Quantifying the Amount of Verboseness *

Richard Beigel† Martin Kummer‡ Frank Stephan§
Yale University Universität Karlsruhe Universität Karlsruhe

Abstract

We study the fine structure of the classification of sets of natural numbers A according to the number of queries which are needed to compute the n-fold characteristic function of A. A complete characterization is obtained, relating the question to finite combinatorics. In order to obtain an explicit description we consider several interesting combinatorial problems.

1 Introduction

In the theory of bounded queries, we measure the complexity of a function by the number of queries to an oracle which are needed to compute it. The field has developed in various directions, both in complexity theory and in recursion theory; see Gasarch [21] for a recent survey. One of the original concerns is the classification of sets A of natural numbers by their "query complexity," i.e., according to the number of oracle queries that are needed to compute the n-fold characteristic function $F_n^A = \lambda x_1, \ldots, x_n \,.\, (\chi_A(x_1), \ldots, \chi_A(x_n))$. In [3, 8] a set A is called *verbose* iff F_n^A is computable by an oracle Turing machine that makes at most $\lceil \log_2(n+1) \rceil$ queries to A, for all $n > 0$. For example, the halting problem K is verbose [9, 8] and all semirecursive sets are verbose [8]. On the other hand, if A is a Kolmogorov-random set, then computing F_n^A requires n queries, regardless of the oracle; such sets are called *superterse*.

We will investigate the number of oracle queries needed in order to compute F_n^A. At one extreme we have the recursive sets A, for which F_n^A requires 0 oracle queries; then the verbose sets, for which F_n^A requires $\lceil \log_2(n+1) \rceil$ queries; and at the other extreme the superterse sets, for which F_n^A requires n queries. By the *Nonspeedup Theorem* [8] at least $\lceil \log_2(n+1) \rceil$ queries are needed for any nonrecursive set A, regardless of the oracle;

*An extended abstract of a preliminary version of this paper appeared in "Logical Foundations of Computer Science — Tver '92," Lecture Notes in Computer Science 620, Springer-Verlag, pp. 21–32, 1992

†Dept. of Computer Science, 51 Prospect St., P.O. Box 2158, Yale Station, New Haven, CT 06520-2158, (email: `beigel@cs.yale.edu`). Supported in part by National Science Foundation grant CCR-8958528.

‡Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe, D-76128 Karlsruhe, Germany, (email: `kummer@ira.uka.de`).

§Institut für Logik, Komplexität und Deduktionssysteme, Universität Karlsruhe, D-76128 Karlsruhe, Germany, (email: `fstephan@ira.uka.de`). Supported by the Deutsche Forschungsgemeinschaft (DFG) grant Me 672/4-1

i.e., there is nothing between the recursive sets and the verbose sets. We will investigate what lies between the verbose sets and the superterse sets.

Along these lines, we call a set A (m,n)-*verbose*[1] iff there exists a recursive function f such that $|W_{f(x_1,\ldots,x_n)}| \leq m$ and $\langle \chi_A(x_1),\ldots,\chi_A(x_n)\rangle \in W_{f(x_1,\ldots,x_n)}$ for all $x_1,\ldots,x_n$. (In this context A is called (m,n)-verbose *via* f.) Here W_i denotes the ith r.e. set in the standard enumeration of all r.e. sets. Let $V(m,n)$ denote the class of all (m,n)-verbose sets. Intuitively, $A \in V(m,n)$ means that we can effectively enumerate for any $x_1,\ldots,x_n$ a set of at most m possibilities for F_n^A such that one of them is correct. We classify a set A according to the least m such that $A \in V(m,n)$.

Clearly, every set is $(2^n,n)$-verbose, every $(1,n)$-verbose set is recursive, and every (m,n)-verbose set is (m',n)-verbose, for all $m' \geq m$. It is easy to see that every r.e. set is $(n+1,n)$-verbose [9, 8]. By a restatement of the Nonspeedup Theorem, every (n,n)-verbose set is recursive. Thus $V(n,n) = V(1,n)$ and all r.e. nonrecursive sets are in $V(n+1,n) - V(n,n)$. If A is superterse then A is $(2^n,n)$-verbose (trivially) but not $(2^n-1,n)$-verbose for any $n \geq 1$ [8]. The connection with bounded queries is given by the fact that $A \in V(2^k,n)$ iff F_n^A is computable by k queries to some oracle [8]. Hence the classification of A according to the least m such that $A \in V(m,n)$ is a refinement of the query complexity of A. For example, if A is verbose then A is $(2^k, 2^k-1)$-verbose for all k (in Corollary 3.2(f) we show that A is in fact $(n+1,n)$-verbose for all n).

We study the *inclusion problem* for verboseness classes; i.e., we ask "For which m,n,m',n' is $V(m,n)$ contained in $V(m',n')$?" Note that this subsumes all questions of the following type: "Is there a set A such that F_n^A is computable by making k queries (to a suitable oracle) but $F_{n'}^A$ is not computable by making k' queries (to any oracle)?" (This question is translated as: "Is $V(2^k,n) \not\subseteq V(2^{k'},n')$?") Our main theorem reduces the inclusion problem to a decidable problem in finite combinatorics; in this sense, the inclusion problem is completely solved.

For the special case of the *equality problem*, i.e., the question "For which m,n,m',n', do $V(m,n)$ and $V(m',n')$ coincide?" we present an explicit solution: for $m < 2n$ the answer depends only on the quotient m/n and we obtain many nontrivial equalities. For $2n \leq m < 2^n$ the behavior changes radically, as there are only the trivial equalities, $V(m,n) = V(m',n')$ iff $(m,n) = (m',n')$. This implies that for all k,n with $\lceil \log_2(n+1) \rceil \leq k \leq n$ there is a set A such that F_n^A can be computed by k queries but not by $k-1$.

We establish an explicit connection between recursion theory and extremal combinatorics. This is illustrated by our answer to the following question (for given k,n,m with $k < n$): "Is there a set A such that $F_{2^k}^A$ can be computed with $k+1$ queries but $F_{2^n}^A$ cannot be computed with m queries?" The answer is yes iff $c(k,n) + 2^n + 1 > 2^m$, where $c(k,n)$ is the largest number of edges in any graph on 2^n vertices that does not contain a cycle of length 2^k or less. An explicit description of $c(k,n)$ is known only for some special cases. This suggests that the inclusion relation on verboseness classes is highly nontrivial, and there is not much hope for a solution in "closed form."

Along the way we obtain new combinatorial results that may be of independent interest (e.g., in connection with the Vapnik-Chervonenkis dimension of concept classes). The necessary combinatorial background is provided in Section 2.

We also consider a polynomial time version of (m,n)-verboseness, classify the complete

[1] This is just a reformulation of the definition "F_n^A is computable by a set of m partial recursive functions" that appears in [3, 8].

sets at each level in Ershov's difference hierarchy, and explore the relationship between verboseness and frequency computation.

1.1 Notation, Definitions, and Basic Theorems

We write $\subseteq$ to denote subset, and $\subset$ to denote proper subset.

Our recursion theoretic notation closely follows the books of Odifreddi [35] and Soare [40]. ω denotes $\{0, 1, 2, \ldots\}$, the set of natural numbers. $\langle \cdot, \cdots, \cdot \rangle$ denotes a recursive tupling function from $\omega^{<\omega}$ to ω (however, when the meaning is clear we will identify the number $\langle x_1, \ldots, x_n \rangle$ with the n-tuple $(x_1, ..., x_n)$). φ_i is the ith function in a standard enumeration of all partial recursive functions of one argument. $\varphi_{i,s}(x)$ denotes the result, if any, of performing s steps in the computation of $\varphi_i(x)$. R_n denotes the set of all total recursive n-place functions. D_i is the ith finite set in the canonical enumeration of all finite sets. $A \oplus B$ denotes $\{2x : x \in A\} \cup \{2x + 1 : x \in B\}$, the *join* of A and B. A set A is called k-r.e. ($k > 0$) iff there exists $h \in \mathrm{R}_2$ such that $\chi_A(x) = \lim_{s \to \infty} h(x, s)$, $|\{s : h(x, s) \neq h(x, s + 1)\}| \leq k$, and $h(x, 0) = 0$ for all x. A set $A \subseteq \omega$ is called *semirecursive* iff A is an initial segment of a recursive linear ordering of ω (cf. [35, p. 295]).

Fact 1.1 ([8]) *A is $(2^k, n)$-verbose iff there is an oracle X such that F_n^A is computable by making at most k queries to X.*

Fact 1.2 (Nonspeedup Theorem [8]) *If A is (n, n)-verbose for some $n \geq 1$ then A is recursive. In particular, if there are an oracle X and an integer $n \geq 1$ such that $F_{2^n}^A$ is computable by making at most n queries to X then A is recursive.*

Definition 1.3 A set $A \subseteq \omega$ is called *strongly (m, n)-verbose* iff there exists a total recursive function f such that $|D_{f(x_1,\ldots,x_n)}| \leq m$ and $(\chi_A(x_1), \ldots, \chi_A(x_n)) \in D_{f(x_1,\ldots,x_n)}$ for all $x_1, \ldots, x_n$. In this context A is called strongly (m, n)-verbose via f. $V_S(m, n)$ denotes the class of strongly (m, n)-verbose sets.

Fact 1.4 ([8]) *A is strongly $(2^k, n)$-verbose iff there exist an oracle X and an oracle Turing machine $M^{(-)}$ that computes a total function for every oracle, such that M^X computes F_n^A by making at most k queries to X.*

2 Combinatorial Tools

In this section we present the combinatorial background that is relevant for an explicit description of inclusion relations among verboseness classes. Here is the intuition behind our combinatorial approach: Before we know anything about a set A, there are 2^n possible values for the n-bit vector $F_n^A(x_1, \ldots, x_n)$. If A is $(r - 1, s)$-verbose, then there are only $r - 1$ possible values for the s-bit vector $F_s^A(x_1, \ldots, x_s)$. Therefore, there are only $r - 1$ possible values for the s-bit vector obtained by projecting $F_n^A(x_1, \ldots, x_n)$ onto s particular components. If we can show that every set of m n-bit vectors has at least r projections onto some s components, then we can conclude that there are only $m-1$ possibilities for the n-bit vector $F_n(x_1,\ldots,x_n)$, and so A is $(m-1,n)$-verbose.

This approach was previously used successfully in the special cases $r=2^s$ and $r=s+1$ [4, 7], where the combinatorics are not too hard. We obtain new bounds by exploiting the extremal theory of finite sets, which has been well studied (see [13]).

Definition 2.1 Let $1 \le k \le n$, $v \in \{0,1\}^n$, and $V \subseteq \{0,1\}^n$.

- If $x_1 < \cdots < x_k$, then $v[x_1,\ldots,x_k]$ denotes the projection of v onto the coordinates $x_1,\ldots,x_k$.
- If $I=\{x_1,\ldots,x_k\}$ where $x_1<\cdots<x_k$, then $P^I(V)$ denotes the projection of the set V onto the coordinates $x_1,\ldots,x_k$, i.e.,
$$P^I(V)=\{v[x_1,\ldots,x_k] : v \in V\}.$$
- $g_V(k)$ denotes the maximum cardinality of any projection of V onto k coordinates; i.e.,
$$g_V(k)=\max\{|P^I(V)| : I \subseteq \{1,\ldots,n\} \wedge |I|=k\}.$$
Note that $g_V(n)=|V|$. By convention, $g_V(0)=1$ and $g_V(m)=|V|$, for $m>n$.

Frankl's arrow relation [20] is very useful when we know $g_V(s)$ and want to bound $|V|$.

Notation 2.2 Let m,n,r,s be positive integers with $n \ge s$.
$$(m,n)\to(r,s) \iff (\forall V \subseteq \{0,1\}^n)[|V|\ge m \Rightarrow g_V(s)\ge r].$$

To restate, $(m,n)\to(r,s)$ means that for every $V\subseteq\{0,1\}^n$ if $g_V(s)\le r-1$ then $|V|\le m-1$. Intuitively, if we have a bound on the number of tuples you can get by projecting V onto s coordinates, then we obtain a bound on $|V|$. The arrow relation will enable us to prove nontrivial upward inclusions among verboseness classes. (Exclusively for the sake of intuition, we call an inclusion $V(r,s)\subseteq V(m,n)$ *upward* if $s\le n$ and *downward* if $s>n$.)

Fact 2.3 *Let m,n,r,s be positive integers with $n\ge s$. If $(m,n)\to(r,s)$ then $V(r-1,s)\subseteq V(m-1,n)$ and $V_S(r-1,s)\subseteq V_S(m-1,n)$.*

Proof: Assume that A is $(r-1,s)$-verbose via $f\in \mathrm{R}_s$ and that $(m,n)\to(r,s)$. On input $(x_1,\ldots,x_n)$ it is easy produce an r.e. set that consists of all n-tuples that are consistent with f. To be precise, there exists a total recursive function g such that for all $x_1,\ldots,x_n$
$$\begin{aligned} W_{g(x_1,\ldots,x_n)} &= \{(a_1,\ldots,a_n)\in\{0,1\}^n : (\forall i_1,\ldots,i_s) \\ &\qquad [1\le i_1<\cdots<i_s\le n \Rightarrow (a_{i_1},\ldots,a_{i_s})\in W_{f(x_{i_1},\ldots,x_{i_s})}]\}.\end{aligned}$$

Clearly $(\chi_A(x_1),\ldots,\chi_A(x_n))\in W_{g(x_1,\ldots,x_n)}$. Let $W=W_{g(x_1,\ldots,x_n)}$. Because A is $(r-1,s)$-verbose via f, $g_W(s)\le r-1$. Because $(m,n)\to(r,s)$ it follows that $|W|<m$. Hence $A\in V(m-1,n)$ via g. The proof for the strongly verbose case is analogous (replace W by D everywhere). ∎

Does every upward inclusion correspond to an arrow relation? This is not the case, e.g., $V(5,5) \subseteq V(1,5)$ by the Nonspeedup Theorem, but of course $(2,5) \not\rightarrow (6,5)$. In many cases however the upward inclusion given by an arrow relation is optimal (see, e.g., Proposition 4.5(b) and Theorem 6.4).

The following arrow relations are well-known (and all of them are best possible):

Fact 2.4 (cf. [13, Section 17]) *Let m, n, k be positive integers.*

(a) $(m,n) \rightarrow (m,n-1)$ *if $m \leq n$. (Bondy)*

(b) $(m,n) \rightarrow (m-1,n-1)$ *if $m \leq \lfloor \frac{3}{2}n \rfloor$. (Bollobás)*

(c) $(m,n) \rightarrow (2^k,k)$ *if $m > \sum_{i=0}^{k-1}\binom{n}{i}$. (Sauer, Perles–Shelah, Vapnik–Chervonenkis)*

(d) $(m,n) \rightarrow (7,3)$ *if $m > \lfloor \frac{n^2}{4} \rfloor + n + 1$. (Alon, Frankl)*

Fact 2.4(a) has been restated in a more attractive form as Owings's "Separation Lemma" [36, p. 764]. Fact 2.4(c) is particularly interesting and has been rediscovered several times [4, 16]. It has important applications in probability theory [42], computational learning theory [10], and computational geometry [22]. For later use, we define

$$S(n,k) = \sum_{i=0}^{k-1}\binom{n}{i},$$

and we explicitly state the inclusion relation provided by Fact 2.4(c):

Fact 2.5 ([3]) *$V(2^k-1,k) \subseteq V(S(n,k),n)$ and $V_S(2^k-1,k) \subseteq V_S(S(n,k),n)$, for all $n,k > 0$.*

Definition

- $\preceq$ denotes the componentwise partial-order on $\{0,1\}^n$, i.e., $(a_1,\ldots,a_n) \preceq (b_1,\ldots,b_n)$ iff $a_i \leq b_i$ for $i = 1,\ldots,n$.
- We say that V is *closed* iff V is closed downwards under $\preceq$, i.e., iff $(v \preceq w \wedge w \in V) \Rightarrow v \in V$.

(Note that 0^n belongs to every closed closed subset of $\{0,1\}^n$.)

As pointed out in [1, 20], Fact 2.4 is a corollary of the following important result which was discovered independently by Alon [1] and Frankl [20]:

Fact 2.6 *For every $V \subseteq \{0,1\}^n$ there is a closed set $W \subseteq \{0,1\}^n$ such that $|W| = |V|$ and*

$$g_W(i) \leq g_V(i), \text{ for all } i \geq 0.$$

The point is that Fact 2.6 greatly simplifies the proof of an arrow relation $(m,n) \rightarrow (r,s)$: It suffices to show that for every *closed* $V \subseteq \{0,1\}^n$ with $|V| \geq m$ we have $g_V(s) \geq r$. By this method, we now prove two new arrow relations that are needed in Section 6.

With each $W \subseteq \{0,1\}^n$, let us associate the undirected graph $G(W) = (V, E)$ defined by

$$\begin{aligned} V &= \{i : 0^{i-1}10^{n-i} \in W\}, \\ E &= \{\{i,j\} : i < j \ \wedge \ 0^{i-1}10^{j-i-1}10^{n-j} \in W\}. \end{aligned}$$

(For intuition, we identify a set $W \subseteq \{0,1\}^n$ with the $|W| \times n$ matrix whose rows are the elements of W. If W is closed then the vertices of $G(W)$ can be identified with the columns that are not equal to $0^{|W|}$; two vertices i, j are connected by an edge iff there is a row that has a 1 in column i and a 1 in column j.)

Conversely, with each graph $G = (V, E)$ such that $V \subseteq \{1, \ldots, n\}$, we associate the set $W(G) \subseteq \{0,1\}^n$ defined by

$$W(G) = \{0^{i-1}10^{n-i} : i \in V\} \ \cup \ \{0^{i-1}10^{j-i-1}10^{n-j} : i < j \ \wedge \ \{i,j\} \in E\} \ \cup \ \{0^n\}.$$

Note that $W(G)$ is closed and $|W(G)| = |V| + |E| + 1$. If W is closed and nonempty then

$$W(G(W)) = \{v \in W : v \text{ contains at most two 1s}\} \subseteq W.$$

We say that I and J *partition* a set S if $I \cup J = S$ and $I \cap J = \emptyset$. We say that I and J *partition a graph* (V, E) if I and J partition E as sets and there is no edge $\{i,j\} \in E$ with $i \in I$ and $j \in J$.

Lemma 2.7 *Assume that $W \subseteq \{0,1\}^n$ is closed, and let $G(W) = (V, E)$.*

(a) $W = W(G(W))$ *iff* W *is nonempty and* $g_W(3) \leq 7$.

(b) If I, J *partition the graph* $G(W)$*, then* $|P^I(W)| + |P^J(W)| = |W| + 1$. *(Recall that* $P^I(V) = \{v[x_1, \ldots, x_k] : v \in V\}$*.)*

(c) If $g_W(k) = g_W(k+1)$ *for some* $1 \leq k < n$*, then* $|W| = g_W(k)$.

(d) If $G(W)$ *is connected, then* $g_W(k+1) \geq g_W(k) + 2$ *for all* $1 \leq k < |V|$.

(e) If $g_W(k) \leq 2k - 1$ *for some* $1 \leq k < n$*, then* $g_W(k+1) \leq g_W(k) + 2$ *and* $|W| \leq g_W(k) + 2(n-k)$.

(f) $(2n, n) \to (2k, k)$ *if* $1 \leq k \leq n$.

Proof:

(a) Because W is closed, $g_W(3) \leq 7$ iff every $v \in W$ contains at most two 1s.

(b) If I, J partition $G(W)$ then W does not contain a vector v with $v[i] = v[j] = 1$, for any $i \in I$, $j \in J$. Let

$$\begin{aligned} X &= \{v \in W : v[I] \neq 0^{|I|}\}, \\ Y &= \{v \in W : v[J] \neq 0^{|J|}\}. \end{aligned}$$

Then X, Y partition $W - \{0^n\}$ as sets, so

$$|W| - 1 = |X| + |Y| = (|P^I(W)| - 1) + (|P^J(W)| - 1).$$

(c) Assume that $g_W(k) = g_W(k+1)$. Choose a k-element set I such that $|P^I(W)|$ is maximum and (without loss of generality) assume that $I = \{1, \ldots, k\}$. Then $|P^I(W)| = g_W(k) = g_W(k+1)$. Because $g_W(k) = g_W(k+1)$, it follows that $P^{I\cup\{j\}}(W)$ is equal to $P^I(W) \times \{0\}$ or $P^I(W) \times \{1\}$ for $j = k+1, \ldots, n$; because W is closed, $P^{I\cup\{j\}}(W) = P^I(W) \times \{0\}$. Hence, $W = P^I(W) \times \{0^{n-k}\}$; therefore, $|W| = g_W(k)$.

(d) Assume that $G(W)$ is connected. For fixed $k < |V|$ choose a k-element set I such that $|P^I(W)|$ is maximum. Then $I \subseteq V$ (because V contains all the nonconstant columns). Without loss of generality, we may assume that $I = \{1, \ldots, k\}$. Because $G(W)$ is connected, there are vertices $i \in I$ and $j \in V - I$ such that $\{i, j\} \in E$. Because W is closed,

$$P^{I\cup\{j\}}(W) \supseteq (P^I(W) \times \{0\}) \cup \{0^k 1, 0^{i-1} 1 0^{k-i} 1\};$$

therefore

$$g_W(k+1) \geq |P^{I\cup\{j\}}(W)| \geq g_W(k) + 2.$$

(e) Assume that $g_W(k) \leq 2k - 1$, for some $k < n$. Choose a $(k+1)$-element set I such that $|P^I(W)| = g_W(k+1)$. Let $U = P^I(W) \subseteq \{0,1\}^{k+1}$. If U has an all-zero column then $g_W(k+1) = g_W(k)$ and we are done. Otherwise $G(U) = (V, E)$ has $k+1$ vertices and we may assume that $k+1$ is a vertex of minimal degree d. Thus, $|E| \geq \frac{1}{2}d(k+1)$ and there are at least $\frac{1}{2}d(k-1)$ edges that do not contain vertex $k+1$. Because U is closed we get

$$2k - 1 \geq g_W(k) \geq g_U(k) \geq |P^{\{1,\ldots,k\}}(U)| \geq 1 + k + \frac{1}{2}d(k-1),$$

so $d \leq 1$. It follows that

$$g_W(k+1) = |U| \leq |P^{\{1,\ldots,k\}}(U)| + 2 \leq g_W(k) + 2.$$

In particular, $g_W(k+1) \leq 2(k+1) - 1$. By iterating this argument, we get $|W| = g_W(n) \leq g_W(k) + 2(n-k)$.

(f) Given $W \subseteq \{0,1\}^n$ with $g_W(k) \leq 2k - 1$. Using Fact 2.6 we may assume that W is closed. By (e), we get $|W| \leq g_W(k) + 2(n-k) \leq 2n - 1$.

∎

The following proposition slightly improves Lemma 2.7(f) and yields an optimal bound.

Proposition 2.8 *For all* $n \geq k > 1$, $(m, n) \to (2k, k)$ *if* $m > 1 + \frac{n(2k-3)}{k-1}$.

Proof: The proof is by double induction on k, n. Suppose the claim is true for all k', n' with $(2 \leq k' < k \ \wedge \ k' \leq n')$ or $(k' = k \leq n' < n)$. From this inductive hypothesis we will prove the claim for k, n. Let $W \subseteq \{0,1\}^n$ be given such that $g_W(k) \leq 2k - 1$. Let $G(W) = (V, E)$. We need to show that $|W| \leq 1 + \frac{n(2k-3)}{k-1}$. By Fact 2.6 we may assume that W is closed. Suppose for a contradiction that

$$|W| > 1 + \frac{n(2k-3)}{k-1}. \tag{1}$$

We claim that

$$g_W(k-1) = 2(k-1). \tag{2}$$

This is obvious if $k=2$, because $|W|>1$. For $k>2$, we get from the inductive hypothesis that $g_W(k-1) \geq 2(k-1)$. If we had $g_W(k-1) > 2(k-1)$, then $g_W(k-1) = g_W(k)$, and, by Lemma 2.7(c), $|W| = 2k-1$. Simple algebra yields a contradiction to (1).

Without loss of generality, we may assume that $P^I(W) = g_W(k-1) = 2(k-1)$, where $I = \{1,\dots,k-1\}$. Let $J = \{k,\dots,n\}$. Because $g_W(k) < 2k$, Lemma 2.7(f) implies that $|W| \leq 2n-1$. Therefore, $2n-1 \geq |W| > 1 + \frac{n(2k-3)}{k-1}$. So, $n > 2k-2$ and, since $2n-1 \leq 1 + \frac{n(2k-3)}{k-1}$ for $n \leq 2k-2$, it follows that $n > 2k-2$. That is, $|J| \geq k$. Thus we can apply the inductive hypothesis with $k' = k$ and $n' = |J|$ to obtain

$$|P^J(W)| \leq 1 + |J|\frac{2k-3}{k-1}. \tag{3}$$

Furthermore, we assert that I and J partition $G(W)$. Why? If there is an edge $\{i,j\} \in E$ with $i \in I$ and $j \in J$ then

$$|P^{I\cup\{j\}}(W)| \geq |P^I(W)| + 2 = 2k > g_W(k),$$

a contradiction. By Lemma 2.7(b), we get

$$|W| = |P^I(W)| + |P^J(W)| - 1 = |P^J(W)| + 2k - 3. \tag{4}$$

By equations (3) and (4) it follows that

$$|W| \leq 1 + |J|\frac{2k-3}{k-1} + (2k-3) = 1 + \frac{n(2k-3)}{k-1},$$

contradicting equation (1). This completes the inductive step. ∎

Remark: The case $k=2$ is contained in Fact 2.4(c). The bounds in Proposition 2.8 are the best possible. Why? For any $n \geq k > 1$ define a graph G of order n as the disjoint union of $\lfloor \frac{n}{k-1} \rfloor$ stars of order $k-1$ and one star of order $(n \bmod (k-1))$. Then $|W(G)| = 1 + \lfloor \frac{n(2k-3)}{k-1} \rfloor$ and $g_{W(G)}(k) = 2k-1$.

The *girth* of a graph is the length of its shortest cycle (by convention the girth is infinite if the graph contains no cycle). Let $c_k(n)$ denote the maximum number of edges in any graph on n vertices with girth greater than k. The following asymptotic bounds for $c_k(n)$ were obtained by easy counting arguments [12, Theorem III.3.7]:

$$c_k(n) = O(n^{1+\frac{2}{k-1}}) \quad \text{and} \quad c_k(n) = \Omega(n^{1+\frac{1}{k+1}}).$$

Matching upper and lower bounds are known for $k = 3,4,5$: $c_3(n) = \lfloor \frac{n^2}{4} \rfloor$ (Turán's Theorem [12, p. 294]), $c_4(n) \asymp \frac{1}{2}n^{\frac{3}{2}}$ ([12, Theorem VI.2.8]), $c_5(n) = \Theta(n^{\frac{3}{2}})$ (this follows from the fact that $c_k(2n) \geq c_{k-1}(n)$ for k odd [12, Theorem III.1.3]). A better general lower bound $c_k(n) = \Omega(n^{1+\frac{4}{3(k+1)}})$ was proved in [32, 33], but the proof is rather complicated. For our main results we will need to know only that $c_k(n)$ has superlinear growth.

Observe that Fact 2.4(d) asserts that $(m,n) \to (7,3)$ if $m > c_3(n) + n + 1$. This is generalized by our next result.

Proposition 2.9 *For all $n \geq k > 2$, $(m,n) \to (2k+1,k)$ if $m > c_k(n)+n+1$.*

Proof: Let $2 < k \leq n$. By Fact 2.6 it suffices to show that for every closed set $W \subseteq \{0,1\}^n$ with $g_W(k) \leq 2k$, we have $|W| \leq c_k(n)+n+1$. We distinguish two cases.

Case 1: $g_W(i) \geq 2i+2$, for some $i < k$. Let i_0 be the least such i, and choose H such that $|H| = i_0$ and $|P^H(W)| = g_W(i_0) \geq 2i_0+2$. Let $U = P^H(W)$.

We assert that $G(U)$ is connected. Why? If H_0, H_1 partition $G(U)$ then, by Lemma 2.7(b), either $|P^{H_0}(W)| \geq 2|H_0|+2$ or $|P^{H_1}(W)| \geq 2|H_1|+2$. By the minimality of i_0, it follows that $G(U)$ cannot be partitioned in a nontrivial way.

Thus, $G(U)$ is connected and there is a maximal set $I \supseteq H$ such that $G(P^I(W))$ is connected. Let $J = \{1,\ldots,n\} - I$, $X = P^I(W)$, $Y = P^J(W)$. I, J partition $G(W)$. Since $G(X)$ is connected, Lemma 2.7(d) implies that $g_X(i) \geq 2i+2$ for $i = i_0,\ldots,|I|$. Hence, $|I| < k$. Let L be any $(k-|I|)$-element subset of J. By Lemma 2.7(b) (with $W' = P^{I\cup L}(W)$) we get $|P^{I\cup L}(W)|+1 = |P^I(W)|+|P^L(W)|$. It follows that $|X|+g_Y(k-|I|) \leq g_W(k)+1 \leq 2k+1$. Because $|X| \geq 2|I|+1$,

$$g_Y(k-|I|) \;\leq\; 2k+1-|X| \;\leq\; 2(k-|I|)-1.$$

By Lemma 2.7(e),

$$|Y| \;\leq\; g_Y(k-|I|)+2(|J|-(k-|I|)) \;\leq\; 2k+1-|X|+2(n-k) \;=\; 2n+1-|X|.$$

Because I, J partition $G(W)$, Lemma 2.7(b) implies that

$$|W| = |X|+|Y|-1 \leq 2n.$$

Obviously, $2n \leq c_k(n)+n+1$ and we are done.

Case 2: $g_W(i) \leq 2i+1$, for all $i < k$. In particular, we have $g_W(3) \leq 7$ (in the case $k=3$ we have $g_W(3) \leq 6$, by our initial hypothesis). Thus, $W(G(W)) = W$. We decompose $G(W)$ into disjoint subgraphs as follows: Let $(V_1,E_1),\ldots,(V_{h-1},E_{h-1})$ be all connected components of $G(W)$ that contain fewer than k vertices, and let (V_h,E_h) be the rest of the graph. Note that, for $i < h$, $|E_i| \leq |V_i|$, because $g_W(|V_i|) \leq 2|V_i|+1$. If (V_h,E_h) contains a cycle of length less than or equal k then it has a connected subgraph $G' = (V',E')$ of k vertices containing a cycle. Thus, $|E'| \geq k$ and $|P^{V'}(W)| = |V'|+|E'|+1 \geq 2k+1$, contradicting $g_W(k) \leq 2k$. Hence (V_h,E_h) has girth greater than k, and we get the required upper bound for $|W|$:

$$\begin{aligned}
|W| = |E|+|V|+1 &= 1+\sum_{i=1}^{h}(|E_i|+|V_i|)\\
&\leq 1+c_k(|V_h|)+|V_h|+\sum_{i=1}^{h-1}(|E_i|+|V_i|)\\
&\leq 1+c_k(|V_h|)+|V_h|+\sum_{i=1}^{h-1}2|V_i|\\
&\leq 1+n+c_k(|V_h|)+(n-|V_h|)\\
&\leq 1+n+c_k(n).
\end{aligned}$$

The last inequality follows from the fact that $c_k(m+1) \geq c_k(m)+1$, for all $m \geq 1$. ∎

Remark: Proposition 2.9 is best possible. Why? For any $n \geq k > 2$, let G be a graph with n vertices, $c_k(n)$ edges, and girth greater than k. Then $|W(G)| = c_k(n) + n + 1$ and $g_{W(G)}(k) \leq 2k$.

3 The Generalized Nonspeedup Theorem

Our first result in this section generalizes the Nonspeedup Theorem and provides a means to obtain downward inclusion results for verboseness classes. These additional inclusions have no counterpart in the "arrow relations".

Theorem 3.1 *Let $1 \leq h < m$, $1 \leq k < n$. Any (m,n)-verbose set is either (h,k)-verbose or $(m-h, n-k)$-verbose (i.e. $V(m,n) \subseteq V(h,k) \cup V(m-h,n-k)$).*

Proof: Suppose that A is (m,n)-verbose via f. We distinguish two cases.

Case 1: There exist $x_1, \ldots, x_{n-k}$ such that for all $y_1, \ldots, y_k$

$$|\{(b_1,\ldots,b_k) : (\chi_A(x_1),\ldots,\chi_A(x_{n-k}), b_1,\ldots,b_k) \in W_{f(x_1,\ldots,x_{n-k},y_1,\ldots,y_k)}\}| \leq h.$$

Then it easy to see that A is (h,k)-verbose.

Case 2: For all $x_1, \ldots, x_{n-k}$ there exist $y_1, \ldots, y_k$ such that

$$|\{(b_1,\ldots,b_k) : (\chi_A(x_1),\ldots,\chi_A(x_{n-k}), b_1,\ldots,b_k) \in W_{f(x_1,\ldots,x_{n-k},y_1,\ldots,y_k)}\}| > h.$$

Then we define uniformly r.e. sets $G(x_1,\ldots,x_{n-k}) = W_{g(x_1,\ldots,x_{n-k})}$ ($g \in \mathrm{R}_{n-k}$) as follows: Search for $y_1,\ldots,y_k$ and $a_1,\ldots,a_{n-k}$ such that

$$|\{(b_1,\ldots,b_k) : (a_1,\ldots,a_{n-k}, b_1,\ldots,b_k) \in W_{f(x_1,\ldots,x_{n-k},y_1,\ldots,y_k)}\}| > h.$$

By hypothesis, the search terminates. Then let

$$\begin{aligned} G(x_1,\ldots,x_{n-k}) &= \{(c_1,\ldots,c_{n-k}) : (\exists b_1,\ldots,b_k) \\ &\qquad [(c_1,\ldots,c_{n-k},b_1,\ldots,b_k) \in W_{f(x_1,\ldots,x_{n-k},y_1,\ldots,y_k)}]\}. \end{aligned}$$

Clearly $(\chi_A(x_1),\ldots,\chi_A(x_{n-k})) \in G(x_1,\ldots,x_{n-k})$. $W_{f(x_1,\ldots,x_{n-k},y_1,\ldots,y_k)}$ contains at least $h+1$ vectors that begin with $(a_1,\ldots,a_{n-k})$. Since $|W_{f(x_1,\ldots,x_{n-k},y_1,\ldots,y_k)}| \leq m$ there are at most $m-h-1$ additional vectors. Therefore the number of projections of these vectors onto the first $n-k$ components is at most $1+(m-h-1) = m-h$, i.e.,

$$|G(x_1,\ldots,x_{n-k})| \leq m-h.$$

Thus, A is $(m-h,n-k)$-verbose via g. ∎

Corollary 3.2

(a) *If A is (m, n_1+n_2)-verbose then there exist m_1, m_2 such that $m_1+m_2 = m+1$ and A is both (m_1,n_1)-verbose and (m_2,n_2)-verbose.*

(*b*) *$V(km, kn) \subseteq V(m, n)$.*

(*c*) *$V(km + c, kn + d) \subseteq V(m, n) \cup V(c, d)$.*

(*d*) *Every (n, n)-verbose set is recursive. (Nonspeedup Theorem)*

(*e*) *Let A be an (m, n)-verbose set. For every partition $n = \sum_{i=1}^{k} n_i$, there exist numbers $m_1, \ldots, m_k$ such that $\sum_{i=1}^{k} m_i = m + k - 1$ and, for every i, A is (m_i, n_i)-verbose.*

(*f*) *If A is verbose (in the sense of [8]) then A is $(n + 1, n)$-verbose for all $n \geq 1$.*

Proof:

(*a*) Suppose that A is $(m, n_1 + n_2)$-verbose, where m, n_1, n_2 are positive integers. Let m_1 be minimal such that A is (m_1, n_1)-verbose. Then A is not $(m_1 - 1, n_1)$-verbose. Let $m_2 = m - (m_1 - 1)$. By Theorem 3.1, A is either $(m_1 - 1, n_1)$-verbose or (m_2, n_2)-verbose. Thus, A is (m_2, n_2)-verbose, and $m_1 + m_2 = m + 1$.

(*b*) This follows by induction on k. It trivially holds for $k = 1$. Suppose it is true for $k - 1$. If A is (km, kn)-verbose then, by Theorem 3.1, either A is (m, n)-verbose or A is $(m(k - 1), n(k - 1))$-verbose. Thus, by the induction hypothesis, A is (m, n)-verbose.

(*c*) By Theorem 3.1, any $(km + c, kn + d)$-verbose set is either (km, kn)-verbose or (c, d)-verbose. The conclusion now follows from (b).

(*d*) By part (b), every (n, n)-verbose set is $(1, 1)$-verbose, i.e., it is recursive.

(*e*) This follows from (a) by induction on k.

(*f*) If A is verbose then A is $(m + 1, m)$-verbose for infinitely many m ($m = 2^k - 1$). Hence it suffices to show that $V(n + 1, n) \subseteq V(n, n - 1)$ for all $n \geq 2$. This follows from Theorem 3.1: $V(n + 1, n) \subseteq V(n, n - 1) \cup V(1, 1) = V(n, n - 1)$.

▮

All of the results mentioned above also hold analogously for the V_S-classes.

As a first application we consider how verboseness is related to the join operator. Any two semirecursive sets are strongly $(n + 1, n)$-verbose and their join is strongly $(p(n), n)$-verbose where $p(n) = (\lceil \frac{n}{2} \rceil + 1)(\lfloor \frac{n}{2} \rfloor + 1)$. By our next result, this bound cannot be improved for the join of *any* two nonrecursive T-incomparable sets.

Theorem 3.3 *For all $n > 0$, if A and B are nonrecursive sets such that $A \oplus B$ is strongly $(p(n) - 1, n)$-verbose then A and B are T-equivalent.*

Proof: Let A and B be given nonrecursive sets. We present the case $n = 2k + 1$. The case $n = 2k$ is similar. Suppose that $A \oplus B$ is strongly $(p(n) - 1, n)$-verbose via f. Let

$$G(x_1, \ldots, x_{k+1}, y_1, \ldots, y_k) = \\ \{(b_1, \ldots, b_k) : (\chi_A(x_1), \ldots, \chi_A(x_{k+1}), b_1, \ldots, b_k) \in D_{f(2x_1, \ldots, 2x_{k+1}, 2y_1+1, \ldots, 2y_k+1)}\}.$$

Note that G is recursive in A. We distinguish two cases.

Case 1: For all $y_1, \ldots, y_k$ there exist $x_1, \ldots, x_{k+1}$ such that

$$|G(x_1, \ldots, x_{k+1}, y_1, \ldots, y_k)| \leq k.$$

Then B is strongly (k,k)-verbose *relative to* A. Thus B is recursive in A by the relativized Nonspeedup Theorem.

Case 2: There exist $y_1, \ldots, y_k$ such that for all $x_1, \ldots, x_{k+1}$,

$$|G(x_1, \ldots, x_{k+1}, y_1, \ldots, y_k)| > k.$$

Fix $y_1, \ldots, y_k$ with that property.

Define a function $g \in \mathrm{R}_{k+1}$ such that

$$D_{g(x_1,\ldots,x_{k+1})} = \{(a_1, \ldots, a_{k+1}) : \\ |\{(b_1, \ldots, b_k) : (a_1, \ldots, a_{k+1}, b_1, \ldots, b_k) \in D_{f(2x_1,\ldots,2x_{k+1},2y_1+1,\ldots,2y_k+1)}\}| > k\}.$$

Clearly, $(\chi_A(x_1), \ldots, \chi_A(x_{k+1})) \in D_{g(x_1,\ldots,x_{k+1})}$. If $|D_{g(x_1,\ldots,x_{k+1})}| \leq k+1$, for all $x_1, \ldots, x_{k+1}$ then A is strongly $(k+1, k+1)$-verbose and hence recursive, by the Nonspeedup Theorem. Thus there exist $x_1, \ldots, x_{k+1}$ such that $|D_{g(x_1,\ldots,x_{k+1})}| > k+1$. Therefore,

$$|D_{f(2x_1,\ldots,2x_{k+1},2y_1+1,\ldots,2y_k+1)}| \geq (k+2)(k+1) = p(n),$$

contradicting the choice of f.

We have shown that B is recursive in A. By symmetry, A is recursive in B, and the proof is complete. ∎

Of course Theorem 3.3 does not hold with "(m,n)-verbose" in place of "strongly (m,n)-verbose," because the join of any two r.e. sets is r.e. and therefore $(n+1,n)$-verbose. However, if K is recursive in A, B and $A \oplus B$ is $(p(n)-1,n)$-verbose then A and B are T-equivalent by virtually the same proof as for Theorem 3.3.

By a similar argument we obtain

Theorem 3.4 *The join of two T-incomparable sets A and B is never verbose (in the sense of [8]).*

Proof: Suppose, for the sake of contradiction, that $F_3^{A\oplus B}$ can be computed with at most 2 queries to $A \oplus B$ by the oracle Turing machine $M^{(-)}$. We distinguish two cases.

Case 1: For all z there are x, y such that $M^{A\oplus B}$ with input $(2x, 2y, 2z+1)$ queries even numbers only.

To compute $\chi_B(z)$ we A-recursively search for x, y as above. Hence $B \leq_T A$, a contradiction.

Case 2: There exists z such that for all x, y, $M^{A\oplus B}$ with input $(2x, 2y, 2z+1)$ queries at least one odd number.

Then, relative to B, F_2^A can be computed with at most one query to A. Thus $A \leq_T B$, by the relativized Nonspeedup Theorem, a contradiction. ∎

A straightforward generalization shows that $F_{2^n+1}^{A\oplus B}$ cannot be computed with fewer than $n+2$ queries to $A \oplus B$, if A and B are T-incomparable.

4 A Decision Procedure for the Inclusion Problem

The following notion is the key to reducing the inclusion problem to finite combinatorics. A similar approach was developed previously by Degtev [17] for the inclusion problem of frequency computation (see Section 9).

Definition 4.1 A nonempty set $V \subseteq \{0,1\}^s$ is called (m,n)-*good* iff for every partition $n_1 + n_2 + \cdots + n_k = n$ with $1 \leq n_j \leq s$ and $k \geq 1$,

$$g_V(n_1) + \cdots + g_V(n_k) \leq m + k - 1.$$

If V is (m,n)-good then $g_V(n) \leq m$ (choose $k = 1$ in Definition 4.1). Thus (m,n)-goodness strengthens the condition that $g_V(n) \leq m$ by adding a certain convexity requirement on g_V. This is motivated by the analogous property of (m,n)-verboseness (cf. Corollary 3.2(e)).

Since $g_V(n_j) \leq 2^{n_j}$ and $2^a + 2^b \leq 2^{a+b-1} + 2$ for $a,b \geq 1$, it follows that

$$g_V(n_1) + \cdots + g_V(n_k) \leq 2^{n_1+\cdots+n_k-(k-1)} + 2(k-1) = 2^{n-k+1} + 2(k-1) \leq 2^n + k - 1.$$

Hence every set V is $(2^n, n)$-good.

If V is (n,n)-good then $g_V(1) = 1$ (choose $k = n$ in Definition 4.1). Hence $|V| = 1$, so V is $(1,n)$-good.

If V is (m,n)-good, $n_1 + \cdots + n_k = n$, and each m_i is minimal such that V is (m_i, n_i)-good then a simple computation shows that $m_1 + \cdots m_k \leq m + k - 1$.

These observations show that the notion (m,n)-goodness shares many properties with the notion (m,n)-verboseness. The precise correspondence will be shown in Theorem 4.3.

Examples:

- $V = \{0^i 1^{a-i}, 1^i 0^{a-i} : 1 \leq i \leq a\}$. Then $g_V(i) = 2i$, for $1 \leq i \leq a$. Hence V is $(2a+1, a+1)$-good but not $(2a-1, a)$-good.

- The set of all vectors of length n which contain at most $a-1$ 1s is $(2^a - 1, a)$-good and $(S(n,a), n)$-good but not $(S(n,a)-1, n)$-good (see Section 5 for a more detailed discussion).

- $W = \{0^n, 10^{n-1}, 010^{n-2}, \ldots, 0^{n-1}1, 110^{n-2}, 0110^{n-3}, \ldots, 0^{n-2}11, 10^{n-2}1\}$, where $n \geq 3$. Note that $W = W(G)$ where G is a cycle of length n. $g_W(n) = 2n+1$ and $g_W(i) = 2i$, for $1 \leq i < n$. Hence W is $(2n', n')$-good but not $(2n, n)$-good, for $0 < n' < n$.

The following observation often simplifies the test of the goodness property.

Lemma 4.2 *Let $V \subseteq \{0,1\}^s$ be a nonempty set of binary vectors. For $n > 0$, let $G_V(n)$ denote the least m such that V is (m,n)-good. If $n > s$, then*

$$G_V(n) = \max\{g_V(i) + G_V(n-i) - 1 : 1 \leq i \leq s\}.$$

Examples:

- $G_V(s+k) = 2^s + 2^k - 1$, if $V = \{0,1\}^s$ and $0 \le k \le s$.
  This follows from Lemma 4.2 by induction on k ($g_V(i) = 2^i$, for $1 \le i \le s$).

- $G_V(n) \le (2^s - 1)\frac{n}{s} + 1$, for any $V \subseteq \{0,1\}^s$ and $n > 0$.
  This follows from Lemma 4.2, because $g_V(i) \le 2^i \le (2^s-1)\frac{i}{s} + 1$, for $i \le s$.
  Thus, for fixed V, $G_V(n)$ is linearly bounded in n.

The next theorem — our main result — characterizes the inclusion problem via goodness and shows that the inclusions for the verbose case and the strongly verbose case are the same.

Theorem 4.3 *For any (m,n) and (m',n') the following statements are equivalent:*

(a) Every (m,n)-good set V is (m',n')-good.

(b) Every (m,n)-good set $V \subseteq \{0,1\}^{n'}$ has at most m' elements.

(c) $V(m,n) \subseteq V(m',n')$.

(d) $V_S(m,n) \subseteq V(m',n')$.

(e) $V_S(m,n) \subseteq V_S(m',n')$.

Proof:

(a) ⇒ (c): Assume that every (m,n)-good set V is (m',n')-good. Suppose that A is (m,n)-verbose. We show that A is (m',n')-verbose, too. For $1 \le i \le n$ let m_i be minimal such that A is (m_i,i)-verbose. By Corollary 3.2(e) it follows that for any partition $n = \sum_{i=1}^{k} n_i$ we have

$$m_{n_1} + \cdots + m_{n_k} \le m + k - 1.$$

Choose recursive functions f_i such that A is (m_i,i)-verbose via $f_i \in \mathrm{R}_i$. We define $f \in \mathrm{R}_{n'}$ for any $x_1 < \cdots < x_{n'}$ as follows:

$$\begin{aligned} W_{f(x_1,\ldots,x_{n'})} &= \{(a_1,\ldots,a_{n'}) : (\forall i \le n)(\forall r_1,\ldots,r_i) \\ &\qquad [1 \le r_1 < \cdots < r_i \le n' \Rightarrow (a_{r_1},\ldots,a_{r_i}) \in W_{f_i(x_{r_1},\ldots,x_{r_i})}]\} \end{aligned}$$

Intuitively, $W_{f(x_1,\ldots,x_{n'})}$ contains all n'-tuples that are consistent with every f_i. Let $V = W_{f(x_1,\ldots,x_{n'})}$. Note that for $1 \le i \le \min(n,n')$, we have $g_V(i) \le m_i$, by the definition of f. Thus for every partition $n_1 + \cdots + n_k = n$ with $1 \le n_j \le n'$ and $k \ge 1$, we have

$$g_V(n_1) + \cdots + g_V(n_k) \le m_{n_1} + \cdots + m_{n_k} \le m + k - 1.$$

Hence V is (m,n)-good. By hypothesis, V is (m',n')-good, thus, $|V| \le m'$. It follows that A is (m',n')-verbose via f.

(a) ⇒ (e): This is analogous to the previous proof.

(c) ⇒ (d), (e) ⇒ (d): Trivial.

(d) ⇒ (b): We will prove the contrapositive. Suppose that $V \subseteq \{0,1\}^{n'}$ is (m,n)-good and $|V| > m'$. Interchanging 0s and 1s in one column does not disturb these properties. If we apply this operation several times we can obtain an all-zero row. Hence we may assume that $0^{n'} \in V$; say $V = \{v_1, \ldots, v_r\}$ where $r > m'$, and $v_1 = 0^{n'}$. We construct a set $A \in V_S(m,n) - V(m',n')$ by recursive approximation; i.e., we define a recursive two place function $A(x,s)$ such that, for all x, $\chi_A(x) = \lim_{s\to\infty} A(x,s)$. (In fact, A will be $(m'+1)$-r.e.)

Before presenting the formal details, we discuss the intuition behind the construction. The following requirements will imply that A is not (m',n')-verbose:

(R_i) $\quad A$ is not (m',n')-verbose via φ_i.

We satisfy (R_i) by the obvious strategy: Specify an interval of n' successive numbers $x_1, \ldots, x_{n'}$. Let $w(s) = (A(x_1,s), \ldots, A(x_{n'},s))$. Initially, we let $w(0) = v_1 = 0^{n'}$ and $c(0) = 1$. Then $W_{\varphi_i(x_1,\ldots,x_{n'})}$ is enumerated. If we discover in stage s that $w(s) = v_{c(s)} \in W_{\varphi_i(x_1,\ldots,x_{n'})}$ then we let $w(s+1) = v_{c(s)+1}$ and $c(s+1) = c(s)+1$. We continue until no further elements are enumerated or $|W_{\varphi_i(x_1,\ldots,x_{n'})}| > m'$. Hence $w(s)$ converges to some element of V, and (R_i) is satisfied.

At each stage s, $l(i,s)$ specifies the left endpoint of the interval on which φ_i is diagonalized. We maintain that all intervals are mutually disjoint and their left endpoints are increasing in i.

It will become clear later on that A is automatically (m,n)-verbose if we combine the different strategies in the standard finite injury fashion: Whenever (R_i) acts at step s (i.e., whenever the ith interval changes), we destroy the previous actions of all (R_j) with $j > i$, i.e., we reset $A(x)$ equal to zero on their intervals and then we start their strategies anew on intervals consisting of numbers greater than s.

CONSTRUCTION

Stage 0: Let $A(x,0) = 0$, for all x. Let $l(i,0) = in'$ and $c(i,0) = 1$, for all i.

Stage $s+1$: If there exists $i \le s$ such that the following two conditions hold —

(1) $|W_{\varphi_{i,s}(x_1,\ldots,x_{n'}),s}| \le m'$

(2) $v_{c(i,s)} \in W_{\varphi_{i,s}(x_1,\ldots,x_{n'}),s}$

— where $x_1 = l(i,s), x_2 = l(i,s)+1, \ldots, x_{n'} = l(i,s)+n'-1$, then choose the least such i. Let $c(i,s+1) = c(i,s)+1$; $A(x_k,s+1) = v_{c(i,s)+1}[k]$, for $1 \le k \le n'$; $A(x,s+1) = 0$, for all $x \ge l(i+1,s)$; and $l(j,s+1) = l(j,s) + sn'$ and $c(j,s+1) = 1$, for all $j > i$. All other parameters $p(-,s)$ remain unchanged; i.e., $p(-,s+1) = p(-,s)$ for $p = c, A, l$.

If no such i exists then all parameters remain unchanged.

END OF CONSTRUCTION

If conditions (1) and (2) are satisfied for $i \le s$ at stage $s+1$ then we say that "i requires attention at stage $s+1$." If i is the least number that requires attention at stage $s+1$ then we say that "i acts at stage $s+1$." In the following we verify that A has all required properties. Claims 0,1,2,3 are standard and Claim 4 is the one of interest.

Claim 0: For all $i, s \geq 0$ and $1 \leq k \leq n'$ the following hold:

- $1 \leq c(i,s) \leq m'+1 \leq r$.
- $A(l(i,s)+k-1,s) = v_{c(i,s)}[k]$.
- If at stage $s+1$ no $j \leq i$ requires attention, then $c(i,s+1) = c(i,s)$, $l(i,s+1) = l(i,s)$, and $A(l(i,s)+k-1,s+1) = A(l(i,s)+k-1,s)$.

Claim 1: Each i requires attention at most finitely often.
Proof: By induction on i. Suppose the claim is true for all $j < i$. Then there exists a least stage s_0 such that no $j < i$ requires attention at any stage $s \geq s_0$. Thus if i requires attention at some stage $s \geq s_0$ then i acts at stage s. Also, $c(i, s_0) = 1$ and $l(i,s) = l(i,s_0)$, for all $s \geq s_0$. Now it is easy to see that there are at most m' stages $s \geq s'$ such that i requires attention at stage s. Thus the claim is true for i. □

Claim 2: A is $(m'+1)$-r.e.
Proof: Fix x, and consider the sequence of values of $A(x,s)$ for $s = 0,1,2,\ldots$. Initially, $A(x,0) = 0$. If there is a least stage s_0+1 such that $A(x,s_0+1) = 1$, then at stage s_0+1 some i acts such that $l(i,s_0) \leq x \leq l(i,s_0)+n'-1$. Consider the later stages until the least stage s_1 where some $j < i$ acts (if there is no such stage, let $s_1 = \infty$). At the intermediate stages s with $s_0+1 < s < s_1$, i acts at most $m'-1$ times. Because $l(i,s) = l(i,s_0)$ for each such s, we see that $A(x,\cdot)$ changes at most $m'-1$ times. At stage s_1, if $s_1 < \infty$, $A(x,s_1)$ is set equal to 0 and stays at this value at all later stages. Thus, A is $(m'+1)$-r.e. □

Claim 3: A is not (m',n')-verbose.
Proof: Suppose, for the sake of contradiction, that A is (m',n')-verbose via φ_i. By claim 1, there exists s_1 such that none of the requirements $0,...,i$ requires attention at any state $s \geq s_1$. Then $l(i,s) = l(i,s_1)$ for all stages $s \geq s_1$. Let $x_k = l(i,s_1)+k-1$ for $k = 1,\ldots,n'$. Then $A(x_k) = A(x_k,s)$, for all stages $s \geq s_1$. Because A is (m',n')-verbose via φ_i, $(A(x_1),\ldots,A(x_{n'})) \in W_{\varphi_i(x_1,\ldots,x_{n'})}$; so, there is a stage $s_2 \geq s_1$ such that $(A(x_1),\ldots,A(x_{n'})) \in W_{\varphi_{i,s_2}(x_1,\ldots,x_{n'}),s_2}$. Thus i requires attention at stage s_2, a contradiction. □

Claim 4: A is strongly (m,n)-verbose.
Proof: Given $X = \{x_1 < \cdots < x_n\}$ we find disjoint intervals

$$I_j = [l_j n',\, l_j(n'+1)-1], \text{ for } j = 1,\ldots,k,$$

such that $l_j < l_{j+1}$ and $\{x_{j1} < \cdots < x_{j,n_j}\} = X \cap I_j$, for each j, and such that $X \subseteq \bigcup_{1 \leq j \leq k} I_j$. Note that if $A(x_{j,h},s+1) \neq A(x_{j,h},s)$ then $A(x_{j',h'}) = A(x_{j',h'},s') = 0$, for all $j' > j, 1 \leq h' \leq n_j, s' > s$. Thus we are led to define a total recursive function f as follows (let $s = x_n+1$):

$$\begin{aligned} D_{f(x_1,\ldots,x_n)} \;=\; \bigcup_{j=1}^{k} \{ & (A(x_{11},s),\ldots,A(x_{1,n_1},s),\ldots,A(x_{j-1,1},s),\ldots,A(x_{j-1,n_{j-1}},s), \\ & v[x_{j1}-l_j n'+1],\ldots,v[x_{j,n_j}-l_j n'+1],0,0,\ldots,0) : v \in V\}. \end{aligned}$$

The jth part of the union is obtained by assuming that the $j-1$ highest priority requirements being dealt with take no action (so the elements involved are as they were at stage s), the jth highest priority requirement does take action (so it can be any $v \in V$ projected onto the relevant coordinates), and the intervals involved with lower priority requirements are all zero. For each j there exists $v \in V$ such that $(A(l_j n', s), \ldots, A(l_j(n'+1)-1, s)) = v$; therefore $(A(x_1), \ldots, A(x_n)) \in D_{f(x_1,\ldots,x_n)}$. Since $0^{n'} \in V$, we get

$$|D_{f(x_1,\ldots,x_n)}| \leq g_V(n_k) + (g_V(n_{k-1}) - 1) + \cdots + (g_V(n_1) - 1) \leq m,$$

because V is (m, n)-good. Thus, A is strongly (m, n)-verbose via f. $\Box$

(b) $\Rightarrow$ (a): Given an (m, n)-good set V that is not (m', n')-good we construct an (m, n)-good set $V' \subseteq \{0, 1\}^{n'}$ having more than m' elements. We need a way to change the size of V while preserving its goodness. Without loss of generality, assume that $0^n \in V$.

Given nonempty sets $V_1, \ldots, V_k$ with $V_i \subseteq \{0, 1\}^{n_i}$, we define

$$\oplus(V_1, V_2, \ldots, V_k) = \bigcup_{i=1}^{k} \{0^{n_1+\cdots+n_{i-1}} v 0^{n_{i+1}+\cdots+n_k} : v \in V_i\}.$$

Fact 1: If $0^{n_i} \in V_i$ for all i, then $|\oplus(V_1, \ldots, V_k)| = |V_1| + \cdots + |V_k| - (k - 1)$.

The next fact follows by an easy computation from the definition of goodness.

Fact 2: Let V be an (m, n)-good set such that $0^s \in V \subseteq \{0, 1\}^s$. Let $k \geq 1$. For $i = 1, \ldots, k$, let $1 \leq n_i \leq s$, let $1 \leq x_{i,1} < \cdots < x_{i,n_i} \leq s$, and let $V_i = \{v[x_{i,1}, \ldots, x_{i,n_i}] : v \in V\}$. Then $\oplus(V_1, \ldots, V_k)$ is (m, n)-good.

Now we are ready to complete the proof. Without loss of generality, we may assume that $0^s \in V \subseteq \{0, 1\}^s$. Because V is not (m', n')-good there exist numbers $1 \leq n_1, \ldots, n_k \leq s$ such that $n_1 + \cdots + n_k = n'$ and $g_V(n_1) + \cdots + g_V(n_k) > m' + k - 1$. Let $V' = \oplus(V_1, \ldots, V_k)$ where V_i is a projection of V onto i components such that $|V_i| = g_V(i)$. By Fact 1, $|V'| = g_V(n_1) + \cdots + g_V(n_k) - (k - 1) > m'$. By Fact 2, V' is (m, n)-good. Clearly, $V' \subseteq \{0, 1\}^{n'}$. ∎

Because condition (b) can be tested by exhaustive search, we have

Corollary 4.4 *The inclusion problem is decidable.*

Remark: By Fact 2.6, condition (b) of Theorem 4.3 is equivalent to

(b') Every *closed* (m, n)-good set $V \subseteq \{0, 1\}^{n'}$ has at most m' elements.

From Theorem 4.3 and the examples mentioned after Definition 4.1 we immediately get

Proposition 4.5

(a) *For all* $k > 0$, $V(2k + 1, k + 1) \not\subseteq V(2k - 1, k)$.

(*b*) *For all* $k, m, n > 0$*,* $V(2^k - 1, k) \subseteq V(m, n)$ *iff* $m \geq S(n, k)$*.*

(*c*) *For all* n, n' *with* $n > 2$*,* $0 < n' < n$*,* $V(2n', n') \not\subseteq V(2n, n)$*.*

Now we will discuss the inclusion relation for $V_S(m, n)$-classes of *recursively enumerable* sets. The following result from [31] is a key ingredient:

Fact 4.6 ([31]) *Assume that* A *is r.e. and strongly* $(2^n - 1, n)$*-verbose via* f*. Then there exists a total recursive function* g *such that, for all* $x_1, \ldots, x_n$*,* $D_{g(x_1,\ldots,x_n)} \subseteq D_{f(x_1,\ldots,x_n)}$*,* $(\chi_A(x_1), \ldots, \chi_A(x_n)) \in D_{g(x_1,\ldots,x_n)}$*, and* $D_{g(x_1,\ldots,x_n)}$ *contains both a least and a greatest element with respect to* $\preceq$*.*

Theorem 4.7 *For any* (m, n) *and* (m', n') *the following statements are equivalent:*

(*a*) *Every r.e. set that is strongly* (m, n)*-verbose is strongly* (m', n')*-verbose.*

(*b*) *Every* (m, n)*-good set* $V \subseteq \{0,1\}^{n'}$ *that contains both* $0^{n'}$ *and* $1^{n'}$ *has at most* m' *elements.*

Remark: It was a challenging problem to find (m, n) and (m', n') such that condition (b) of Theorem 4.7 is satisfied but condition (b) of Theorem 4.3 is not satisfied. The noninclusions corresponding to Proposition 4.5(a,b) can be transferred to the r.e. case. (The analog of Proposition 4.5(a) goes through because the set of vectors contains two diametrically opposite points. For the analog of Proposition 4.5(b) use the set $V \subseteq \{0,1\}^n$ containing all vectors v such that $0v$ has at most $k - 1$ alternations from 0 to 1 or 1 to 0.) Only by exhaustive search (using a computer) could we verify that Proposition 4.5(c) cannot be transferred: Every $(8, 4)$-good set $V \subseteq \{0,1\}^5$ with $0^5, 1^5 \in V$ contains at most 10 elements. Hence, by Theorem 4.7, any r.e. set that is strongly $(8, 4)$-verbose is also strongly $(10, 5)$-verbose. On the other hand, by Proposition 4.5(c), there is a strongly $(8, 4)$-verbose set that is not strongly $(10, 5)$-verbose. Thus the inclusion relation for V_S-classes is properly contained in the inclusion relation for V_S-classes restricted to r.e. sets.

Proof of Theorem 4.7:

(a) $\Rightarrow$ (b): We will prove the contrapositive. Assume that there exists $V \subseteq \{0,1\}^{n'}$, such that $\{0^{n'}, 1^{n'}\} \subseteq V$, $|V| > m'$, and V is (m, n)-good. Then we need to construct an *r.e.* set $A \in V_S(m, n) - V_S(m', n')$. This is done by a slight modification of the construction in the proof of "(d) $\Rightarrow$ (b)" in Theorem 4.3. We change (R_i) to require that A is not *strongly* (m', n')-verbose via φ_i, and the basic strategy for satisfying requirement (R_i) is now as follows:

Specify an interval of n' successive numbers $x_1, \ldots, x_{n'}$, and wait for $\varphi_i(x_1, \ldots, x_{n'})$ to converge. When this happens and $|D_{\varphi_i(x_1,\ldots,x_{n'})}| \leq m'$, then choose $v \in V - D_{\varphi_i(x_1,\ldots,x_{n'})}$ and enumerate each x_i with $v[i] = 1$ into A. Hence $(A(x_1), \ldots, A(x_{n'})) = v \notin D_{\varphi_i(x_1,\ldots,x_{n'})}$, and (R_i) is satisfied.

We combine the different strategies as in the previous proof. The only (and essential) difference is that whenever we reset an interval, we do not reset $A(x)$ equal to zero, but equal to one. This amounts to enumerating x into A and guarantees that A becomes an r.e. set. Because $0^{n'}, 1^{n'} \in V$ we will have $(A(x_1, s), \ldots, A(x_{n'}, s)) \in V$, at each stage s and for all intervals $\{x_1 < \cdots < x_{n'}\}$. Therefore, the proof of Claim 4 can be easily adapted. The formal construction and its verification are left to the reader.

(*b*) $\Rightarrow$ (*a*): Assume that every (m,n)-good set $V \subseteq \{0,1\}^{n'}$ with $0^{n'}, 1^{n'} \in V$ has at most m' elements. Further assume that A is r.e. and strongly (m,n)-verbose. We need to show that $A \in V_S(m', n')$. By the proof of "(b) $\Rightarrow$ (e)" in Theorem 4.3, there exists $f \in \mathrm{R}_{n'}$ such that, for all $x_1, \ldots, x_{n'}$, $(A(x_1), \ldots, A(x_{n'})) \in D_{f(x_1,\ldots,x_{n'})}$ and $D_{f(x_1,\ldots,x_{n'})}$ is (m,n)-good. If $|D_{f(x_1,\ldots,x_{n'})}| = 2^{n'}$ for some $x_1, \ldots, x_{n'}$ then $0^{n'}, 1^{n'} \in D_{f(x_1,\ldots,x_{n'})}$ and, by hypothesis $m' \geq 2^{n'}$. Then trivially $A \in V_S(m', n')$. For the following suppose that $|D_{f(x_1,\ldots,x_{n'})}| < 2^{n'}$, for all $x_1, \ldots, x_{n'}$. Then we may assume by Fact 4.6 that $D_{f(x_1,\ldots,x_{n'})}$ contains a least and a greatest element.

Suppose for a contradiction that there exist $x_1, \ldots, x_{n'}$ such that $|D_{f(x_1,\ldots,x_{n'})}| > m'$. Let $V = D_{f(x_1,\ldots,x_{n'})}$. Let $v_{\min}$ and $v_{\max}$ be the least and greatest elements of V, respectively. Let $I = \{i : v_{\min}[i] < v_{\max}[i]\}$, let $s = |I|$, and let $U = P^I(V)$. Then $|U| = |V|$, $0^s \in U$, $1^s \in U$, U is (m,n)-good, and U is not (m',n')-good. Now we argue as in the proof of "(b) $\Rightarrow$ (a)" in Theorem 4.3: Because U is not (m',n')-good there exist numbers $1 \leq n_1, \ldots, n_k \leq s$ such that $n_1 + \cdots + n_k = n'$ and $g_U(n_1) + \cdots + g_U(n_k) > m' + k - 1$. Let U_i be a projection of U onto i components such that $|U_i| = g_U(n_i)$. Define $W \subseteq \{0,1\}^{n'}$ as follows:

$$W = \bigcup_{i=1}^{k} \{0^{n_1+\cdots+n_{i-1}} v 1^{n_{i+1}+\cdots+n_k} : v \in U_i\}.$$

An easy computation shows that W is (m,n)-good. Because $0^{n_i}, 1^{n_i} \in V_i$ it follows that $0^{n'}, 1^{n'} \in W$. Hence, by hypothesis, W has at most m' elements. But in fact $|W| > m'$, a contradiction.

It follows that $|D_{f(x_1,\ldots,x_{n'})}| \leq m'$ for all $x_1, \ldots, x_n$. Thus, A is strongly (m',n')-verbose via f. ∎

5 Infinite Versions of Theorem 4.3

In this section we will give necessary and sufficient conditions on $f : \omega \to \omega$ so that there exists a set A which is $(f(n), n)$-verbose but not $(f(n) - 1, n)$-verbose for all $n \geq 1$. To produce such an A, we combine infinitely many constructions of the kind used in the proof of Theorem 4.3, which is possible iff a certain convexity condition is met. The following example shows that not even two such constructions can be combined in general:

Example: Let

$$\begin{aligned} V_1 &= \{000, 001, 010, 100, 011, 110, 101\} \\ V_2 &= \{0000, 0001, 0010, 0100, 1000, 0011, 0110, 1100, 1001\}. \end{aligned}$$

It is not hard to see that V_1 is $(14,7)$-good but not $(6,3)$-good and that V_2 is $(14,7)$-good but not $(8,4)$-good. By Theorem 4.3 it follows that there exists a $(14,7)$-verbose that is not $(6,3)$-verbose and another $(14,7)$-verbose set that is not $(8,4)$-verbose. However, by Theorem 3.1, every $(14,7)$-verbose set is either $(8,4)$-verbose or $(6,3)$-verbose.

We will now discuss some infinite versions of Theorem 4.3. First, we extend the notion of (m,n)-goodness to finite sets $V \subseteq \{0,1\}^{<\omega}$, by appending strings of zeros to each vector in V so that all of them have the same length and then applying Definition 4.1. Second, we extend that notion to arbitrary sets $V \subseteq \{0,1\}^{<\omega}$, by calling V (m,n)-good iff every finite subset of V is (m,n)-good.

Notation 5.1 For $V \subseteq \{0,1\}^{<\omega}$, let $G_V(n)$ denote the smallest m such that V is (m,n)-good. Let $g_V(n)$ denote the maximum value of $g_{V'}(n)$ where V' ranges over all finite subsets of V. For $A \subseteq \omega$ let $\Gamma_A(n)$ denote the smallest m such that A is (m,n)-verbose.

By an easy modification of the construction from the "(d) $\Rightarrow$ (b)"-part of the proof of Theorem 4.3 we obtain

Theorem 5.2 *If V is an r.e. subset of $\{0,1\}^{<\omega}$, then there is a set $A \leq_T K$ such that $\Gamma_A = G_V$.*

The theorem above is also true under the weaker hypothesis that $V \in \Sigma_2^0$. In many applications the computation of G_V is simplified by the following fact:

Fact 5.3 *Let $V \subseteq \{0,1\}^{<\omega}$. If $g_V(i) + g_V(j) \leq g_V(i+j) + 1$, for all $i,j \geq 1$ then $g_V = G_V$. A sufficient condition for this to happen is that g_V is* convex *in the sense that $g_V(i+2) - g_V(i+1) \geq g_V(i+1) - g_V(i)$.*

Proof: For any convex function $f : \omega \to \omega$ with $f(0) = 1$ we have

$$f(i+j)-f(j)+1 = 1+\sum_{k=1}^{i}(f(i+j-k+1)-f(i+j-k)) \geq 1+\sum_{k=1}^{i}(f(i-k+1)-f(i-k)) = f(i).$$

∎

Beigel et al. [8] observed that for any $k > 0$ there is a set A (the join of k semirecursive sets) such that $\Gamma_A(n) = \Theta(n^k)$. Such a set can easily be obtained using our method: Let V be the set of all vectors containing at most k 1s; then $g_V(n) = S(n, k+1)$. Since g_V is convex (this follows easily from the identity $S(i+1,k) = S(i,k) + S(i-1,k-1)$), we conclude that for any k there is a set A such that $\Gamma_A(n) = S(n,k+1) = \Theta(n^k)$.

Beigel et al. [8] noticed that, by Fact 2.5, for every A either $\Gamma_A = \lambda n\ .\ 2^n$ or $\Gamma_A = O(n^k)$, for some k. In this statement we cannot replace $O(n^k)$ by $\Theta(n^k)$, as demonstrated by the following examples:

- Let V be the set of all vectors $0^{i-1}10^{j-i-1}1$ such that the binary representations of i and j differ in exactly one bit. Each vector in V can be identified with an edge of the infinite hypercube. It is easy to see that g_V is convex. By a result of Burtin [15] (see also [14, Theorem 3]) any n-vertex subgraph of the infinite hypercube has at most $\frac{1}{2}n\ \log_2 n$ edges. Thus $g_V = \Theta(n\ \log_2 n)$ and we conclude that there is a set A with $\Gamma_A = \Theta(n\ \log_2 n)$.

- Consider the infinite graph $C^{(k)}$, $k > 2$, which is the disjoint union of all $C_{k,n}$, $n > 0$, where $C_{k,n}$ is a graph with n vertices, $c_k(n)$ edges, and girth greater than k. Let $V \subseteq \{0,1\}^{<\omega}$ be the associated closed set of vectors. Then $g_V(n) = 1 + n + c_k(n)$. Because $c_k(i) + c_k(j) \leq c_k(i+j)$, we conclude by Fact 5.3 and Theorem 5.2 that there is a set A_k with $\Gamma_{A_k} = \lambda n\ .\ 1 + n + c_k(n)$. Hence, $\Gamma_{A_k} = \Theta(c_k)$ and the asymptotic bounds for c_k from Section 2 apply to Γ_{A_k}.

All sets A considered so far were recursive in K (because V was r.e.). Which functions occur as Γ_A when A ranges over *all* subsets of ω? The next theorem gives a general answer:

Theorem 5.4 *Let $f : \omega \to \omega$ be any function. There exists a set $A \subseteq \omega$ with $\Gamma_A = f$ iff there exists a set $V \subseteq \{0,1\}^{<\omega}$ with $G_V = f$.*

The proof uses a modification of the "k-branch construction" from [29] and is omitted here. Using Theorem 5.4 one can show, for instance, that there are continuum many different functions of the form Γ_A.

6 The Equality Problem

Define $E_{m,n} = \{(m',n') : V(m,n) = V(m',n')\}$ ($= \{(m',n') : V_S(m,n) = V_S(m',n')\}$, by Theorem 4.3). In this section we want to explicitly determine $E_{m,n}$. The Nonspeedup Theorem and the existence of nonrecursive verbose sets imply that $E_{1,1} = \{(m,n) : 0 < \frac{m}{n} \leq 1\}$. The next theorem, an upward inclusion result, is the key to an explicit solution for $m < 2n$. In this case $E_{m,n}$ is infinite and depends only on the quotient $\frac{m}{n}$.

Theorem 6.1 *$V(2k-1,k) \subseteq V(1 + \lfloor \frac{2k-3}{k-1} n \rfloor, n)$, for all $n > 0, k > 1$.*

Proof: By Fact 2.3 and Proposition 2.8 the theorem follows for all $n \geq k$. It remains to prove the case $n < k$. This is done by induction on k. Let $T(n,k) = 1 + \lfloor \frac{2k-3}{k-1} n \rfloor$. If A is $(2k-1,k)$-verbose and $1 \leq n < k$ then, by Theorem 3.1, either A is $(T(n,k),n)$-verbose or A is $(2k-1-T(n,k), k-n)$-verbose. If the former is true, we are done. Suppose the latter. Then

$$\begin{aligned} 2k-1-T(n,k) &= 2k-1-(2n+1-\lceil \frac{n}{k-1} \rceil) \\ &= 2(k-n)-1. \end{aligned}$$

Hence, A is $(2(k-n)-1, k-n)$-verbose. If $k-n=1$ then A is recursive. If $k-n>1$ then, by the inductive hypothesis, A is $(T(n,k-n),n)$-verbose. Because $T(n,k)$ is nondecreasing in k the theorem follows. ∎

Using Corollary 3.2, Proposition 4.5, and Theorem 6.1, we are able to characterize of $E_{m,n}$ for $m < 2n$.

Corollary 6.2 *For all $k > 1$,*

$$E_{2k-1,k} = \{(m,n) : \frac{2k-3}{k-1} < \frac{m}{n} \leq \frac{2k-1}{k}\}.$$

Proof: It suffices to verify the following properties:

(1) $\frac{m}{n} > \frac{2k-3}{k-1} \Rightarrow V(2k-1,k) \subseteq V(m,n)$.

(2) $\frac{m}{n} \leq \frac{2k-1}{k} \Rightarrow V(m,n) \subseteq V(2k-1,k)$.

(3) $V(2k-1,k) \not\subseteq V(2k-3,k-1)$.

Proof of (1): If $\frac{m}{n} > \frac{2k-3}{k-1}$ then $m \geq 1 + \lfloor \frac{2k-3}{k-1} n \rfloor$. Thus $V(2k-1,k) \subseteq V(m,n)$, by Theorem 6.1.

Proof of (2): Suppose that $\frac{m}{n} \leq \frac{2k-1}{k}$, and $n = ak + b$, with $0 \leq b < k$. If $b = 0$, then $V(m,n) \subseteq V(a(2k-1), ak) \subseteq V(2k-1,k)$, by Corollary 3.2(b). If $b > 0$ then, by Corollary 3.2(c),

$$V(m,n) \ \subseteq \ V(a(2k-1) + 2b - 1, ak + b) \ \subseteq \ V(2b-1,b) \cup V(2k-1,k).$$

If $1 < b < k$ then $\frac{2b-3}{b-1} < \frac{2k-1}{k}$, and hence it follows from (1) that $V(2b-1,b) \subseteq V(2k-1,k)$, for all $b > 1$. If $b = 1$ this inclusion holds trivially. Thus we get $V(m,n) \subseteq V(2k-1,k)$.

Proof of (3): This is Proposition 4.5. ▮

By properties (2) and (3) in the previous proof we obtain

Corollary 6.3 *(a)* $V(2k-1,k) \subset V(2k+1,k+1)$.

(b) $0 < \frac{m}{n} \leq \frac{m'}{n'} < 2 \Rightarrow V(m,n) \subseteq V(m',n')$.

Thus, the inclusion structure of the $V(m,n)$-classes for $\frac{m}{n} < 2$ forms an infinite hierarchy linearly ordered by $\frac{m}{n}$ and with breakpoints at $\frac{2k-1}{k}$. Also note that Corollary 6.3(b) gives many new upward inclusions that do not correspond to an arrow relation.

For $m \geq 2n$ we get a radically different picture. Let $h_k(n)$ denote the least m such that $V(2k,k) \subseteq V(m,n)$. Recall that $c_k(n)$ is the maximum number of edges in any graph on n vertices with girth greater than k.

Theorem 6.4 *For $n \geq k > 2$, $h_k(n) = c_k(n) + n + 1$.*

Proof: By Fact 2.3 and Proposition 2.9 we get $h_k(n) \leq c_k(n) + n + 1$. For the lower bound let G be a graph with n vertices, $c_k(n)$ edges, and girth greater than k. Let $V = W(G)$. It is easy to see that V is $(2k,k)$-good. Because $|V| = c_k(n) + n + 1$, Theorem 4.3 implies that $V(2k,k) \not\subseteq V(c_k(n) + n, n)$. ▮

From the asymptotic bounds for $c_k(n)$ (see Section 2) we obtain

Corollary 6.5 *For all $k > 2$, $h_k(n) = O(n^{1+\frac{2}{k-1}})$ and $h_k(n) = \Omega(n^{1+\frac{4}{3(k+1)}})$.*

In particular, h_k has superlinear but subquadratic growth for all $k > 3$. Because $h_k(n)$ is superlinear in n for all $k > 1$, we have

Corollary 6.6 *If $2n \leq m < 2^n$, then $E_{m,n}$ is finite.*

Proof: Suppose that $2n_0 \leq m_0 < 2^{n_0}$. Then, by Corollary 6.5, there are a rational number $c > 0$ and an integer n_1 such that $V(m_0, n_0) \not\subseteq V(cn^{1+\frac{1}{n_0+1}}, n)$, for all $n > n_1$.

On the other hand there exists n_2 such that $V(cn^{1+\frac{1}{n_0+1}}, n) \not\subseteq V(m_0, n_0)$ for all $n > n_2$. Why? Let $V = \{0,1\}^{n_0}$. V is not (m_0, n_0)-good, but it is $(cn^{1+\frac{1}{n_0}}, n)$-good for almost all n, because $G_V(n)$ is linearly bounded in n. Hence, by Theorem 4.3, $V(cn^{1+\frac{1}{n_0+1}}, n) \not\subseteq V(m_0, n_0)$ for almost all n.

Therefore,

$$E_{m_0,n_0} \subseteq \{(m,n) : n \leq \max(n_1, n_2) \text{ and } m \leq 2^n\},$$

so E_{m_0,n_0} is finite. ▮

With more effort it can be shown that $E_{m,n} = \{(m,n)\}$ if $2n \leq m < 2^n$, so there are no nontrivial equalities when $\frac{m}{n} \geq 2$. We defer the lengthy proof to the appendix.

The results of this section are summarized in the following theorem:

Theorem 6.7 *$V(m,n) = V(r,s)$ iff one of the following conditions holds:*

(a) $\frac{m}{n} \leq 1$ *and* $\frac{r}{s} \leq 1$.

(b) $(\exists k > 1)\left[\frac{2k-3}{k-1} < \frac{m}{n} \text{ and } \frac{r}{s} \leq \frac{2k-1}{k-1}\right]$.

(c) $(m,n) = (r,s)$.

(d) $m \geq 2^n$ *and* $r \geq 2^s$.

Remark: It follows from this theorem that $V(2^k - 1, n)$ is a proper subset of $V(2^k, n)$ for $n \leq 2^k \leq 2^n$ (note that there is exactly one k with $n \leq 2^k < 2n$). In other words, by Fact 1.1, for all k, n with $\lceil \log_2(n+1) \rceil \leq k \leq n$ there is a set A such that F_n^A can be computed by k queries (to some oracle) but not by $k-1$ queries (to any oracle).

7 Polynomial Time Verboseness

The general study of bounded queries with polynomial time computation was initiated by Amir and Gasarch [2] and Beigel [3, 4, 6, 7]. In this section we consider the polynomial-time analogue of (m,n)-verboseness and obtain an explicit solution to the corresponding inclusion problem. Some early results of a similar nature are due to Kinber [27].

Definition 7.1 A set A is *polynomial time (m,n)-verbose* (denoted $A \in V_P(m,n)$), iff there is a function $f : \omega^n \to \omega$ which is computable in polynomial time (i.e., the binary representation of the output is computable within time polynomial in the length of the binary representation of the input) such that A is strongly (m,n)-verbose via f.

Definition 7.2 $V \subseteq \{0,1\}^s$ is called *$(m,n)_P$-good* iff $g_V(n) \leq m$.

Theorem 7.3 *$V_P(m,n) \subseteq V_P(m',n')$ iff every $(m,n)_P$-good set $V \subseteq \{0,1\}^{n'}$ is $(m',n')_P$-good.*

Proof:

($\Leftarrow$) Assume that every $(m,n)_P$-good set $V \subseteq \{0,1\}^{n'}$ is $(m',n')_P$-good. Let A be polynomial time (m,n)-verbose via f. If $n' < n$ then $A \in V_p(m',n')$ via the function g defined by

$$D_{g(x_1,\ldots,x_{n'})} = \\ \{(b_1,\ldots,b_{n'}) : (\exists c_1,\ldots,c_{n-n'})[(b_1,\ldots,b_{n'},c_1,\ldots,c_{n-n'}) \in D_{f(x_1,\ldots,x_{n'},0,\ldots,0)}]\}.$$

$D_{g(x_1,\ldots,x_{n'})}$ is a projection of an $(m,n)_P$-good set onto n' coordinates. Hence, by hypothesis, its cardinality is at most m'.

If $n \leq n'$ then $A \in V_p(m',n')$ via h defined as follows:

$$D_{h(x_1,\ldots,x_{n'})} = \\ \{(b_1,\ldots,b_{n'}) : (\forall i_1,\ldots,i_n \in \{1,\ldots,n'\})[(b_{i_1},\ldots,b_{i_n}) \in D_{f(x_{i_1},\ldots,x_{i_n})}]\}.$$

$D_{h(x_1,\ldots,x_{n'})}$ is by construction an $(m,n)_P$-good set. Hence, by hypothesis, it is $(m',n')_P$-good; so, its cardinality is at most m'.

Note that g and h are computable in polynomial time since one needs to examine only a constant number of possibilities (the constant depends on n, but n is fixed).

($\Rightarrow$) Suppose that $V \subseteq \{0,1\}^{n'}$ is $(m,n)_P$-good but not $(m',n')_P$-good. We construct a set $A \in V_P(m,n) - V_P(m',n')$ by a diagonalization over all polynomial time computable $(m',n')_P$-verboseness operators. The diagonalization intervals are set up at a superexponential distance from each other, and the restriction of the characteristic function of A to any interval equals a string in V. In this way, we can ensure that $A \in V_P(m,n)$. The formal implementation is straightforward (cf. [6, Theorem 6.1]) and therefore omitted. A can be constructed to be an arbitrarily complex recursive set. ∎

Note that, in contrast to Theorem 3.1, there are only trivial downward inclusions: If $n' \leq n$, then $V_P(m,n) \subseteq V_P(m',n')$ iff either $m \leq m'$ or $m' \geq 2^{n'}$. Also note that the upward inclusion relation on the V_P-classes corresponds exactly to the arrow relation. Hence the inclusion relation on the V_P-classes is properly contained in the inclusion relation on the V_S-classes.

Furthermore, it coincides with the *effective* inclusion relation on the V_S-classes, defined as follows: $V_S(m,n)$ is effectively included in $V_S(m',n')$ iff for any $f \in \mathrm{R}_n$ there exists $g \in \mathrm{R}_{n'}$ such that $\{A : A \text{ is strongly } (m,n)\text{-verbose via } f\} \subseteq \{B : B \text{ is strongly } (m',n')\text{-verbose via } g\}$, and an index for g can be computed uniformly from an index for f. This is easy to see: If every $(m,n)_P$-good set is $(m',n')_P$-good then $V_S(m,n)$ is effectively included in $V_S(m',n')$, either trivially, for $n' \leq n$, or by the proof of Fact 2.3, for $n \leq n'$. For the converse suppose that there exists $V \subseteq \{0,1\}^{n'}$ such that $g_V(n) \leq m$ and $|V| > m'$. Let $M = \{A : A \subseteq \{0,\ldots,n'-1\} \wedge \chi_A \restriction n' \in V\}$. Then there exists $f \in R_n$ such that every $A \in M$ is strongly (m,n)-verbose via f. But there is no g such that every $A \in M$ is strongly (m',n')-verbose via g.

Because of the lack of nontrivial downward inclusions we can easily characterize the equality relation on the V_P-classes in an explicit way. (The converse direction of the following corollary is implicit in [7].)

Corollary 7.4 *$V_P(m,n) = V_P(r,s)$ iff one of the following conditions holds:*

(a) $\frac{m}{n} \leq 1$, $\frac{r}{s} \leq 1$, and $m = r$.

(b) $(m,n) = (r,s)$.

(c) $m \geq 2^n$ and $r \geq 2^s$.

Proof:

($\Leftarrow$) If (b) or (c) is satisfied then $V_P(m,n) = V_P(r,s)$ for trivial reasons. Suppose, for the sake of contradiction, that (a) is satisfied and $V_P(m,n) \neq V_P(r,s)$. Without loss of generality, we may assume that $s > n$. Thus, $V_P(r,s) \subseteq V_P(m,n)$. By hypothesis, $V_P(m,n) \not\subseteq V_P(r,s)$. By Theorem 7.3 there is an $(m,n)_P$-good set $V \subseteq \{0,1\}^s$ that is not $(r,s)_P$-good. In particular $|V| \geq r+1 = m+1$. By Fact 2.4(a) there is a projection of V onto m coordinates containing at least $m+1$ strings. Because $m \leq n$ it follows that V is not $(m,n)_P$-good, a contradiction.

($\Rightarrow$) By Theorem 7.3, it suffices to exhibit an $(m,n)_P$-good set $V \subseteq \{0,1\}^s$ that is not $(r,s)_P$-good, in each of the following three cases.

Case 1: $m \geq 2^n$ and $r < 2^s$. Let $V = \{0,1\}^s$.

Case 2: $r < 2^s$ and $r < m$. Choose any $V \subseteq \{0,1\}^r$ with $|V| = r+1$.

Case 3: $m = r$, $n < s$, and $n < m < 2^n$. Let $V' \subseteq \{0,1\}^{n+1}$ be any set of cardinality $m+1$ that contains all strings of length $n+1$ that contain at most one 1. Let $V = \{v0^{s-n-1} : v \in V'\}$. Because $|V| > r$, V is not $(r,s)_P$-good. Any projection of V onto n coordinates omits at least one coordinate i with $1 \leq i \leq n$. Thus 0^s and $0^{i-1}10^{s-i}$ are identified and $g_V(n) \leq m$; i.e., V is $(m,n)_P$-good. ∎

Note that the classes $V_P(m,n)$ with $m \leq n$ consist entirely of "cheatable sets" (in the sense of [7, Def. 5.2]), and all nontrivial equalities come from the cheatability phenomenon.

8 Verboseness and the Difference Hierarchy

The difference hierarchy is the Boolean algebra generated by the r.e. sets. Equivalently, it is the class of all k-r.e. sets, $k > 0$, cf. [40, p. 58]. We refer the reader to Ershov [19] and Epstein et. al. [18] for a detailed study. Beigel, Gasarch and Hay [9] studied the relationship between the difference hierarchy and the bounded query hierarchy with respect to the fixed oracle K.

Let $d_k(n)$ denote the least m such that every k-r.e. set is (m,n)-verbose. Let $f_k(n)$ denote the maximum length s of a sequence $v_1, \ldots, v_s$ of pairwise distinct binary vectors of length n such that $v_1 = 0^n$ and each sequence $v_1[i], \ldots, v_s[i]$ $(1 \leq i \leq n)$ changes its value at most k times. Clearly, $d_k(n) \leq f_k(n)$. By a straightforward diagonalization there is an $(f_k(n), n)$-verbose k-r.e. set A which is not $(f_k(n)-1, n)$-verbose (in fact, $\Gamma_A = f_k$ can be achieved in this way). Thus we get the following:

Fact 8.1 *For all $k, n > 0$, $d_k(n) = f_k(n)$.*

It is easy to see that $f_k(n) \leq kn + (k \bmod 2)$ (because every element of the sequence differs from the previous one, and each column changes at most k times). We have developed a scheme for producing sequences that give $f_k(n) \geq kn + (k \bmod 2)$, for all $n > k+1$. As the correctness proof is lengthy and tedious we omit further details. We conjecture that $d_k(n) = \min\{2^n, kn + (k \bmod 2)\}$, for $k > 0$. This is easy to see for $k = 1$; it has been verified with the assistance of a computer for $2 \leq k \leq 6$. The next result shows how lower bounds for $d_k(n)$ can be "lifted."

Proposition 8.2 *For all $k \geq 1$, $d_k(n+1) \geq d_k(n) + k$.*

Proof: By Fact 8.1, it is sufficient to show that $f_k(n+1) \geq f_k(n) + k$. Let $m = f_k(n)$ and let $v_1, \ldots, v_m$ be a sequence of vectors that proves $f_k(n) \geq m$. Then the sequence

$$0v_1, 1v_1, 1v_2, 0v_2, 0v_3, 1v_3, \ldots, av_k, bv_k, bv_{k+1}, \ldots, bv_m,$$

where $b = k \bmod 2$ and $a = 1 - b$, proves that $f_k(n+1) \geq f_k(n) + k$. ∎

By the remarks above, there exist 2-r.e. sets that are not $(3,2)$-verbose (in contrast, every r.e. set is $(3,2)$-verbose). There is also a 2-r.e. $(3,2)$-verbose set that does not have r.e. degree (use a slight modification of the standard construction of a properly 2-r.e. degree as presented in [40, VII.2.4]). On the other hand, it is shown in [31] that every k-r.e. strongly $(2^n - 1, n)$-verbose set has r.e. degree.

The numbers $d_k(n)$ also appear in quite different contexts. For $A \subseteq \omega$, Beigel et al. [8] defined

$$\text{Parity}_n^A = \{(x_1, \ldots, x_n) : \chi_A(x_1) + \cdots + \chi_A(x_n) \text{ is odd}\}.$$

A set is weakly semirecursive [25] iff it is an initial segment of an r.e. partial ordering on ω (see [30], originally Jockusch and Owings [25] gave a different but equivalent definition). An r.e. partial ordering on ω is an r.e. irreflexive, transitive relation $\sqsubset$ on $\omega \times \omega$, and $A \subseteq \omega$ is an initial segment of $\sqsubset$ iff $x \sqsubset y$, for all $x \in A$ and $y \in \overline{A}$.

Let $r_k(n)$, $s_k(n)$, and $w_k(n)$ respectively denote the least m such that Parity_k^A is (m,n)-verbose for every r.e., semirecursive, and weakly semirecursive set A. By our next theorem all of these numbers are equal.

Theorem 8.3 *For all $k, n > 0$, $d_k(n) = f_k(n) = r_k(n) = s_k(n) = w_k(n)$.*

Proof: We know that $d_k(n) = f_k(n)$. First we show that $r_k(n), s_k(n) \geq f_k(n)$. For ease of readability we identify Parity_k^A with its characteristic function.

Claim: There is an r.e. semirecursive set A such that Parity_k^A is not $(f_k(n) - 1, n)$-verbose.

Proof: Let $m = f_k(n)$ and let $v_1, \ldots, v_m$ be a sequence of vectors of length n that proves $f_k(n) \geq m$. We want to enumerate a set A, make A semirecursive, and satisfy the following requirements:

(R_e) Parity_k^A is not $(m-1, n)$-verbose via φ_e.

Let M be a set of kn distinct numbers $\{x_{i,j}\}_{1\le i\le k, 1\le j\le n}$. It is easy to see that there is an increasing sequence $S_1, \ldots, S_m$ of subsets of M such that

$$v_l = (\text{Parity}_k^{S_l}(x_{11}, \ldots, x_{1k}), \ldots, \text{Parity}_k^{S_l}(x_{n1}, \ldots, x_{nk})), \text{ for } l = 1, ..., m. \qquad (*)$$

Thus, for any interval I consisting of kn numbers there is a numbering $\{x_{i,j}\}_{1\le i\le k, 1\le j\le n}$ of the elements of I such that there exists an increasing sequence $I_1 \subset I_2 \subset \cdots \subset I_m$ of end segments of I such that $(*)$ is satisfied by taking $S_l = I_l$.

Now it is clear how our basic strategy for satisfying the requirement (R_e) works: We set up an interval I consisting of kn numbers as above, and we enumerate I_l into A at stage s if l is minimal such that $v_l \notin W_{\varphi_e(y_1,\ldots,y_n)}$ where each $y_j = (x_{j1}, \ldots, x_{jk})$.

It is well-known and easy to see that A is semirecursive (even co-retraceable) if we establish the "dump property," which says that if at stage s a number x is enumerated into A_s then all y such that $x < y \le s \ \wedge \ y \notin A_{s-1}$ are enumerated into A_s.

The basic strategies for (R_e) can be combined with dumping in an easy finite injury construction (of the kind we encountered in Section 4). We leave the remaining details to the reader. This completes the proof of the claim. $\Box$

Because any r.e. or semirecursive set is weakly semirecursive we complete the proof of the theorem by showing that $w_k(n) \le f_k(n)$. Suppose that A is an initial segment of the r.e. partial ordering $\sqsubset$. Consider the ordering restricted to a finite set X. We call $E \subseteq X$ an initial segment of X, iff $x \sqsubset y$ for all $x \in E$ and $y \in X - E$. Initial segments are linearly ordered by inclusion and can be enumerated uniformly given a canonical index of X. Note that $A \cap X$ is an initial segment of X.

Now we can enumerate a sequence $W_{f(y_1,\ldots,y_n)}$ (with f total recursive) as follows: Let

$$X = \{x_1, \ldots, x_k : (\exists i)[y_i = (x_1, \ldots, x_k)]\}.$$

Run an algorithm that enumerates all initial segments E of X, and for each such E enumerate $(\text{Parity}_k^E(y_1), \ldots, \text{Parity}_k^E(y_n))$ into $W_{f(y_1,\ldots,y_n)}$.

Because $A \cap X$ is an initial segment of X, we have $(\text{Parity}_k^A(y_1), \ldots, \text{Parity}_k^A(y_n)) \in W_{f(y_1,\ldots,y_n)}$. Let $E_1 = \emptyset \subset E_2 \subset \cdots \subset E_r = X$ be the sequence of all initial segments of X, and define $v_i = (\text{Parity}_k^{E_i}(y_1), \ldots, \text{Parity}_k^{E_i}(y_n))$, for $1 \le i \le r$. Then $v_i[j] = \chi_{E_i}(x_1) + \cdots + \chi_{E_i}(x_k) \bmod 2$, where $y_j = (x_1, \ldots, x_k)$. Because the E_i's are linearly ordered by inclusion, it follows that $|\{i : 1 \le i < r \ \wedge \ v_i[j] \ne v_{i+1}[j]\}| \le k$. Also $v_1 = 0^n$. Thus there are at most $f_k(n)$ different such vectors, and it follows that Parity_k^A is $(f_k(n), n)$-verbose via f.

Therefore, $w_k(n) \le f_k(n)$, and the proof of the theorem is completed. ∎

9 Verboseness and Frequency Computation

Frequency computation was introduced by Rose [37] in the early sixties. For $m, n \ge 0$, a set $A \subseteq \omega$ is called (m,n)-*recursive* iff there exists a total recursive function f such that for any sequence of pairwise distinct numbers $x_1, \ldots, x_n$

$$f(x_1, \ldots, x_n) = (y_1, \ldots, y_n) \Rightarrow |\{i : \chi_A(x_i) = y_i\}| \ge m.$$

The class of (m,n)-recursive sets is denoted $\Omega(m,n)$. Informally, the characteristic function of an (m,n)-recursive set A can be approximated on each n-tuple of pairwise distinct numbers in such a way that at least m components of the approximation are correct.

Myhill (see [34, p. 393]) asked if A had to be recursive if m was close to n (i.e., if χ_A can be approximated with frequency $\frac{m}{n}$ close to 1). Trakhtenbrot [41] answered this question by proving that A is recursive if $\frac{m}{n} > \frac{1}{2}$. Kinber [26] studied the aspect of uniformity in Trakhtenbrot's Theorem (see [23, 28] for a recent survey and generalizations of Trakhtenbrot's and Kinber's work). Degtev [17] began the study of the inclusion problem for frequency classes $\Omega(m,n)$ with $\frac{m}{n} \leq \frac{1}{2}$. For a solution of the equality problem and new results on the inclusion problem as well as several applications we refer the reader to [31]. We would like to mention that the decidability of the inclusion problem for frequency classes is still open.

An obvious connection with verboseness is given by the fact that $\Omega(1,n) = V_S(2^n - 1, n)$. Therefore $\bigcup_{m,n\geq 1} \Omega(m,n) = \bigcup_{n\geq 1, m<2^n} V_S(m,n)$. In [31] this relationship has been determined more precisely:

Fact 9.1 *For $1 \leq 2h \leq k$, let $t_{h,k}(n)$ denote the least m such that every (h,k)-recursive set is strongly (m,n)-verbose. Then, $t_{h,k}(n) = S(n, k-h+1) = \Theta(n^{k-h})$.*

Combining our results with a result of [31], we obtain the following characterization of the r.e. semirecursive sets:

Corollary 9.2 *An r.e. set is semirecursive iff it is strongly $(\lfloor\frac{3}{2}n\rfloor, n)$-verbose for some $n > 1$. If $m > \lfloor\frac{3}{2}n\rfloor$ then there is an r.e. strongly (m,n)-verbose set which is not semirecursive.*

Proof: In [31, Corollary 15] it is shown that the r.e. $(1,2)$-recursive sets coincide with the r.e. semirecursive sets (this follows also from our Fact 4.6). Because $\Omega(1,2) = V_S(3,2)$ and, by Corollary 6.2, $V_S(3,2) = V_S(m,n)$ for all m,n such that $1 < \frac{m}{n} \leq \frac{3}{2}$, the first part of the corollary follows. For the second part note that, by Theorem 4.7 (using the fact that $\{0,1\}^2$ is $(5,3)$-good), there is an r.e. set which is strongly $(5,3)$-verbose but not strongly $(3,2)$-verbose, and hence not semirecursive. By Corollary 6.2 this set is strongly (m,n)-verbose for all $m > \lfloor\frac{3}{2}n\rfloor$. ∎

Appendix: $E_{m,n} = \{(m,n)\}$ for all $m \in \{2n, \ldots, 2^n - 1\}$

To prove this result, it suffices to construct suitable (m,n)-good sets as specified in the following proposition. Recall from Lemma 4.2 that $G_V(n)$ denotes the least m such that V is (m,n)-good.

Proposition A.1 *If $2n \leq m < 2^n$, then there exist $V, W \subseteq \{0,1\}^n$ such that $|V| = G_V(n) = m$, $|W| = G_W(n) = m+1$, and $g_V(i) = g_W(i)$ for all $i < n$.*

Proof: The basic idea in this construction is to choose V and W such that $W = V \cup \{w\}$ for some vector $w \in \{0,1\}^n - V$. This vector w will be selected such that any vector v having Hamming distance 1 from w is in V. It follows that for any nontrivial projection P there exists $v \in V$ such that $P(v) = P(w)$, and therefore $P(V) = P(W)$. So, $g_V(i) = g_W(i)$ for all $i < n$. It remains to construct a suitable (m,n)-good set V consisting of m elements. We take three cases.

Case 1: $n \geq 5$ and $2n \leq m \leq 4n$.

Case 2: $n \geq 7$ and $4n+1 \leq m \leq 2^{n-2}+3$.

Case 3: $n \geq 3$ and $2^{n-2}+4 \leq m \leq 2^n - 1$.

If $2n \leq m < 2^n$, then (m,n) is covered by at least one case. In cases 1 and 3 we will explicitly write down the sets V and W. Case 2 is the interesting one; there we will use an inductive construction that once more exploits Fact 2.6.

For fixed n, let $v_{i,0}$ denote the vector $1^i 0^{n-i}$, and let $v_{i,j}$ denote the jth rotational right-shift of $v_{i,0}$. Examples with $n=5$: $v_{0,j} = 00000$ for all j, $v_{2,2} = 00110$, $v_{2,4} = 10001$, $v_{4,0} = 11110$ and $v_{1,3} = 00010$.

Construction for case 1: Let

$$\begin{aligned} m &= an + b \text{ with } 0 \leq b < n, \\ V &= \{v_{i,j} : (1 \leq i \leq a) \text{ or } (i = a+1 \text{ and } (\exists k < b)[j = \lfloor kn/b \rfloor])\}, \\ W &= V \cup \{0^n\}. \end{aligned}$$

Note that $a \geq 2$. Here it is convenient to number the coordinates from 0 to $n-1$. Let $\oplus$ denote addition modulo n.

We show that V is (m,n)-good. To do this we need to bound $g_V(k)$. Consider a projection onto $I \subset \{0,\ldots,n-1\}$ with $k = |I| < n$. For any $j \notin I$ and any $i > 0$, the projections $P^I(v_{i,j})$ and $P^I(v_{i-1,j\oplus 1})$ are equal. Therefore an upper bound for $|P^I(V)|$ is established by counting only those vectors $v_{i,j} \in W$ for which $j \in I$.

I is the union of contiguous, disjoint sets $J_0,\ldots,J_h$ such that $J_{h'} = \{x_{h'}, x_{h'} \oplus 1, \ldots, x_{h'} \oplus y_{h'}\}$ and $x_{h'} \oplus y_{h'} \oplus 1 \notin I$ for $h' = 0,\ldots,h$. The number $c_{h'}$ of vectors $v_{i,j}$ with $i > 0$ and $j \in J_{h'}$ is bounded by $\left\lceil (y_{h'}+1)\cdot(a+\frac{b}{n})\right\rceil$ (note that for $x, l \in \omega$ there are at most $\lceil (l+1)\frac{b}{n}\rceil$ rational numbers of the form $\frac{kn}{b}$ in the interval $[x, x+l+1)$; hence there are at most $\lceil (l+1)\frac{b}{n}\rceil$ integers of the form $\lfloor \frac{kn}{b} \rfloor$ in the interval $[x, x+l]$). Since $P^I(v_{2,x_{h'}\oplus y_{h'}}) = P^I(v_{1,x_{h'}\oplus y_{h'}})$ we even get $c_{y_{h'}} < \left\lceil (y_{h'}+1)\cdot(a+\frac{b}{n})\right\rceil$. Thus we obtain the following upper bound (the 1 counts the null vector):

$$\begin{aligned} g_V(k) &\leq c_{y_0} + \cdots + c_{y_h} + 1 \\ &< ((y_0+1) + \cdots + (y_h+1))\cdot(a+\frac{b}{n}) + 1 \\ &= k\cdot(a+\frac{b}{n}) + 1. \end{aligned}$$

If $k = n$ the same upper bound holds: $|V| = m < n \cdot (a + \frac{b}{n}) + 1$. Therefore for any j and any partition $n_1 + \cdots + n_j = n$ we get

$$g_V(n_1) + \cdots + g_V(n_j) < (n_1 + \cdots + n_j) \cdot (a + \frac{b}{n}) + j \leq an + b + j = m + j.$$

Hence V is (m, n)-good (note that the first inequality is strict).

Construction for case 2: This is shown by induction on m. By case 1 or by the inductive hypothesis, there is an $(m - 2n, n)$-good set U consisting of $m - 2n$ elements. By Fact 2.6. we may assume that U is closed. Since $m - 2n < 2^{n-2}$, U contains no vector $v_{i,j}$ for $i \geq n - 2$. Now let $V = U \cup \{v_{i,j} : n - 2 \leq i \leq n - 1, 1 \leq j \leq n\}$ and let $W = V \cup \{1^n\}$.

$V - U$ has $2n$ elements, therefore $|V| = m$. Furthermore, $g_{V-U}(k) = 2k$ and $g_V(k) \leq g_U(k) + 2k$ for all $k \in \{1, \ldots, n-1\}$. Now G_V satisfies

$$\begin{aligned} m = |V| &\leq G_V(n) \\ &\leq \max\{g_V(n_1) + \cdots + g_V(n_j) + 1 - j : n_1 + \cdots + n_j = n\} \\ &\leq \max\{g_U(n_1) + 2n_1 + \cdots + g_U(n_j) + 2n_j + 1 - j : n_1 + \cdots + n_j = n\} \\ &\leq 2n + \max\{g_U(n_1) + \cdots + g_U(n_j) + 1 - j : n_1 + \cdots + n_j = n\} \\ &= 2n + G_U(n) \;=\; 2n + (m - 2n) \;=\; m. \end{aligned}$$

Hence V is (m, n)-good.

Construction for case 3: Let V be any set of cardinality m that contains $v_{i,1}$ and $v_{i,2}$ for all i but does not contain $v_{0,0}$. Such a set V exists because $2n \leq 2^{n-2} + 4 \leq m \leq 2^n - 1$. Let $W = V \cup \{0^n\}$.

Let P be a nontrivial projection that omits coordinate i. Thus $P(v_{2,i}) = P(v_{1,i\oplus 1})$ and $|P(V)| \leq m - 1$. Therefore $g_V(n-1) \leq m - 1$, $g_V(n-1) + g_V(1) \leq m + 1$ and $g_V(n - i) + g_V(i) \leq 2^{n-i} + 2^i \leq 2^{n-2} + 4 \leq m + 1$, where $2 \leq i \leq n - 2$. So $g_V(n_1) + g_V(n_2) \leq m + 1$ for $n_1 + n_2 = n$. In the same way as it was shown that every set is $(2^n, n)$-verbose one verifies that $g_V(n_1) + \cdots + g_V(n_k) \leq m + k - 1$, for $k \geq 3$ and $n_1 + \cdots + n_k = n$. Hence V is (m, n)-good. ▮

Theorem A.2 *If $2n \leq m \leq 2^n - 1$, then $E_{m,n} = \{(m, n)\}$.*

Proof: The theorem follows from the statement

If $n' \leq n$, $2n \leq m \leq 2^n - 1$, and $(m, n) \neq (m', n')$, then $V(m, n) \neq V(m', n')$.

Suppose that m, n, m', n' satisfy the premise. Then, by Proposition A.1, there exist sets V and W such that $|V| = m = |W| - 1$, $G_V(k) = G_W(k)$ for all $k < n$, and V is (m, n)-good. If W is (m', n')-good, then, by Theorem 4.3, W witnesses the existence of an (m', n')-verbose non-(m, n)-verbose set. If W is not (m', n')-good then V is not (m', n')-good either, because $n' \leq n$ and $(m', n') \neq (m, n)$. So, by Theorem 4.3, V witnesses the existence of an (m, n)-verbose set that is not (m', n')-verbose. ▮

Acknowledgments: We would like to thank Bill Gasarch for helpful comments and for proofreading a preliminary version of this paper. Thanks are also due to Jim Owings for encouragement at an early stage of this work. We are grateful to both referees for many detailed suggestions that improved the presentation.

References

[1] N. Alon. On the density of sets of vectors. *Discrete Mathematics*, 46:199–202, 1983.

[2] A. Amir, W. I. Gasarch. Polynomial terse sets. *Information and Computation*, 77:37–56, 1988.

[3] R. Beigel. Query-limited reducibilities. Ph.D. thesis, Stanford University, Stanford, CA, USA, 1987.

[4] R. Beigel. A structural theorem that depends quantitatively on the complexity of SAT. In *Proceedings of the Second Annual Structure in Complexity Theory Conference*. IEEE Computer Society Press, pp. 28–32, 1987.

[5] R. Beigel. NP-hard Sets are P-Superterse unless R = NP. Technical report 88-04, Dept. of Computer Science, The Johns Hopkins University, 1988.

[6] R. Beigel. Bi-immunity results for cheatable sets. *Theoretical Computer Science*, 73:249–263, 1990.

[7] R. Beigel. Bounded queries to SAT and the Boolean hierarchy. *Theoretical Computer Science*, 83:199–223, 1991.

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

[9] R. Beigel, W. I. Gasarch, L. Hay. Bounded query classes and the difference hierarchy. *Arch. Math. Logic*, 29:69–84, 1989.

[10] A. Blumer, A. Ehrenfeucht, D. Haussler, M. K. Warmuth. Learnability and the Vapnik-Chervonenkis dimension. *J. of the ACM*, 36:929-965, 1989.

[11] J. A. Bondy. Induced subsets. *J. of Combinatorial Theory*, 12:201–202, 1972.

[12] B. Bollobás. *Extremal Graph Theory*. Academic Press, London, 1978.

[13] B. Bollobás. *Combinatorics*. Cambridge University Press, Cambridge, 1986.

[14] B. Bollobás, I. Leader. Exact face-isoperimetric inequalities. *Europ. J. Combinatorics*, 11:335-340, 1990.

[15] Yu. D. Burtin. Connection probability of a random subgraph of an n-dimensional cube. *Problems of Information Transmission*, 13:147–152, 1977.

[16] S. Clarke, J. C. Owings, Jr., J. Spriggs. Trees with full subtrees. In *Proceedings of the Sixth Southeastern Conference on Combinatorics, Graph Theory, and Computing.* Congressus Numerantium, No. XIV, Utilitas Math., Winnipeg, Man., pp. 169–172, 1975.

[17] A. N. Degtev. On (m,n)-computable sets. In *Algebraic Systems (Edited by D.I. Moldavanskij).* Ivanova Gos. Univ. pp. 88–99, 1981. (Russian) (MR 86b:03049)

[18] R. L. Epstein, R. Haas, R. L. Kramer. Hierarchies of sets and degrees below $0'$. In *Logic Year 1979-80 (Edited by M. Lerman, J.H. Schmerl, R.I. Soare).* Lecture Notes in Mathematics 859, Springer-Verlag, Berlin, pp. 32–48, 1981.

[19] Yu. L. Ershov. On a hierarchy of sets I. *Algebra and Logic*, 7:25–43, 1968.

[20] P. Frankl. On the trace of finite sets. *J. of Combinatorial Theory, Ser. A*, 34, 41–45, 1983.

[21] W. I. Gasarch. Bounded queries in recursion theory: a survey. In *Proceedings of the Sixth Annual Structure in Complexity Theory Conference.* IEEE Computer Society Press, pp. 62–78, 1991.

[22] D. Haussler, E. Welzl. ϵ-nets and simplex range queries. *Discrete Comput. Geom.*, 2:127–151, 1987.

[23] V. Harizanov, M. Kummer, J. C. Owings, Jr. Frequency computation and the cardinality theorem. *J. Symb. Log.*, 57:682–687, 1992.

[24] C. G. Jockusch, Jr. Semirecursive sets and positive reducibility. *Trans. Amer. Math. Soc.*, 131:420–436, 1968.

[25] C. G. Jockusch, Jr., J. C. Owings, Jr. Weakly semirecursive sets. *J. Symb. Log.*, 55:637–644, 1990.

[26] E. B. Kinber. On frequency calculations of general recursive predicates. *Sov. Math. Dokl.*, 13:873–876, 1972.

[27] E. B. Kinber. On frequency real-time computations. In *Teoriya Algorithmov i Programm*, Vol. 2 (Edited by Ya. M. Barzdin). Latv. Valst. Gos. Univ. pp. 174–182, 1975. (Russian) (MR 58:3624, Zbl 335:02023)

[28] M. Kummer. A proof of Beigel's cardinality conjecture. *J. Symb. Log.*, 57:677–681, 1992.

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

[30] M. Kummer, F. Stephan. Weakly semirecursive sets and r.e. orderings. *Annals of Pure and Applied Logic* 60:133–150, 1993.

[31] M. Kummer, F. Stephan. Recursion theoretic properties of frequency computation and bounded queries. Submitted for publication. An extended abstract appeared in *Proceedings of the Third Kurt Gödel Colloquium, KGC'93*, Lecture Notes in Computer Science 713, pp. 243–254, Springer-Verlag, Berlin, 1993.

[32] A. Lubotzky, R. Phillips, P. Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988.

[33] G. A. Margulis. Explicit Group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problems of Information Transmission*, 24:39–46, 1988.

[34] R. McNaughton. The theory of automata, a survey. *Advances of Computers*, 2:379–421, 1961.

[35] P. Odifreddi. *Classical Recursion Theory.* North-Holland, Amsterdam, 1989.

[36] J. C. Owings, Jr. A cardinality version of Beigel's nonspeedup theorem. *J. Symb. Log.*, 54:761–767, 1989.

[37] G. F. Rose. An extended notion of computability. In *Abstr. Intern. Congr. for Logic, Meth., and Phil. of Science*, Stanford, California, 1960.

[38] N. Sauer. On the density of families of sets. *J. of Combinatorial Theory, Ser. A*, 13:145–147, 1972.

[39] S. Shelah. A combinatorial problem: Stability and order for models and theories in infinitary languages. *Pacific J. of Mathematics*, 41:247–261.

[40] R. I. Soare. *Recursively Enumerable Sets and Degrees.* Springer-Verlag, Berlin, 1987.

[41] B. A. Trakhtenbrot. On frequency computation of functions. *Algebra i Logika*, 2:25–32, 1963. (Russian)

[42] V. N. Vapnik, A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16:264–280, 1971.