From multivalued to Boolean functions: preservation of soft nested canalization

Élisabeth Remy, Paul Ruet

From multivalued to Boolean functions: preservation of soft nested canalization

Élisabeth Remy[a], Paul Ruet[b]

[a] *Aix Marseille Univ, CNRS, I2M, Marseille, France*
[b] *CNRS, Université Paris Cité, Paris, France*

Abstract

Nested canalization (NC) is a property of Boolean functions which has been recently extended to multivalued functions. We study the effect of the Van Ham mapping (from multivalued to Boolean functions) on this property. We introduce the class of softly nested canalizing (SNC) multivalued functions, and prove that the Van Ham mapping sends SNC multivalued functions to NC Boolean functions. Since NC multivalued functions are SNC, this preservation property holds for NC multivalued functions as well. We also study the relevance of SNC functions in the context of gene regulatory network modelling.

Keywords: nested canalizing functions, Boolean functions, multivalued functions, regulatory network modelling

1. Introduction

Dynamical properties of biological systems, such as stability and robustness, are often associated with the canalyzing property. This notion is at the basis of the work of Waddington in embryology, which describes, in the 1940s, epigenetic landscapes representing embryogenesis by canalizing configurations [26]. In the 1990's, S. Kauffman introduced the class of canalizing Boolean functions [4, 5] to formalize the canalizing behaviour observed in gene regulatory networks. In short, canalizing Boolean functions are functions $f : (\mathbb{Z}/2\mathbb{Z})^n \to \mathbb{Z}/2\mathbb{Z}$ (or $\mathbb{R}$) such that at least one input variable, say

Email addresses: elisabeth.remy@univ-amu.fr (Élisabeth Remy), ruet@irif.fr (Paul Ruet)

URL: https://mabios.math.cnrs.fr/ (Élisabeth Remy), https://www.irif.fr/~ruet/ (Paul Ruet)

x_i ($1 \leqslant i \leqslant n$), has a value $a = 0$ or 1 which determines the value of $f(x)$. Nested canalizing (NC) functions provide a "recursive" version of canalizing functions: an NC function f is canalizing and, moreover, its restriction $f\restriction_{x_i \neq a}$ is itself NC. NC functions are particularly interesting because of their "low complexity": in particular, their average sensitivity (a measure of complexity related to spectral concentration, learning properties, decision tree complexity [13] and stability properties [4, 6]) is bounded above by a constant [8, 7].

A natural question is the extension of this canalizing property to multivalued functions, involving more than two expression levels, which are often needed in the modeling of biological systems to circumvent the too restrictive Boolean representation [23]. The present paper investigates the property of nested canalization for multivalued functions $f : (\mathbb{Z}/k\mathbb{Z})^n \to \mathbb{Z}/k\mathbb{Z}$ or $\mathbb{R}$ for some $k \geqslant 2$. A definition of nested canalizing (NC) multivalued function has been proposed in [10, 11]. In [16], we introduced a more general class of multivalued functions, which we called weakly nested canalizing (WNC), for which we proved that the average sensitivity is bounded above by a constant.

Multivalued modelling allows to distinguish different thresholds of action of a gene on its targets, and is of course more complex to analyse than the Boolean case. Boolean mappings have been proposed, associating one Boolean variable to each threshold. This gives access to a large number of theoretical works and numerical tools specific to the analysis of Boolean models. Thus we are interested in the relationship between the notions of canalization for multivalued and Boolean functions. Clearly an NC Boolean function is NC multivalued, but what can we say about the other direction? More precisely, a mapping β from multivalued functions to Boolean functions have been defined by Van Ham in [25], and it have been proved in [2, 24] to be in some sense unique (it is the only injective such mapping preserving neighbours in the state space, and for which the regulatory graphs are unchanged).

We prove here that β maps NC multivalued functions to NC Boolean functions. To this end, we introduce in Section 2 the notion of softly nested canalizing (SNC) multivalued functions, and prove that NC $\Rightarrow$ SNC $\Rightarrow$ WNC (Proposition 2 and Section 2.3). Then we actually prove in Theorem 3 that SNC functions are mapped by β to NC Boolean functions. On the other hand, we show that Theorem 3 does not extend to WNC functions: we shall see in Section 3.5 an example of a WNC function which is mapped by β to a non NC Boolean function.

NC functions appear as appropriate rules in Boolean models of gene regulatory networks. Indeed, Boolean networks have been proposed as models

for gene regulatory networks [22], with different nodes corresponding to different genes. The activity of a gene is regulated by the activity of other genes to which it is connected through a Boolean function. These networks are built from biological data and knowledge from the literature, and their associated Boolean functions are far from random. It turns out that NC functions are predominant in the large databases of Boolean gene networks [21]. In Section 4 we extract the multivalued genes from this database, and we study the canalizing properties of their multivalued functions and their corresponding Boolean functions.

2. Nested canalizing multivalued functions

The concept of nested canalization, originally considered for Boolean functions [4], has been extended to multivalued functions in [10, 11, 3]. We first recall here this definition, before introducing soft nested canalization in Section 2.1.

Let k, n be positive integers, $k \geqslant 2$. $\mathbb{Z}/k\mathbb{Z}$ is the ring of integers modulo k. A function $f : (\mathbb{Z}/k\mathbb{Z})^n \to \mathbb{Z}/k\mathbb{Z}$ is said to be *canalizing with respect to coordinate i and $(a, b) \in \mathbb{Z}/k\mathbb{Z} \times \mathbb{Z}/k\mathbb{Z}$* if there exists a function $g : (\mathbb{Z}/k\mathbb{Z})^n \to \mathbb{Z}/k\mathbb{Z}$ different from the constant b such that

$$f(x) = \begin{cases} b & \text{if } x_i = a \\ g(x) & \text{if } x_i \neq a. \end{cases}$$

We shall simply say that f is *canalizing* if it is canalizing with respect to some i, a, b.

A *segment* is a subset of $\mathbb{Z}/k\mathbb{Z}$ of the form $\{0, \ldots, i\}$ or $\{i, \ldots, k-1\}$. Let $\sigma \in \mathfrak{S}_n$ be a permutation, $A_1, \ldots, A_n$ be segments, and $c_1, \ldots, c_{n+1} \in \mathbb{Z}/k\mathbb{Z}$ be such that $c_n \neq c_{n+1}$. Then f is said to be *nested canalizing (NC) with respect to σ, $A_1, \ldots, A_n$, $c_1, \ldots, c_{n+1}$* if

$$f(x) = \begin{cases} c_1 & \text{if } x_{\sigma(1)} \in A_1 \\ c_2 & \text{if } x_{\sigma(1)} \notin A_1, x_{\sigma(2)} \in A_2 \\ \vdots & \quad \vdots \\ c_n & \text{if } x_{\sigma(1)} \notin A_1, \ldots, x_{\sigma(n-1)} \notin A_{n-1}, x_{\sigma(n)} \in A_n \\ c_{n+1} & \text{if } x_{\sigma(1)} \notin A_1, \ldots, x_{\sigma(n-1)} \notin A_{n-1}, x_{\sigma(n)} \notin A_n. \end{cases}$$

We shall simply say that f is *NC* if it is NC with respect to some σ, $A_1, \ldots, A_n$, $c_1, \ldots, c_{n+1}$.

Note that the segment A_i plays the role of a threshold for the variable $x_{\sigma(i)}$. For an NC multivalued function f as defined in [10, 11, 3], each variable $x_{\sigma(i)}$ is used exactly once in the process of canalization (and the value of f if determined by the threshold corresponding to A_i). It is possible to relax the definition by allowing canalization along successive thresholds for a same variable. This is essentially what we shall do in defining softly nested canalizing functions below. In doing so, we shall also remove the condition $c_n \neq c_{n+1}$.

2.1. Softly nested canalizing multivalued functions

Let n be a positive integer. For each $i \in \{1, \ldots, n\}$, let $k_i > 0$, Ω_i be a set of k_i integers, $\Omega = \prod_i \Omega_i$, and $f : \Omega \to \mathbb{R}$. Note that we do not require $k_i \geqslant 2$ for all i. If $k_j = 1$ for some j, f could be viewed as a function with one less variable, *i.e.* as a function on $\prod_{i \neq j} \Omega_i$, but we still consider it as a function defined on $\prod_i \Omega_i$.

We shall say that f is *softly canalizing with respect to coordinate i and* $(a, b) \in \Omega_i \times \mathbb{R}$ if $f(x) = b$ whenever $x_i = a$, and simply that it is *softly canalizing* if it is softly canalizing with respect to some i, a, b.

Note that this definition differs slightly from the usual definition by the absence of condition on the values of f for $x_i \neq a$: we do not require the existence of some x such that $x_i \neq a$ and $f(x) \neq b$. In particular, constant functions are softly canalizing, though not canalizing.

If f is canalizing with respect to i, a, b and $k_i \geqslant 2$, we shall consider

$$f\restriction_{x_i \neq a} : \Omega \cap \{x \mid x_i \neq a\} \to \mathbb{R},$$

the restriction of f to the set of $x \in \Omega$ such that $x_i \neq a$.

The class of softly nested canalizing on $\Omega = \prod_i \Omega_i$ is then defined by induction on the cardinality $|\Omega| = \prod_i k_i$ of Ω. If $|\Omega| = 1$, *i.e.* $k_i = 1$ for all i, any $f : \Omega \to \mathbb{R}$ is *softly nested canalizing (SNC) on* Ω. If $|\Omega| > 1$, $f : \Omega \to \mathbb{R}$ is *SNC on* Ω if there exist i, a, b such that

- $k_i \geqslant 2$,
- a is either the min or the max of Ω_i,
- f is softly canalizing with respect to i, a, b,
- $f\restriction_{x_i \neq a}$ is SNC on $\Omega \cap \{x \mid x_i \neq a\}$, a strict subset of Ω.

Intuitively, a function $f : \Omega \to \mathbb{R}$ is SNC if its domain Ω can be "peeled" by successively removing coordinate hyperplanes (defined by equations of the form $x_i = a$ with a minimal or maximal) whose points are mapped by f to the same value, whence the following characterization:

Proposition 1. *Letting $K = \sum_i k_i - n$, f is SNC if and only if there exist a function $v : \{1, \dots, K\} \to \{1, \dots, n\}$, numbers $a_i \in \Omega_{v(i)}$ and $b_i \in \mathbb{R}$ for each $i \in \{1, \dots, K\}$ such that:*

$$f(x) = \begin{cases} b_1 & \text{if } x_{v(1)} = a_1 \\ b_2 & \text{if } x_{v(1)} \neq a_1, x_{v(2)} = a_2 \\ \vdots & \quad \vdots \\ b_K & \text{if } x_{v(1)} \neq a_1, \dots, x_{v(K-1)} \neq a_{K-1}, x_{v(K)} = a_K \end{cases}$$

and for each $i \in \{1, \dots, K\}$, $\Omega'_{v(i)} = \Omega_{v(i)} \setminus \{a_j \mid j < i, v(i) = v(j)\}$ is not a singleton and a_i is either the min or the max of $\Omega'_{v(i)}$.

In decomposing an NC function $f : (\mathbb{Z}/k\mathbb{Z})^n \to \mathbb{Z}/k\mathbb{Z}$, each coordinate $i \in \{1, \dots, n\}$ is considered exactly once (in some order prescribed by a permutation σ) and the value of f is fixed for $x_{\sigma(i)}$ in some segment A_i. This can be realized by successively fixing the value of f for each $\alpha \in A_i$, and therefore, the class of SNC functions contains the class of NC functions, as stated in the following Proposition:

Proposition 2. *If $f : (\mathbb{Z}/k\mathbb{Z})^n \to \mathbb{Z}/k\mathbb{Z}$ is NC, then it is SNC.*

Proof. Assume f is NC with respect to σ, $A_1, \dots, A_n$, $c_1, \dots, c_{n+1}$. For each $i \in \{1, \dots, n\}$, let

$$A_i = \{\alpha_i^1, \dots, \alpha_i^{|A_i|}\}$$
$$(\mathbb{Z}/k\mathbb{Z}) \setminus A_i = \{\alpha_i^{1+|A_i|}, \dots, \alpha_i^{k-1}\}.$$

More precisely, the $K = n(k-1)$ numbers $\alpha_i^j \in \mathbb{Z}/k\mathbb{Z}$ are defined according to the following ordering which depends on the nature of the segment A_i:

- either $0 \in A_i$, and then we choose the numbers α_i^j so that $0 = \alpha_i^1 < \cdots < \alpha_i^{|A_i|}$ and $k-1 = \alpha_i^{1+|A_i|} > \cdots > \alpha_i^{k-1}$;
- or $k-1 \in A_i$, and we choose the α_i^j so that $k-1 = \alpha_i^1 > \cdots > \alpha_i^{|A_i|}$ and $0 = \alpha_i^{1+|A_i|} < \cdots < \alpha_i^{k-1}$.

If both 0 and $k-1 \in A_i$, *i.e.* if $A_i = \mathbb{Z}/k\mathbb{Z}$, we choose any of the two orders. Now, for each $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, k-1\}$, let

$$\beta_i^j = \begin{cases} c_i & \text{if } j \leqslant |A_i| \\ c_{n+1} & \text{otherwise.} \end{cases}$$

To comply with the characterization of SNC functions (Proposition 1), we relabel the numbers α_i^j, β_i^j by identifying the list

$$\alpha_1^1, \ldots, \alpha_1^{|A_1|}, \ldots, \alpha_n^1, \ldots, \alpha_n^{|A_i|}, \alpha_1^{1+|A_1|}, \ldots, \alpha_1^k, \ldots, \alpha_n^{1+|A_n|}, \ldots, \alpha_n^{k-1}$$

as the list $a_1, \ldots, a_K$, and by identifying similarly the list

$$\beta_1^1, \ldots, \beta_1^{|A_1|}, \ldots, \beta_n^1, \ldots, \beta_n^{|A_n|}, \beta_1^{1+|A_1|}, \ldots, \beta_1^k, \ldots, \beta_n^{1+|A_n|}, \ldots, \beta_n^{k-1}$$

as the list $b_1, \ldots, b_K$. Call φ this relabelling, which maps $r \in \{1, \ldots, K\}$ to the pair $\varphi(r) = (i, j)$ such that $a_r = \alpha_i^j$ and $b_r = \beta_i^j$. For instance, $\varphi(1) = (1, 1)$ and $\varphi(K) = (n, k-1)$. Then finally, a function $v : \{1, \ldots, K\} \to \{1, \ldots, n\}$ is defined by $v(r) = \sigma(i)$ if $\varphi(r) = (i, j)$. Then f clearly enjoys the characterization of SNC functions, with the choice of function v and numbers a_r, b_r. □

2.2. Some examples

- We have already noticed that constant functions from $(\mathbb{Z}/k\mathbb{Z})^n$ to $\mathbb{Z}/k\mathbb{Z}$ are SNC but not NC.

- An easy induction on k shows that the functions min and max : $(\mathbb{Z}/k\mathbb{Z})^2 \to \mathbb{Z}/k\mathbb{Z}$ are SNC. For instance, min = $\min_k : \{0, \ldots, k-1\}^2 \to \{0, \ldots, k-1\}$ is softly canalizing with respect to $1, 0, 0$, then $\min_k{\restriction}_{x_1 \neq 0}$ is softly canalizing with respect to $2, 0, 0$, and $\min_k{\restriction}_{x_1 \neq 0, x_2 \neq 0}$ is identical to the function $\min_{k-1} : \{1, \ldots, k-1\}^2 \to \{1, \ldots, k-1\}$, which is SNC.

  However, they are not NC [3]. Intuitively, there is more freedom in the construction of SNC functions, which can be built by successively "peeling" coordinate hyperplanes defined on some coordinate i (*i.e.* by some equation $x_i = a$), then a coordinate hyperplane defined on some other coordinate j, and later a coordinate hyperplane defined on i again.

- For the same reason, the identity from $(\mathbb{Z}/k\mathbb{Z})^n$ to $(\mathbb{Z}/k\mathbb{Z})^n$ is SNC but not NC. As we shall see in Section 4, this applies to the functions governing the regulation of genes such as **Raf**, **Dsor**, **Drk**, **Stat92E**.

- In Section 4, we shall see other examples of SNC and non SNC functions occurring in multivalued gene regulatory networks found in the literature.

2.3. WNC multivalued functions and average sensitivity

It is possible to consider a slightly more general notion of multivalued canalization, by removing the condition on canalizing values in the definition of SNC functions. This is what we do in [16]: *weakly nested canalizing (WNC) functions* $f : (\mathbb{Z}/k\mathbb{Z})^n \to \mathbb{Z}/k\mathbb{Z}$ are defined like SNC functions, but the values a used to define $f(x)$ for $x_i = a$ need not be extreme values (initially 0 or $k-1$), they can be intermediate values: $0 < a < k-1$.

By Proposition 2, NC $\Rightarrow$ SNC $\Rightarrow$ WNC, and we prove in [16] that WNC (hence NC and SNC) multivalued functions have "low complexity" in the sense that their *average sensitivity* (see [13, Chapter 8]) is bounded above by a constant (independent of n), while the average sensitivity of an arbitrary multivalued function is of order $\mathcal{O}(n)$.

It is worth noticing that the above implications are strict. For instance, the function $f : \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$ defined by the following table:

$$f(x,y) = \begin{array}{c|ccc} x\backslash y & 0 & 1 & 2 \\ \hline 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 2 & 0 & 2 \end{array}$$

is WNC but not SNC: the first canalization has to take the intermediate value $x = 1$ (and determines $f(x,y) = 1$), then the second canalization takes $y = 0$ (determining $f(x,y) = 2$) or $y = 1$ (determining $f(x,y) = 0$), and finally $x = 0$ or $x = 2$.

3. Preservation of nested canalization under Boolean mapping

3.1. Mapping multivalued functions to Boolean functions

Let n, $\Omega_i = \mathbb{Z}/k_i\mathbb{Z}$ and $\Omega = \prod_{i=1}^n \Omega_i$ be as above.

The injective mapping $\beta : \Omega \hookrightarrow \{0,1\}^k$, with $k = \sum_i k_i - n$, proposed by Van Ham [25] is defined as follows. For $i \in \{1,\dots,n\}$ and $1 \leqslant a \leqslant k_i - 1$, let $\beta_{i,a} : \Omega \to \{0,1\}$ be defined by

$$\beta_{i,a}(x) = \begin{cases} 1 & \text{if } x_i \geqslant a \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\beta_i(x) = (\beta_{i,1}(x), \dots, \beta_{i,k_i-1}(x)) = 1^{x_i} 0^{k_i-1-x_i},$$

where $1^\ell = (1, \dots, 1)$ is the ℓ-tuple of 1's (and similarly for 0^ℓ), and tuples of 0's and 1's are represented by words (without commas and parentheses). The maps can be combined to define $\beta : \Omega \hookrightarrow \{0,1\}^k$ by

$$\begin{aligned}\beta(x) &= (\beta_{1,1}(x), \dots, \beta_{1,k_1-1}(x), \dots, \beta_{n,1}(x), \dots, \beta_{n,k_n-1}(x)) \\ &= 1^{x_1}0^{k_1-1-x_1} \cdots 1^{x_n}0^{k_n-1-x_n}.\end{aligned}$$

The image $\beta(\Omega)$ of β is a strict subset of $\{0,1\}^k$ (unless $k_i = 2$ for all i), and a point $x \in \{0,1\}^k$ is said *admissible* when $x \in \beta(\Omega)$. Given $F : \Omega \to \Omega$, by injectivity of β, the equality

$$F^\beta \circ \beta = \beta \circ F,$$

defines $F^\beta : \beta(\Omega) \to \beta(\Omega)$, the *Booleanization* of F.

Clearly, if F is Boolean (*i.e.* $k_i = 2$ for all i), then $\beta(\Omega) = \Omega = (\mathbb{Z}/2\mathbb{Z})^n$ and $F^\beta = F$.

Let us mention that other injective mappings from Ω to $\{0,1\}^k$ may be defined but β is the only one preserving neighbours and regulatory graphs [2, 24]:

- two neighbouring states (*i.e.* states $x, y \in \Omega$ whose Hamming distance equals 1) are mapped by β to neighbouring states in $\beta(\Omega)$,
- the global regulatory graphs $\mathcal{G}(F)$ and $\mathcal{G}(F^\beta)$ underlying the dynamics F and F^β are isomorphic. We refer to, e.g., [15, 17, 18] for the definition of $\mathcal{G}(F)$.

The purpose of the remainder of this section is to prove that β maps SNC multivalued functions to NC (Boolean) functions. Before doing this in Theorem 3 we need to recall the definition of nested canalization for Boolean functions.

3.2. Nested canalizing Boolean functions

Let k be a positive integer, $\sigma \in \mathfrak{S}_k$ be a permutation, and $a_1, \dots, a_k$, $b_1, \dots, b_k \in \{0,1\}$. We recall that a Boolean function $f : \{0,1\}^k \to \{0,1\}$ is said to be *nested canalizing (NC) with respect to* $\sigma, a_1, \dots, a_k, b_1, \dots, b_k$ if

$$f(x) = \begin{cases} b_1 & \text{if } x_{\sigma(1)} = a_1 \\ b_2 & \text{if } x_{\sigma(1)} \neq a_1, x_{\sigma(2)} = a_2 \\ \vdots & \quad \vdots \\ b_k & \text{if } x_{\sigma(1)} \neq a_1, \dots, x_{\sigma(k-1)} \neq a_{k-1}, x_{\sigma(k)} = a_k. \end{cases} \qquad (*)$$

Note that we slightly modify the usual definition by not requiring $f(x) = 1 - b_k$ when $x_{\sigma(1)} \neq a_1, \ldots, x_{\sigma(k-1)} \neq a_{k-1}, x_{\sigma(k)} \neq a_k$.

We shall simply say that f is *NC* if it is NC with respect to some $\sigma, a_1, \ldots, a_k, b_1, \ldots, b_k$, and that a function $F = (F_1, \ldots, F_k) : \{0,1\}^k \to \{0,1\}^k$ is *NC* if $F_j : \{0,1\}^k \to \{0,1\}$ is NC for all j.

3.3. Nested canalizing partial Boolean functions

In Section 3.4, we shall consider functions $f : X \to \{0,1\}$ defined on a subset $X \subseteq \{0,1\}^k$. We extend the notion of nested canalization to such partial functions, by simply saying that $f : X \to \{0,1\}$ is *NC* with respect to the above data if condition $(*)$ holds for all $x \in X$.

3.4. Relation between β and nested canalization

Theorem 3. *If $F = (F_1, \ldots, F_n) : \Omega \to \Omega$ is SNC, then $F^\beta : \beta(\Omega) \to \beta(\Omega)$ is NC.*

Proof. As in the definition of β, we shall use double indices for Boolean functions, and let

$$\begin{aligned} F^\beta &= (F^\beta_{1,1}, \ldots, F^\beta_{1,k_1-1}, \ldots, F^\beta_{n,1}, \ldots, F^\beta_{n,k_n-1}) \\ &= (F^\beta_1, \ldots, \ldots, F^\beta_n), \end{aligned}$$

where $F^\beta_{j,a} = \beta_{j,a} \circ F : \beta(\Omega) \to \{0,1\}$ and

$$F^\beta_j = \beta_j \circ F = (F^\beta_{j,1}, \ldots, F^\beta_{j,k_j-1}) : \beta(\Omega) \to \{0,1\}^{k_j-1}.$$

Therefore $F^\beta_j(\beta(x)) = 1^{F_j(x)}0^{k_j-1-F_j(x)}$ and $F^\beta_{j,a}(\beta(x)) = 1 \Leftrightarrow F_j(x) \geqslant a$.

Let us fix $j \in \{1, \ldots, n\}$ and $1 \leqslant a \leqslant k_j - 1$ and prove that $F^\beta_{j,a}$ is NC.

By assumption, $F_j : \Omega \to \Omega_j$ is SNC, hence by Proposition 1, there exist a function $v : \{1, \ldots, K\} \to \{1, \ldots, n\}$, with $K = \sum_i k_i - n$, and numbers $a_i \in \Omega_{v(i)}$ and $b_i \in \Omega_j$ for each $i \in \{1, \ldots, K\}$ such that, for all $x \in \Omega$:

$$F_j(x) = \begin{cases} b_1 & \text{if } x \in \Omega^{(1)} \text{ and } x_{v(1)} = a_1 \\ b_2 & \text{if } x \in \Omega^{(2)} \text{ and } x_{v(2)} = a_2 \\ \vdots & \quad \vdots \\ b_K & \text{if } x \in \Omega^{(K)} \text{ and } x_{v(K)} = a_K, \end{cases}$$

where for all $i \in \{1, \ldots, K\}$, $\Omega^{(i)} \subseteq \Omega$ is the set of x such that

$$x_{v(1)} \neq a_1, \ldots, x_{v(i-1)} \neq a_{i-1}.$$

Note that $\Omega = \Omega^{(1)} \supset \Omega^{(2)} \supset \cdots \supset \Omega^{(K)}$ form a decreasing sequence of subsets of Ω.

Moreover, by Proposition 1, the numbers a_i are assumed to satisfy the following constraint:

$$\text{either } a_i = \min \Omega^{(i)}_{v(i)} \text{ or } a_i = \max \Omega^{(i)}_{v(i)}.$$

Let $\varepsilon_i = 0, a'_i = a_i + 1$ in the first case (*min* case) and $\varepsilon_i = 1, a'_i = a_i$ in the second case (*max* case). Letting $y = \beta(x)$, we observe that in the min case, when $x \in \Omega^{(i)}$ we have

$$x_{v(i)} = a_i \Leftrightarrow x_{v(i)} \leqslant a_i \Leftrightarrow y_{v(i),a_i+1} = \beta_{v(i),a_i+1}(x) = 0,$$

and that in the max case, we have

$$x_{v(i)} = a_i \Leftrightarrow x_{v(i)} \geqslant a_i \Leftrightarrow y_{v(i),a_i} = \beta_{v(i),a_i}(x) = 1.$$

These two equivalences are summerized in the following property:

$$x_{v(i)} = a_i \Leftrightarrow y_{v(i),a'_i} = \varepsilon_i \text{ whenever } x \in \Omega^{(i)}. \tag{P}$$

Let us prove by induction on i that

$$x \in \Omega^{(i)} \Leftrightarrow \begin{cases} y_{v(1),a'_1} \neq \varepsilon_1 \\ \vdots \\ y_{v(i-1),a'_{i-1}} \neq \varepsilon_{i-1}. \end{cases} \tag{Q_i}$$

Q_1 is trivial because $\Omega^{(1)} = \Omega$, and by property P, we have

$$\begin{aligned} x \in \Omega^{(i+1)} &\Leftrightarrow \left(x \in \Omega^{(i)} \text{ and } x_{v(i)} \neq a_i \right) \\ &\Leftrightarrow \left(x \in \Omega^{(i)} \text{ and } y_{v(i),a'_i} \neq \varepsilon_i \right), \end{aligned}$$

hence Q_i entails Q_{i+1}, and we have shown that Q_i holds for any i. Combining properties P and Q_i, we then obtain:

$$\begin{cases} x \in \Omega^{(i)} \\ x_{v(i)} = a_i \end{cases} \Leftrightarrow \begin{cases} y_{v(1),a'_1} \neq \varepsilon_1 \\ \vdots \\ y_{v(i-1),a'_{i-1}} \neq \varepsilon_{i-1} \\ y_{v(i),a'_i} = \varepsilon_i. \end{cases}$$

On the other hand, $F^\beta_{j,a}(y) = 1 \Leftrightarrow F_j(x) \geqslant a$. Hence, letting $\chi_{b \geqslant a} = 1$ if $b \geqslant a$ and 0 otherwise, we have $F^\beta_{j,a}(y) = \chi_{F_j(x) \geqslant a}$. In particular

$$F_j(x) = b_i \Rightarrow F^\beta_{j,a}(y) = \chi_{b_i \geqslant a}.$$

Therefore, the condition that F_j is SNC gives, for any $y \in \beta(\Omega)$:

$$F^\beta_{j,a}(y) = \begin{cases} \chi_{b_1 \geqslant a} & \text{if } y_{v(1),a'_1} = \varepsilon_1 \\ \chi_{b_2 \geqslant a} & \text{if } y_{v(1),a'_1} \neq \varepsilon_1, y_{v(2),a'_2} = \varepsilon_2 \\ \vdots & \quad \vdots \\ \chi_{b_K \geqslant a} & \text{if } y_{v(1),a'_1} \neq \varepsilon_1, \ldots, y_{v(K-1),a'_{K-1}} \neq \varepsilon_{K-1}, y_{v(K),a'_K} = \varepsilon_K. \end{cases}$$

This means that $F^\beta_{j,a}$ is NC in the sense of Section 3.3. Since this holds for all j, a, F^β is NC. □

3.5. Counterexample for WNC functions

Let $f : \Omega = \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \to \mathbb{Z}/3\mathbb{Z}$ be the function defined in Section 2.3. It is WNC but not SNC, and $\beta(\Omega) \subseteq \{0,1\}^2 \times \{0,1\}^2$. The function $f^\beta : \beta(\Omega) \to \{0,1\}^2$ is given by the following table:

$f^\beta(x,y) =$

	00	10	11
00	11	00	00
10	10	10	10
11	11	00	11

where pairs in $\{0,1\}^2$ are written without commas and parentheses. This table is simply obtained from the one defining f by the substitutions $0 \mapsto 00, 1 \mapsto 10, 2 \mapsto 11$.

Now, the two Boolean functions $f^\beta_1, f^\beta_2 : \beta(\Omega) \to \{0,1\}$ such that $f^\beta = (f^\beta_1, f^\beta_2)$ are therefore:

$f^\beta_1(x,y) =$

	00	10	11
00	1	0	0
10	1	1	1
11	1	0	1

$f^\beta_2(x,y) =$

	00	10	11
00	1	0	0
10	0	0	0
11	1	0	1

and it is easy to see that f^β_1 is NC but that f^β_2 is not NC. Therefore Theorem 3 does not extend to WNC functions.

4. Examples in the regulatory network literature

Boolean NC functions appear predominant in databases of Boolean genetic networks [21]. If this canalyzing property reflects physical characteristics of biological systems, we wonder whether this property is also present in multivalued functions that model these same systems. Thus, in order to have an insight into the characteristics of multivalued functions specified in the context of biological systems, we have selected the multivalued genes present in the logical models explored in [21] (45 genes endowed with a ternary variable) and studied the canalizing properties (NC, SNC, WNC) of their logical functions. We also applied the Boolean mapping β and studied the canalizing property of the resulting Boolean functions. The results of the analysis are collected in Table 1, reflecting four different qualitative situations:

(a) functions F which are NC (hence SNC, WNC, and with NC Booleanization F^β);

(b) functions which are not NC, but SNC (hence WNC, and with NC Booleanization);

(c) functions which are not WNC (hence neither NC nor SNC) but with NC Booleanization;

(d) functions which are not WNC and with non NC Booleanizations.

Remark that in this set of functions, all WNC functions are actually SNC.

The logical rules corresponding to these genes are collected in the Appendix. To simplify the notation in the rules, the name of the gene is written to represent its activity variable (for instance $\mathbf{A} : 1$ stands for $x_{\mathbf{A}} = 1$), and the value of $F_{\mathbf{A}}$ is defined by means of the usual logical connectives (conjonction $\wedge$, disjunction $\vee$ and negation $\neg$). For example, the first and third lines of Table A.2 mean:

$$F_{\mathbf{Drk}}(x) = \begin{cases} 1 & \text{if } x_{\mathbf{Der}} = 1 \\ 2 & \text{if } x_{\mathbf{Der}} = 2 \\ 0 & \text{otherwise} \end{cases} \qquad F_{\mathbf{Rl}}(x) = \begin{cases} 1 & \text{if } x_{\mathbf{Dsor1}} = 1 \text{ and } x_{\mathbf{Msk}} = 1 \\ 2 & \text{if } x_{\mathbf{Dsor1}} = 2 \text{ and } x_{\mathbf{Msk}} = 1 \\ 0 & \text{otherwise} \end{cases}$$

This means that gene **Drk** can be activated at level 1 (resp. 2) when its (unique) regulator **Der** is at level 1 (resp. 2). Gene **Rl** has two regulators, **Msk** (which is Boolean) and **Dsor1**, and it can be activated at level 1 (resp. 2) if **Msk** is present ($x_{\mathbf{Msk}} = 1$) and **Dsor1** is at level 1 (resp. 2).

Some situations are more complex. For instance, gene **MadMed** (extracted from the model described in [9]) has three regulators: a multilevel activator **Tkv**, a Boolean activator **Sax** and a Boolean inhibitor **Dad**. The rule for **MadMed** is the following (cf Table A.5):

$$F_{\mathbf{MadMed}}(x) = \begin{cases} 1 & \text{if } (x_{\mathbf{Tkv}} = 1 \text{ or } x_{\mathbf{Sax}} = 1) \text{ and } x_{\mathbf{Dad}} = 0 \text{ and } x_{\mathbf{Tkv}} < 2 \\ 2 & \text{if } x_{\mathbf{Tkv}} = 2 \text{ and } x_{\mathbf{Dad}} = 0 \\ 0 & \text{otherwise.} \end{cases}$$

Thus, in the absence of the inhibitor **Dad**, the presence of one of the two activators at level 1 (max level for **Sax** and intermediate for **Tkv**) allows **MadMed** to be activated at its level 1. **MadMed** can be activated at its maximum level as soon as its activator **Tkv** is at its maximum level 2 and the inhibitor **Dad** is absent.

The structure of multivalued logical rules varies greatly. A quite current situation is when the level of only one regulator determines the target value of the gene:

$$\begin{cases} \mathbf{A} = 1 & \text{if } \mathbf{B} \wedge \varphi \\ \mathbf{A} = 2 & \text{if } \neg\mathbf{B} \wedge \varphi \end{cases} \qquad (S)$$

with **B** a regulator of **A**, and φ a "context" (conditions on the presence or absence of other regulators). This concerns the genes listed in Tables A.2, A.3, A.4. We can see in Table 1 that the rules satisfying this structure (S) generally have strong canalyzing properties: situations (a) or (b). In particular, their Booleanizations are all NC. Then, depending on the context, we distinguish three situations:

- if φ depends on multivalued variables and is expressed only with $\wedge$, then the function behaves according to situation (b); this is the case of, e.g., **Stat92E**; these functions are listed in Table A.2;
- if φ is expressed with at least one $\vee$, then the function behaves according to situation (c) (functions listed in Table A.3);
- if φ depends on Boolean variables and is expressed only with $\wedge$, then the function behaves according to situation (a) (functions listed in Table A.4). Note that nodes **IL4RA** and **E2F3** are controlled by a multivalued gene (resp. **STAT5** and **CHEK**), but their level depends only on one threshold (to be above or below level 2).

Gene	NC	SNC	WNC	Bool. NC	Struct.
E_Spl	Yes	Yes	Yes	Yes	
mQH2_Q	Yes	Yes	Yes	Yes	(S)
mdH	Yes	Yes	Yes	Yes	(S)
mGR	Yes	Yes	Yes	Yes	
mGSH_GSSG	Yes	Yes	Yes	Yes	(S)
mTRX	Yes	Yes	Yes	Yes	(S)
cGSH_GSSG	Yes	Yes	Yes	Yes	(S)
cGR	Yes	Yes	Yes	Yes	
E2F3	Yes	Yes	Yes	Yes	(S)
IL12RB1	Yes	Yes	Yes	Yes	(S)
IL4RA	Yes	Yes	Yes	Yes	(S)
Drk	No	Yes	Yes	Yes	(S)
Dsor1	No	Yes	Yes	Yes	(S)
Pnt	No	Yes	Yes	Yes	(S)
Stat92E	No	Yes	Yes	Yes	(S)
Raf	No	Yes	Yes	Yes	(S)
RI	No	Yes	Yes	Yes	(S)
Sos	No	Yes	Yes	Yes	(S)
Tkv	No	Yes	Yes	Yes	
mNNT	No	Yes	Yes	Yes	
mCa	No	Yes	Yes	Yes	(S)
mGPX	No	Yes	Yes	Yes	(S)
mTR	No	Yes	Yes	Yes	
cGPX	No	Yes	Yes	Yes	(S)
cTR	No	Yes	Yes	Yes	
cTRX	No	Yes	Yes	Yes	(S)
STAT5	No	Yes	Yes	Yes	

Gene	NC	SNC	WNC	Bool. NC	Struct.
Ras	No	No	No	Yes	
MadMed	No	No	No	Yes	
Hop	No	No	No	Yes	
mNADPH_NADP	No	No	No	Yes	
mNADH_NAD	No	No	No	Yes	(S)
cCa	No	No	No	Yes	
KrebsCycle	No	No	No	Yes	(S)
VIM	No	No	No	Yes	
CDH1	No	No	No	Yes	
EMT	No	No	No	Yes	
IL2R	No	No	No	Yes	(S)
IL4R	No	No	No	Yes	(S)
mROS	No	No	No	No	
Der	No	No	No	No	
cROS	No	No	No	No	
cNADPH_NADP	No	No	No	No	
E2F1	No	No	No	No	
Spi1	No	No	No	No	

Table 1: Multivalued genes occurring in logical models from database considered in [21]. The two rightmost columns indicate whether the Booleanized functions are nested canalizing, and which rules satisfy structure (S).

The rules that do not satisfy structure (S) are collected in Tables A.5 and A.6. In particular, Table A.6 gathers all the rules in situation (d), for which even the Booleanized functions are not NC.

The study of these multivalued functions suggests a certain relevance of canalization, notably SNC, among models arising in the modelling of gene networks. But it should be noted that we have not specified the assumptions for updating the system while they play an important role in the dynamical behavior of the system and its properties. Implicitly, we here considered that the gene directly reached its target value: for instance, **Drk** can be activated at level 2 when **Der** is at level 2, and if the current state of **Drk** is 0, it will update directly to value 2. However, in such context of biological systems, it is not a realistic hypothesis, steps of 1 are more common: in the previous example, if **Der** is at level 2 and **Drk** is not active, its activity will first reach level 1. The introduction of the "steps of 1" hypothesis in the definition of the function would obviously change the canalizing properties.

Appendix A. Logical rules

Gene i	x_i	Rule	Ref
Drk	1	**Der** : 1	[9]
	2	**Der** : 2	
Dsor1	1	**Raf** : 1	
	2	**Raf** : 2	
Rl	1	**Dsor1** : 1 $\wedge$ **Msk**	
	2	**Dsor1** : 2 $\wedge$ **Msk**	
Pnt	1	**Rl** : 1	
	2	**Rl** : 2	
Sos	1	**Drk** : 1	
	2	**Drk** : 2	
Stat92E	1	**Hop** : 1 $\wedge \varphi$	[9]
	2	**Hop** : 2 $\wedge \varphi$ where $\varphi = \neg$**Su$_{\text{var}}$** $\wedge \neg$**Ptp61F** $\wedge$ $\neg$**Ken** $\wedge \neg$**Brwd3** $\wedge \neg$**Socs44A**	
Raf	1	**Ras** : 1 $\wedge$ **Cnk** $\wedge$ **Src42** $\wedge$ **Ksr**	
	2	**Ras** : 2 $\wedge$ **Cnk** $\wedge$ **Src42** $\wedge$ **Ksr**	
mCa	1	**cCa** : 1	[19]
	2	**cCa** : 2	
mGPX	1	**mGSH_GSSG** : 1 $\wedge \neg$**mROS**	
	2	**mGSH_GSSG** : 2 $\wedge \neg$**mROS**	
cTRX	1	**cTR** $\wedge$ **cROS** : 1	
	2	**cTR** $\wedge \neg$**cROS**	
cGPX	1	**cGSH_GSSG** : 1 $\wedge \neg$**cROS**	
	2	**cGSH_GSSG** : 2 $\wedge \neg$**cROS**	

Table A.2: Rules associated to multivalued genes from databases considered in [21], satisfying structure (S) and behaving according to situation (b).

Gene i	x_i	Rule	Ref
mNADH_NAD	1	(**KrebsCycle** $\vee$ **mPDH** $\vee$ **FAO**) $\wedge$ **ETC**	[19]
	2	(**KrebsCycle** $\vee$ **mPDH** $\vee$ **FAO**) $\wedge \neg$**ETC**	
KrebsCycle	1	$\neg$**mCa** $\wedge \neg$**mNADH_NAD** $\wedge$ (**AcetylCoA** $\vee$ **GLUTAMINOLYSIS**)	
	2	**mCa** $\wedge \neg$**mNADH_NAD** $\wedge$ (**AcetylCoA** $\vee$ **GLUTAMINOLYSIS**)	
IL2R	1	$\neg$**IL2RA** $\wedge$ **CGC** $\wedge$ **IL2RB** $\wedge$ (**IL2** $\vee$ **IL2_e**)	[12]
	2	**IL2RA** $\wedge$ **CGC** $\wedge$ **IL2RB** $\wedge$ (**IL2** $\vee$ **IL2_e**)	

Gene i	x_i	Rule (continued)	Ref
IL4R	1	**IL4RA** : $1 \wedge$ **CGC** $\wedge$ (**IL4** $\vee$ **IL4_e**)	[12]
	2	**IL4RA** : $2 \wedge$ **CGC** $\wedge$ (**IL4** $\vee$ **IL4_e**)	

Table A.3: Rules associated to multivalued genes from databases considered in [21], satisfying structure (S) and behaving according to situation (c).

Gene i	x_i	Rule	Ref
IL12RB1	1	$\neg$**IRF1**	[12]
	2	**IRF1**	
mGSH_GSSG	1	**mGR** $\wedge$ **mGPX**	[19]
	2	**mGR** $\wedge \neg$**mGPX**	
mdH	1	**ETC** $\wedge$ **ATPSyn**	
	2	**ETC** $\wedge \neg$**ATPSyn**	
mQH2_Q	1	**ETC** $\wedge \neg$**mdH**	[19]
	2	**ETC** $\wedge$ **mdH**	
mTRX	1	**mTR** $\wedge$ **mROS**	
	2	**mTR** $\wedge \neg$**mROS**	
cGSH_GSSG	1	**cGR** $\wedge$ **cGPX**	
	2	**cGR** $\wedge \neg$**cGPX**	
IL4RA	1	$\neg$**STAT5** : 2	[12]
	2	**STAT5** : 2	
E2F3	1	$\neg$**CHEK** : $2 \wedge \neg$**RB1** $\wedge$ **RAS**	[14]
	2	**CHEK** : $2 \wedge \neg$**RB1** $\wedge$ **RAS**	

Table A.4: Rules associated to multivalued genes from databases considered in [21], satisfying structure (S) and behaving according to situation (a).

Gene i	x_i	Rule	Ref
Ras	1	(**Sos** : $1 \wedge \neg$(**Sty** $\wedge$ **Gap1**)) $\vee$ (**Gap1** $\wedge$ **Sty** $\wedge$ **Sos** : 2)	[9]
	2	**Sos** : $2 \wedge \neg$(**Sty** $\wedge$ **Gap1**)	
Tkv	1	(**Dpp** : $1 \vee$ **Scw** $\vee$ **Gbb**) $\wedge$ **Punt** $\wedge \neg$(**Sog** $\vee$ **Tsg**)	
	2	**Dpp** : $2 \wedge$ **Punt** $\wedge \neg$(**Sog** $\vee$ **Tsg**)	
mNNT	1	**mNADH_NAD** : $1 \wedge$ **mdH** : 1	[19]
	2	(**mNADH_NAD** : $2 \wedge$ **mdH** : 1) $\vee$ (**mNADH_NAD** : $2 \wedge$ **mdH** : 2) $\vee$ (**mNADH_NAD** : $1 \wedge$ **mdH** : 2)	
mNADPH_NADP	1	(**mNNT** $\wedge \neg$**mIDH2**) $\vee$ (**mNNT** : $1 \wedge \neg$**mIDH2**)	
	2	**mNNT** : $2 \vee$ (**mNNT** : $1 \wedge$ **mIDH2**)	

Gene i	x_i	Rule (continued)	Ref
mGR	1	(**mNADPH_NADP** : 1 ∨ **mNADPH_NADP** : 2) ∧ (¬**mGSH_GSSG** ∨ **mGSH_GSSG** : 1)	[19]
	2	(**mNADPH_NADP** : 1 ∨ **mNADPH_NADP** : 2) ∧ **mGSH_GSSG** : 2	
mTR	1	(**mNADPH_NADP** : 2 ∧ (¬**mTRX** ∨ **mTRX** : 1)) ∨ **mNADPH_NADP** : 1	
	2	**mTRX** : 2 ∧ **mNADPH_NADP** : 2	
cCa	1	(**IP3R** ∧ ¬**ORAI1** ∧ ¬**TRPM2** ∧ ¬**PMCA**) ∨ (¬**IP3R** ∧ ¬**ORAI1** ∧ **TRPM2** ∧ ¬**PMCA**) ∨ (¬**IP3R** ∧ **ORAI1** ∧ **TRPM2** ∧ ¬**PMCA**) ∨ (¬**IP3R** ∧ **ORAI1** ∧ ¬**TRPM2** ∧ ¬**PMCA**)	
	2	(**IP3R** ∧ **ORAI1** ∧ **TRPM2** ∧ ¬**PMCA**) ∨ (**IP3R** ∧ **ORAI1** ∧ ¬**TRPM2** ∧ ¬**PMCA**) ∨ (**IP3R** ∧ ¬**ORAI1** ∧ **TRPM2** ∧ ¬**PMCA**)	
cGR	1	(**cNADPH_NADP** : 1 ∨ **cNADPH_NADP** : 2) ∧ (¬**cGSH_GSSG** ∨ **cGSH_GSSG** : 1)	
	2	(**cNADPH_NADP** : 1 ∨ **cNADPH_NADP** : 2) ∧ **cGSH_GSSG** : 2	
cTR	1	(**cNADPH_NADP** : 2 ∧ (¬**cTRX** ∨ **cTRX** : 1)) ∨ **cNADPH_NADP** : 1	
	2	**cTRX** : 2 ∧ **cNADPH_NADP** : 2	
Hop	1	**Dome** ∧ ¬**ET** ∧ [(**Stam** ∧ **Hrs** ∧ **Socs36E**) ∨ (¬(**Stam** ∧ **Hrs**) ∧ ¬**Socs36E**)]	[9]
	2	**Dome** ∧ ¬**ET** ∧ **Stam** ∧ **Hrs** ∧ ¬**Socs36E**	
MadMed	1	((**Tkv** : 1 ∨ **Sax** : 1) ∧ ¬**Dad** : 1 ∧ ¬**Tkv** : 2) ∨ (**Tkv** : 2 ∧ **Dad** : 1)	
	2	**Tkv** : 2 ∧ ¬**Dad** : 1	
E_Spl	1	¬(**Nicd** ∧ **Mam**)	
	2	**Nicd** ∧ **Mam** ∧ ¬**H** ∧ ¬**Gro** ∧ ¬**CtBP**	
STAT5	1	¬(**IL4R** : 2 ∨ **IL2R** : 2) ∧ (**IL4R** : 1 ∨ **IL2R** : 1 ∨ **IL15R**)	[12]
	2	**IL4R** : 2 ∨ **IL2R** : 2	
VIM	1	(**SNAIL1** ∨ **ZEB1**) ∧ ¬(**ZEB1** ∧ **SNAIL1**)	[20]
	2	**ZEB1** ∧ **SNAIL1**	
CDH1	1	¬**ZEB1** ∧ ¬**SNAIL1**	
	2	(**SNAIL1** ∨ **ZEB1**) ∧ ¬(**ZEB1** ∧ **SNAIL1**)	
EMT	1	(**VIM** ∨ ¬**CDH1**) ∧ ¬(**VIM** ∧ ¬**CDH1**)	
	2	**VIM** ∧ ¬**CDH1**	

Table A.5: Rules associated to multivalued genes from databases considered in [21], not satisfying structure (S) and behaving according to situations (a), (b) or (c).

Gene i	x_i	Rule	Ref
mROS	1	$((\neg$**mPRX** $\wedge$ **mGPX**$) \vee ($**mPRX** $\wedge \neg$**mGPX**$)) \wedge$ (**ETC** $\vee$ **mGR** $\vee$ **mTR**$) \wedge \neg$**cROS**	[19]
	2	$[($**ETC** $\vee$ **mGR** $\vee$ **mTR** $\vee$ **cROS**$) \wedge \neg($**mPRX** $\vee$ **mGPX**$)] \vee [$**cROS** $\wedge$ (**mPRX** $\vee$ **mGPX** $\vee$ **ETC** $\vee$ **mGR** $\vee$ **mTR**$)]$	
cROS	1	$(($**cPRX** $\wedge \neg$**cGPX**$) \vee (\neg$**cPRX** $\wedge$ **cGPX**$)) \wedge$ (**NOX2** $\vee$ **DUOX1** $\vee$ **cTR** $\vee$ **cGR**$) \wedge \neg$**mROS**	
	2	$[\neg($**cPRX** $\vee$ **cGPX**$) \wedge$ (**NOX2** $\vee$ **mROS** $\vee$ **DUOX1** $\vee$ **cTR** $\vee$ **cGR**$)] \vee$ [**mROS** $\wedge$ (**cPRX** $\vee$ **cGPX** $\vee$ **NOX2** $\vee$ **DUOX1** $\vee$ **cTR** $\vee$ **cGR**$)]$	
cNADPH_NADP	1	(**mShuttle** $\wedge \neg$**PPP** $\wedge \neg$**GLUTAMINOLYSIS**$) \vee$ (**mShuttle** $\wedge \neg$**PPP** $\wedge \neg$**GLUTAMINOLYSIS**$) \vee$ (**mShuttle** $\wedge \neg$**PPP** $\wedge$ **GLUTAMINOLYSIS**)	[19]
	2	(**mShuttle** $\wedge$ **PPP** $\wedge$ **GLUTAMINOLYSIS**$) \vee$ (**mShuttle** $\wedge \neg$**PPP** $\wedge$ **GLUTAMINOLYSIS**$) \vee$ (**mShuttle** $\wedge$ **PPP** $\wedge \neg$**GLUTAMINOLYSIS**)	
Der	1	$[($**Spi** $: 1 \vee$ **Vein**$) \wedge \neg$**Aos** $: 1 \wedge \neg$**Kek** $: 1 \wedge \neg$**Cbl** $\wedge \neg$**Spi** $: 2 \wedge$ **Shc**$] \vee [$**Spi** $: 2 \wedge$ **Shc** $\wedge ($**Kek** $: 1 \vee$ **Aos** $: 1 \vee$ **Cbl**$)]$	[9]
	2	**Spi** $: 2 \wedge \neg$**Kek** $: 1 \wedge \neg$**Aos** $: 1 \wedge \neg$**Cbl** $\wedge$ **Shc**	
E2F1	1	$[(\neg($**CHEK** $: 2 \wedge$ **ATM** $: 2) \wedge ($**RAS** $\vee$ **E2F3**$)) \vee$ (**CHEK** $: 2 \wedge$ **ATM** $: 2 \wedge \neg$**RAS** $\wedge$ **E2F3** $: 1)] \wedge \neg$**RB1** $\wedge \neg$**RBL2**	[14]
	2	**¬RB1** $\wedge \neg$**RBL2** $\wedge$ **ATM** $: 2 \wedge$ **CHEK** $: 2 \wedge$ (**RAS** $\vee$ **E2F3** $: 2)$	
Spi1	1	[**Spi1** $\wedge$ **Runx1** $\wedge \neg(($(**Cebpa** $\vee$ **Cebpb**$) \wedge$ **Csf1r**$) \vee ($(**Foxo1** $\wedge \neg$**Ikzf1**$) \vee \neg$**Gfi1**$)] \vee ($**Foxo1** $\wedge$ **Ebf1** $\wedge$ **Ikzf1** $\wedge \neg($**Spi1** $\vee$ **Runx1**$))$	[1]
	2	(**Spi1** $\wedge$ **Runx1** $\wedge ($**Cebpa** $\vee$ **Cebpb**$) \wedge$ **Csf1r**$) \vee$ (**Spi1** $\wedge$ **Runx1** $\wedge \neg($**Gfi1** $\vee ($**Foxo1** $\wedge$ **Ikzf1**$)))$	

Table A.6: Rules associated to multivalued genes from databases considered in [21], not satisfying structure (S) and behaving according to situation (d).

References

[1] S. Collombet, C. van Oevelen, J. L. Sardina Ortega, W. Abou-Jaoudé, B. Di Stefano, M. Thomas-Chollier, T. Graf, and D. Thieffry. Logical

modeling of lymphoid and myeloid cell specification and transdifferentiation. *Proc. Natl. Acad. Sci.*, 114(23):5792–5799, 2017.

[2] G. Didier, É. Remy, and C. Chaouiya. Mapping multivalued onto Boolean dynamics. *J. Theoret. Biol.*, 270(1):177–184, 2011.

[3] C. Kadelka, Y. Li, J. Kuipers, J. O. Adeyeye, and R. Laubenbacher. Multistate nested canalizing functions and their networks. *Theoret. Comput. Sci.*, 675:1–14, 2017.

[4] S. A. Kauffman. *The origins of order: Self organization and selection in evolution.* Oxford University Press, 1993.

[5] S. Kauffman, C. Peterson, B. Samuelsson, and C. Troein. Random Boolean network models and the yeast transcriptional network. *Proc. Natl. Acad. Sci.*, 100(25), 2003.

[6] S. Kauffman, C. Peterson, B. Samuelsson, and C. Troein. Genetic networks with canalyzing Boolean rules are always stable. *Proc. Natl. Acad. Sci.*, 101(49), 2004.

[7] J. G. Klotz, R. Heckel, and S. Schober. Bounds on the average sensitivity of nested canalizing functions. *Plos One*, 8(5), 2013.

[8] Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, and R. Laubenbacher. Boolean nested canalizing functions: A comprehensive analysis. *Theoret. Comput. Sci.*, 481:24–36, 2013.

[9] A. Mbodj, G. Junion, C. Brun, E. Furlong, and D. Thieffry. Logical modelling of Drosophila signalling pathways. *Mol. Biosyst.*, 9(9):2248–58, 2013.

[10] D. Murrugarra and R. Laubenbacher. Regulatory patterns in molecular interaction networks. *J. Theor. Biol.*, 288:66–72, 2011.

[11] D. Murrugarra and R. Laubenbacher. The number of multistate nested canalyzing functions. *Physica D: Nonlinear Phenomena*, 241(10):929–938, 2012.

[12] A. Naldi, J. Carneiro, C. Chaouiya, and D. Thieffry. Diversity and plasticity of Th cell types predicted from regulatory network modelling. *PLoS Comput. Bio.*, 6(9), 2010.

[13] R. O'Donnell. *Analysis of Boolean functions.* Cambridge University Press, 2014.

[14] É. Remy, S. Rebouissou, C. Chaouiya, A. Zinovyev, F. Radvany, and L. Calzone. A modeling approach to explain mutually exclusive and co-occurring genetic alterations in bladder tumorigenesis. *Cancer Res.*, 75(19): 4042–4052, 2015.

[15] É. Remy and P. Ruet. On differentiation and homeostatic behaviours of Boolean dynamical systems. BioConcur 2005, in *Transactions on Computational Systems Biology VIII*, Springer LNCS 4780: 92–101, 2007.

[16] É. Remy and P. Ruet. Average sensitivity of nested canalizing multi-valued functions. *arXiv:2211.14649*, 2022.

[17] É. Remy, P. Ruet, and D. Thieffry. Graphic requirements for multistability and attractive cycles in a Boolean dynamical framework. *Adv. Appl. Math.*, 41(3):335–350, 2008.

[18] P. Ruet. Local cycles and dynamical properties of Boolean networks. *Math. Struct. Comput. Sci.*, 26(4):702–718, 2016.

[19] J.A. Sánchez-Villanueva, O. Rodríguez-Jorge, O. Ramírez-Pliego, G.R. Salgado, W. Abou-Jaoudé, C. Hernandez, A. Naldi, D. Thieffry, and M.A. Santana. Contribution of ROS and metabolic status to neonatal and adult CD8+ T cell activation. *PLoS ONE* 14(12): e0226388, 2019.

[20] D.A. Silveira, and J.C. Mombach. Dynamics of the feedback loops required for the phenotypic stabilization in the epithelial-mesenchymal transition. *FEBS J.*, 287(3):578–588, 2020.

[21] A. Subbaroyan, O. C. Martin, and A. Samal. Minimum complexity drives regulatory logic in Boolean models of living systems. *PNAS Nexus*, 1:1–12, 2022.

[22] R. Thomas. Boolean formalization of genetic control circuits. *J. Theor. Biol.*, 42:563–585, 1973.

[23] R. Thomas. Regulatory networks seen as asynchronous automata: a logical description. *J. Theor. Biol.*, 153:1–23, 1991.

[24] E. Tonello. On the conversion of multivalued gene regulatory networks to Boolean dynamics. *Discrete Appl. Math.*, 259(C):193–204, 2019.

[25] P. Van Ham. How to deal with variables with more than two levels. In *Kinetic logic: a Boolean approach to the analysis of complex regulatory systems*. Springer, 326–343, 1979.

[26] C.H. Waddington. Canalization of development and the inheritance of acquired characters. *Nature*, 150:563–565, 1942.