Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

A characterization of nested canalyzing functions with maximum average sensitivity

Richard E. Stearns [a], Daniel J. Rosenkrantz [a], S.S. Ravi [b,a,*], Madhav V. Marathe [b,c]

[a] Department of Computer Science, University at Albany – State University of New York, Albany, NY 12222, United States
[b] Network Dynamics and Simulation Sciences Laboratory, Biocomplexity Institute, Virginia Tech, Blacksburg, VA 24061, United States
[c] Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, United States

ARTICLE INFO

Article history:
Received 16 April 2017
Received in revised form 29 March 2018
Accepted 6 May 2018
Available online xxxx

Keywords:
Boolean functions
Nested canalyzing functions
Maximum average sensitivity
Characterization
Counting

ABSTRACT

Nested canalyzing functions (NCFs) are a class of Boolean functions which are used to model certain biological phenomena. We derive a complete characterization of NCFs with the largest average sensitivity, expressed in terms of a simple structural property of the NCF. This characterization provides an alternate, but elementary, proof of the tight upper bound on the average sensitivity of any NCF established by Klotz et al. (2013). We also utilize the characterization to derive a closed form expression for the number of NCFs that have the largest average sensitivity.

1. Definitions, problem formulation and prior work

1.1. Nested canalyzing functions

Boolean functions arise in many different application areas (see, for example [2,3]). A class of Boolean functions, called **nested canalyzing functions** (NCFs), was introduced in [7] to model the behavior of certain biological systems. We follow the presentation in [10] in defining such a Boolean function. (For a Boolean value b, the complement is denoted by $\overline{b}$.)

Definition 1.1. Let $X = \{x_1, x_2, \ldots, x_n\}$ denote a set of n Boolean variables. Let π be a permutation of $\{1, 2, \ldots, n\}$. A Boolean function $f(x_1, x_2, \ldots, x_n)$ over X is **nested canalyzing** in the variable order $x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(n)}$ with **canalyzing values** $a_1, a_2, \ldots, a_n$ and **canalyzed values** $b_1, b_2, \ldots, b_n$ if f can be expressed in the following form:

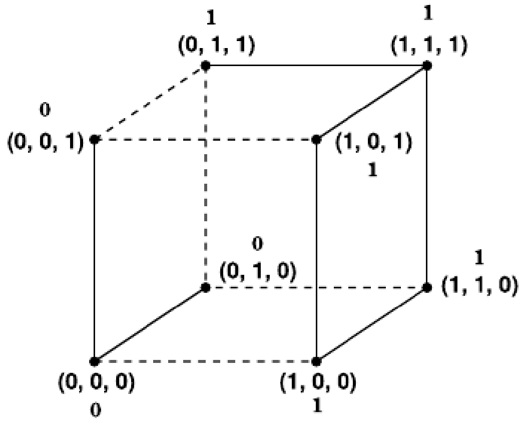
$$f(x_1, x_2, \ldots, x_n) = \begin{cases} b_1 & \text{if } x_{\pi(1)} = a_1 \\ b_2 & \text{if } x_{\pi(1)} \neq a_1 \text{ and } x_{\pi(2)} = a_2 \\ \vdots & \vdots \\ b_n & \text{if } x_{\pi(1)} \neq a_1 \text{ and } \ldots x_{\pi(n-1)} \neq a_{n-1} \text{ and } x_{\pi(n)} = a_n \\ \overline{b_n} & \text{if } x_{\pi(1)} \neq a_1 \text{ and } \ldots x_{\pi(n)} \neq a_n \end{cases}$$

* Corresponding author at: Network Dynamics and Simulation Sciences Laboratory, Biocomplexity Institute, Virginia Tech, Blacksburg, VA 24061, United States.

E-mail addresses: thestearns2@gmail.com (R.E. Stearns), drosenkrantz@gmail.com (D.J. Rosenkrantz), ssravi@vt.edu (S.S. Ravi), mmarathe@vt.edu (M.V. Marathe).

https://doi.org/10.1016/j.dam.2018.05.014
0166-218X/

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.



Assignment	Sensitivity
$(0,0,0)$	1
$(0,0,1)$	2
$(0,1,0)$	2
$(0,1,1)$	2
$(1,0,0)$	1
$(1,0,1)$	1
$(1,1,0)$	1
$(1,1,1)$	0

Fig. 1. The hypercube for the 3-variable Boolean function $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$. The value of the function f for each of the eight assignments is also shown. Edges of the hypercube where the end points have different values for f are shown as dashed lines. The table shows the sensitivity of each assignment to function f.

For convenience, we will use a computational notation to represent NCFs. For $1 \le i \le n$, line i of our representation has the following form:

$x_{\pi(i)} : a_i \longrightarrow b_i$

We say that $x_{\pi(i)}$ is the **canalyzing variable** that is **tested** in line i, with a_i and b_i denoting respectively the canalyzing and canalyzed values in line i as before, $1 \le i \le n$. We refer to each such line as a **rule**. When none of the conditions "$x_{\pi(i)} = a_i$" is satisfied, we have line $n+1$ with the "Default" rule for which the canalyzed value is $\overline{b_n}$:

Default: $\overline{b_n}$

In the remainder of this paper, we will refer to the above specification of an NCF as the **simplified representation** and assume (without loss of generality) that each NCF is specified in this manner. The simplified representation provides the following convenient computational view of an NCF. Lines defining an NCF are considered sequentially in a top-down manner. The computation stops at the first line where the specified condition is satisfied, and the value of the function is the canalyzed value on that line. We now present an example of an NCF using the two representations mentioned above.

Example 1. Consider the function $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$. This function is nested canalyzing using the identity permutation π on $\{1, 2, 3\}$ with canalyzing values 1, 0, 0 and canalyzed values 1, 0, 0. We first show how this function can be expressed using the syntax of Definition 1.1.

$$f(x_1, x_2, x_3) = \begin{cases} 1 & \text{if } x_1 = 1 \\ 0 & \text{if } x_1 \neq 1 \text{ and } x_2 = 0 \\ 0 & \text{if } x_1 \neq 1 \text{ and } x_2 \neq 0 \text{ and } x_3 = 0 \\ 1 & \text{if } x_1 \neq 1 \text{ and } x_2 \neq 0 \text{ and } x_3 \neq 0 \end{cases}$$

A simplified representation of the same function is as follows.

$x_1 : 1 \longrightarrow 1$
$x_2 : 0 \longrightarrow 0$
$x_3 : 0 \longrightarrow 0$
Default: 1

Many researchers have studied mathematical properties of NCFs and pointed out the importance of NCFs in modeling biological phenomena (e.g., [7,8,10–14]). Since our focus is on the sensitivity of NCFs, we now introduce the relevant concepts.

1.2. *Sensitivity of a Boolean Function*

Consider a Boolean function $f(x_1, x_2, \ldots, x_n)$ of n variables. An **assignment** α is a vector $(a_1, a_2, \ldots, a_n)$, with $a_i \in \{0, 1\}$ being the value assigned to variable x_i, $1 \le i \le n$. Let H_n denote the hypercube formed by the 2^n different assignments in the following manner: each node of H_n represents an assignment, and there is an edge between two nodes if the corresponding assignments differ in *exactly* one bit (i.e., the **Hamming distance** between the two assignments is equal to 1). An example of the hypercube for the 3-variable Boolean function $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$ (used in Example 1) is shown in Fig. 1.

With a slight abuse of notation, we let $f(v)$ denote the value of the function f for the assignment represented by the node $v \in H_n$. We can now define the necessary concepts related to the **sensitivity** of the Boolean function f as follows.

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

Definition 1.2. Consider a Boolean function $f(x_1, x_2, \ldots, x_n)$ of n variables. Let H_n denote the hypercube formed by the 2^n assignments to the variables of f.

(i) For any assignment v, the **sensitivity** of v, denoted by $q(v)$, is the number of neighbors w of v in H_n such that $f(v) \neq f(w)$. Formally, $q(v) = |\{w : w$ is a neighbor of v in H_n and $f(w) \neq f(v)\}|$.
(ii) The **sensitivity** of the function f, denoted by $\sigma(f)$, is the largest sensitivity value over all the assignments to f. Formally, $\sigma(f) = \max\{q(v) : v \in H_n\}$.
(iii) The **total sensitivity** of the function f, denoted by $\gamma(f)$, is the sum of the sensitivity values over all the assignments to f. Formally, $\gamma(f) = \sum_{v \in H_n} q(v)$.
(iii) The **average sensitivity** of the function f, denoted by $\widehat{\sigma}(f)$, is the ratio of the total sensitivity of f to 2^n (the total number of assignments to f). Formally, $\widehat{\sigma}(f) = \gamma(f)/2^n$.

Example 2. Consider again the 3-variable NCF $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$. The hypercube corresponding to this function is shown in Fig. 1. The figure also shows the sensitivity of each assignment to the function f. From the table, it can be seen that the sensitivity of f is 2 and its total sensitivity is 10. Hence, the average sensitivity of f is $10/2^3 = 5/4$.

1.3. Related work and our contributions

There is extensive literature on the sensitivity of various classes of Boolean functions (e.g., [1,15,18]). For a discussion on how the stability of a Boolean network is related to the sensitivities of the update functions used in the network, the reader is referred to [8,10,11]. Observations regarding the relationship between the sensitivity and the computational complexity of a Boolean function are presented in [1,15]. Li and Adeyeye [12] present lower and upper bounds on the sensitivity of any NCF. Li et al. [13,14] conjectured that the average sensitivity of any NCF is strictly less than 4/3. This conjecture was proved by Klotz et al. [9] by establishing a tight upper bound on the average sensitivity of any NCF. Their methods rely on Fourier analysis of Boolean functions [16].

In this paper, we derive a complete characterization of NCFs with the largest average sensitivity, expressed in terms of a simple structural property of the NCF. In particular, we prove that an NCF has the largest average sensitivity iff the canalyzed value in every even canalyzing line in its simplified representation is the complement of the canalyzed value in the preceding odd line, except for the last line when the number of variables is even. (It is permissible for the canalyzed value in an odd numbered line to be the same as the one in the preceding even numbered line.) A formal statement of this property is provided in Theorem 2.9. This characterization leads to an alternate, but elementary, proof of the tight upper bound established in [9] on the average sensitivity of any NCF. We also utilize the characterization to derive a closed form expression for the number of NCFs that have the largest average sensitivity.

Researchers have also studied the class of k-canalyzing functions, which generalize the class of nested canalyzing functions [4,5]. A k-canalyzing function of $n \geq k$ variables specifies canalyzing rules for k of the variables, and the default line specifies a Boolean function of the remaining $n - k$ variables. The parameter k is referred to as the *canalyzing depth*. Thus, a nested canalyzing function of n variables has a canalyzing depth of n. He and Macauley [4] develop techniques that provide an algebraic characterization of all Boolean functions in terms of their canalyzing depth; they use this characterization to obtain a closed form expression for the number of n-variable Boolean functions with a canalyzing depth of k. Kadelka et al. [5] study a more general notion of sensitivity (called c-sensitivity) for k-canalyzing functions. They show that the stability of a Boolean network whose update functions are k-canalyzing functions can be expressed as a weighted sum of the c-sensitivities of the update functions. In another paper, Kadelka et al. [6] study a different generalization of nested canalyzing functions where the values of the variables and the functions are from a finite field whose number of elements is a prime. They present a parameterized polynomial form for such functions and show how the representation is useful in computing several characteristics (e.g., the c-sensitivity and network stability) of generalized nested canalyzing functions.

2. NCFs with maximum average sensitivity: a structural characterization

2.1. Notation and terminology

Let $f(x_1, x_2, \ldots, x_n)$ be a Boolean function of n variables specified as an NCF. Throughout this section, we will assume that f is specified using the simplified representation for NCFs introduced in Section 1. Without loss of generality, we assume that the nested canalyzing order is $\langle x_1, x_2, \ldots, x_n \rangle$ so that x_i is the canalyzing variable being tested in line i, $1 \leq i \leq n$. Further, let $a_i \longrightarrow b_i$ be the rule on line i, $1 \leq i \leq n$. As in Section 1.2, an assignment α is a vector $(a_1, a_2, \ldots, a_n)$, with $a_i \in \{0, 1\}$ being the value assigned to variable x_i, $1 \leq i \leq n$. Given an assignment α, the value assigned by α to the variable x_i is denoted by $\alpha(x_i)$.

2.2. Proof outline for our results

We show that for any n, any NCF f with n variables has maximum total sensitivity (and hence maximum average sensitivity) if and only if each even numbered rule (with the possible exception of the last rule when n is even) has a different

Table 1
Values of sets from Definition 2.1 for the NCF $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$.

Set	Value
S_1	{(0, 0, 0), (1, 0, 0), (0, 1, 0), (1, 1, 0), (0, 0, 1), (1, 0, 1)}
S_2	{(0, 0, 1), (0, 1, 1)}
S_3	{(0, 1, 0), (0, 1, 1)}
W_1	{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)}
W_2	{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1)}
W_3	{(0, 1, 0), (0, 1, 1)}
Z_2	{(0, 0, 0), (0, 0, 1), (0, 1, 0)}
Z_3	{(0, 1, 1)}

canalyzed value compared to the preceding rule. This result (Theorem 2.9) provides a characterization of NCFs with the largest average sensitivity. To prove this result, we define certain subsets of assignments to f and establish some properties of these sets. In turn, these properties allow us to prove some relationships (Lemmas 2.5–2.7) between the rules used to define an NCF f and its total sensitivity $\gamma(f)$. Once Theorem 2.9 is established, we obtain closed form expressions for the maximum total sensitivity (Theorem 2.11) for odd and even values of n using simple summations.

2.3. *Definitions and lemmas used to establish our characterization*

We begin with the definitions of some subsets of assignments to an NCF f.

Definition 2.1.

(a) For $1 \le i \le n$, S_i denotes the set of assignments α such that complementing the value of $\alpha(x_i)$ changes the value of f.
(b) For $1 \le i \le n$, W_i denotes the set of assignments α such that for all j, $1 \le j < i$, $\alpha(x_j) = \overline{a_j}$.
(c) For $2 \le i \le n$, Z_i denotes the set of assignments $\alpha \in W_i$ such that $f(\alpha) = \overline{b_{i-1}}$.

As will be shown (Lemma 2.2), sets S_i, $1 \le i \le n$, determine the total sensitivity of f. Set W_1 contains all the 2^n assignments to f. For $1 \le i \le n$ and for any assignment $\alpha \notin W_i$, complementing any of the bits in positions i through n cannot change the value of f. Sets Z_i, $2 \le i \le n$, help in establishing some relationships among the sizes of sets S_i, $1 \le i \le n$ (Lemmas 2.5–2.7).

Example 3. Consider the function $f(x_1, x_2, x_3) = x_1 \vee (x_2 \wedge x_3)$ from Examples 1 and 2. For this function, the values of the various sets are shown in Table 1. These values can be verified using the NCF representation of f presented in Example 1 and the hypercube shown in Fig. 1.

The following lemma points out some properties of sets S_i, $1 \le i \le n$.

Lemma 2.2. *(i) The total sensitivity of f is given by $\gamma(f) = \sum_{i=1}^{n} |S_i|$. (ii) $|S_n| = 2$.*

Proof. To prove Part (i), consider the hypercube representation of the function f. Each edge $\{u, v\}$ of the hypercube where $f(u) \neq f(v)$ contributes 2 to the total sensitivity $\gamma(f)$, and the other edges do not contribute to $\gamma(f)$. For any edge $\{u, v\}$ where $f(u) \neq f(v)$, the assignments corresponding to u and v differ in exactly in one position, say position i. By the definition of S_i, the assignments corresponding to u and v appear in set S_i. Thus, $\gamma(f) = \sum_{i=1}^{n} |S_i|$.

To prove Part (ii), we consider the given NCF representation of f and observe that the two assignments

$$(\overline{a_1}, \overline{a_2}, \ldots, \overline{a_{n-1}}, 0) \text{ and } (\overline{a_1}, \overline{a_2}, \ldots, \overline{a_{n-1}}, 1)$$

are both in S_n. Any other assignment has the value a_i in position i for some $i < n$, so the value of f is determined before line n of the NCF representation is reached. Thus, for any such assignment, the value of f remains unchanged when the bit in position n is complemented. Therefore, such an assignment is not in S_n, and hence $|S_n| = 2$. □

We now prove some lemmas that point out properties of sets S_i, W_i and Z_i, $1 \le i \le n$.

Lemma 2.3. *(i) For $1 \le i \le n$, $|W_i| = 2^{n-i+1}$. (ii) For $1 \le i \le n$, $S_i \subseteq W_i$.*

Proof. Part (i) follows from the fact that W_i has one member for each assignment of values to the last $n - i + 1$ variables.

For Part (ii), consider any assignment $\alpha \notin W_i$. We will show that $\alpha \notin S_i$. Since $\alpha \notin W_i$, there is a $j < i$ such that $\alpha(x_j) = a_j$. Let j' be the smallest such value. Then, $f(\alpha) = b_{j'}$, and the value of f remains unchanged if the value of variable x_i is complemented. So, $\alpha \notin S_i$. □

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

Lemma 2.4. *For* $1 \le i < n$, $|S_i| = 2\,|Z_{i+1}|$.

Proof. From Part (ii) of Lemma 2.3, $S_i \subseteq W_i$. Consider an assignment α in W_i. There are two cases.

Case 1: $\alpha(x_i) = a_i$. Hence, $f(\alpha) = b_i$. Let $\hat{\alpha}$ be the assignment obtained from α by changing x_i to $\overline{a_i}$. Note that $\hat{\alpha} \in W_{i+1}$. Changing the value of x_i in α changes the value of f iff $f(\hat{\alpha}) = \overline{b_i}$, that is, iff $\hat{\alpha} \in Z_{i+1}$. Thus, the number of assignments α in S_i with $\alpha(x_i) = a_i$ equals $|Z_{i+1}|$.

Case 2: $\alpha(x_i) = \overline{a_i}$. In this case, note that $\alpha \in W_{i+1}$. Changing the value of x_i in α makes f have value b_i, and so changes the value of f iff $f(\alpha) = \overline{b_i}$, that is, iff $\alpha \in Z_{i+1}$. Thus, the number of assignments in S_i with $\alpha(x_i) = \overline{a_i}$ is also equal to $|Z_{i+1}|$.

The lemma follows. □

Lemma 2.5. *For any* $n \ge 1$ *and any* i, $1 \le i \le n$, *let* f *and* g *be two NCFs with the same sequence of* n *canalyzing variables, and identical rules for variables* $x_i, x_{i+1}, \ldots, x_n$. *Let* S_i^f *denote set* S_i *for* f *and* S_i^g *denote set* S_i *for* g. *Then,* $|S_i^f| = |S_i^g|$.

Proof. If $i = n$, then from Part (ii) of Lemma 2.2, $|S_i^f| = |S_i^g| = 2$. So, assume that $i < n$. For $1 \le j \le n$, let W_j^f and Z_j^f denote sets W_j and Z_j for f, and let a_j^f denote the canalyzing value of the rule for x_j for f. Similarly, let W_j^g and Z_j^g denote sets W_j and Z_j for g, and let a_j^g denote the canalyzing value of the rule for x_j for g. Consider an assignment α^f in Z_{i+1}^f. Let α^g be the assignment where $x_j = \overline{a_j^g}$ for $1 \le j < i$, and x_j has the same value as in α^f for $i \le j \le n$. Since f and g have identical rules for variables $x_i, x_{i+1}, \ldots, x_n$, assignment α^g is in W_{i+1}^g and $g(\alpha^g) = f(\alpha^f)$. Thus, $\alpha^g \in Z_{i+1}^g$. Therefore, for each assignment in Z_{i+1}^f, there is a unique corresponding assignment in Z_{i+1}^g. Hence, $|Z_{i+1}^g| \ge |Z_{i+1}^f|$. Similarly, $|Z_{i+1}^f| \ge |Z_{i+1}^g|$. Thus, $|Z_{i+1}^f| = |Z_{i+1}^g|$. The result now follows from Lemma 2.4. □

Lemma 2.6. *For any* i, $1 \le i < n$, *if* $b_i \ne b_{i+1}$, *then* $|S_i| + |S_{i+1}| = 2^{n-i+1}$.

Proof. Assume $b_i \ne b_{i+1}$. Consider an assignment α in $S_i \cup S_{i+1}$. From Part (ii) of Lemma 2.3, we can assume that α is of the form

$$(\overline{a}_1,\ \ldots,\ \overline{a}_{i-1},\ c_i,\ c_{i+1},\ c_{i+2},\ \ldots,\ c_n).$$

We partition the 2^{n-i+1} assignments of this form into four groups, based on the values of c_i and c_{i+1}. Each of these four groups contains 2^{n-i-1} assignments. For each such group, we now compute the number of assignments contributed by the group to sets S_i and S_{i+1}.

Group 1: $c_i = a_i$ and $c_{i+1} = a_{i+1}$. Then $f(\alpha) = b_i$, $\alpha \in S_i$, and $\alpha \notin S_{i+1}$. Group 1 adds 2^{n-i-1} assignments to S_i, and none to S_{i+1}.

Group 2: $c_i = a_i$ and $c_{i+1} = \overline{a_{i+1}}$. Then $f(\alpha) = b_i$, $\alpha \notin S_{i+1}$, and $\alpha \in S_i$ iff for the Group 4 assignment $\alpha' = (\overline{a}_1, \ldots, \overline{a}_{i-1}, \overline{a}_i, \overline{a_{i+1}}, c_{i+2}, \ldots, c_n)$, $f(\alpha') = \overline{b_i}$. Group 2 adds some assignments to S_i, whose number is discussed below, and none to S_{i+1}.

Group 3: $c_i = \overline{a_i}$ and $c_{i+1} = a_{i+1}$. Then $f(\alpha) = b_{i+1}$, $\alpha \in S_i$, and $\alpha \in S_{i+1}$ iff for the Group 4 assignment $\alpha' = (\overline{a}_1, \ldots, \overline{a}_{i-1}, \overline{a}_i, \overline{a_{i+1}}, c_{i+2}, \ldots, c_n)$, $f(\alpha') = b_i$. Group 3 adds 2^{n-i-1} assignments to S_i, and some assignments to S_{i+1}, whose number is discussed below.

Group 4: $c_i = \overline{a_i}$ and $c_{i+1} = \overline{a_{i+1}}$. Then $\alpha \in S_i$ iff $f(\alpha) = \overline{b_i}$, and $\alpha \in S_{i+1}$ iff $f(\alpha) = b_i$. Each of the 2^{n-i-1} Group 4 assignments is added to either S_i or S_{i+1}, but not both. Thus, Group 4 adds a total of 2^{n-i-1} assignments to the sum $|S_i| + |S_{i+1}|$.

Corresponding to each of the 2^{n-i-1} Group 4 assignments, there is either a Group 2 assignment added to S_i, or a Group 3 assignment added to S_{i+1}, but not both. Thus, the number of Group 2 assignments added to S_i plus the number of Group 3 assignments added to S_{i+1} is 2^{n-i-1}.

When all four groups are considered, the total value of the sum $|S_i| + |S_{i+1}|$ is 2^{n-i+1}. □

Lemma 2.7. *Consider a given* n *variable NCF* f *such that there is a* k, $1 \le k < n/2$, *with* $b_{2k-1} = b_{2k}$. *Then there is another NCF* g *with the same number of variables such that the total sensitivity of* g *is greater than that of* f.

Proof. Let k be the smallest value such that $k < n/2$ and $b_{2k-1} = b_{2k}$. Then the total sensitivity $\gamma(f) = \sum_{i=1}^{n} |S_i|$ is the sum of the following three quantities:

1. $\sum_{i=1}^{2k-2} |S_i|$
2. $|S_{2k-1}|$
3. $\sum_{i=2k}^{n} |S_i|$.

Let g be the NCF obtained from f by changing the canalyzed output in the rule for variable x_{2k-1} from b_{2k-1} to $\overline{b_{2k-1}}$.

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

We will show that the sums in Items 1 and 3 above have the same values for f and g. Subsequently, we will show that value of the term in Item 2 is larger for g.

Consider the sum in Item 1 above. Note that $\sum_{i=1}^{2k-2}|S_i| = \sum_{j=1}^{k-1}(|S_{2j-1}| + |S_{2j}|)$. For each $j < k$, $b_{2j-1} \neq b_{2j}$, so from Lemma 2.6, the value of $|S_{2j-1}| + |S_{2j}|$ is independent of the canalyzing output value in the rule for variable x_{2k-1}, and so is the same for f and g. Thus, the value of the sum in Item 1 above is the same for f and g.

Consider the sum in Item 3 above. By Lemma 2.5, this sum is also independent of the canalyzing output value in the rule for variable x_{2k-1}, and so is the same for f and g.

Now, consider the term in Item 2 above. Note that W_{2k} is the same for f and g. For half the assignments in W_{2k}, we have $x_{2k} = a_{2k}$, so the value of f on all of these assignments is b_{2k}. Among the half of assignments in W_{2k} for which $x_{2k} = \overline{a_{2k}}$, there are two that are also in W_n. Since f has complementary output values on these two assignments, there is at least one assignment in W_{2k} with $x_{2k} = \overline{a_{2k}}$ for which f has value b_{2k}. Thus, for more than half of the assignments in W_{2k}, the value of f is b_{2k}. These assignments are not in Z_{2k} for f, but are in Z_{2k} for g. Therefore, the value of $|Z_{2k}|$ for g is greater than $|Z_{2k}|$ for f. Thus, from Lemma 2.4, $|S_{2k-1}|$ for g is greater than $|S_{2k-1}|$ for f.

Considering all the three quantities above, we can conclude that the total sensitivity of g is greater than that of f. □

We end this section with a simple lemma which shows a property of the canalyzed value in the last rule of a simplified representation of an NCF.

Lemma 2.8. *Suppose an NCF f with n variables is specified using the simplified representation. For any $z \in \{0, 1\}$, there is a representation for f that satisfies both of the following conditions: (i) the rule in line n has z as the canalyzed value and (ii) the rules in lines 1 through $n - 1$ remain unchanged.*

Proof. Let line n of the representation of f be

$$x_n : \ a_n \longrightarrow b_n.$$

If $z \neq b_n$, we change line n to

$$x_n : \ \overline{a_n} \longrightarrow z$$

and the default line to

$$\text{Default} : \overline{z}.$$

It can be seen that these modifications leave the function f unchanged. □

2.4. A characterization of NCFs with maximum average sensitivity

We can now state and prove our characterization of NCFs with the largest total (and hence average) sensitivity.

Theorem 2.9. *Let f be an NCF with n variables specified using the simplified representation. Then f has the largest total sensitivity iff the canalyzed value on each computational rule with a line number of the form $2k$ with $2k < n$ is different from the canalyzed value on the rule which precedes it.*

Proof. Lemma 2.7 implies the "only if" part. From Lemma 2.8, we can assume without loss of generality that the canalyzed values on rules $n - 1$ and n are different. If we pair odd numbered rules with their subsequent even numbered rules (if any), Lemma 2.6 says the sensitivity due to the variables in this pair of rules is independent of the actual rules. If n is odd, then variable x_n is unpaired, but from Part (ii) of Lemma 2.2, $|S_n|$ is always 2. The theorem follows. □

Our next theorem uses the above characterization to derive an expression for the maximum total sensitivity of NCFs. We use the following observation in proving that theorem.

Observation 2.10. *For any $k \geq 1$, $\sum_{p=1}^{k} 2^{2p} = (4^{k+1} - 4)/3$.* □

Theorem 2.11. *The total sensitivity $\gamma(f)$ of an n-variable NCF f is at most $\frac{4}{3}(2^n - 1)$ if n is even and at most $\frac{4}{3}\left(2^n - \frac{1}{2}\right)$ if n is odd.*

Proof. Let f be an NCF with n variables and the largest total sensitivity. By Theorem 2.9, we may assume that in the definition of f, the canalyzed value on each computational rule with a line number of the form $2k$ with $2k < n$ is different from the canalyzed value on the rule which precedes it. We have two cases.

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

Case 1: Suppose n is even, so $n = 2k$ for some k. From Lemma 2.8, we can assume without loss of generality that the canalyzed values on rules $n-1$ and n are different. The total sensitivity $\gamma(f)$ is given by

$$\begin{aligned}\gamma(f) &= \sum_{i=1}^{2k} |S_i| \\ &= \sum_{j=1}^{k} (|S_{2j-1}| + |S_{2j}|) \\ &= \sum_{j=1}^{k} 2^{2k-2j+2} \quad \text{(by Lemma 2.6).}\end{aligned}$$

Reindexing the last summation by letting $p = k + 1 - j$ gives

$$\begin{aligned}\gamma(f) &= \sum_{p=1}^{k} 2^{2p} \\ &= (4^{k+1} - 4)/3 \quad \text{(by Observation 2.10)} \\ &= \frac{4}{3}\left(2^n - 1\right) \quad (\text{since } n = 2k).\end{aligned}$$

Case 2: Suppose that n is odd, so $n = 2k + 1$ for some k. The variables of f can be paired except for the last. From Part (ii) of Lemma 2.2, $|S_n| = 2$. So, the total sensitivity $\gamma(f)$ is given by

$$\begin{aligned}\gamma(f) &= \sum_{i=1}^{2k+1} |S_i| \\ &= 2 + \sum_{j=1}^{k} (|S_{2j-1}| + |S_{2j}|) \\ &= 2 + \sum_{j=1}^{k} 2^{2k+1-2j+2} \quad \text{(by Lemma 2.6).}\end{aligned}$$

Reindexing the last summation by letting $p = k + 1 - j$ gives

$$\begin{aligned}\gamma(f) &= 2 + \sum_{p=1}^{k} 2^{2p+1} \\ &= 2 + 2\sum_{p=1}^{k} 2^{2p} \\ &= 2 + 2(4^{k+1} - 4)/3 \quad \text{(by Observation 2.10)} \\ &= (4 \cdot 2^n - 2)/3 \quad (\text{since } n = 2k + 1) \\ &= \frac{4}{3}\left(2^n - \frac{1}{2}\right).\end{aligned}$$

This completes the proof of Theorem 2.11. □

The following corollaries are immediate consequences of Theorem 2.11.

Corollary 2.12. *The average sensitivity $\widehat{\sigma}(f)$ of an n variable NCF f is at most $\frac{4}{3}\left(1 - \frac{1}{2^n}\right)$ if n is even and at most $\frac{4}{3}\left(1 - \frac{1}{2^{n+1}}\right)$ if n is odd.* □

Corollary 2.13. *The average sensitivity of any NCF is strictly less than $4/3$.* □

2.5. Additional observations

Klotz et al. [9] obtain the following upper bound on average sensitivity $\widehat{\sigma}(f)$ of an NCF f:

$$\widehat{\sigma}(f) \leq \frac{4}{3} - 2^{-n} - \frac{1}{3}2^{-n}(-1)^n.$$

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

When n is even, the above expression becomes $4(1-1/2^n)/3$. When n is odd, the expression becomes $4(1-1/2^{n+1})/3$. Thus, our upper bound on $\widehat{\sigma}(f)$, stated in Corollary 2.12, exactly matches the one derived in [9].

To show that the upper bound is tight, two lower bound examples are presented in [9,14]. These examples use respectively the two alternating sequences $\langle 0, 1, 0, 1, \ldots, \rangle$ and $\langle 1, 0, 1, 0, \ldots, \rangle$ for the canalyzed values on consecutive rules of the definition of the corresponding NCF. It readily follows from our characterization (Theorem 2.9) that these two sequences define NCFs with the largest average sensitivity. Our characterization, which captures *all* sequences of canalyzing values that define NCFs with the largest average sensitivity, allows us to construct many other sequences to define such NCFs. For example, when n is a multiple of 4, each of the two sequences of canalyzed values, namely $\langle 0, 1, 1, 0, 0, 1, 1, 0, \ldots, 0, 1, 1, 0\rangle$ and $\langle 1, 0, 0, 1, 1, 0, 0, 1, \ldots, 1, 0, 0, 1\rangle$, defines an NCF with the largest average sensitivity.

2.6. Counting the number of NCFs with maximum average sensitivity

We now derive a closed form expression for the number of NCFs with n variables and the largest average sensitivity. In proving this result (Theorem 2.15), we use Theorem 2.9 (our characterization theorem), Lemma 2.8 and the following observation.

Observation 2.14. *Suppose an NCF f with n variables is specified using the simplified representation. For any $q \geq 2$ and for any i, $1 \leq i \leq n-q+1$, if the q consecutive lines $i, i+1, \ldots, i+q-1$ have the same canalyzed value, then the function remains unchanged if these q lines are permuted in any order without changing the other lines.* □

We can now state and prove the main result of this section.

Theorem 2.15. *For any $n \geq 1$, let $\Gamma(n)$ denote the number of NCFs with n variables and maximum average sensitivity. Then,*

$$\Gamma(n) = \begin{cases} 2 & \text{if } n = 1 \\ 8 & \text{if } n = 2 \\ \dfrac{4}{3}\, n!\, 6^{\lfloor n/2 \rfloor} & \text{if } n \text{ is odd and } \geq 3 \\ \dfrac{16}{27}\, n!\, 6^{n/2} & \text{if } n \text{ is even and } \geq 4. \end{cases}$$

Proof. We will consider the above four cases separately.

Case 1: $n = 1$.

For $n = 1$, it can be seen from Corollary 2.12 that the maximum average sensitivity is 1. Of the four possible Boolean functions of one variable, it can be verified that there are exactly two NCFs with average sensitivity of 1: the **identity** function defined by the rule $0 \longrightarrow 0$ (with default value 1) and the **complement** function defined by the rule $0 \longrightarrow 1$ (with default value 0).

Case 2: $n = 2$.

For $n = 2$, it can be seen from Corollary 2.12 that the maximum average sensitivity is 1. By Lemma 2.8, we may assume that the canalyzed values on lines 1 and 2 are equal. Thus, for the first line, there are two choices for the canalyzing value and two choices for the canalyzed value, giving a total of four choices. For each such choice, there are two choices for the canalyzing value on the second line but only one choice for the canalyzed value on that line. This gives a total of eight choices for the two lines. It can be verified that each of these eight choices leads to a distinct function with the maximum average sensitivity of 1.

Case 3: n is odd and ≥ 3.

Let $n = 2r + 1$ for some integer r. Suppose we partition lines 1 through n of the simplified representation of an NCF into $r + 1$ blocks, where Block 0 consists only of line 1 and each of the remaining r blocks (numbered 1 through r) consists of two consecutive lines numbered $2k$ and $2k + 1$, $1 \leq k \leq r$. We now evaluate the number of possible choices of rules for each of these blocks in three stages: choices for Block 0, Blocks 1 through $r - 1$ and Block r. (The last block needs to be considered separately since by Lemma 2.8, the two lines of the block can be assumed to have the same canalyzed values. For $1 \leq i \leq r-1$, the two lines in Block i need not have the same canalyzed value.)

(i) Block 0: Recall that this block consists only of line 1. There are n ways to choose the variable tested in line 1. For each such choice, there are two ways to choose the canalyzing value and two ways to choose the canalyzed value on that line. Thus, there are $4n$ choices for the rule in Block 0. In other words, Block 0 contributes the factor $4n$ towards the required number of functions $\Gamma(n)$.

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

(ii) Block k, where $1 \le k \le r-1$: Recall that Block k consists of lines $2k$ and $2k+1$.

When this block is considered, $2k-1$ test variables have been chosen for lines 1 through $2k-1$. Thus, lines $2k$ and $2k+1$ use two of the remaining $n-2k+1$ variables. Hence, there are $C(n-2k+1, 2)$ choices[1] for the two test variables on lines $2k$ and $2k+1$. Consider one such choice and let x_α and x_β denote the two test variables used in Block k. We have two subcases.

Subcase 3.(ii).1: Lines $2k$ and $2k+1$ have *different* canalyzed values (i.e., $b_{2k} \neq b_{2k+1}$).

Here, the two variables x_α and x_β can be permuted in two ways between lines $2k$ and $2k+1$. For each such permutation, there are two choices each for the canalyzing values in lines $2k$ and $2k+1$. However, there is only one choice for the canalyzed values on these lines since the canalyzed value on line $2k$ must be $\overline{b_{2k-1}}$ (by Theorem 2.9) and that on line $2k+1$ must be b_{2k-1} (by our assumption for this subcase). So, we get $2 \times 2 \times 2 = 8$ choices in this subcase.

Subcase 3.(ii).2: Lines $2k$ and $2k+1$ have the *same* canalyzed value (i.e., $b_{2k} = b_{2k+1}$).

Here, by Observation 2.14, permuting the two variables does not produce different functions. There are two choices each for the canalyzing values on the two lines and only one choice for the canalyzed values on these lines (since they must both be $\overline{b_{2k-1}}$). So, we get $2 \times 2 = 4$ choices in this subcase.

Combining the two subcases, we conclude that for $1 \le k \le r-1$, Block k contributes the factor $12\, C(n-2k+1, 2) = 6\,(n-2k+1)(n-2k)$ towards $\Gamma(n)$.

(iii) Block r: Recall that this block consists of lines $2r = n-1$ and $2r+1 = n$.

By Theorem 2.9, the canalyzed value on line $n-1$ must be $\overline{b_{n-2}}$, the complement of the canalyzed value on line $n-2$. From Lemma 2.8, we may assume that the canalyzed value on line n is also $\overline{b_{n-2}}$. Thus, from Observation 2.14, it follows that the function remains the same when lines $n-1$ and $n-2$ are permuted. Thus, we have two choices each for the canalyzing values on lines n and $n-1$ and only one choice for the canalyzed values on these lines. Therefore, Block r contributes the factor of 4 towards $\Gamma(n)$ in this case.

In summary, when n is odd and ≥ 3, the contributions of the various blocks towards $\Gamma(n)$ are as follows: (a) Block 0 contributes the factor $4n$, (b) for each k, $1 \le k \le r-1$, Block k contributes the factor $6\,(n-2k+1)(n-2k)$ and (c) Block r contributes the factor 4. Therefore, for this case,

$$
\begin{aligned}
\Gamma(n) &= 4n \times \left[\prod_{k=1}^{r-1} 6\,(n-2k+1)(n-2k) \right] \times 4 \\
&= 16 \times 6^{r-1} \times [n(n-1)(n-2)\ \ldots\ (n-2r+3)(n-2r+2)] \\
&= 16 \times 6^{\lfloor n/2 \rfloor - 1} \times [n(n-1)(n-2)\ \ldots\ 4 \cdot 3] \quad (\text{since } n = 2r+1) \\
&= (4/3) \times 6^{\lfloor n/2 \rfloor} \times n!
\end{aligned}
$$

as indicated in the statement of the theorem.

Case 4: n is even and ≥ 4.

Let $n = 2r$ for some integer $r \ge 2$. We partition the n-line simplified representation of an NCF into r blocks, numbered 0 through $r-1$ as follows: Block 0 consists only of line 1, each of the next $r-2$ blocks (numbered 1 through $r-2$) consists of two consecutive lines numbered $2k$ and $2k+1$, $1 \le k \le r-2$ and the last block consists of three lines, namely lines $n-2$, $n-1$ and n. As in Case 3, we evaluate the number of possible choices of rules for each of these blocks in three stages.

(i) Block 0: As in Case 3(i), the number of choices contributed by this block is $4n$.

(ii) Block k, where $1 \le k \le r-2$: As in Case 3(ii), the number of choices contributed by Block k is $6\,(n-2k+1)(n-2k)$.

(iii) Block $r-1$: This block consists of three lines, namely $n-2$, $n-1$ and n. Let b_{n-2}, b_{n-1} and b_n denote the respective canalyzed values. By Lemma 2.8, we may assume that $b_{n-1} = b_n$. By Theorem 2.9, $b_{n-3} \neq b_{n-2}$. We have two subcases depending on the values of b_{n-2} and b_{n-1}.

Subcase 4.(iii).1: Lines $n-2$ and $n-1$ have *different* canalyzed values (i.e., $b_{n-2} \neq b_{n-1}$).

There are three choices for the test variable on line $n-2$. There are two choices for the canalyzing value on each of the three lines in Block $r-1$, but only one choice for the canalyzed value on each line (since $b_{n-3} \neq b_{n-2}$ and $b_{n-2} \neq b_{n-1}$). So, this subcase contributes $3 \times 2^3 = 24$ choices.

Subcase 4.(iii).2: Lines $n-2$ and $n-1$ have the *same* canalyzed value (i.e., $b_{n-2} = b_{n-1}$).

Here, since all the three lines have the same canalyzed value, by Observation 2.14, permuting test variables has no effect on the function. There are two choices for the canalyzing value on each of the three lines in Block $r-1$, but only one choice for the canalyzed value on each line (since $b_{n-3} \neq b_{n-2}$ and $b_{n-2} = b_{n-1}$). So, this subcase contributes $2^3 = 8$ choices.

Hence, the two subcases together contribute $24 + 8 = 32$ choices.

In summary, when n is even and ≥ 4, the contributions of the various blocks towards $\Gamma(n)$ are as follows: (a) Block 0 contributes the factor $4n$, (b) for each k, $1 \le k \le r-2$, Block k contributes the factor $6\,(n-2k+1)(n-2k)$ and (c) Block r contributes the factor 32. Therefore, for this case,

[1] For nonnegative integers p and q, where $p \ge q$, we use $C(p, q)$ to denote the Binomial coefficient $\binom{p}{q}$, whose value is given by $p!/[(p-q)!q!]$.

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.

$$
\begin{aligned}
\Gamma(n) &= 4n \times \left[\prod_{k=1}^{r-2} 6\,(n-2k+1)(n-2k) \right] \times 32 \\
&= 128 \times 6^{r-2} \times [n(n-1)(n-2)\ \ldots\ (n-2r+5)(n-2r+4)] \\
&= 128 \times 6^{(n/2)-2} \times [n(n-1)(n-2)\ \ldots\ 5 \cdot 4] \quad (\text{since } n = 2r) \\
&= 128 \times 6^{(n/2)-3} \times n! \\
&= (16/27) \times 6^{n/2} \times n!.
\end{aligned}
$$

This completes the proof of Theorem 2.15. □

3. Concluding remarks

We presented an elementary proof of the conjecture by Li et al. [13,14] that the average sensitivity of any NCF is strictly less than 4/3. Our approach provides a characterization of NCFs with the largest average sensitivity. The upper bound resulting from our method exactly matches the one derived in [9] using Fourier analysis of Boolean functions. We also derived an expression for the number of NCFs with the largest average sensitivity. Our current work [17] focuses on the analysis of discrete dynamical systems whose local functions are specified as NCFs.

Acknowledgments

We thank the reviewers for providing valuable suggestions and corrections which have greatly improved the paper. This work has been partially supported by DTRA CNIMS (Contract HDTRA1-11-D-0016-0001), NSF DIBBS Grant ACI-1443054 and NSF BIG DATA Grant IIS-1633028. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon.

Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government.

References

[1] H. Buhrman, R. de Wolf, Complexity measures and decision tree complexity: A survey, Theoret. Comput. Sci. 288 (2002) 21–43.
[2] Y. Crama, P.L. Hammer (Eds.), Boolean Models and Methods in Mathematics, Computer Science, and Engineering, in: Encyclopedia of Mathematics and Applications, vol. 134, Cambridge University Press, Cambridge, UK, 2010.
[3] Y. Crama, P.L. Hammer, Boolean Functions: Theory, Algorithms, and Applications, Cambridge University Press, New York, NY, 2011.
[4] Q. He, M. Macauley, Stratification and enumeration of boolean functions by canalizing depth, Physica D 314 (2016) 1–8.
[5] C. Kadelka, J. Kuipers, R. Laubenbacher, The influence of canalization on the robustness of Boolean networks, Physica D 353–354 (2017) 39–47.
[6] C. Kadelka, Y. Li, J. Kuipers, J.O. Adeyeye, R. Laubenbacher, Multistate nested canalyzing functions and their networks, Theoret. Comput. Sci. 675 (2017) 1–14.
[7] S. Kauffman, C. Peterson, B. Samuelsson, C. Troein, Random boolean network models and the yeast transcriptional network, Proc. Natl. Acad. Sci. 100 (25) (2003) 14796–14799.
[8] S. Kauffman, C. Peterson, B. Samuelsson, C. Troein, Genetic networks with canalyzing Boolean rules are always stable, Proc. Natl. Acad. Sci. 101 (49) (2004) 17102–17107.
[9] J.G. Klotz, R. Heckel, S. Schober, Bounds on the average sensitivity of nested canalyzing functions, PLoS One 8 (2013) 8.
[10] L. Layne, Biologically Relevant Classes of Boolean Functions(Ph.D. thesis), Department of Mathematics, Clemson University, 2011.
[11] L. Layne, E. Dimitrova, M. Macauley, Nested canalyzing depth and network stability, Bull. Math. Biol. 74 (2) (2012) 422–433.
[12] Y. Li, J.O. Adeyeye, Sensitivity and block sensitivity of nested canalyzing functions, Sept. 2012. arXiv:1209.1597v1 [cs.DM].
[13] Y. Li, J.O. Adeyeye, R.C. Laubenbacher, Nested canalyzing functions and their average sensitivities, Nov. 2011. arXiv:1111.7217v1 [cs.DM].
[14] Y. Li, J.O. Adeyeye, D. Murrugarra, B. Aguilar, R.C. Laubenbacher, Boolean nested canalizing functions: A comprehensive analysis, Theoret. Comput. Sci. 481 (2013) 24–36.
[15] N. Nisan, M. Szegedy, On the degree of Boolean functions as real polynomials, Comput. Complexity 4 (1994) 301–313.
[16] R. O'Donnell, Analysis of Boolean Functions, Cambridge University Press, New York, NY, 2014.
[17] D.J. Rosenkrantz, M.V. Marathe, S.S. Ravi, R.E. Stearns, Testing phase space properties of synchronous dynamical systems with nested canalyzing local functions, in: Proceedings of the 17th International Conference on Autonomous Agents and Multi-Agent Systems, AAMAS, 2018, in press.
[18] S. Zhang, Note on the average sensitivity of monotone Boolean functions, Preprint, 2011, p. 4.

Please cite this article in press as: R.E. Stearns, et al., A characterization of nested canalyzing functions with maximum average sensitivity, Discrete Applied Mathematics (2018), https://doi.org/10.1016/j.dam.2018.05.014.