THE JOURNAL OF SYMBOLIC LOGIC
Volume 57, Number 2, June 1992

A PROOF OF BEIGEL'S CARDINALITY CONJECTURE

MARTIN KUMMER

In 1986, Beigel [Be87] (see also [Od89, III.5.9]) proved the nonspeedup theorem: if $A, B \subseteq \omega$, and as a function of 2^n variables $(\chi_A(x_1), \ldots, \chi_A(x_{2^n}))$ can be computed by an algorithm which makes at most n queries to B, then A is recursive (informally, 2^n parallel queries to a nonrecursive oracle A cannot be answered by making n sequential (or "adaptive") queries to an arbitrary oracle B). Here, 2^n cannot be replaced by $2^n - 1$. In subsequent papers of Beigel, Gasarch, Gill, Hay, and Owings the theory of "bounded query classes" has been further developed (see, for example, [BGGOta], [BGH89], and [Ow89]). The topic has also been studied in the context of structural complexity theory (see, for example, [AG88], [Be90], and [JY90]).

If $A \subseteq \omega$ and $n \geq 1$, let $\#_n^A(x_1, \ldots, x_n) = \#\{i: x_i \in A\} = \sum_{i=1}^n \chi_A(x_i)$. Beigel [Be87] stated the powerful "cardinality conjecture" (CC): if $A, B \subseteq \omega$, and $\#_{2^n}^A$ can be computed by an algorithm which makes at most n queries to B, then A is recursive. Owings [Ow89] verified CC for $n = 1$, and, for $n > 1$, he proved that A is recursive in the halting problem. We prove that CC is true for all n.

NOTATION. $\omega = \{0, 1, 2, 3, \ldots\}$. $W_i \subseteq \omega$ is the ith r.e. set in the standard enumeration of all r.e. sets. For unexplained recursion theoretic notation the reader is referred to [Od89]. $\#A$ denotes the cardinality of the set A. χ_A is the characteristic function of A. $\{0,1\}^{<\omega}$ is the set of all finite strings of zeros and ones. λ is the empty string, $|s|$ denotes the length of string s, and $|\lambda| = 0$. $s \sqsubseteq t$ means that s is an initial segment of t. $s(n) = b$ iff $n < |s|$ and b is the $(n+1)$st symbol of s. A *tree* T is a subset of $\{0,1\}^{<\omega}$ which is closed under initial segments; $s \in T$ is called a *node* of T. $t \in \{0,1\}^\omega$ is a *branch* of T iff every finite initial segment of t is a node of T; t is the characteristic function of the set $\{n \in \omega: t(n) = 1\}$.

We wish to prove that if $\#_{2^n}^A$ can be computed with n queries to some set B, then A is recursive. As in the proof of Beigel's nonspeedup theorem we need to view functions computed by bounded queries in a different light.

LEMMA 0 ([Be87], [BGGOta]). *If a function f can be computed with n queries to some set B, then there exists a set S of at most 2^n partial recursive functions such that for each argument x there is a function $h \in S$ such that $g(x) = f(x)$.*

PROOF. Assume that f is computed by the oracle Turing machine M^B such that, for any oracle X and any input, M^X makes at most n queries to X. For each string

Received April 2, 1991; revised June 13, 1991.

$w \in \{0,1\}^n$ define a partial recursive function h_w and let $S = \{h_w: w \in \{0,1\}^n\}$. $h_w(x)$ is computed by running $M^{(\cdot)}$ with input x using the bits of w consecutively for the query answers. Since one of the sequences is correct (i.e. would be the sequence of answers if B was used for the oracle), $h_w(x)$ is equal to $f(x)$ for some w. □

By Lemma 0 Beigel's conjecture is reduced to the following theorem:

CARDINALITY THEOREM (CT). *Let $A \subseteq \omega$ and $m \geq 1$. Assume that there exists a recursive function $g(x_1,\ldots,x_m)$ such that, for any m-tuple $(x_1,\ldots,x_m)$ of distinct natural numbers, (1) $W_{g(x_1,\ldots,x_m)} \subset \{0,1,\ldots,m\}$ (note that it is a proper inclusion), and (2) $\#_m^A(x_1,\ldots,x_m) \in W_{g(x_1,\ldots,x_m)}$. Then A is recursive.*

PROOF. Corresponding to a set A and a recursive function g satisfying the antecedent of CT we have an r.e. tree of possibilities defined as follows:

$$T_g = \left\{ t \in \{0,1\}^{<\omega}: \forall x_1,\ldots,x_m \left[x_1 < \cdots < x_m < |t| \rightarrow \sum_{i=1}^{m} t(x_i) \in W_{g(x_1,\ldots,x_m)} \right] \right\}.$$

Note that, by (2), χ_A is a branch of T_g. Our proof will proceed in two phases: First, we prove a general lemma about conditions under which r.e. trees actually have every branch recursive. We then show that T_g satisfies these conditions. The latter proof is purely combinatorial and uses a Ramsey-type theorem.

Let $B_n = \{0,1\}^{\leq n}$ (the full binary tree of height n). $*$ denotes concatenation of strings. $f: B_n \rightarrow T$ is an *embedding* of B_n into T iff $\forall s[|s| < n \rightarrow [f(s) * 0 \sqsubseteq f(s * 0) \wedge f(s) * 1 \sqsubseteq f(s * 1)]]$. B_n is embeddable into T *above* $e \in T$ iff there exists an embedding f of B_n into T such that $e \sqsubseteq f(\lambda)$. The *rank* of T, denoted by $\mathrm{rk}(T)$, is the supremum of all n such that B_n is embeddable into T.

LEMMA 1. *If T is an r.e. tree of finite rank, then every branch of T is recursive.*

PROOF. Suppose that T is r.e., $\mathrm{rk}(T)$ is finite, and t is a branch of T. Let k_0 be the supremum of all n such that B_n is embeddable above every node $e \sqsubseteq t$. Then $k_0 \leq \mathrm{rk}(T)$. Choose a node $e_0 \sqsubseteq t$ such that B_{k_0+1} is not embeddable above e_0. We claim that t is recursive via the following algorithm:

$t(x)$ is computed for $x > |e_0|$ by enumerating T and searching for an embedding f of B_{k_0} into T above e_0 such that $|f(\lambda)| > x$. Output $(f(\lambda))(x)$. If $x \leq |e_0|$, then $t(x)$ is looked up in a finite table.

By choice of k_0 the algorithm terminates. We show that $f(\lambda) \sqsubseteq t$, for every embedding f of B_{k_0} into T above e_0; thus the algorithm is correct. Suppose for a contradiction that there exists an embedding f of B_{k_0} into T above e_0 such that $f(\lambda) \not\sqsubseteq t$. Let $d \in T$ be the maximal common prefix of $f(\lambda)$ and t. Clearly $e_0 \sqsubseteq d$. Since B_{k_0} is embeddable above $d * (t(|d|))$, it follows that d is the root of an embedded B_{k_0+1}, contradicting the choice of e_0 (Figure 1 visualizes this case). □

Now we turn to the combinatorial part of the proof. First we need a Ramsey-theoretic lemma. For a generalization see Deuber [D75].

LEMMA 2 ([D75]). *For any 2-coloring of B_{2k} there exists an embedding g of B_k into B_{2k} such that all nodes of $g(B_k)$ are monochromatic.*

PROOF. By induction on $m + n$, any coloring of B_{m+n} with two colors, green and red, contains either an embedded green B_m, or an embedded red B_n. □

Note that the bound in Lemma 2 is tight. In our final lemma we show that for any tree of large rank there exist n distinct numbers $x_1,\ldots,x_n$ and $n + 1$ nodes

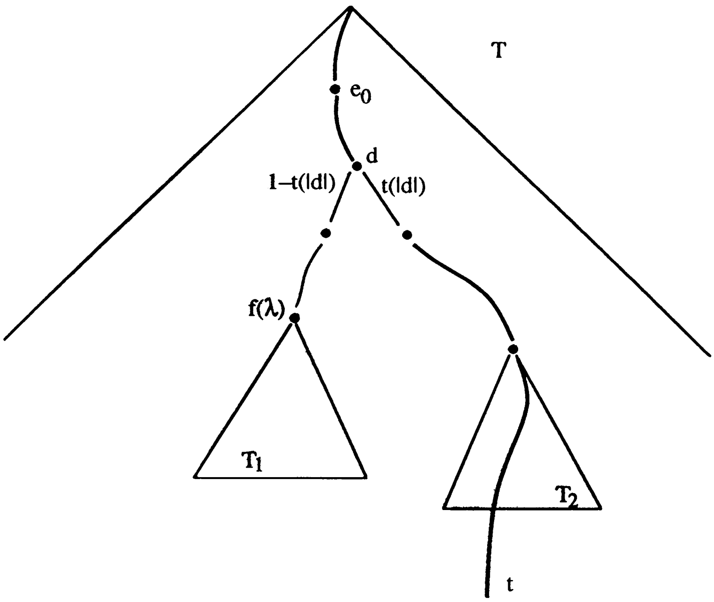

FIGURE 1

$t_1, \dots, t_{n+1}$ such that $\sum_{i=1}^{n} t_j(x_i)$ takes on all values from $\{0, 1, \dots, n\}$, for $1 \le j \le n+1$. See Figure 2 for a typical example.

For $n \ge 1$ and $1 \le i \le 2n-1$, let $h(n, 2n-1) = 0$ and $h(n, i-1) = 2(h(n,i)+1)$, and let $k(n) = h(n, 0) = 4^n - 2$.

LEMMA 3. *For each $n \ge 1$ and any tree T such that $B_{k(n)}$ is embeddable into T there exist nodes $t_1, \dots, t_{n+1}$ of T and numbers $x_1 < \cdots < x_n$, and $b \in \{0, 1\}$, such that:*

(+) *for $j = 1, \dots, n+1$ and $i = 1, \dots, n$:*
$$t_j(x_i) = 1 - b \text{ for } i < j \quad \text{and} \quad t_j(x_i) = b \text{ for } i \ge j.$$

In particular, $\{\sum_{i=1}^{n} t_j(x_i) \colon 1 \le j \le n+1\} = \{0, 1, \dots, n\}$.

PROOF. Let the tree T be given, and suppose that $B_{k(n)}$ is embeddable into T via f_0.

For $i = 1, \dots, 2n-1$ we define inductively $w_i, s_i \in T$, $b_i \in \{0, 1\}$, and $f_i \colon B_{h(n,i)} \to T$ as follows: By the induction hypothesis we are given $f_{i-1} \colon B_{h(n,i-1)} \to T$. Let s be a leaf node of $B_{h(n,i-1)}$ such that $f_{i-1}(s)$ has *maximal* length. Let $s_i = f_{i-1}(s)$. Then s_i induces a 2-coloring of $B_{h(n,i-1)}$, as follows: each inner node e is colored by $s_i(|f_{i-1}(e)|)$, and each leaf node is colored arbitrarily, say by 0. By Lemma 2 and the definition of h, there exists an embedding g of $B_{h(n,i)+1}$ into $B_{h(n,i-1)}$ such that $g(B_{h(n,i)+1})$ is monochromatic. Let $w_i = f_{i-1}(g(\lambda))$, and let $b_i = s_i(|w_i|)$. Define f_i by $f_i(s) = f_{i-1}(g((1-b_i) * s))$, for $|s| \le h(n,i)$.

Note that $f_i(B_{h(n,i)}) \subseteq f_{i-1}(B_{h(n,i-1)})$, $w_i * (1-b_i) \sqsubseteq w_{i+1}, s_{i+1}$, and $s_j(|w_i|) = b_j$ for $i \ge j$. By the pigeonhole principle, there exist $b \in \{0, 1\}$ and n indices $1 \le i_1 < i_2 < \cdots < i_n \le 2n-1$ such that $s_{i_m}(|w_{i_m}|) = b$. Let $t_m = s_{i_m}$ and $x_m = |w_{i_m}|$ for $1 \le m \le n$, and $t_{n+1} = w_{i_n} * (1-b)$. Now (+) follows immediately.

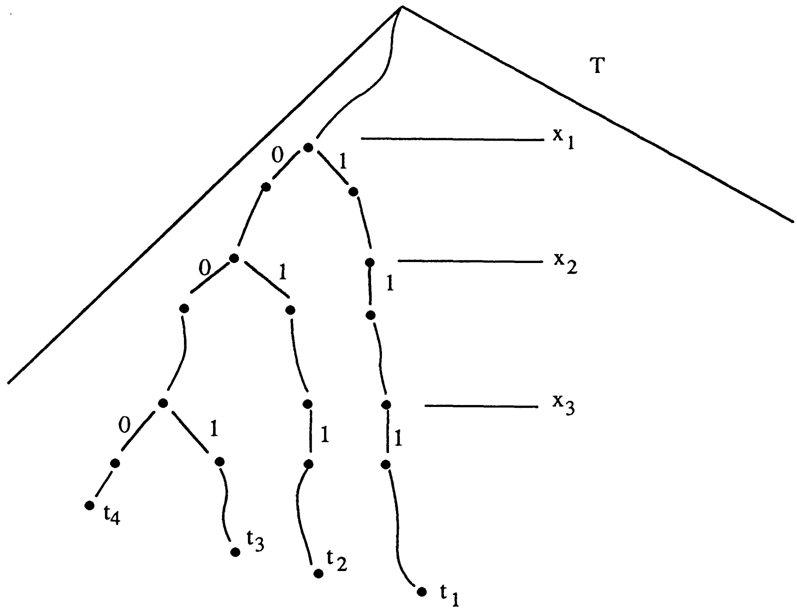

FIGURE 2

REMARKS. (1) Owings observed that already $\{\sum_{i=1}^{n} s_j(|w_i|)\colon 1 \leq j \leq n+1\} = \{0, 1, \ldots, n\}$, which is sufficient for the proof of CT. However, the more general property $(+)$ may be useful for other applications.

(2) Note that, for each m-element subset $\{x_{i_1} < \cdots < x_{i_m}\}$ $(m \geq 1)$ of $\{x_1, \ldots, x_n\}$,

$$\left\{\sum_{k=1}^{m} t_j(x_{i_k})\colon j \in \{i_1, \ldots, i_m, i_m + 1\}\right\} = \{0, 1, \ldots, m\}. \qquad \square$$

Now we are ready to finish the proof of CT. Suppose that $\mathrm{rk}(T_g) \geq k(m)$. Then, by Lemma 3, there exist $t_1, \ldots, t_{m+1} \in T_g$ and numbers $x_1 < \cdots < x_m$ such that

$$\left\{\sum_{i=1}^{m} t_j(x_i)\colon 1 \leq j \leq m+1\right\} = \{0, 1, \ldots, m\}.$$

From the definition of T_g it follows that $\{0, 1, \ldots, m\} \subseteq W_{g(x_1, \ldots, x_m)}$, which contradicts hypothesis (1) of CT.

Thus $\mathrm{rk}(T_g) < k(m)$. Therefore, by Lemma 1, each branch of T_g is recursive. Since χ_A is a branch of T_g, A is recursive. $\square$

REMARKS. (1) In the same way, some interesting variants of CT can be obtained:

(a) Hypothesis (2) can be replaced by:

(2′) $W_{g(x_1, \ldots, x_m)}$ contains the length of the maximal block of consecutive 1's occurring in the string $\chi_A(x_1) \cdots \chi_A(x_m)$.

The proof uses property $(+)$ of Lemma 3.

(b) Given $A \subseteq \omega$, call a finite set $D \subseteq \omega$ *A-biased* iff

$$\lceil \#D/2 \rceil < \max(\#(D \cap A), \#(D \cap \bar{A})).$$

(∗) Suppose $n \geq 3$, and one can compute for all numbers $x_1 < \cdots < x_n$ the canonical index of an A-biased subset of $\{x_1, \ldots, x_n\}$. Then A is recursive.

The proof uses Remark (2), above.

For any r.e. semirecursive set A there exists a recursive 4-place function g such that $W_{g(x_1,\ldots,x_4)}$ is an A-biased subset of $\{x_1 < \cdots < x_4\}$. As there exist nonrecursive r.e. semirecursive sets, (∗) does not hold if canonical indices are replaced by Σ_1-indices.

(2) A special case of CT arises if condition (1) is replaced by:

(1′) $\{0, n\} \not\subseteq W_{g(x_1,\ldots,x_n)}$.

Precursors of this variant are due to Trakhtenbrot [Tr63] and Kinber [Ki72, Theorem 5]. It has a direct proof which is much easier than the proof of Theorem 1. For more details and additional information the reader is referred to [HKO92].

Acknowledgement. I would like to thank James C. Owings, Jr., for introducing me to the cardinality conjecture, and for some stimulating discussions via mail and e-mail. Thanks are also due to Valentina Harizanov for information on the work of Trakhtenbrot and Kinber. Finally, I would like to thank the referee for suggestions that improved the exposition.

REFERENCES

[AG88] A. Amir and W. I. Gasarch, *Polynomial terse sets*, ***Information and Computation***, vol. 77 (1988), pp. 37–56.

[Be87] R. Beigel, ***Query-limited reducibilities***, Ph.D. thesis, Department of Computer Science, Stanford University, Stanford, California, 1987.

[Be90] ———, *Bi-immunity results for cheatable sets*, ***Theoretical Computer Science***, vol. 73 (1990), pp. 249–263.

[BGGOta] R. Beigel, W. I. Gasarch, J. T. Gill, and J. C. Owings, Jr., *Terse, superterse and verbose sets*, ***Information and Computation*** (to appear).

[BGH89] R. Beigel, W. I. Gasarch, and L. Hay, *Bounded query classes and the difference hierarchy*, ***Archive for Mathematical Logic***, vol. 29 (1989), pp. 69–84.

[D75] W. Deuber, *A generalization of Ramsey's theorem for regular trees*, ***Journal of Combinatorial Theory***, ser. B, vol. 18 (1975), pp. 18–23.

[HKO92] V. Harizanov, M. Kummer, and J. C. Owings, Jr., *Frequency computations and the cardinality theorem*, this Journal, vol. 57 (1992), pp. 682–687.

[JY90] D. Joseph and P. Young, *Self-reducibility: Effects of internal structure on computational complexity*, ***Complexity theory retrospective: in honor of Juris Hartmanis on the occasion of his sixtieth birthday*** (A. L. Selman, editor), Springer-Verlag, Berlin, 1990, pp. 82–107.

[Ki72] E. B. Kinber, *Frequency calculations of general recursive predicates and frequency enumeration of sets*, ***Soviet Mathematics Doklady***, vol. 13 (1972), pp. 873–876.

[Od89] P. Odifreddi, ***Classical recursion theory***, North-Holland, Amsterdam, 1989.

[Ow89] J. C. Owings, Jr., *A cardinality version of Beigel's nonspeedup theorem*, this Journal, vol. 54 (1989), pp. 761–767.

[Tr63] B. A. Trakhtenbrot, *On the frequency computability of functions*, ***Algebra i Logika***, vol. 2 (1963), pp. 25–32. (Russian)

INSTITUT FÜR LOGIK, KOMPLEXITÄT UND DEDUKTIONSSYSTEME
UNIVERSITÄT KARLSRUHE
W-7500 KARLSRUHE 1, GERMANY

E-mail: kummer@ira.uka.de