarch, J. Gill, and J. Owings Jr. Terse, superterse and verbose sets. *Information and Computation*, 103:68–85, 1993.

[BGS75] T. Baker, J. Gill, and R. Solovay. Relativizations of the P =? NP question. *SIAM J. Comput.*, 4(4):431–441, Dec. 1975.

[BH77] L. Berman and H. Hartmanis. On isomorphisms and density of NP and other complete sets. *SIAM J. Comput.*, 6:305–322, 1977.

[BH91] S.R. Buss and L. Hay. On truth-table reducibility to SAT. *Information and Computation*, 90(2):86–102, February 1991.

[BKS] R. Beigel, M. Kummer, and F. Stephan. Quantifying the ammount of verboseness. *Information and Computation.* to appear.

[BKS95] R. Beigel, M. Kummer, and F. Stephan. Approximable sets. *Information and Computation*, 120(2):304–314, 1995.

[BS79] Baker and Selman. A second step toward the polynomial hierarchy. *Theoretical Computer Science*, 8, 1979.

[BT96a] H. Buhrman and T. Thierauf. The complexity of generating and checking proofs of membership. In C. Puech and R. Reischuk, editors, *13th Annual Symposium on Theoretical Aspects of Computer Science*, volume 1046 of *Lecture Notes in Computer Science*, pages 75–86. Springer, 1996.

[BT96b] H. Buhrman and L. Torenvliet. P-selective self-reducible sets: A new characterization of P. *J. Computer and System Sciences*, 53(2):210–217, 1996.

[CNS96] J. Cai, V. Naik, and D. Sivakumar. On the existence of hard sparse sets under weak reductions. In C. Puech and R. Reischuk, editors, *13th annual symposium on theoretical aspects of computer science*, volume 1046 of *Lecture Notes in Computer Science*, pages 307–318. Springer, 1996.

[FK96] L. Fortnow and M. Kummer. Resource-bounded instance complexity. *Theoretical Computer Science A*, 161:123–140, 1996.

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

[GHJY91] J. Goldsmith, L. Hemachandra, D. Joseph, and P. Young. Near-testable sets. *SIAM J. Comput.*, 20(3):506–523, 1991.

[GJ79] M. Garey and D. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, San Francisco, 1979.

[GJ93] J. Goldsmith and D. Joseph. Relativized isomorphisms of NP-complete sets. *Computational Complexity*, 3:186–205, 1993.

[HHN+95] L. Hemaspaandra, A. Hoene, A. Naik, M. Ogiwara, A. Selman, T. Thierauf, and J. Wang. Nondeterministically selective sets. *International Journal of Foundations of Computer Science*, 6(4):403–416, 1995.

[HL94] S. Homer and L. Longpré. On reductions of NP sets to sparse sets. *J. Computer and System Sciences*, 48:324–336, 1994.

[HN93] A. Hoene and A. Nickelsen. Counting, selecting, and sorting by query-bounded machines. In *STACS 93*, volume 665 of *Lecture Notes in Computer Science*, pages 196–205, 1993.

[HU79] J. Hopcroft and J. Ullman. *Introduction to Automata Theory, Languages, and Computation*. Addison-Wesley, Reading, Massachusetts, 1979.

[JT95] Jenner and Toran. Computing functions with parallel queries to NP. *Theoretical Computer Science*, 141, 1995.

[JT97] B. Jernner and J. Toran. *Complexity Theory Retrospective II*, chapter The Complexity of Obtaining Solutions for Problems in NP and NL. Springer-Verlag, 1997. to appear.

[KF80] C. Kintala and P. Fischer. Refining nondeterminism in relativized poslynomial-time bounded computations. *SIAM J. Comput.*, 9(1):46–53, 1980.

[KL80] R. Karp and R. Lipton. Some connections between nonuniform and uniform complexity classes. In *Proc. 12th ACM Symposium on Theory of Computing*, pages 302–309, 1980.

[Ko83] K.-I. Ko. On self-reducibility and weak P-selectivity. *J. Comput. System Sci.*, 26:209–211, 1983.

[KS91] M. Kummer and F. Stephan. Some aspects of frequency computation. Technical Report 21/91, Fakultät für Informatik, Universität Karlsruhe, Karlsruhe, 1991.

[KW95] Köbler and Watanabe. New collapse consequences of NP having small circuits. In *Annual International Colloquium on Automata, Languages and Programming*, 1995.

[Mah82] S. Mahaney. Sparse complete sets for NP: solution of a conjecture of Berman and Hartmanis. *J. Comput. System Sci.*, 25:130–143, 1982.

[Ogi95] M. Ogihara. Polynomial-time membership comparable sets. *SIAM Journal on Computing*, 24(5):1168–1181, 1995.

[OW91] M. Ogiwara and O. Watanabe. On polynomial time bounded truth-table reducibility of NP sets to sparse sets. *SIAM J. Comput.*, 20:471–483, 1991.

[Sau72] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory (A)*, 13:145–147, 1972.

[Sel79] A. Selman. P-selective sets, tally languages, and the behavior of polynomial time reducibilities on NP. *Math. Systems Theory*, 13:55–65, 1979.

[Sel82] A. Selman. Analogues of semicursive sets and effective reducibilities to the study of NP complexity. *Information and Control*, 52(1):36–51, January 1982.

[Sel96] A. Selman. Much ado about functions. In *Proceedings of 11th anual conference on Computational Complexity*, pages 198–212. IEEE Computer Society Press, 1996.

[She72] S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific Journal of Mathematics*, 41:241–261, 1972.

[Tod91a] S. Toda. On polynomial-time truth-table reducibility of intractable sets to p-selective sets. *Math. Systems Theory*, 24:69–82, 1991.

[Tod91b] S. Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991.

[VC71] V.N. Vapnik and A.Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory of Probability and its Applications*, 16(2):264–280, 1971.

[VV86] L. Valiant and V. Vazirani. NP is as easy as detecting unique solutions. *Theoretical Computer Science*, 47:85–93, 1986.

[Wag90] K. Wagner. Bounded query classes. *SIAM Journal on Computing*, 19(5):833–846, 1990.

[WT93] O. Watanabe and S. Toda. Structural analysis on the complexity of inverse functions. *Mathematical Systems Theory*, 26:203–214, 1993.