THEORY OF COMPUTING, Volume 7 (2011), pp. 45–48
www.theoryofcomputing.org

NOTE

Tight Bounds on the Average Sensitivity of k-CNF

Kazuyuki Amano*

Received: November 25, 2010; published: March 15, 2011.

Abstract: The average sensitivity of a Boolean function is the expectation, given a uniformly random input, of the number of input bits which when flipped change the output of the function. Answering a question by O'Donnell, we show that every Boolean function represented by a k-CNF (or a k-DNF) has average sensitivity at most k. This bound is tight since the parity function on k variables has average sensitivity k.

ACM Classification: F.1.3

AMS Classification: 68R05, 68Q15

Key words and phrases: Boolean functions, sensitivity, influence, conjunctive normal form

1 Introduction and Results

For $x \in \{0,1\}^n$ and $i \in \{1,\dots,n\}$, let x^i denote x with the i-th bit flipped. Let $f : \{0,1\}^n \to \{0,1\}$ be a Boolean function on n variables. The *sensitivity* of f at x, denoted by $s(f,x)$, is the number of bits i for which $f(x) \neq f(x^i)$. The *average sensitivity* (also known as *total influence*) of f, denoted by $S(f)$, is the expected sensitivity at a random input:

$$S(f) \quad = \quad \frac{1}{2^n} \sum_{x \in \{0,1\}^n} s(f,x)\,.$$

The average sensitivity is one of the most studied concepts in the analysis of Boolean functions (see, e. g., [3, 4, 5, 7]).

*Supported in part by KAKENHI (No. 21500005) from JSPS, Japan.

DOI: 10.4086/toc.2011.v007a004

A *literal* is a Boolean variable or its negation. Let k be a nonnegative integer. A *k-clause* is a disjunction of at most k literals, and a *k-term* is a conjunction of at most k literals. A *k-CNF* function is a conjunction of k-clauses, and a *k-DNF* function is a disjunction of k-terms.

Boppana [1] proved that $S(f) \leq 2k$ for every k-DNF (as well as k-CNF) function f. Recently, Traxler [9] improved this upper bound to $S(f) \leq 1.062k$. This is nearly optimal, since the parity function on k variables, which is obviously a k-CNF function, has average sensitivity k.

In this note, we close the gap by showing:

Theorem 1.1. *If f is a k-DNF or k-CNF function, then $S(f) \leq k$.*

This solves an open problem posed by O'Donnell in 2007 (see [6]).

2 Proof of Theorem 1.1

Our proof is a small modification of the proof of the $1.062k$ upper bound by Traxler [9], which is based on a clever use of the Paturi-Pudlák-Zane algorithm (PPZ algorithm, in short) for k-SAT [8].

Let f be a k-CNF function. (The k-DNF case is dual.) Note that Traxler's bound is in fact

$$S(f) \leq 2z\log_2(1/z)k$$

where z is the probability that f outputs 1. This upper bound is larger than k when $0.25 < z < 0.5$.

We introduce the distribution D_f over $\{0,1\}^n \cup \{\perp\}$, which is essentially identical to the distribution used in Traxler's proof.

Consider the algorithm $eppz(f)$ that takes f as input and tries to choose a satisfying assignment for f. The algorithm first chooses uniformly at random some permutation π on the index set $\{1,\ldots,n\}$ of the variables. Then, for $j = 1,\ldots,n$, it does the following: it sets the variable $x_{\pi(j)}$ to 1 if the single-variable clause $(x_{\pi(j)})$ is in f and to 0 if the single-variable clause $(\overline{x_{\pi(j)}})$ is in f. We say that in these two cases "$x_{\pi(j)}$ is forced." Otherwise $x_{\pi(j)}$ is set to 0 or 1 uniformly at random. Each time, the formula is syntactically simplified, i. e., all clauses which became satisfied are deleted. At the end, the algorithm outputs x. If the algorithm ever produces two contradictory unit clauses, then it just "gives up" and outputs "$\perp$."

Define D_f as

$$D_f(x) = \Pr[eppz(f) \text{ outputs } x]\,,$$

where the probability is over all the random choices made in $eppz$. In what follows, we are only interested in the value of $D_f(x)$ for $x \in f^{-1}(1)$. Note that a similar algorithm was introduced in [2] for obtaining a lower bound on the success probability of the PPZ algorithm.

For $x \in f^{-1}(1)$, let $t(f,\pi,x,i)$ denote the indicator variable for whether x_i is forced or not, given that π was chosen and x output.[1] Note that given that π is chosen and x is output, all of the other random choices of $eppz(f)$ are fixed; i. e., there is only one outcome that leads to a given π and x.

[1] If we borrow Traxler's notation [9], $t(f,\pi,x,i)$ is defined as $t(f,\pi,x,i) = 1-(\ell_0(f,\pi,x,i)+\ell_1(f,\pi,x,i))$.

The key observation that relates the distribution D_f to the sensitivity of f is (essentially from [8]) that, for every $x \in f^{-1}(1)$, if $f(x) \neq f(x^i)$, i. e., f is sensitive at x for the i-th bit, then

$$\mathbf{E}_\pi[t(f,\pi,x,i)] \geq \frac{1}{k}\,. \tag{1}$$

We include the proof of Eq. (1) for completeness. If $1 = f(x) \neq f(x^i) = 0$, then there must be a clause C such that the only literal in C set to 1 by x is the literal of the i-th variable. The variable x_i is forced by *eppz* if i appears last in π among all variable indices occurring in C. This happens with probability at least $1/k$ since C has at most k literals. This establishes Eq. (1).

In order to show $S(f) \leq k$, it is enough to show that $D_f(x) \geq s(f,x)/2^{n-1}k$ for every $x \in f^{-1}(1)$. This is because we may combine $\sum_{x\in f^{-1}(1)} D_f(x) \leq 1$ (since D_f is a distribution) with the elementary fact

$$S(f) \quad = \quad \frac{1}{2^{n-1}} \sum_{x\in f^{-1}(1)} s(f,x)$$

(see, e. g., [1, Lemma 1(b)]).

The proof is finished by observing

$$\begin{aligned}
D_f(x) &= \mathbf{E}_\pi\left[\prod_{i=1}^{n}\left(\frac{1}{2}\right)^{1-t(f,\pi,x,i)}\right] \\
&= \frac{1}{2^n}\mathbf{E}_\pi\left[2^{\sum_{i=1}^n t(f,\pi,x,i)}\right] \\
&\geq \frac{1}{2^n}\mathbf{E}_\pi\left[2\sum_{i=1}^{n} t(f,\pi,x,i)\right] && \text{(since } 2^a \geq 2a \text{ for all \textit{integers} } a \geq 0\text{)} \\
&= \frac{2}{2^n}\sum_{i=1}^{n}\mathbf{E}_\pi[t(f,\pi,x,i)] && \text{(linearity of expectation)} \\
&\geq \frac{s(f,x)}{2^{n-1}k} && \text{(by Eq. (1)).}
\end{aligned}$$

This completes the proof of the theorem.

Acknowledgments

The author would like to thank the anonymous referee for his/her helpful comments to improve the presentation of the paper.

References

[1] RAVI B. BOPPANA: The average sensitivity of bounded-depth circuits. *Inform. Process. Lett.*, 63(5):257–261, 1997. [doi:10.1016/S0020-0190(97)00131-2] 46, 47

[2] CHRIS CALABRO, RUSSELL IMPAGLIAZZO, VALENTINE KABANETS, AND RAMAMOHAN PATURI: The complexity of unique k-SAT: An isolation lemma for k-CNFs. *J. Comput. System Sci.*, 74(3):386–393, 2008. (Conference version in CCC '03, pp. 135–141, 2003). [doi:10.1016/j.jcss.2007.06.015] 46

[3] RONALD DE WOLF: *A Brief Introduction to Fourier Analysis on the Boolean Cube*. Number 1 in Graduate Surveys. Theory of Computing Library, 2008. [doi:10.4086/toc.gs.2008.001] 45

[4] EHUD FRIEDGUT: Boolean functions with low average sensitivity depend on few coordinates. *Combinatorica*, 18(1):27–35, 1998. [doi:10.1007/PL00009809] 45

[5] NATHAN LINIAL, YISHAY MANSOUR, AND NOAM NISAN: Constant depth circuits, Fourier transform and learnability. *J. ACM*, 40(3):607–620, 1993. [doi:10.1145/174130.174138] 45

[6] RYAN O'DONNELL: The lecture notes of the course "analysis of Boolean Functions": Lecture 29: Open problems. `http://www.cs.cmu.edu/~odonnell/boolean-analysis/lecture29.pdf`, 2007. 46

[7] RYAN O'DONNELL: Some topics in analysis of Boolean functions. In *Proc. 40th STOC*, pp. 569–578. ACM Press, 2008. [doi:10.1145/1374376.1374458] 45

[8] RAMAMOHAN PATURI, PAVEL PUDLÁK, AND FRANCIS ZANE: Satisfiability coding lemma. *Chicago Journal of Theoretical Computer Science*, 1999(115):1–18, 1999. (Conference version in FOCS '97, pp. 566–574, 1997). [doi:10.4086/cjtcs.1999.011] 46, 47

[9] PATRICK TRAXLER: Variable influences in conjunctive normal forms. In *Proc. 12th Internat. Conf. on Theory and Appl. of Satisfiability (SAT'09), volume 5584 of LNCS*, pp. 101–113. Springer, 2009. [doi:10.1007/978-3-642-02777-2_12] 46

AUTHOR

Kazuyuki Amano
associate professor
Gunma University, Japan
amano@gunma-u.ac.jp
`http://www.cs.gunma-u.ac.jp/~amano/index-e.html`

ABOUT THE AUTHOR

KAZUYUKI AMANO received his Ph. D. in 1996 from Tohoku University under the supervision of Akira Maruoka. Since 2006, he has been an associate professor at Gunma University. His research interests include computational complexity and combinatorics. He enjoys writing computer programs to solve puzzles and problems of discrete mathematics.