NESTED CANALYZING FUNCTIONS AND THEIR AVERAGE SENSITIVITIES

YUAN LI1∗, JOHN O. ADEYEYE 2∗, REINHARD LAUBENBACHER3

Abstract. In this paper, we obtain complete characterization for nested canalyzing functions (NCFs) by obtaining its unique algebraic normal form (polynomial form). We introduce a new concept, LAYER NUMBER for NCF. Based on this, we obtain explicit formulas for the following important parameters: 1) Number of all the nested canalyzing functions, 2) Number of all the NCFs with given LAYER NUMBER, 3) Hamming weight of any NCF, 4) The activity number of any variable of any NCF, 5) The average sensitivity of any NCF. Based on these formulas, we show the activity number is greater for those variables in outer layer and equal in the same layer. We show the average sensitivity attains minimal value when the NCF has only one layer. We also prove the average sensitivity for any NCF (No matter how many variables it has) is between 0 and 2. Hence, theoretically, we show why NCF is stable since a random Boolean function has average sensitivity $\frac{n}{2}$. Finally we conjecture that the NCF attain the maximal average sensitivity if it has the maximal LAYER NUMBER $n - 1$. Hence, we guess the uniform upper bound for the average sensitivity of any NCF can be reduced to $\frac{4}{3}$ which is tight.

1. Introduction

Canalyzing function were introduced by Kauffman [19] as appropriate rules in Boolean network models or gene regulatory networks. Canalyzing functions are known to have other important applications in physics, engineering and biology. In [30] it was shown that the dynamics of a Boolean network which operates according to canalyzing rules is robust with regard to small perturbations. In [18], W. Just, I. Shmulevich and J. Konvalina derived an exact formula for the number of canalyzing functions. In [28], the definition of canalyzing functions was generalized to any finite fields $\mathbb{F}_q$, where $q$ is a power of a prime. Both the exact formulas and the asymptotes of the number of the generalized canalyzing functions were obtained.

Nested Canalyzing Functions (NCFs) were introduced recently in [20]. One important characteristic of (nested) canalyzing functions is that they exhibit a stabilizing effect on the dynamics of a system. That is, small perturbations of an initial state should not grow in time and must eventually end up in the same attractor of the initial state. The stability is typically measured using so-called Derrida plots which monitor the Hamming distance between a random initial state and its perturbed state as both evolve over time. If the Hamming distance decreases over time, the system is considered stable. The slope of the Derrida curve is used as a numerical measure of stability. Roughly speaking, the phase space of a stable system has few components and the limit cycle of each component is short.

In [21], the authors studied the dynamics of nested canalyzing Boolean networks over a variety of dependency graphs. That is, for a given random graph on $n$ nodes, where the in-degree of each node is chosen at random between 0 and $k$, where $k \leq n$, a nested canalyzing function is assigned to each node in terms of the in-degree variables of that node. The dynamics of these networks were then analyzed and the stability measured using Derrida plots. It is shown that

* Supported by an award from the USA DoD # W911NF-11-10166.
nested canalyzing networks are remarkably stable regardless of the in-degree distribution and that the stability increases as the average number of inputs of each node increases.

An extensive analysis of available biological data on gene regulations (about 150 genes) showed that 139 of them are regulated by canalyzing functions [13]. In [21, 33], it was shown that 133 of the 139 are in fact nested canalyzing.

Most published molecular networks are given in the form of a wiring diagram, or dependency graph, constructed from experiments and prior published knowledge. However, for most of the molecular species in the network, little knowledge, if any, could be deduced about their regulatory mechanisms, for instance in the gene transcription networks in yeast [14] and E. Coli [3]. Each one of these networks contains more than 1000 genes. Kauffman et. al [20] investigated the effect of the topology of a sub-network of the yeast transcriptional network where many of the transcriptional rules are not known. They generated ensembles of different models where all models have the same dependency graph. Their heuristic results imply that the dynamics of those models which used only nested canalyzing functions were far more stable than the randomly generated models. Since it is already established that the yeast transcriptional network is stable, this suggests that the unknown interaction rules are very likely nested canalyzing functions. In a recent article [2], the whole transcriptional network of yeast, which has 3459 genes as well as the transcriptional networks of E. Coli (1481 genes) and B. subtilis (840 genes) have been analyzed in a similar fashion, with similar findings.

These heuristic and statistical results show that the class of nested canalyzing functions is very important in systems biology. It is shown in [15] that this class is identical to the class of so-called unate cascade Boolean functions, which has been studied extensively in engineering and computer science. It was shown in [8] that this class produces the binary decision diagrams with the shortest average path length. Thus, a more detailed mathematical study of this class of functions has applications to problems in engineering as well.

In [16], the authors provided a description of nested canalyzing function. As a corollary of the equivalence, a formula in the literature for the number of unate cascade functions also provides such a formula the number of nested canalyzing functions. Recently, in [32], those results were generalized to the multi-state nested canalyzing functions on finite fields $\mathbb{F}_p$, where $p$ is a prime. They obtained the formula for the number of the generalized NCFs, as a recursive relation.

In [12], Cook et al. introduced the notion of sensitivity as a combinatorial measure for Boolean functions providing lower bounds on the time needed by CREW PRAM (concurrent read, but exclusive write (CREW) parallel random access machine (PRAM)). It was extended by Nisan [34] to block sensitivity. It is still open whether sensitivity and block sensitivity are polynomially related (they are equal for monotone Boolean functions). Although the definition is straightforward, the sensitivity is understood only for a few classes function. For monotone functions, Ilya Shmulevich [37] derived asymptotic formulas for a typical monotone Boolean functions. Recently, Shengyu Zhang [43] find a formula for the average sensitivity of any monotone Boolean functions, hence, a tight bound is derived. In [38], Ilya Shmulevich and Stuart A. Kauffman considered the activities of the variables of Boolean functions with only one canalyzing variable. They obtained the average sensitivity of this kind of Boolean function.

In this paper, we revisit the NCF, obtaining a more explicit characterization of the Boolean NCFs than those in [16]. We introduce a new concept, the LAYER NUMBER in order to classify all the variables. Hence, the dominance of the variable can be quantified. As a consequence, we obtain an explicit formula for the number of NCFs. Thus, a nonlinear recursive relation (the original formula) is solved, which maybe of independent mathematical interest. Using our unique algebraic normal form of NCF, for any NCF, we get the formula of activity for its variables. We show that the variables in a more dominant layer have greater activity number. Variables in the same layer have the same activity numbers. Consequently, we obtain
the formula of any NCF’s average sensitivity, its lower bound is $\frac{n}{2}$ and its upper bound is 2 ($\text{No matter what } n \text{ is}$) which is much less than $\frac{n}{2}$, the average sensitivity of a random Boolean function. So, theoretically, we proved why NCF is “stable”. We also find the formula of the Hamming weight of each NCF. Finally, we conjecture that the NCF attains its maximal value if it has the maximal LAYER NUMBER $n-1$. Hence, we guess the tight upper bound is $\frac{4}{3}$. In the next section, we introduce some definitions and notations.

2. Preliminaries

In this section we introduce the definitions and notations. Let $F = F_2$ be the Galois field with 2 elements. If $f$ is a $n$ variable function from $F^n$ to $F$, it is well known [29] that $f$ can be expressed as a polynomial, called the algebraic normal form (ANF):

$$f(x_1, x_2, \ldots, x_n) = \bigoplus_{0 \leq k_i \leq 1, i=1, \ldots, n} a_{k_1 k_2 \ldots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

where each coefficient $a_{k_1 k_2 \ldots k_n} \in F$ is a constant. The number $k_1 + k_2 + \cdots + k_n$ is the multivariate degree of the term $a_{k_1 k_2 \ldots k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ with nonzero coefficient $a_{k_1 k_2 \ldots k_n}$. The greatest degree of all the terms of $f$ is called the algebraic degree, denoted by $\text{deg}(f)$.

**Definition 2.1.** $f(x_1, x_2, \ldots, x_n)$ is essential in variable $x_i$ if $x_i = 1$ and $f(x_1, x_2, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n) \neq f(x_1, x_2, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n).$

**Definition 2.2.** A function $f(x_1, x_2, \ldots, x_n)$ is < $a : b >$ canalyzing if $f(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n) = b$, for all $x_j$, $j \neq i$, where $i \in \{1, \ldots, n\}$, $a, b \in F$.

The definition is reminiscent of the concept of canalisation introduced by the geneticist C. H. Waddington [41] to represent the ability of a genotype to produce the same phenotype regardless of environmental variability.

**Definition 2.3.** Let $f$ be a Boolean function in $n$ variables. Let $\sigma$ be a permutation on $\{1, 2, \ldots, n\}$. The function $f$ is nested canalyzing function (NCF) in the variable order $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$ with canalyzing input values $a_1, \ldots, a_n$ and canalyzed values $b_1, \ldots, b_n$, if it can be represented in the form

$$f(x_1, \ldots, x_n) = \begin{cases} b_1 & x_{\sigma(1)} = a_1, \\ b_2 & x_{\sigma(1)} = a_1, x_{\sigma(2)} = a_2, \\ b_3 & x_{\sigma(1)} = a_1, x_{\sigma(2)} = a_2, x_{\sigma(3)} = a_3, \\ \vdots & \\ b_n & x_{\sigma(1)} = a_1, x_{\sigma(2)} = a_2, \ldots, x_{\sigma(n-1)} = a_{n-1}, x_{\sigma(n)} = a_n, \\ b_n & x_{\sigma(1)} = a_1, x_{\sigma(2)} = a_2, \ldots, x_{\sigma(n-1)} = a_{n-1}, x_{\sigma(n)} = a_n. \end{cases}$$

Where $\overline{a} = a \oplus 1$. The function $f$ is nested canalyzing if $f$ is nested canalyzing in the variable order $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$ for some permutation $\sigma$.

Let $\alpha = (a_1, a_2, \ldots, a_n)$ and $\beta = (b_1, b_2, \ldots, b_n)$, we say $f$ is $\{\sigma : \alpha : \beta\}$ NCF if it is NCF in the variable order $x_{\sigma(1)}, \ldots, x_{\sigma(n)}$ with canalyzing input values $\alpha = (a_1, \ldots, a_n)$ and canalyzed values $\beta = (b_1, \ldots, b_n)$.

Given vector $\alpha = (a_1, a_2, \ldots, a_n)$, we define $\alpha^{i_1 \cdots i_k} = (a_1, \ldots, a_{i_1}, \ldots, a_{i_k}, \ldots, a_n)$

From the above definition, we immediately have the following

**Proposition 2.4.** $f$ is $\{\sigma : \alpha : \beta\}$ NCF $\iff$ $f$ is $\{\sigma : \alpha^n : \beta^n\}$ NCF

**Example 2.5.** $f(x_1, x_2, x_3) = x_1(x_2 \oplus 1)x_3 \oplus 1$ is $\{(1, 2, 3) : (0, 1, 0) : (1, 1, 1)\}$ NCF.

Actually, one can check this function is nested canalyzing in any variable order.
**Example 2.6.** \( f(x_1, x_2, x_3) = (x_1 + 1)(x_2(x_3 + 1) + 1) + 1. \) This function is \( \{1, 2, 3\} : (1, 0, 1) : (1, 0, 0) \) NCF. It is also \( \{1, 3, 2\} : (1, 1, 1) : (1, 0, 1) \) NCF.

One can check this function can be nested canalyzing in only two variable orders \((x_1, x_2, x_3)\) and \((x_1, x_3, x_2)\).

From the above definitions, we know a function is NCF, all the \( n \) variable must be essential. However, a constant function \( b \) can be \( < i : a : b > \) canalyzing for any \( i \) and \( a \).

### 3. A Complete Characterization for NCF

In [23], the author introduced Partially Nested Canalyzing Functions (PNCFs), a generalization of the NCFs, and the nested canalyzing depth, which measures the extent to which it retains a nested canalyzing structure. In [17], the author introduced the extended monomial system.

As we will see, in a Nested Canalyzing Function, some variables are more dominant than the others. We will classify all the variables of a NCF into different levels according to the extent of their dominance. Hence, we will give description about NCF with more detail. Actually, we will obtain clearer description about NCF by introducing a new concept: LAYER NUMBER. As a by-product, we also obtain some enumeration results. Eventually, we will find an explicit formula of the number of all the NCFs.

**First, we have**

**Definition 3.1.** [17] \( M(x_1, \ldots, x_n) \) is an extended monomial of essential variables \( x_1, \ldots, x_n \) if \( M(x_1, \ldots, x_n) = (x_1 + a_1)(x_2 + a_2) \ldots (x_n + a_n) \), where \( a_i \in \mathbb{F}_2 \).

**Lemma 3.2.** \( f(x_1, x_2, \ldots, x_n) \) is \( < i : a : b > \) canalyzing iff \( f(X) = f(x_1, x_2, \ldots, x_n) = (x_i + a)Q(x_1, \ldots, x_{i−1}, x_{i+1} \ldots x_n) + b \).

**Proof.** From the algebraic normal form of \( f \), we rewrite it as \( f = x_i g_1(X_i) + g_0(X_i) \), where \( X_i = (x_1, \ldots, x_{i−1}, x_{i+1}, \ldots, x_n) \). Hence, \( f(X) = f(x_1, x_2, \ldots, x_n) = (x_i + a)g_1(X_i) + a g_1(X_i) + g_0(X_i) \). Let \( g_1(X_i) = Q(x_1, \ldots, x_{i−1}, x_{i+1} \ldots x_n) \) and \( r(X_i) = ag_1(X_i) + g_0(X_i) \). Then \( f(X) = f(x_1, \ldots, x_n) = (x_i + a)Q(x_1, \ldots, x_{i−1}, x_{i+1} \ldots x_n) + r(X_i) \).

Since \( f(X) \) is \( < i : a : b > \) canalyzing, we get \( f(X) = f(x_1, \ldots, x_{i−1}, a, x_{i+1}, \ldots, x_n) = b \) for any \( x_1, \ldots, x_{i−1}, x_{i+1} \ldots x_n \), i.e., \( r(X_i) = b \) for any \( X_i \). So \( r(X_i) \) must be the constant \( b \). We finished the necessity. The sufficiency is obvious. \( \square \)

**Remark 3.3.** 1) When we contrast this lemma to the first part of Theorem 3.1 in [16], we make clear that here, the \( x_i \) is not essential in \( Q \). 2) In [23], there is a general version of this Lemma over any finite fields. 3) In the above lemma, if \( f \) is constant, then \( Q = 0 \).

**From Definition 2.3** we have the following

**Proposition 3.4.** If \( f(x_1, \ldots, x_n) \) is \( \{\sigma : a : \beta\} \) NCF, i.e., if it is NCF in the variable order \( x_{\sigma(1)}, \ldots, x_{\sigma(n)} \) with canalyzing input values \( \alpha = (a_1, \ldots, a_n) \) and canalyzed values \( \beta = (b_1, \ldots, b_n) \).

Then, for \( 1 \leq k \leq n−1 \), let \( x_{\sigma(1)} = a_1, \ldots, x_{\sigma(k)} = a_k \), then the function \( f(x_1, \ldots, a_1, \ldots, a_k, \ldots, x_n) \) is \( \{\sigma^* : a^* : \beta^*\} \) NCF on those remaining variables, where \( \sigma^* = x_{\sigma(k+1)}, \ldots, x_{\sigma(n)} \), \( a^* = (a_{k+1}, \ldots, a_n) \) and \( \beta^* = (b_{k+1}, \ldots, b_n) \).

**Definition 3.5.** If \( f(x_1, \ldots, x_n) \) is a NCF. We call variable \( x_i \) the most dominant variable of \( f \), if there is an order \( \alpha = (x_1, \ldots) \) such that \( f \) is NCF with this variable order (In other words, if \( f \) is also \( < i : a : b > \) canalyzing for some \( a \) and \( b \)).
In Example 2.5, all the three variables are most dominant, in Example 2.6 only $x_1$ is the most dominant variable. We have

**Theorem 3.1.** Given NCF $f(x_1,\ldots,x_n)$, all the variables are most dominant iff

$$f = M(x_1,\ldots,x_n) \oplus b,$$

where $M$ is an extended monomial, i.e.,

$$M = (x_1 \oplus a_1)(x_2 \oplus a_2)\ldots(x_n \oplus a_n).$$

*Proof.* $x_1$ is the most dominant, from Lemma 3.2, we know there exist $a_1$ and $b$ such that

$$f(x_1, x_2,\ldots,x_n) = (x_1 \oplus a_1)Q(x_2,\ldots,x_n) \oplus b,$$

i.e., $(x_1 \oplus a_1))(f \oplus b)$. Now, $x_2$ is also the most dominant, we have $a_2$ and $b'$ such that

$$f(x_1, x_2,\ldots,x_n) = b'$$

for any $x_1, x_3,\ldots,x_n$. Specifically, let $x_1 = a_1$, we get

$$f(a_1, a_2, x_3,\ldots,x_n) = b = b'.$$

Hence, we also get $(x_2 \oplus a_2)Q(x_2,\ldots,x_n)$, since $x_1 \oplus a_1$ and $x_2 \oplus a_2$ are coprime, we get $(x_2 \oplus a_2)Q(x_2,\ldots,x_n)$, hence, $f(x_1, x_2,\ldots,x_n) = (x_1 \oplus a_1)(x_2 \oplus a_2)Q(x_3,\ldots,x_n) \oplus b$. With induction principle, the necessity is proved. The sufficiency if evident.

We are ready to prove the following main result of this section.

**Theorem 3.2.** Given $n \geq 2$, $f(x_1, x_2,\ldots,x_n)$ is nested canalyzing iff it can be uniquely written as

$$f(x_1, x_2,\ldots,x_n) = M_1(M_2(\ldots(M_{r-1}(M_r \oplus 1) \oplus 1)\ldots) \oplus 1) \oplus b. \quad (3.1)$$

Where each $M_i$ is an extended monomial of a set of disjoint variables. More precisely, $M_i = \prod_{j=1}^{k_i} (x_{ij} \oplus a_{ij})$, $i = 1,\ldots,r$, $k_i \geq 1$ for $i = 1,\ldots,r-1$, $k_r \geq 2$, $k_1 \oplus \ldots \oplus k_r = n$, $a_{ij} \in \mathbb{F}_2$, \{i\}_{j=1} = 1,\ldots,k_i, i = 1,\ldots,r = \{1,\ldots,n\}$.

*Proof.* We use induction on $n$.

When $n = 2$, there are 16 boolean functions, 8 of them are NCFs, namely

$$(x_1 \oplus a_1)(x_2 \oplus a_2) \oplus c = M_1 \oplus 1 \oplus b,$$

where $b = 1 \oplus c$ and $M_1 = (x_1 \oplus a_1)(x_2 \oplus a_2)$.

If $(x_1 \oplus a_1)(x_2 \oplus a_2) \oplus c = (x_1 \oplus a_1')(x_2 \oplus a_2') \oplus c'$, by equating the coefficients, we immediately obtain $a_1 = a_1'$, $a_2 = a_2'$ and $c = c'$. So, uniqueness is true.

We have proved that equation 3.1 is true for $n = 2$, where $r = 1$.

Let’s assume that equation 3.1 is true for any nested canalyzing function which has at most $n-1$ essential variables.

Now, consider NCF $f(x_1,\ldots,x_n)$.

Suppose $x_{\sigma(1)},\ldots,x_{\sigma(k_1)}$ are all the most dominant canalyzing variables of $f$, $1 \leq k_1 \leq n$.

Case 1: $k_1 = n$, by Theorem 3.1, the conclusion is true with $r = 1$.

Case 2: $k_1 < n$, with the same arguments to Theorem 3.1, we can get $f = M_1g \oplus b$, where $M_1 = (x_{\sigma(1)} \oplus a_{\sigma(1)})\ldots(x_{\sigma(k_1)} \oplus a_{\sigma(k_1)})$. Let $x_{\sigma(1)} = a_{\sigma(1)}',\ldots,x_{\sigma(k)} = a_{\sigma(k)}'$ in $f$, the function $g \oplus b$, hence, $g$, of the remaining variables will also be nested canalyzing by Proposition 3.1.

Since $g$ has $n - k_1 \leq n - 1$ variables, by induction assumption, we get

$$g = M_2(M_3(\ldots(M_{r-1}(M_r \oplus 1) \oplus 1)\ldots \oplus 1) \oplus b_1),$$

at this time, $b_1$ must be 1. Otherwise, all the variables in $M_2$ will also be the most dominant variables of $f$. Hence, we are done.

Because each NCF can be uniquely written as 3.1 and the number $r$ is uniquely determined by $f$, we have

**Definition 3.6.** For a NCF written as equation 3.1, the number $r$ will be called its LAYER NUMBER. Essential variables of $M_1$ will be the most dominant variables(canalyzing variable), they belong to the first layer of this NCF. Essential variables of $M_2$ will be called the second most dominant variables and belong to the second layer of this NCF and etc.

The function in example 2.5 has LAYER NUMBER 1 and the function in example 2.6 has LAYER NUMBER 2.
Remark 3.7. In Theorem 3.2, 1) $k_r \geq 2$. It is impossible that $k_r = 1$. Otherwise, $M_r \oplus 1$ will be a factor of $M_{r-1}$, which means LAYER NUMBER is $r-1$. 2) If variable $x_i$ is in the first layer, and $x_i \oplus a_i$ is a factor of $M_i$, then this NCF is $< a_i : b >$ canalyzing, we simply say $x_i$ is a canalyzing variable of this NCF.

Let $\text{NCF}(n, r)$ stands for the set of all the $n$ variable nested canalyzing functions with LAYER NUMBER $r$ and $\text{NCF}(n)$ stands for the set of all the $n$ variable nested canalyzing functions. We have

Corollary 3.8. Given $n \geq 2$,

$$|\text{NCF}(n, r)| = 2^{n+1} \sum_{\substack{k_1, \ldots, k_r = n \geq 1, i = 1, \ldots, r-1, k_r \geq 2}} \binom{n}{k_1, \ldots, k_{r-1}}$$

and

$$|\text{NCF}(n)| = 2^{n+1} \sum_{r=1}^{n-1} \sum_{\substack{k_1, \ldots, k_r = n \geq 1, i = 1, \ldots, r-1, k_r \geq 2}} \binom{n}{k_1, \ldots, k_{r-1}}$$

Where the multinomial coefficient $\binom{n}{k_1, \ldots, k_{r-1}} = \frac{n!}{k_1! \cdots k_{r-1}!}$

Proof. From Equation 3.1 for each choice $k_1, \ldots, k_r$, with condition $k_1 + \ldots + k_r = n$, $k_i \geq 1$, $i = 1, \ldots, r-1$ and $k_r \geq 2$,

there are $2^{k_1} \binom{n}{k_1}$ many ways to form $M_1$,

there are $2^{k_2} \binom{n-k_1}{k_2}$ many ways to form $M_2$,

...,

there are $2^{k_r} \binom{n-k_1-\ldots-k_{r-1}}{k_r}$ many ways to form $M_r$,

$b$ has two choices.

Hence,

$$|\text{NCF}(n, r)| = 2^{n+1} \sum_{\substack{k_1, \ldots, k_r = n \geq 1, i = 1, \ldots, r-1, k_r \geq 2}} 2^{k_1+\ldots+k_r} \binom{n}{k_1} \binom{n-k_1}{k_2} \cdots \binom{n-k_1-\ldots-k_{r-1}}{k_r}$$

$$= 2^{n+1} \sum_{\substack{k_1, \ldots, k_r = n \geq 1, i = 1, \ldots, r-1, k_r \geq 2}} \frac{n!}{(k_1)!(n-k_1)!} \frac{(n-k_1)!}{(k_2)!(n-k_1-k_2)!} \cdots \frac{(n-k_1-\ldots-k_{r-1})!}{k_r!(n-k_1-\ldots-k_r)!}$$

$$= 2^{n+1} \sum_{\substack{k_1, \ldots, k_r = n \geq 1, i = 1, \ldots, r-1, k_r \geq 2}} \frac{n!}{k_1!k_2! \cdots k_r!} = 2^{n+1} \sum_{\substack{k_1, \ldots, k_r = n \geq 1, i = 1, \ldots, r-1, k_r \geq 2}} \binom{n}{k_1, \ldots, k_{r-1}}.$$ 

Since $\text{NCF}(n) = \bigcup_{r=1}^{n-1} \text{NCF}(n, r)$ and $\text{NCF}(n, i) \cap \text{NCF}(n, j) = \phi$ when $i \neq j$, we get the formula of $|\text{NCF}(n)|$. \hfill \square

One can check that $|\text{NCF}(2)| = 8$, $|\text{NCF}(3)| = 64$, $|\text{NCF}(4)| = 736$, $|\text{NCF}(5)| = 10624$...

These results are consistent with those in [4, 36].

By equating our formula to the recursive relation in [4, 36], we have the following

Corollary 3.9. The solution of the nonlinear recursive sequence

$$a_2 = 8, a_n = \sum_{r=2}^{n-1} \binom{n}{r-1} 2^{r-1} a_{n-r+1} + 2^{n+1}, n \geq 3$$
other variables. To formalize this, a concept called activity is defined as a measure of the influence of a variable on the output of a Boolean function. For example, if a function has no canalyzing variables, then only one value is nonzero. But biased functions may have no canalyzing variables. For example, the function $f(x_1, x_2, x_3) = x_1x_2x_3 \oplus x_1x_2 \oplus x_1x_3 \oplus x_2x_3$ is biased but without canalyzing variables.

In Boolean functions, some variables have greater influence over the output of the function than others. To formalize this, a concept called activity was introduced. Let $\alpha_i^f = 1 \frac{\partial f(x_1, \ldots, x_n)}{\partial x_i}$ be the activity of variable $x_i$ of $f$. The activity of any variables of constant functions is 0. For affine function $f(x_1, \ldots, x_n) = x_1 \oplus \ldots \oplus x_n \oplus b$, $\alpha_i^f = 1$ for any $i$. It is clear, for any $f$ and $i$, we have $0 \leq \alpha_i^f \leq 1$.

Another important quantity is the sensitivity of a Boolean function, which measures how sensitive the output of the function is if the input changes. This was introduced in \cite{12}. The sensitivity $s_i^f(x_1, \ldots, x_n)$ of $f$ on vector $(x_1, \ldots, x_n)$ is defined as the number of Hamming neighbors of $(x_1, \ldots, x_n)$ on which the function value is different from $f(x_1, \ldots, x_n)$.

The average sensitivity of function $f$ is defined as

$$s^f = E[s_i^f(x_1, \ldots, x_n)] = \frac{1}{2^n} \sum_{(x_1, \ldots, x_n) \in \mathbb{F}_2^n} s_i^f(x_1, \ldots, x_n) = \sum_{i=1}^n \alpha_i^f.$$ 

It is clear that $0 \leq s^f \leq n$.

The average sensitivity is one of the most studied concepts in the analysis of Boolean functions. Recently, it receives a lot of attention. See \cite{11, 5, 6, 7, 10, 11, 22, 24, 26, 35, 37, 38, 39, 40, 42}. Bernasconi \cite{5} has showed that a random Boolean function has average sensitivity $\frac{n}{2}$. It means the average value of the average sensitivities of all the $n$ variables Boolean functions is $\frac{n}{2}$. In \cite{38}, Ilya Shmulevich and Stuart A. Kauffman calculated the activity of all the variables of a Boolean functions with exactly one canalyzing variable and unbiased input for the other variable. Add all the activities, the average sensitivity of this function was also obtained.

In the following, using Equation 3.1, we will obtain the formula of the Hamming weight of any NCF, the activities of all the variables of any NCF and the average sensitivity (which is bounded by constant) of any NCF.
First, we have

**Lemma 4.1.** \((x_1 \oplus a_1) \ldots (x_k \oplus a_k) = \begin{cases} 1, & (x_1, \ldots, x_k) = (\overline{a_1}, \ldots, \overline{a_k}) \\ 0, & \text{otherwise.} \end{cases}\) i.e., only one value is 1 and all the other \(2^k - 1\) values are 0.

**Theorem 4.1.** Given \(n \geq 2\). Let \(f_r = M_1(M_2(\ldots (M_{r-1}(M_r \oplus 1) \oplus 1) \ldots) \oplus 1)\), \(r \geq 2\), where \(M_i\) is same as that in the in Theorem 3.2. Then the Hamming weight of \(f_r\) is

\[
W(f_r) = \sum_{j=1}^{r} (-1)^{j-1} 2^{n - \sum_{i=1}^{j} k_i} \tag{4.3}
\]

The Hamming weight of \(f_r \oplus 1\) is

\[
W(f_r \oplus 1) = \sum_{j=0}^{r} (-1)^{j} 2^{n - \sum_{i=1}^{j} k_i} \tag{4.4}
\]

Where \(\sum_{i=1}^{0} k_i\) should be explained as 0.

**Proof.** First, let’s consider the Hamming weight of \(f_r\).

When \(r = 1\), we know the result is true by Lemma 4.1.

When \(r > 1\), we consider two cases:

Case A: \(r\) is odd, \(r = 2t + 1\).

All the vectors make \(f = 1\) will be divided into the following disjoint groups.

Group 1: \(M_1 = 1, M_2 = 0\);

Group 2: \(M_1 = 1, M_2 = 1, M_3 = 1, M_4 = 0\);

\[ \ldots \]

Group \(j\): \(M_1 = 1, M_2 = 1, \ldots, M_{2j-1} = 1, M_{2j} = 0\);

\[ \ldots \]

Group \(t\): \(M_1 = 1, M_2 = 1, \ldots, M_{2t-1} = 1, M_{2t} = 0\);

Group \(t + 1\): \(M_1 = 1, M_2 = 1, \ldots, M_{2t} = 1, M_{2t+1} = M_r = 1\).

In Group 1, the number of vectors is \((2^{k_2 - 1}) 2^{n-k_1-k_2} = 2^{n-k_1} - 2^{n-k_1-k_2}\).

In Group 2, the number of vector is \((2^{k_4 - 1}) 2^{n-k_1-k_2-k_3-k_4} = 2^{n-k_1-k_2-k_3} - 2^{n-k_1-k_2-k_3-k_4}\).

\[ \ldots \]

In Group \(t\), the number of vector is \((2^{k_2t - 1}) 2^{n-k_1-k_2-\ldots-k_2t} = 2^{n-k_1-\ldots-k_2t-1} - 2^{n-k_1-\ldots-k_2t}\).

In Group \(t + 1\), the number of vectors is \(2^{n-k_1-\ldots-k_r}\).

Add all of them, we get the formula Equation 4.3 again.

Case B: \(r\) is even, \(r = 2t\).

All the vectors make \(f = 1\) will be divided into the following disjoint groups.

Group 1: \(M_1 = 1, M_2 = 0\);

Group 2: \(M_1 = 1, M_2 = 1, M_3 = 1, M_4 = 0\);

\[ \ldots \]

Group \(j\): \(M_1 = 1, M_2 = 1, \ldots, M_{2j-1} = 1, M_{2j} = 0\);

\[ \ldots \]

Group \(t - 1\): \(M_1 = 1, M_2 = 1, \ldots, M_{2t-3} = 1, M_{2t-2} = 0\);

Group \(t\): \(M_1 = 1, M_2 = 1, \ldots, M_{2t-1} = 1, M_{2t} = M_r = 0\).

In Group 1, the number of vectors is \((2^{k_2 - 1}) 2^{n-k_1-k_2} = 2^{n-k_1} - 2^{n-k_1-k_2}\).

In Group 2, the number of vector is \((2^{k_4 - 1}) 2^{n-k_1-k_2-k_3-k_4} = 2^{n-k_1-k_2-k_3} - 2^{n-k_1-k_2-k_3-k_4}\).

\[ \ldots \]

In Group \(t - 1\), the number is \((2^{k_{2t-2} - 1}) 2^{n-k_1-\ldots-k_{2t-2}} = 2^{n-k_1-\ldots-k_{2t-3}} - 2^{n-k_1-\ldots-k_{2t-2}}\).

In Group \(t\), the number of vectors is \(2^{n-k_1-\ldots-k_{2t-1}} - 2^{n-k_1-\ldots-k_{2t}} = 2^{k_{2t}} - 1\).

Add all of them, we get the formula Equation 4.3 again.
Because \(|\{(x_1, \ldots, x_n) | f(x_1, \ldots, x_n) = 0\}| + |\{(x_1, \ldots, x_n) | f(x_1, \ldots, x_n) = 1\}| = 2^n\), we know the Hamming weight of \(f_r \oplus 1\) is

\[
W(f_r \oplus 1) = 2^n - W(f_r) = 2^n - \sum_{j=1}^{r} (-1)^{j-1} \frac{1}{2^{n-1-j}} \sum_{i=1}^{j} k_i = \sum_{j=0}^{r} (-1)^{j} 2^{n-j} \sum_{i=1}^{j} k_i.
\]

Where \(\sum_{i=1}^{j} k_i\) should be explained as 0.

In the following, we will calculate the activities of the variables of any NCF.

Let \(f\) be a NCF and written as the form in Theorem 3.2. Without loss of generality (to avoid the complicated notation), we assume \(M_1 = (x_1 \oplus a_1)(x_2 \oplus a_2) \ldots (x_k \oplus a_k)\) and \(m_1 = (x_1 \oplus a_1) \ldots (x_i \oplus a_{i-1})(x_i+1 \oplus a_i+1) \ldots (x_k \oplus a_k)\), i.e., \(M_1 = (x_i \oplus a_i)m_1\).

If \(r = 1\), i.e., \(k_1 = n\), then

\[
\alpha_i^f = \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_i-1, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} (f(x_1, \ldots, 0, \ldots, x_n) \oplus f(x_1, \ldots, 1, \ldots, x_n))
\]

\[
= \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_i-1, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} m_1 = \frac{1}{2^{n-1}} W(m_1) = \frac{1}{2^{n-1}}.
\]

by Lemma 4.2.

If \(1 < r \leq n - 1\),

Let’s consider the activity of \(x_i\) in the first layer, i.e., \(1 \leq i \leq k_1\). We have

\[
\alpha_i^f = \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_i-1, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} (f(x_1, \ldots, 0, \ldots, x_n) \oplus f(x_1, \ldots, 1, \ldots, x_n))
\]

\[
= \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_i-1, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} m_1(M_2(\ldots (M_{r-1}(M_r \oplus 1) \oplus 1) \ldots) \oplus 1).
\]

By Theorem 4.1. Note, in the above, \(k_1 = 1\) means \(m_1 = 1\), we used the Equation 4.4 with layer number \(r - 1\) and the first layer is \(M_2\) for \(n - 1\) variables functions.

Now let’s consider the variables in the second layer, i.e., \(x_i\) is an essential variable of \(M_2\). We have \(M_2 = (x_i \oplus a_i)m_2\) and

\[
\alpha_i^f = \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_i-1, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} (f(x_1, \ldots, 0, \ldots, x_n) \oplus f(x_1, \ldots, 1, \ldots, x_n))
\]

\[
= \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_i-1, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} M_1(m_2(\ldots (M_{r-1}(M_r \oplus 1) \oplus 1) \ldots)).
\]
\[
\frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} M_1 m_2 \ldots (M_{r-1}(M_r \oplus 1) \oplus 1) \ldots .
\]

\[
= \frac{1}{2^{n-1}} \sum_{j=1}^{r-1} (-1)^{j-1} 2^{n-1} - ((k_1 + k_2 - 1) + \ldots + k_{j+1})) = \frac{1}{2^{n-1}} \sum_{j=1}^{r-1} (-1)^{j-1} 2^{n-1} - \sum_{i=1}^{j+1} k_i
\]

by Equation 4.3 in Theorem 4.1. Note, \( M_1 m_2 \) is the first layer, \( M_3 \) is the second layer and etc.

Now let’s consider the variables in the \( l \)th layer, i.e., \( x_i \) is an essential variable of \( M_l \), \( 2 \leq l \leq r - 1 \). We have \( M_l = (x_i + a_i) m_l \) and

\[
\alpha_i^f = \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} (f(x_1, \ldots, 0, \ldots, x_n) \oplus f(x_1, \ldots, 1, \ldots, x_n))
\]

\[
= \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} M_1 \ldots M_{l-1} m_l (M_{l+1}(\ldots(M_r \oplus 1)\ldots) \oplus 1).
\]

\[
= \frac{1}{2^{n-1}} \sum_{j=1}^{r-l+1} (-1)^{j-1} 2^{n-1} - ((k_1 + \ldots + k_l - 1) + k_{l+1} + \ldots + k_{j+1} - 1)) = \frac{1}{2^{n-1}} \sum_{j=1}^{r-l+1} (-1)^{j-1} 2^{n-1} - \sum_{i=1}^{j+1} k_i
\]

by Equation 4.3 in Theorem 4.1. Note, \( M_1 \ldots M_{l-1} m_l \) is the first layer, \( M_{l+1} \) is the second layer, and etc.

Let \( x_i \) be the variable in the last layer \( M_r \), we have

\[
= \frac{1}{2^{n-1}} \sum_{(x_1, \ldots, x_{l-1}, x_{l+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}} M_1 M_2 \ldots M_{r-1} m_r = \frac{1}{2^{n-1}}
\]

by Lemma 4.1.

Variables in the same layer have the same activities, so we use \( A_i^f \) to stand for the activity number of each variable in the \( l \)th layer \( M_l \), \( 1 \leq l \leq r \). We find the formula of \( A_i^f \) for \( 2 \leq l \leq r - 1 \) is also true when \( l = r \) or \( r = 1 \). Hence, we write all the above as the following

**Theorem 4.2.** Let \( f \) be a NCF and written as in the Theorem 3.2 then the activity of each variable in the \( l \)th layer, \( 1 \leq l \leq r \), is

\[
A_i^f = \frac{1}{2^{n-1}} \sum_{j=1}^{r-l+1} (-1)^{j-1} 2^{n-1} - \sum_{i=1}^{j+1} k_i
\]

The average sensitivity of \( f \) is

\[
s_f = \frac{r}{\sum_{l=1}^{r} k_l A_i^f} = \frac{1}{2^{n-1}} \sum_{l=1}^{r} k_l \sum_{j=1}^{r-l+1} (-1)^{j-1} 2^{n-1} - \sum_{i=1}^{j+1} k_i
\]

We do some analysis about the formulas in Theorem 4.2, we have

**Corollary 4.2.** \( n \geq 3 \), \( A_1^f > A_2^f > \ldots > A_l^f \) and \( \frac{n}{2^{n-1}} \leq s_f < 2 - \frac{1}{2^{n-2}} \)

**Proof.**

\[
A_i^f = \frac{1}{2^{n-1}} \sum_{j=1}^{r-l+1} (-1)^{j-1} 2^{n-1} - \sum_{i=1}^{j+1} k_i \]

(4.5)

(4.6)
Since the sum is an alternate decreasing sequence and $k_{l+1} \geq 1$, we have
\[
\frac{1}{2n-1} (2^{n-k_1-\ldots-k_l} - 2^{n-k_1-\ldots-k_{l+1}}) < A_f < \frac{1}{2n-1} (2^{n-k_1-\ldots-k_l})
\]
Hence,
\[
A_{l+1}^f < \frac{1}{2n-1} (2^{n-k_1-\ldots-k_{l+1}}) \leq \frac{1}{2n-1} (2^{n-k_1-\ldots-k_{l-1}}) < A_f.
\]
We have
\[
k_1 A_1^f = \frac{k_1}{2^{n-1}} (2^{n-k_1} - 2^{n-k_1-k_2} + 2^{n-k_1-k_2-k_3} - \ldots (-1)^{r-1})
\]
\[
k_2 A_2^f = \frac{k_2}{2^{n-1}} (2^{n-k_1-k_2} - 2^{n-k_1-k_2-k_3} + 2^{n-k_1-k_2-k_3-k_4} - \ldots (-1)^{r-2})
\]
\[
\ldots \ldots
\]
\[
k_l A_l^f = \frac{k_l}{2^{n-1}} (2^{n-k_1-\ldots-k_l} - 2^{n-k_1-\ldots-k_l-k_{l+1}} - \ldots (-1)^{r-1})
\]
\[
\ldots \ldots
\]
\[
k_r A_r^f = \frac{k_r}{2^{n-1}}
\]
Hence, $s^f = \sum_{l=1}^{r} k_l A_l^f \geq \frac{k_1}{2^{n-1}} + \frac{k_2}{2^{n-1}} + \ldots + \frac{k_r}{2^{n-1}} = \frac{n}{2^{n-1}}$, so we know the NCF with LAYER NUMBER 1 has the minimal average sensitivity.

On the other hand, $s^f = \sum_{l=1}^{r} k_l A_l^f < \frac{k_1}{2^{n-1}} 2^{n-k_1} + \frac{k_2}{2^{n-1}} 2^{n-k_1-k_2} + \ldots + \frac{k_r}{2^{n-1}} 2^{n-k_1-\ldots-k_l} + \ldots + \frac{k_r}{2^{n-1}} = U(k_1, \ldots, k_r)$, where $k_1 + \ldots + k_r = n$, $k_i \geq 1$, $i = 1, \ldots, r-1$ and $k_r \geq 2$. We will find the maximal value of $U(k_1, \ldots, k_r)$ in the following.

First, we claim $k_r = 2$ if $U(k_1, \ldots, k_r)$ reach maximal value. Because if $k_r$ is increased by 1, and the last term makes $\frac{k_r}{2^{n-1}}$ more contributions to $U(k_1, \ldots, k_r)$, then there exists $l$, $k_l$ will be decreased by 1 $(k_1 + \ldots + k_r = n)$, hence
\[
\frac{k_l}{2^{n-1}} 2^{n-k_1-\ldots-k_l}
\]
will be decreased more than $\frac{1}{2^{n-1}}$.

Now, Look at $\frac{k_1}{2^{n-1}} 2^{n-k_1}$, it is obvious it attains the maximal value only when $k_1 = 1$ or 2 but obviously $k_1 = 1$ will be the choice since it also make all the other terms greater..

Now look at $\frac{k_2}{2^{n-1}} 2^{n-k_1-k_2}$, it attains the maximal value when $k_1 = k_2 = 1$ or $k_1 = 1$ and $k_2 = 2$, again, $k_2 = 1$ is the best choice to make all the other terms greater.

In general, if $k_1 = \ldots = k_{l-1} = 1$, then $\frac{k_l}{2^{n-1}} 2^{n-k_1-\ldots-k_l}$ attains its maximal value when $k_l = 1$, where $1 \leq l \leq r-1$.

In summary, we have showed that $U(k_1, \ldots, k_r)$ reaches maximal value when $r = n-1$, $k_1 = \ldots = k_{n-2} = 1$, $k_{n-1} = 2$ and
\[
\text{Max } U(k_1, \ldots, k_r) = U(1, \ldots, 1, 2) = \frac{1}{2^{n-1}} (2^{n-1} + 2^{n-2} + \ldots + 2^2 + 2) = 2 - \frac{1}{2^{n-2}}.
\]

**Remark 4.3.** So, we know the average sensitivity is bounded by constants for any NCF with any number of variables since the minimal value approaches to 0 and the maximal value of $U(k_1, \ldots, k_r)$ approaches to 2 as $n \to \infty$. Hence, $0 < s^f < 2$ for any NCF with arbitrary number of variables.

In the following, we evaluate the formula Equation 4.6 for some parameters $k_1, \ldots, k_r$, we have
Lemma 4.4. 1) When \( r = n - 1 \), \( k_1 = \ldots = k_{n-2} = 1 \), \( k_{n-1} = 2 \), \( s^f = \frac{4}{3} - \frac{3 + (-1)^n}{3 \times 2^n} \).

2) Given \( n \geq 4 \), \( r = n - 2 \), \( k_1 = \ldots = k_{n-3} = 1 \), \( k_{n-2} = 3 \), \( s^f = \frac{4}{3} - \frac{9 + 5(-1)^{n-1}}{3 \times 2^{n-1}} \).

3) If \( n \) is even and \( n \geq 6 \), \( r = \frac{n}{2} \), \( k_1 = 1 \), \( k_2 = \ldots = k_{\frac{n}{2}-1} = 2 \), \( k_{\frac{n}{2}} = 3 \), \( s^f = \frac{4}{3} - \frac{4}{3 \times 2^n} \).

Hence, these three cardinalities are equal if \( n \) is even.

Proof. When \( r = n - 1 \), \( k_1 = \ldots = k_{n-2} = 1 \), \( k_{n-1} = 2 \) by Equation 4.6, we have

\[
s^f = \sum_{i=1}^{r} k_i a_i^f = \frac{1}{2^{n-1}} \sum_{i=1}^{n-1} k_i (-1)^{j-1} 2^n - \sum_{i=1}^{j-1} k_i
\]

\[
= \frac{1}{2^{n-1}} \sum_{i=1}^{n-1} k_i (-1)^{j-1} 2^n - \sum_{i=1}^{j-1} k_i
\]

\[
= \frac{1}{2^{n-1}} \sum_{i=1}^{n-1} \left( \frac{1}{3} 2^{n-1} + \frac{1}{3} (-1)^{n-1} + 2 \right) = \frac{4}{3} - \frac{3 + (-1)^n}{3 \times 2^n}
\]

The other two formulas are also routine simplifications of Equation 4.6. \( \square \)

Based on our numerical calculation, Lemma 4.4 and the proof of Corollary 4.2, we have the following.

Conjecture 4.5. The maximal value of \( s^f \) is \( s^f = \frac{4}{3} - \frac{3 + (-1)^n}{3 \times 2^n} \). It will be reached if the NCF has the maximal LAYER NUMBERS \( n - 1 \), i.e., if \( r = n - 1 \), \( k_1 = \ldots = k_{n-2} = 1 \), \( k_{n-1} = 2 \). When \( n \) is even, this maximal value is also reached by NCF with parameters \( n \geq 4 \), \( r = n - 2 \), \( k_1 = \ldots = k_{n-3} = 1 \), \( k_{n-2} = 3 \) or \( n \geq 6 \), \( r = \frac{n}{2} \), \( k_1 = 1 \), \( k_2 = \ldots = k_{\frac{n}{2}-1} = 2 \) and \( k_{\frac{n}{2}} = 3 \).

Remark 4.6. When \( n = 6 \), the NCF with \( k_1 = 1 \), \( k_2 = 2 \), \( k_3 = 1 \) and \( k_4 = 2 \) also has the maximal average sensitivity \( \frac{4}{10} \). But this can not be generalized. If the above conjecture is true, then we have \( 0 < s^f < \frac{4}{3} \) for any NCF with arbitrary number of variables. In other words, both \( 0 \) and \( \frac{4}{3} \) are uniform tight bounds for any NCF.

We point out, given the algebraic normal form of \( f \), it is easy to find all of its canalyzing variables (the first layer \( M_1 \)), then write \( f = M_1 g + b \), repeating the schedule to \( g \), we can easily to determine if \( f \) is NCF, if yes, then write it as the form in Theorem 3.2.

We end this section by the following example.

Example 4.7. Let \( N(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 \oplus x_2 x_3 x_4 \oplus x_1 x_3 \oplus x_3 x_4 \oplus 1 \) and

\[
Y(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 x_4 \oplus x_1 x_2 x_3 \oplus x_1 x_2 x_4 \oplus x_1 x_3 x_4 \oplus x_1 x_3 \oplus x_1 x_4 \oplus x_1 x_4 \oplus x_1 x_4 \oplus x_1 x_4 \oplus x_1 x_4 .
\]

For \( N(x_1, x_2, x_3, x_4) \), for all the 4 variables, we found when \( x_2 = 1 \) or \( x_3 = 0 \), then functions becomes constant 1, so we know \( N(x_1, x_2, x_3, x_4) = (x_2 \oplus 1) (x_3) N_1 \oplus 1 \).

Actually,

\[
N(x_1, x_2, x_3, x_4) = x_1 x_2 x_3 \oplus x_2 x_3 x_4 \oplus x_1 x_3 \oplus x_3 x_4 \oplus 1 = x_3 (x_1 x_2 \oplus x_2 x_4 \oplus x_1 x_4) \oplus 1
\]

\[
= x_2 (x_1 x_4 \oplus x_1 x_4) \oplus 1 = x_3 (x_2 x_4 \oplus x_1 x_4) \oplus 1 .
\]

Since \( x_1 \oplus x_4 \) has no canalyzing variable, we know \( N \) is not NCF, but a partially NCF.

For \( Y(x_1, x_2, x_3, x_4, x_5) \), we find \( x_1 = 0 \) or \( x_3 = 1 \), the function will be reduced to 0, so we know \( Y = x_1 (x_3 \oplus 1) Y_1 \).

Where \( Y_1 = x_3 x_4 x_5 \oplus x_2 x_4 \oplus x_4 \oplus 1 \), for this function we find only when \( x_4 = 0 \), \( Y_1 \) will be reduced to 0, so \( Y_1 = x_4 Y_2 \oplus 1 \), where \( Y_2 = x_2 x_5 \oplus x_2 \oplus 1 \), and finally, we have \( Y_2 = x_2 (x_5 \oplus 1) \oplus 1 \), so \( Y \) is NCF with \( n = 5 \), \( r = 3 \) and \( k_1 = 2 \), \( k_2 = 1 \), \( k_3 = 2 \), \( M_1 = x_1 (x_3 \oplus 1) \), \( M_2 = x_4 \) and \( M_3 = x_2 (x_5 \oplus 1) \); hence its Hamming weight is 5 by Equation 4.3, and its average sensitivity is \( \frac{1}{10} \) by Equation 4.6.
5. Conclusion

We obtain a complete characterization for nested canalyzing functions (NCFs) by deriving its unique algebraic normal form (polynomial form). We introduced a new invariant, LAYER NUMBER for nested canalyzing function. So, the dominance of nested canalyzing variables is quantified. Consequently, we obtain the explicit formula of the number of nested canalyzing functions. Based on the polynomial form, we also obtain the formula of the Hamming weight of each NCF. The activity number of each variable of a NCF is also provided with an explicit formula. Consequently, we proved the average sensitivity of any NCF is less than 2, hence, we proved why NCF is stable theoretically. Finally, we conjecture that the tight upper bound for the average sensitivity of any NCF is $\frac{4}{3}$.

References

[1] Kazuyuki Amano, “Tight bounds on the average sensitivity of k-CNF,” Theory of Computing, Vol 7 (2011), pp. 45-48.
[2] E. Balleza, E. R. Alvarez-Buylla, A. Chaos, S. Kauffman, I. Shmulevich, and M. Aldana, “Critical dynamics in genetic regulatory networks: Examples from four kingdoms,” PLoS ONE, 3 (2008), p. e2456.
[3] C. Barrett, C. Herring, J. Reed, and B. Palsson, “The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states,” Proc Natl Acad Sci USA, 102 (2005), pp. 19103-19108.
[4] E. A. Bender, J. T. Butler, “Asymptotic approximations for the number of fanout-tree functions,” IEEE Trans. Comput. 27 (12) (1978) 1180-1183.
[5] A. Bernasconi, “Mathematical techniques for the analysis of Boolean functions,” Ph.D. thesis, Dipartimento di Informatica, Universita di Pisa (March, 1998).
[6] A. Bernasconi, “Sensitivity vs. block sensitivity (an average-case study)” Information processing letters 59 (1996) 151-157.
[7] Ravi B. Boppana, “The average sensitivity of bounded-depth circuits” Information processing letters 63 (1997) 257-261.
[8] J. T. Butler, T. Sasao, and M. Matsuura, “Average path length of binary decision diagrams,” IEEE Transactions on Computers, 54 (2005), pp. 1041-1053.
[9] David Cunrigh, Sugata Gangopadhyay, Subhajoy Maitra, Pantelimon Stănică, “Laced Boolean functions and subset sum problems in finite fields,” Discrete applied mathematics, 159 (2011), pp. 1059-1069.
[10] Shijian Chen and Yiguang Hong, “Control of random Boolean networks via average sensitivity of Boolean functions,” Chin. Phys. B Vol. 20, No 3 (2011) 036401.
[11] Demetres Christofides, “Influences of Monotone Boolean Functions,” Preprint 2009.
[12] S. A. Cook, C. Dwork, R. Reischuk, “Upper and lower time bounds for parallel random access machines without simultaneous writes,” SIAM J. Comput, 15 (1986), pp. 87-89.
[13] S. E. Harris, B. K. Sawhill, A. Wuensch, and S. Kauffman, “A model of transcriptional regulatory networks based on biases in the observed regulation rules,” Complex, 7 (2002), pp. 2340.
[14] M. Herrgard, B. Lee, V. Portnoy, and B. Palsson, “Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in saccharomyces cerevisiae,” Genome Res, 16 (2006), pp. 627635.
[15] A. Jarrah, R. Laubenbacher, and A. Veliz-Cuba, “A polynomial framework for modeling and analyzing logical models.” In Preparation, 2008.
[16] A. Jarrah, B. Ropasa and R. Laubenbacher, “Nested Canalyzing, Unate Cascade, and Polynomial Functions”, Physica D 233 (2007), pp. 167-174.
[17] Winfried Just, “The steady state system problem is NP-hard even for monotone quadratic Boolean dynamical systems ” Preprint,2006
[18] Winfried Just, Ilya Shmulevich, John Konvalina, “The number and probability of canalyzing functions”, Physica D 197 (2004), pp. 211-221.
[19] S. A. Kauffman, “The Origins of Order: Self-Organization and Selection in Evolution”, Oxford University Press, New York, Oxford (1993).
[20] S. A. Kauffman, C. Peterson, B. Samueleson, C. Troein, “Random Boolean Network Models and the Yeast Transcription Network”, Proc. Natl. Acad. Sci 100 (25) (2003), pp. 14796-14799.
[21] S. A. Kauffman, C. Peterson, B. Samuelsson, and C. Troein, “Genetic networks with canalyzing Boolean rules are always stable,” PNAS, 101 (2004), pp. 1710217107.
[22] N. Keller and H. Pilpel, “Linear transformations on monotone functions on the discrete cube,” Discrete Math. 309 (2009), 4210-4214.
[23] R. Laubenbacher and B. Pareigis, “Equivalence relations on finite dynamical systems,” Advances in applied mathematics 26 (2001), pp. 237-251.
[24] W. Liu, H. Lähdevirta, Edward R. Dougherty and I. Shmulevich, “Inference of Boolean Networks Using Sensitivity Regularization,” EURASIP Journal of Bioinformatics and System Biology Volume 2008, Article ID 780541, 12 pages.
[25] Lori Layne, “Biologically Relevant Classes of Boolean Functions”, Ph.D Thesis, Clemson University (2011).
[26] Jiyou Li, “On the average sensitivity of the weighted sum function,” arXiv:1108.3198v2 [cs.IT] 18 Aug 2011.
[27] Yuan Li, “Results on Rotation Symmetric Polynomials Over GF(p),” Information Sciences 178 (2008), pp. 280-286.
[28] Yuan Li, David Murrugarra, John O Adeyeye and Reinhard Laubenbacher “Multi-State Canalizing Functions over Finite Fields,” Preprint (2010)
[29] R. Lidl and H. Niederreiter, “Finite Fields”, Cambridge University Press, New York (1977).
[30] A. A. Moreira and L. A. Amaral, “Canalizing Kauffman networks: Nonergodicity and its effect on their critical behavior,” Phys. Rev. lett. 94 (21) (2005), 218702.
[31] D. Murrugarra and R. Laubenbacher, “Generalized nested canalizing functions capture the logic of gene regulation,” (2010), p. under review.
[32] D. Murrugarra and R. Laubenbacher, “The number of multistate nested canalizing functions,” arXiv 1108.0206v2 [math.AG] 7 Aug 2011.
[33] S. Nikolajewa, M. Friedel, and T. Wilhelm, “Boolean networks with biologically relevant rules show ordered behavior star, open,” Biosystems, 90 (2007), pp. 4047.
[34] N. Nisan, “CREW PRAMs and decision tree,” SIAM J. Comput, 20 (6) (1991), PP. 999-1070.
[35] Xiaoning Qian and Edward R. Dougherty, “A comparative study on sensitivities of Boolean networks,” 978-1-6184-792-4/10 2011 IEEE.
[36] T. Sasao, K. Kinoshita, “On the number of fanout-tree functions and unate cascade functions,” IEEE Trans. Comput. 28 (1) (1979) 66-72.
[37] Ilya Shmulevich, “Average sensitivity of typical monotone Boolean functions,” arXiv:math/0507030v1 [math. Co] 1 July 2005.
[38] Ilya Shmulevich and Stuart A. Kauffman, “Activities and sensitivities in Boolean network models,” Physical Review Letters, Vol 93, Number 4 (2004), 048701.
[39] Igor. E. Shparlinski, “Bounds on the Fourier coefficients of the weighted sum function,” Information Processing Letters. 103 (2007), 83-87.
[40] Steffen Schober and Martin Bossert, “Analysis of random Boolean networks using the average sensitivity,” arXiv: 0704.0197v1 [nlin.CG] 2 Apr 2007.
[41] C. H. Waddington, “Canalisation of development and the inheritance of acquired characters,” Nature, 150 (1942), pp. 563-566.
[42] Madars Virza, “Sensitivity versus block sensitivity of Boolean functions,” arXiv:1008.0521v2 [cs.CC] 8 Dec 2010.
[43] Shengyu Zhang, “Note on the average sensitivity of monotone Boolean functions,” Preprint 2011.

1Department of Mathematics, Winston-Salem State University, NC 27110,USA, email: liyu@wssu.edu, 2Department of Mathematics, Winston-Salem State University, NC 27110,USA, email: adeyeyej@wssu.edu, 3Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA 24061,USA, email: reinhard@vbi.vt.edu