ON k-STABLE FUNCTIONS

by

Andrzej Ehrenfeucht
Department of Computer Science
University of Colorado at Boulder
Boulder, Colorado 80309

Jan Mycielski
Department of Mathematics
University of Colorado at Boulder
Boulder, Colorado 80309

#CU-CS-111-77 August, 1977

Abstract

We prove that a k-continuous or a k-stable function cannot depend on more than $k4^{k-1}$ variables and related facts.

A function $f: \{0,1\}^n \to R$, where R is any set is called k-continuous iff for every $x = (x_1,\ldots,x_n) \in \{0,1\}^n$ there exists a sequence $1 \le i_1 < \ldots < i_p \le n$, where $p \le k$, such that for every $y = (y_1,\ldots,y_n) \in \{0,1\}^n$ if $(y_{i_1},\ldots,y_{i_p}) = (x_{i_1},\ldots,x_{i_p})$ then $f(y) = f(x)$. This property was studied in [1,2,3,5].

Now we will study a larger class of functions $f: \{0,1\}^n \to R$ called k-stable. To explain this property, for every $x = (x_1,\ldots,x_n) \in \{0,1\}^n$ and every i with $1 \le i \le n$ we put

$$x^i = (x_1,\ldots,x_{i-1},1-x_i,x_{i+1},\ldots,x_n) .$$

Now f is called k-stable iff for every $x = (x_1,\ldots,x_n) \in \{0,1\}^n$ there exist $1 \le i_1 < \ldots < i_p \le n$, where $p \le k$, such that for every $i \notin \{i_1,\ldots,i_p\}$, $1 \le i \le n$, we have $f(x^i) = f(x)$. Thus, of course, k-continuity implies k-stability.

Examples. 1. The function $f: \{0,1\}^4 \to \{0,1\}$ defined by $f(x) = 0$ if $x \in \{(0,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1), (1,0,0,0), (1,1,0,0), (1,0,1,0), (1,0,0,1)\}$, and $f(x) = 1$ otherwise, is 2-continuous.

2. The function $f: \{0,1\}^{10} \to \{0,1\}$ defined by $f(x_1,\ldots,x_{10}) = x_1$ if $x_1 = x_2$, $f(x_1,\ldots,x_{10}) = 0$ if $x_1 = x_3 = x_4 = 0$ or $x_1 = x_5 = x_6 = 0$ or $x_2 = x_7 = x_8 = 0$ or $x_2 = x_9 = x_{10} = 0$, and

$f(x_1,\ldots,x_{10}) = 1$ otherwise, is 3-continuous (see fig. 1). For other examples of k-continuous Boolean functions see [2], Notes 3, 4, and 5, and [3].

3. The function $f\colon \{0,1\}^4 \to \{0,1\}$ defined by $f(x) = 0$ if $x \in \{(0,0,0,0), (1,0,0,0), (1,1,0,0), (1,1,1,0), (1,1,1,1), (0,1,1,1), (0,0,1,1), (0,0,0,1)\}$, and $f(x) = 1$ otherwise, is 2-stable but not 2-continuous. (see fig. 2)

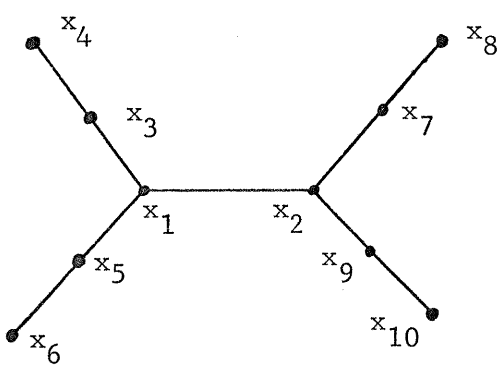

Fig. 1

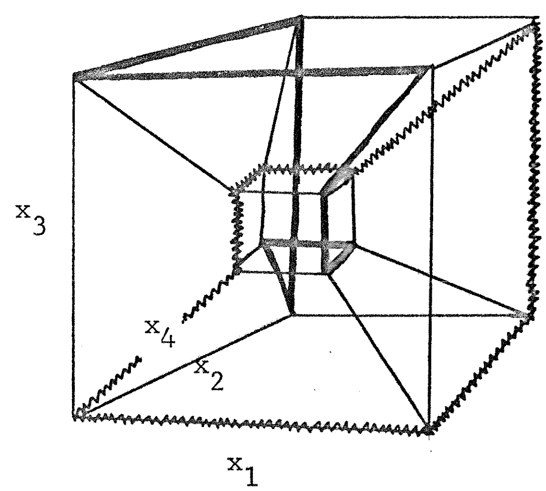

Fig 2.

A function $f\colon \{0,1\}^n \to R$ is said to <u>depend</u> on the variable x_i iff there exists a sequence $y = (y_1,\ldots,y_n) \in \{0,1\}^n$ such that $f(y) \neq f(y^i)$. And a function $f\colon \{0,1\}^n \to R$ is called <u>Boolean</u> iff $R \subseteq \{0,1\}$.

E.g.: the functions of Examples 1 and 3 depend on 4 variables and the function of Example 2 depends on 10 variables, and all are Boolean.

In [2] we have studied the maximum number of variables on which a k-continuous Boolean function can depend. It turns out that such a maximum exists and we will denote it here (unlike in [2]) by $\varphi_2(k)$. The following problem is still unsolved

(P_1) Does there exist for every $n < \varphi_2(k)$ a k-continuous Boolean function which depends just on n variables?

It is not hard to prove that $\varphi_2(1) = 1$ and $\varphi_2(2) = 4$ (see Example 1). By Example 2 we have $\varphi_2(3) \geq 10$. It seems that $\varphi_2(3) = 10$.

We shall also study functions $f: X \to R$, where X can be a proper subset of $\{0,1\}^n$. We shall say that f is <u>total</u> if $X = \{0,1\}^n$ and <u>partial</u> if $X \neq \{0,1\}^n$. For a partial f we shall say that f <u>depends</u> on the variable x_i if there exists a $y \in X$ such that $y^i \in X$ and $f(y) \neq f(y^i)$. Also f is called <u>Boolean</u> if $R \subseteq \{0,1\}$. It is called k-<u>continuous</u> if for every $x \in X$ there exists $1 \leq i_1 < \ldots < i_p \leq n$ such that $p \leq k$ and for every $y \in \{0,1\}^n$ if $(y_{i_1},\ldots,y_{i_p}) = (x_{i_1},\ldots,x_{i_p})$ then $y \in X$ and $f(y) = f(x)$. (In [2] this property was called regular k-continuity.) f is called k-<u>stable</u> if for every $x \in X$ there exists $1 \leq i_1 < \ldots < i_p \leq n$ such that $p \leq k$ and for all $i \notin \{i_1,\ldots,i_p\}$, $1 \leq i \leq n$, we have $x^i \in X$ and $f(x^i) = f(x)$.

(P_2) For which k, n, ℓ is it true that k-stability of $f: \{0,1\}^n \to \{0,1\}$ implies ℓ-continuity of f ? (For $k = \ell = n-1$ it is so.)

(P_3) What is the maximum height (see [3]) of a total k-stable function? (The maximum height of a total k-continuous function is k^2 as proven in [3].)

Let now $\varphi(k)$, $\varphi^*(k)$, or $\varphi_2^*(k)$, denote the maximum number of variables on which a k-continuous function which is total, partial or partial Boolean, respectively, can depend. Also let $\psi(k)$, or $\psi^*(k)$, denote the maximum number of variables on which a k-stable function which is total, or partial, respectively, can depend.

We shall prove that all these maxima exist. We have of course

$$\varphi_2(k) \le \varphi_2^*(k) \le \varphi^*(k) \le \psi^*(k) ,$$

$$\varphi_2(k) \le \varphi(k) \le \psi(k) \le \psi^*(k) \quad \text{and} \quad \varphi(k) \le \varphi^*(k) .$$

The main result of this paper is that $\psi^*(k) \le k4^{k-1}$.

(P_4) Is any of the above inequalities sharp for large enough k ?

In [2] (Theorem 17A and Note 6) we have proven that

$$2(k-2) + 4\binom{2(k-2)}{k-2} \le \varphi_2(k) \le \varphi_2^*(k) \le (2k-1)\binom{2(k-1)}{k-1} ,$$

and we gave (Theorem 23) a different combinatorial interpretation of the quantity $\varphi_2^*(k)$ (see also [4]). Again it is easy to prove that $\varphi_2^*(1) = 1$ and $\varphi_2^*(2) = 4$ and it seems that $\varphi_2^*(3) = 10$. The analogs of problem (P_1) for φ_2^* , φ , φ^* , ψ and ψ^* are also open.

Now we will prove that $\psi^*(1) = 1$. (Concerning $\psi(2)$ and $\psi^*(2)$ we know only that $4 \le \psi(2) \le \psi^*(2) \le 8$ (by Example 3 and the general fact $\psi^*(k) \le k4^{k-1}$ proved below)). First we need an auxiliary proposition. Let I be the interval [0,1].

<u>Proposition</u>. If H is a nonempty set of edges of the n-cube I^n such that every vertex of the graph H has valency not less than n-1 , then either the union $\bigcup H$ is connected or H consists of all the edges of two disjoint (n-1)-faces of I^n .

<u>Proof</u>. We proceed by induction on n. For $n = 1$ the Proposition is obvious. Suppose that it is true for $n - 1$. If $\bigcup H$ is connected we are done, thus suppose that it is disconnected. Let F_0 and F_1 be two disjoint $(n-1)$-faces of I^n. Let $A_0 = F_0 \cap \bigcup H$ and $A_1 = F_1 \cap \bigcup H$. If A_0 is connected and A_1 is connected then all vertices of A_0 which are of valency $n-2$ must be connected in $\bigcup H$ to some vertices in F_1. Those are in A_1 and hence $\bigcup H$ is connected contrary to our assumption. Thus A_0 has no vertices of valency $n-2$ and hence it is the union of all the edges of F_0. Similarly A_1 must be the union of all the edges of F_1 and the conclusion of the Proposition follows. Now suppose that A_0 is connected but A_1 is not. Then, by the inductive assumption, A_1 is a union of all the edges of two disjoint $(n-2)$-faces of F_1. Then every vertex of A_1 must be connected in $\bigcup H$ to a vertex of F_0. It follows that $\bigcup H$ is connected, contrary to our assumption. By symmetry, there remains only the case when both A_0 and A_1 are disconnected. Then, by the inductive assumption both are unions of all the edges of two disjoint $(n-2)$-faces of F_0 and F_1 respectively and every edge from F_0 to F_1 is in H. Thus again H consists of all the edges of two disjoint $(n-1)$-faces of I^n.

<u>Remark</u>: Recently James Fickett refined the above Proposition proving that if all vertices of $\bigcup H$ have at least $n-k$ edges then $\bigcup H$ has at least 2^{n-k} vertices and hence at most 2^k connected components and related results (to appear).

<u>Corollary</u>. $\psi^*(1) = 1$, i.e., a 1-stable function $f: X \to R$ depends on one variable at most.

Proof. If f is a constant function the conclusion is trivially true. Thus let us assume that u and v are two different values of f. Let H_0 be the set of all edges of I^n with both vertices in $f^{-1}(u)$ and H_1 the set of all edges of I^n with both vertices in $f^{-1}(v)$. Then let $H = H_0 \cup H_1$. Of course $\cup H$ is disconnected. Since f is 1-stable H satisfies the assumption of Proposition, and the Corollary follows from the Proposition.

Now we shall prove the main result of this paper.

Theorem 1. $\psi^*(k) \le k4^{k-1}$.

Proof. Let $X \subseteq \{0,1\}^n$ and $f: X \to R$ be k-stable. For each i, $1 \le i \le n$, we put

$$A_i = \{x \in X: x^i \in X \text{ and } f(x^i) \neq f(x)\} ,$$

and, for $j \neq i$, $1 \le j \le n$ and $b \in \{0,1\}$,

$$A_{ijb} = \{x \in A_i: x_j = b\} .$$

We shall prove by induction on n the following lemma.

(L_1) If $n \ge 2k$ and $|A_i| > 0$ then $|A_i| \ge 2^{n-2k+2}$.

Step I. $n = 2k$. Let $x \in A_i$. Since f is k-stable there exist $1 \le i_1 < \ldots < i_k \le n$ such that $x^j \in X$ and $f(x^j) = f(x)$ for every $j \notin \{i_1,\ldots,i_k\}$. Hence $i \in \{i_1,\ldots,i_k\}$. Also there exist $1 \le j_1 < \ldots < j_k \le n$ such that $(x^i)^j \in X$ and $f((x^i)^j) = f(x^i)$ for every $j \notin \{j_1,\ldots,j_k\}$. Hence $i \in \{j_1,\ldots,j_k\}$. Thus $|\{i_1,\ldots,i_k, j_1,\ldots,j_k\}| < 2k$, and, since $n \ge 2k$, there exists some $s \notin \{i_1,\ldots,i_k, j_1,\ldots,j_k\}$, $1 \le s \le n$. Hence $x, x^i, x^s, (x^i)^s \in A_i$

and $|A_i| \geq 4$ follows.

Step II. $n > 2k$ and (L_1) is valid for $n-1$. Choose s as in the proof of Step I. Then $|A_i \cap A_{isb}| > 0$ for $b = 0,1$. Hence, by the inductive supposition, $|A_i \cap A_{isb}| \geq 2^{n-1-2k+2}$ for $b = 0,1$. Therefore, since $A_{is0} \cap A_{is1} = \emptyset$, we have $|A_i| \geq 2^{n-2k+2}$ as required in (L_1).

Now we can conclude the proof of Theorem 1. By the Corollary we can assume without loss of generality that $k > 1$ and also that f depends on all its n variables and $n \geq 2k$. Let p_i be the probability that $x \in A_i$, x being uniformly distributed over X, i.e., $p_i = |A_i|/|X|$. Since f depends on n variables $|A_i| > 0$ for all i. Hence, by (L_1), we have $|A_i| \geq 2^{n-2k+2}$. Since $|X| \leq 2^n$ we get

$$(1) \qquad p_i \geq 4^{-k+1} .$$

Notice that

$$(2) \qquad \sum_1^n p_i = \frac{1}{|X|} \sum_{x \in X} |\{i: x \in A_i\}| ,$$

and, since f is k-stable,

$$|\{i: x \in A_i\}| \leq k$$

for all $x \in X$. Hence, by (1) and (2),

$$n4^{-k+1} \leq \sum_1^n p_i \leq k$$

which implies $n \leq k4^{k-1}$, and Theorem 1 follows.

Let $\delta(r)$ be the minimal number n such that there exists a function $f: \{0,1\}^n \to R$, where $|R| = r$, which has the following property

(*) f depends on all its n variables, but for every function $g: R \to S$, where $|S| < r$, $g \circ f$ depends on less than n variables.

For any real number ξ we let $\lceil \xi \rceil$ be the least integer not less than ξ .

<u>Theorem 2</u>. $\delta(r) = \binom{r}{2} + \lceil \log_2 \binom{r}{2} \rceil$.

<u>Proof</u>. We put $s = \binom{r}{2}$ and $t = \lceil \log_2 \binom{r}{2} \rceil$. First we show that

$$\delta(r) \geq s + t \; . \tag{3}$$

(This inequality was conjectured by Mycielski and proved first by Ralph McKenzie.) Let f have the property (*) . Then for every pair $u,v \in R$, $u \neq v$ there exists $1 \leq i\{u,v\} \leq n$ such that $g \circ f$ does not depend on the variable $x_{i\{u,v\}}$ whenever $g(u) = g(v)$. Clearly, if $u',v' \in R$, $u' \neq v'$ and $\{u',v'\} \neq \{u,v\}$ then $i\{u',v'\} \neq i\{u,v\}$. (This already proves that $\delta(r) \geq s$.) Let $I = \{i\{u,v\}: u,v \in R$, $u \neq v\}$. Hence

$$|I| = s \; . \tag{4}$$

We need the following lemma.

(L_2) If $f(x^{i\{u,v\}}) \neq f(x)$, then, for every $y \in \{0,1\}^n$ such that $y_j = x_j$ for $j \notin I$ and for $j = i\{u,v\}$, we have $f(y) = f(x)$.

To prove this we put $\tilde{x} = x^{i\{u,v\}}$. It is enough to check that for all $j \in I - \{i\{u,v\}\}$ we have $f(x^j) = f(x)$; in fact, by symmetry, the same will then be true about $\tilde{x}$ and hence the point x^j will also satisfy the supposition of (L_2) and (L_2) follows. Then suppose to the contrary that $f(x^j) \neq f(x)$. By our choice of j we have $j = i\{u',v'\}$ for some $u',v' \in R$, $u' \neq v'$, $\{u',v'\} \neq \{u,v\}$. Thus $f(x) \in \{u',v'\}$ and we can assume without loss of generality that $f(x) = u' = u$ and $f(x^j) = v' \notin \{u,v\}$. Hence $f(\tilde{x}^j) = v'$ and $f(\tilde{x}^{jj}) \in \{u,v'\}$. But $f(\tilde{x}^{jj}) = f(\tilde{x}) = v \notin \{u,v'\}$. This contradiction completes the proof of (L_2).

Now, by (L_2), for every pair $u,v \in R$, $u \neq v$ there exists an $x \in \{0,1\}^n$ such that $x_i = 0$ for all $i \in I$ and $\{f(x), f(x^{i\{u,v\}})\} = \{u,v\}$. Then by (4) there are at least s elements $x \in \{0,1\}^n$ with $x_i = 0$ for all $i \in I$. Thus $2^{n-s} \geq s$, i.e., $n \geq s+t$ and (3) follows.

Now we prove the converse inequality

$$\delta(r) \leq s + t .$$

It is enough to define some $f: \{0,1\}^n \to R$ with $n = s+t$, $|R| = r$ and the property (*). Let $P = \{\{i,j\}: i,j \in \{1,\dots,r\}, i \neq j\}$. Thus $|P| = s$. Let $h: P \to \{0,1\}^t$ be one-to-one and $\iota: P \to \{1,\dots,s\}$ be one-to-one. For any sequences $x \in \{0,1\}^s$ and $y \in \{0,1\}^t$ we put $xy = (x_1,\dots,x_s,y_1,\dots,y_t)$. It is clear that there exists an $f: \{0,1\}^{s+t} \to \{1,\dots,r\}$ such that $\{f(xh(p)), f(x^{\iota(p)}h(p))\} = p$ for all $x \in \{0,1\}^s$ and $p \in P$ and $f(xy) = 1$ if $x \in \{0,1\}^s$ and $y \in \{0,1\}^t - \text{range}(h)$. It is easy to check that all such f have the required properties.

(P_5) What are the analogs of Theorem 2 if we restrict f's to be k-continuous or k-stable functions?

References

[1] A. Ehrenfeucht, Practical decidability, Journal of Computer and System Sciences 11 (1975), 392-396.

[2] A. Ehrenfeucht and J. Mycielski, Interpolation of functions over a measure space and conjectures about memory, The Journal of Approximation Theory 9 (1973), 218-236.

[3] A. Ehrenfeucht and J. Mycielski, Learnable functions, Proceedings of the Fifth International Congress of Logic, Methodology and Philosophy of Science, London, Ontario, 1975 (to appear).

[4] A. Ehrenfeucht and J. Mycielski, On families of intersecting sets, Journal of Combinatorial Theory 17 (1974), 259-260.

[5] R. McKenzie, J. Mycielski and D.Thompson, On Boolean functions and connected sets, Mathematical Systems Theory 5 (1971), 259-270.

Acknowledgment. We are indebted to Ralph McKenzie for a part of the proof of Theorem 2.