Brief paper

Symbolic dynamics of Boolean control networks☆

Gal Hochma [a], Michael Margaliot [a,1], Ettore Fornasini [b], Maria Elena Valcher [b]

[a] *School of Electrical Engineering Systems, Tel Aviv University, Israel*
[b] *Dipartimento di Ingegneria dell'Informazione, Università di Padova, Italy*

ABSTRACT

We consider Boolean control networks (BCNs), and in particular Boolean networks (BNs), in the framework of symbolic dynamics (SD). We show that the set of state-space trajectories of a BCN is a shift space of finite type (SFT). This observation allows us to extend two important analysis tools from SD, namely, the Artin–Mazur zeta function and the topological entropy, to BNs and BCNs. Some of the theoretical results are illustrated using a BCN model of the core network regulating the mammalian cell cycle.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Boolean network (BNs) are useful modeling tools for dynamical systems whose state variables can attain two possible values. Examples include artificial neural networks with threshold function type neurons (see, e.g., Hassoun, 1995), and models for the interactions and the emergence of social consensus between simple agents (see, e.g., Green, Leishman, & Sadedin, 2007). BNs have recently been attracting considerable attention as computational tools in systems biology, and, in particular, as models for genetic regulation networks. Here each gene is either expressed (ON) or not expressed (OFF) (Chaos et al., 2006; Kauffman, Peterson, Samuelsson, & Troein, 2003; Li, Long, Lu, Ouyang, & Tang, 2004). Kauffman (1969) has studied the behavior of large, randomly constructed nets of these binary genes. His pioneering ideas stimulated research in the theoretical analysis of the dynamics of large-scale BNs, especially using tools from the fields of complex systems and statistical physics (see, e.g., Albert & Barabasi, 2000, Aldana, 2003, Drossel, Mihaljev, & Greil, 2005, Kauffman, 1993).

BNs have also been used to model various cellular processes. Specific examples include: the complex cellular signaling network controlling stomatal closure in plants (Li, Assmann, & Albert, 2006), the molecular pathway between two neurotransmitter systems, the dopamine and glutamate receptors (Gupta, Bisht, Kukreti, Jain, & Brahmachari, 2007), carcinogenesis, and the effects of therapeutic intervention (Szallasi & Liang, 1998).

Many biological systems have exogenous inputs and it is natural to extend BNs to *Boolean control networks* (BCNs) by adding Boolean inputs. For example, in a BCN modeling the progression of a disease, a binary input may represent whether a certain medicine is administered or not at each time step.

Cheng, Qi, and Li (2011) have developed an *algebraic state-space representation* (ASSR) of BCNs (and, in particular, of BNs). This representation has proved useful for studying control-theoretic questions, as they reduce BCNs to linear switched systems whose input, state and output variables are canonical vectors. A drawback of the ASSR is its computational complexity, as the ASSR of a BN with n state variables includes a $2^n \times 2^n$ matrix. Thus, any algorithm based on the ASSR has an exponential time complexity. A natural question is that of whether better algorithms exist. Zhao (2005) has shown that determining whether a BN has a fixed point is NP-complete. Akutsu, Hayashida, Ching, and Ng (2007) has shown that several control problems for BCNs are NP-hard. Laschov, Margaliot, and Even (2013) have shown that the observability problem for BCNs is also NP-hard. Thus, unless $P = NP$, these analysis problems for BCNs cannot be solved in polynomial time.

We develop a new approach to the analysis of BCNs based on *symbolic dynamics* (SD) (Lind & Marcus, 1995). The main objects of

study in SD are shift spaces. We show that the set of all possible trajectories of a BCN is a shift space. Consequently, many results and analysis tools from SD are immediately applicable to BCNs. We demonstrate this by defining and computing the zeta function and the topological entropy of a BCN. The zeta function stores the number of limit cycles of a dynamical system and their lengths, while the topological entropy is a nonnegative real number that measures how rich the control is. We illustrate some of the theoretical results using a BCN model of an important biological process: the regulation of the mammalian cell cycle.

The remainder of this note is organized as follows. Section 2 reviews BNs, BCNs, and some definitions and tools from SD. Section 3 includes our main results. Section 4 details the biological example.

Notation. We consider Boolean vectors and matrices, with entries in $\mathscr{S} := \{0, 1\}$, and the usual logical operations (AND: $\wedge$; OR: $\vee$; negation: $\bar{\cdot}$). $[A]_{\ell j}$ is the (ℓ, j)th entry of the matrix A. The canonical vector $\delta_N^i \in \mathscr{S}^N$, $i = 1, \ldots, N$, is the ith column of the identity matrix I_N. A matrix L whose columns are canonical vectors is called a *logical matrix*. Note that L maps any canonical vector into a canonical vector. A *permutation matrix* P is a nonsingular square logical matrix. An $N \times N$ permutation matrix is *cyclic* if it takes the form $\begin{bmatrix} \delta_N^2 & \delta_N^3 & \cdots & \delta_N^N & \delta_N^1 \end{bmatrix}$.

The *semi-tensor product* (STP) (Cheng et al., 2011) of $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$ is

$$A \ltimes B := (A \otimes I_{\alpha/n})(B \otimes I_{\alpha/p}) \in \mathbb{R}^{(m\alpha/n) \times (q\alpha/p)},$$

where $\otimes$ denotes the Kronecker (or tensor) product, and α is the least common multiple of n and p. If $n = p$, then $A \ltimes B = (A \otimes I_1)(B \otimes I_1) = AB$. Hence, the semi-tensor product is a generalization of the standard matrix product that provides a way to multiply two matrices with arbitrary dimensions.

A matrix $A \in \mathscr{S}^{N \times N}$ is associated with a *directed graph* $G(A) = (V, E)$ in the following way (Brualdi & Ryser, 1991). $V = \{1, \ldots, N\}$ is the set of vertices, while $E \subseteq V \times V$ is the set of edges (or arcs). There is an arc (j, ℓ) from j to ℓ if and only if $[A]_{\ell j} = 1$. A sequence $j_1 \to j_2 \to \cdots \to j_r \to j_{r+1}$ in $G(A)$ is a *walk* of length r from j_1 to j_{r+1} provided that $(j_1, j_2), \ldots, (j_r, j_{r+1})$ are arcs of $G(A)$. A closed walk is called a *cycle*. A cycle γ with no repeated vertices is called *elementary*, and its *length* $|\gamma|$ coincides with the number of (distinct) vertices appearing in it.

Conversely, with every directed graph $G = (V, E)$, with $V = \{1, 2, \ldots, N\}$, we associate an *adjacency matrix* $A \in \mathscr{S}^{N \times N}$ with $[A]_{\ell j} = 1$ if and only if $(j, \ell) \in E$.

2. Preliminaries

A BCN is a discrete-time logical dynamical system

$$X_1(k + 1) = f_1(X_1(k), \ldots, X_n(k), U_1(k), \ldots, U_m(k)),$$
$$\vdots$$
$$X_n(k + 1) = f_n(X_1(k), \ldots, X_n(k), U_1(k), \ldots, U_m(k)),$$

where $X_i, U_i \in \mathscr{S}$, and each f_i is a Boolean function, i.e. $f_i : \mathscr{S}^{n+m} \to \mathscr{S}$. It is useful to write this in vector form as

$$X(k + 1) = f(X(k), U(k)). \tag{1}$$

A BN is a BCN without inputs, i.e.

$$X(k + 1) = f(X(k)). \tag{2}$$

D. Cheng et al. have developed an algebraic state-space representation of BCNs using the semi-tensor product of matrices. In this setup, any Boolean variable X_i, taking values in $\mathscr{S}$, is associated with the vector $x_i := \begin{bmatrix} X_i & \bar{X}_i \end{bmatrix}^\top$, taking values in $\{\delta_2^1, \delta_2^2\}$. The definition of the STP implies that $x := x_1 \ltimes x_2 \ltimes \cdots \ltimes x_n$ is a vector in $\mathscr{S}^{2^n}$
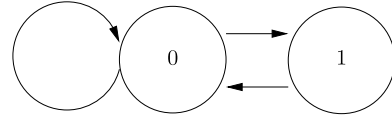

Fig. 1. Vertex graph of the golden mean shift.

that includes all the minterms of the X_is. Note that, being a vector of distinct minterms, x is a canonical vector. Any Boolean function $f : \mathscr{S}^n \to \mathscr{S}$ can be represented as a sum of minterms, and this implies that the STP can be used to provide an ASSR of BCNs.

Theorem 1 (*Cheng & Qi, 2010*). *Consider the BCN* (1). *Set* $x(k) := x_1(k) \ltimes \cdots \ltimes x_n(k)$, *and* $u(k) := u_1(k) \ltimes \cdots \ltimes u_m(k)$. *There exists a unique logical matrix* $L \in \mathscr{S}^{2^n \times 2^{n+m}}$, *called the* transition matrix *of the BCN, such that*

$$x(k + 1) = L \ltimes u(k) \ltimes x(k). \tag{3}$$

Algorithms for converting a BCN from the form (1) to its ASSR (3), and vice versa, may be found in Cheng et al. (2011). Similarly, the BN (2), with n Boolean variables, may be represented in the ASSR

$$x(k + 1) = Lx(k), \tag{4}$$

where $x(k) \in \mathscr{S}^{2^n}$ and $L \in \mathscr{S}^{2^n \times 2^n}$. Note that the fact that a BN may be represented in a linear form using the vector of minterms has been known for a long time (see, e.g., Cull, 1971, 1975), but the ASSR provides an explicit algebraic form that is particularly suitable for control-theoretic analysis. For example, it can be used to derive an ASSR of the adjoint control system (see Laschov & Margaliot, 2011, in press).

Given the ASSR (4) of a BN, we can associate it with the directed graph $G(L) = G(V, E)$, where $V = \{\delta_{2^n}^1, \ldots, \delta_{2^n}^{2^n}\}$, and there is a directed edge from vertex $\delta_{2^n}^j$ to vertex $\delta_{2^n}^i$ if and only if $[L]_{ij} = 1$.

We now describe some basic ideas from SD (Lind & Marcus, 1995). Given an alphabet $\mathcal{A}$ (e.g. the binary alphabet $\mathscr{S}$), and a set of strings $\mathcal{F}$ with symbols in $\mathcal{A}$, let $\mathcal{X}_{\mathcal{F}}$ denote the set of (one-sided) infinite sequences of symbols that do *not* contain any string in $\mathcal{F}$. The *shift operator* $\sigma : \mathcal{X}_{\mathcal{F}} \to \mathcal{X}_{\mathcal{F}}$, defined by $\sigma(.x_0 x_1 x_2 \ldots) := .x_1 x_2 x_3 \ldots$, shifts any (one-sided) infinite sequence one position to the left. If $\mathcal{F}$ is a finite set, then the dynamical system $(\mathcal{X}_{\mathcal{F}}, \sigma)$ is called a *shift of finite type* (SFT). If $\mathcal{X}_{\mathcal{F}}$ can be defined by means of a finite set of forbidden strings, all of which have length $k + 1$, then $(\mathcal{X}_{\mathcal{F}}, \sigma)$ is called a *k-step SFT*.

Alternatively, a *k*-step SFT may be described by its set of *allowed strings* of length $k + 1$. This leads to a useful graph-theoretic representation, called the *vertex graph*. By restricting our attention to one-step SFTs, we can associate with any such SFT a vertex graph whose vertices correspond to the possible symbols in $\mathcal{A}$, and there is a directed edge from vertex j to vertex i if and only if ji is an allowed string. Each vertex graph with N vertices can be represented by its adjacency matrix $A \in \mathscr{S}^{N \times N}$.

For example, if $\mathcal{A} = \mathscr{S}$ then $\mathcal{X}_{\{11\}}$ is the set of all binary sequences that do not contain the string 11. $(\mathcal{X}_{\{11\}}, \sigma)$ is called the *golden mean shift* (GMS) (Williams, 2004). Alternatively, it can be characterized by the set of its allowed strings of length 2, namely, $\{00, 01, 10\}$. The associated vertex graph has two vertices (0 and 1, or, equivalently δ_2^2 and δ_2^1), and is depicted in Fig. 1. Every element of $\mathcal{X}_{\{11\}}$ corresponds to an (infinite) walk on this vertex graph, and vice versa. The adjacency matrix corresponding to the vertex graph of $\mathcal{X}_{\{11\}}$ is $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$.

For an SFT $(\mathcal{X}_{\mathcal{F}}, \sigma)$, let p_i denote the number of *period i sequences*, i.e. sequences x such that $\sigma^i(x) = x$. The *Artin–Mazur*

zeta function $\zeta : \mathbb{R} \to \mathbb{R}$ is a bookkeeping device for storing all the p_is, defined by

$$\zeta(t) := \exp\left(\sum_{i=1}^{\infty} p_i \frac{t^i}{i}\right). \tag{5}$$

This implies that given ζ, one can easily obtain every p_i as

$$p_i = \frac{1}{(i-1)!} \frac{d^i}{dt^i} \ln(\zeta(t))\Big|_{t=0}. \tag{6}$$

The next result provides an algebraic expression for the zeta function in terms of the adjacency matrix of the vertex graph.

Theorem 2 (*Bowen–Lanford formula* (*Lind & Marcus, 1995*)). *Suppose that $(\mathcal{X}_{\mathcal{F}}, \sigma)$ is an SFT over an alphabet with N symbols. Let G be its vertex graph and $A \in \mathscr{S}^{N \times N}$ the associated adjacency matrix. Then*

$$\zeta(t) = (t^N P_A(1/t))^{-1}, \tag{7}$$

where $P_A(s) := \det(sI_N - A)$.

The *topological entropy* of a shift space $(\mathcal{X}_{\mathcal{F}}, \sigma)$ is

$$h := \lim_{\ell \to \infty} \frac{1}{\ell} \log(N_\ell), \tag{8}$$

where N_ℓ is the number of allowed strings of length ℓ. In other words, h is the "growth rate" of the number of allowed blocks of a given length. Existence of the limit in (8) follows from combining the fact that $N_{\ell+h} \le N_\ell N_h$ with Fekete's Lemma (see e.g. van Lint & Wilson, 2001).

Example 1. Recall that the GMS consists of all the binary sequences with no consecutive 1's. The number of allowed strings of length ℓ satisfies the recursion

$$N_{\ell+2} = N_{\ell+1} + N_\ell, \tag{9}$$

with $N_1 = 2, N_2 = 3$. Indeed, we can form an allowed string of length $(\ell + 2)$ by either adding 01 to an allowed string of length ℓ, or by concatenating 0 to every allowed $(\ell + 1)$ string. Eq. (9) implies that $N_{\ell-2} = F_\ell$, the ℓth Fibonacci number. It is well-known that F_ℓ grows exponentially as $c\gamma^\ell$, for some constant c, where $\gamma := (1 + \sqrt{5})/2 \approx 1.618$ is the golden mean (Williams, 2004). Hence, the topological entropy of the GMS is

$$\lim_{\ell \to \infty} \frac{1}{\ell} \log(c\gamma^{\ell+2}) = \log(\gamma).$$

Suppose that $(\mathcal{X}_{\mathcal{F}}, \sigma)$ is an SFT over an alphabet with N symbols, and let A be the $N \times N$ adjacency matrix of the associated vertex graph. The number of allowed ℓ-strings of the SFT is $N_\ell = \sum_{i,j=1}^{N} [A^{\ell-1}]_{ij}$. Combining this with the Perron–Frobenius theory (Lind & Marcus, 1995, Chapter 4) yields

$$h = \log \lambda_A, \tag{10}$$

where λ_A is the Perron root of A.

3. Main results

Given a BCN, define its *set of state trajectories* as

$$\mathcal{A}_S := \{X(0)X(1)\cdots : X(k+1) = f(X(k), U(k)),$$
$$U(k) \in \mathscr{S}^m, X(0) \in \mathscr{S}^n\},$$

i.e., the state trajectories over all possible controls and initial conditions. Note that for a BN this becomes

$$\{X(0)X(1)\cdots : X(k+1) = f(X(k)), \ X(0) \in \mathscr{S}^n\}.$$

The next result shows that the set of trajectories of a BCN (and hence that of a BN) is a one-step SFT.

Theorem 3. *In the ASSR (3), the set of state trajectories of a BCN is a one-step SFT over the alphabet $\{\delta_{2^n}^1, \ldots, \delta_{2^n}^{2^n}\}$.*

Proof. Set $L_i := L \ltimes \delta_{2^m}^i, i = 1, \ldots, 2^m$, where L is the transition matrix of the BCN, and define

$$M := L_1 \vee L_2 \vee \cdots \vee L_{2^m}. \tag{11}$$

Consider the SFT $(\mathcal{X}_{\mathcal{F}_S}, \sigma)$, where

$$\mathcal{F}_S := \{\delta_{2^n}^i \delta_{2^n}^j : [M]_{ji} = 0\}. \tag{12}$$

Note that $\mathcal{F}_S = \{\delta_{2^n}^i \delta_{2^n}^j : [M \ltimes v]_{ji} = 0, \forall v \in \{\delta_{2^m}^1, \ldots, \delta_{2^m}^{2^m}\}\}$. Suppose that $w = .\delta_{2^n}^{i_0} \delta_{2^n}^{i_1} \delta_{2^n}^{i_2} \cdots \in \mathcal{A}_S$. Then for any k there exists a $j = j(k)$ such that $\delta_{2^n}^{i_{k+1}} = L_j \delta_{2^n}^{i_k}$. Thus, $[L_j]_{i_{k+1}i_k} = 1$, so $[M]_{i_{k+1}i_k} = 1$. By (12), this implies that w is a string of the SFT. Conversely, suppose that $w = .\delta_{2^n}^{i_0} \delta_{2^n}^{i_1} \delta_{2^n}^{i_2} \ldots$ is a string of the SFT. By the definition of $\mathcal{F}$, this implies that $[M]_{i_{k+1}i_k} = 1$ for all k. Thus, there exists a $j = j(k)$ such that $[L_j]_{i_{k+1}i_k} = 1$, i.e. $\delta_{2^n}^{k+1} = L \ltimes \delta_{2^m}^j \ltimes \delta_{2^n}^k$, so w is a trajectory of the BCN. $\quad\square$

Note that M is the adjacency matrix of the graph associated with the allowed strings of length 2 in the SFT. The next result follows from Theorem 3 on replacing M with L.

Corollary 1. *The set of trajectories of a BN in the ASSR (4) is a one-step SFT over the alphabet $\{\delta_{2^n}^1, \ldots, \delta_{2^n}^{2^n}\}$.*

Note that not all one-step SFTs over a finite alphabet can be represented as the set of trajectories of a BN. For example, the GMS includes an infinite number of distinct sequences, whereas any BN has a finite number of distinct trajectories.

3.1. The zeta function of a Boolean control network

Since a BCN (1) induces an SFT and the matrix M is the adjacency matrix of the corresponding graph, the Bowen–Lanford formula yields the following result.

Corollary 2. *The zeta function of a BCN satisfies $\zeta(t) = (t^{2^n} P_M (1/t))^{-1}$, where $P_M(s) := \det(sI_{2^n} - M)$ is the characteristic polynomial of M.*

Corollary 3. *The zeta function of a BN with ASSR (4) satisfies*

$$\zeta(t) = (t^{2^n} P_L(1/t))^{-1}. \tag{13}$$

Example 2. Consider the BN

$$X_1(k+1) = (X_1(k) \wedge X_2(k)) \vee (X_1(k) \wedge X_3(k))$$
$$\vee (\bar{X}_1(k) \wedge \bar{X}_2(k) \wedge \bar{X}_3(k)),$$
$$X_2(k+1) = X_1(k) \vee (X_2(k) \wedge X_3(k)),$$
$$X_3(k+1) = (X_1(k) \wedge \bar{X}_2(k)) \vee (X_2(k) \wedge \bar{X}_3(k)).$$

Here $n = 3$ and $L = [\delta_8^2 \ \delta_8^1 \ \delta_8^1 \ \delta_8^5 \ \delta_8^6 \ \delta_8^7 \ \delta_8^8 \ \delta_8^4] \in \mathscr{S}^{8 \times 8}$. A calculation yields

$$P_L(s) = s^8 - s^6 - s^3 + s. \tag{14}$$

So, by (13),

$$\zeta(t) = \frac{1}{t^8(t^{-8} - t^{-6} - t^{-3} + t^{-1})} = \frac{1}{t^7 - t^5 - t^2 + 1}.$$

Hence, $\ln \zeta(t) = -\ln(t^7 - t^5 - t^2 + 1)$. Thus,

$$p_1 = \frac{1}{0!} \frac{d \ln \zeta(t)}{dt}\Big|_{t=0} = 0, \qquad p_2 = \frac{1}{1!} \frac{d^2 \ln \zeta(t)}{dt^2}\Big|_{t=0} = 2,$$

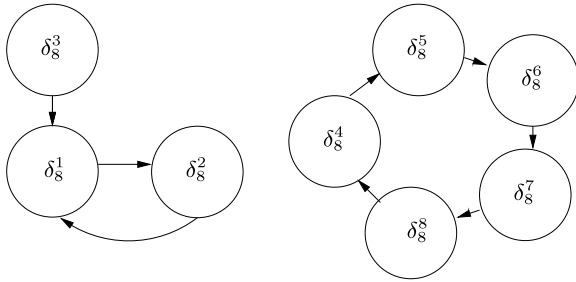and proceeding in this fashion yields $p_3 = 0, p_4 = 2,$ and $p_5 = 5$.

Fig. 2. State-transition graph for the BN in Example 2.

Fig. 2 depicts the graph associated with this BN. It may be seen that there are no equilibrium points (corresponding to period 1 sequences), so $p_1 = 0$. Also, there are two period 2 sequences, namely, $\cdot \delta_8^1 \delta_8^2 \delta_8^1 \delta_8^2 \ldots$ and $\cdot \delta_8^2 \delta_8^1 \delta_8^2 \delta_8^1 \ldots$. Each of them is also a period 4 sequence. Finally, each vertex in the cycle of length 5 is the initial state of a period 5 sequence.

The concept of period ν sequences is closely related to that of limit cycles. An ordered ν-tuple of *distinct* Boolean vectors $(X^{i_1}, X^{i_2}, \ldots, X^{i_\nu})$ is called a *limit cycle of length ν* of the BN if: (1) $X(0) = X^{i_j}$ implies that $X(1) = X^{i_{j+1}}$ for $j = 1, 2, \ldots, \nu - 1$; and (2) $X(0) = X^{i_\nu}$ implies that $X(1) = X^{i_1}$. Clearly, a limit cycle of length ν corresponds to ν sequences of period ν, and also to ν sequences of period 2ν, ν sequences of period 3ν, etc. For an integer $\nu > 0$, let $D(\nu)$ denote the set of proper divisors of ν (i.e., excluding ν itself). Let q_ν denote the number of *distinct* limit cycles of length ν. Then $q_1 = p_1$, and for $\nu > 1$,

$$q_\nu = \left(p_\nu - \sum_{j \in D(\nu)} p_j \right) \Big/ \nu. \tag{15}$$

For the BN in Example 2, (15) yields $q_1 = p_1 = 0$, $q_2 = (p_2 - p_1)/2 = 1$, $q_3 = (p_3 - p_1)/3 = 0$, $q_4 = (p_4 - p_1 - p_2)/4 = 0$, and $q_5 = (p_5 - p_1)/5 = 1$.

The state-transition graph of a BN can be partitioned into isolated components, each of them consisting of a limit cycle and a number of states accessing it. On the basis of this observation, the following useful expression for $P_L(s)$ was derived.

Proposition 1 (*Fornasini & Valcher, 2013*). *Given a BN, let $L \in \mathscr{S}^{2^n \times 2^n}$ be the transition matrix of the associated ASSR (4). There exist $r \in \mathbb{N}$ and a permutation matrix P such that*

$$P^\top L P = \text{blockdiag}\{D_1, D_2, \ldots, D_r\},$$

with $D_i := \begin{bmatrix} N_i & 0 \\ T_i & C_i \end{bmatrix} \in \mathscr{S}^{n_i \times n_i}$, $\tag{16}$

where $C_i \in \mathscr{S}^{k_i \times k_i}$, $k_i \geq 1$, is a cyclic matrix and $N_i \in \mathscr{S}^{(n_i - k_i) \times (n_i - k_i)}$ is a nilpotent matrix. Consequently,

$$P_L(s) = \prod_{i=1}^{r} P_{D_i}(s) = s^{\left(2^n - \sum_{i=1}^{r} k_i \right)} \cdot \prod_{i=1}^{r} (s^{k_i} - 1). \tag{17}$$

Example 3. Consider the BN in Example 2. The ASSR is given by $n = 3$, $L = [\delta_8^2 \ \delta_8^1 \ \delta_8^1 \ \delta_8^5 \ \delta_8^6 \ \delta_8^7 \ \delta_8^8 \ \delta_8^4]$. For the permutation matrix $P = [\delta_8^3 \ \delta_8^2 \ \delta_8^1 \ \delta_8^4 \ \delta_8^5 \ \delta_8^6 \ \delta_8^7 \ \delta_8^8]$, $P^\top L P = [\delta_8^3 \ \delta_8^3 \ \delta_8^2 \ \delta_8^5 \ \delta_8^6 \ \delta_8^7 \ \delta_8^8 \ \delta_8^4]$, which is in the form (16) with $r = 2$, $n_1 = 3$, $k_1 = 2$, $n_2 = 5$, and $k_2 = 5$.

Combining Proposition 1 with Corollary 3 yields the following result.

Corollary 4. *The zeta function of a BN satisfies*

$$\zeta(t) = \prod_{i=1}^{r} \zeta_i(t), \tag{18}$$

where $\zeta_i(t) := (1 - t^{k_i})^{-1}$ is the zeta function of the ith (isolated) component of the BN (consisting of all states that access the ith limit cycle), and $k_1, k_2, \ldots, k_r$ are the lengths of the distinct limit cycles of the BN. Moreover, as $q_v = |\{i : k_i = v\}|$, $v \in \mathbb{N}$, then

$$\zeta(t) = \prod_{i=1}^{\max\{k_j\}} (1 - t^i)^{(-q_i)}.$$

Example 4. Consider again the BN in Example 2. The characteristic polynomial of L factorizes as $P_L(s) = s(s^2 - 1)(s^5 - 1)$. By Proposition 1, the BN has one limit cycle of length 2, and one limit cycle of length 5. Equivalently, its zeta function is $\zeta(t) = \zeta_1(t)\zeta_2(t)$, with $\zeta_1(t) := \frac{1}{1-t^2}$ and $\zeta_2(t) := \frac{1}{1-t^5}$.

3.2. Topological entropy of a Boolean control network

Definition 1. The *topological entropy* of the BCN (1) is

$$h_S := \lim_{j \to \infty} \frac{1}{j} \log |\mathscr{A}_S^j|, \tag{19}$$

where $\mathscr{A}_S^j$ is the set of state trajectories of length j.

In a BCN with n state variables, the number of distinct state trajectories of length j is bounded above by 2^{nj}. Hence,

$$h_S \leq \lim_{j \to \infty} \frac{1}{j} \log 2^{nj} = n \log 2.$$

This upper bound is attained, for example, by the (trivial) BCN $X(k + 1) = U(k)$.

Example 5. Consider the BCN

$$X(k + 1) = U(k) \vee (\bar{U}(k) \wedge \bar{X}(k)).$$

For $U(k) = 1$ [$U(k) = 0$], we have $X(k + 1) = 1$ [$X(k + 1) = \bar{X}(k)$]. The possible state trajectories of length 1 are, of course, 0 and 1, so $|\mathscr{A}_S^1| = 2$. To determine $\mathscr{A}_S^2$, we calculate all possible sequences of length 2. This yields $\mathscr{A}_S^2 = \{11, 10, 01\}$, so $|\mathscr{A}_S^2| = 3$. More generally, from the two possible subsystems we see that all sequences, except for those that contain two consecutive zeros, can appear. This is analogous to the GMS case, so $|\mathscr{A}_S^{j+2}| = |\mathscr{A}_S^{j+1}| + |\mathscr{A}_S^j|$, and $h_S = \log \gamma$.

Our main result in this subsection, obtained by combining Theorem 3 and (10), provides a simple way of computing h_S using the ASSR.

Corollary 5. *The topological entropy of a BCN satisfies*

$$h_S = \log \lambda_M, \tag{20}$$

with M defined as in (11).

Example 6. Consider the BCN in Example 5. The ASSR is given by (3) with $n = m = 1$, and $L = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$. Thus, $L_1 = L \ltimes \delta_2^1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $L_2 = L \ltimes \delta_2^2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $M = L_1 \vee L_2 = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. The eigenvalues of this matrix are $(1 \pm \sqrt{5})/2$, so (20) yields $h_S = \log \gamma$.

Corollary 6. *The topological entropy of a BN is $h = 0$.*

Proof. By (17), any eigenvalue λ of L satisfies either $\lambda = 0$ or $|\lambda| = 1$. Since the state space of a BN is finite, there must be at least one limit cycle, so r in (17) satisfies $r \geq 1$. Hence, 1 is always an eigenvalue, so $\lambda_L = 1$. $\square$

Indeed, the entropy measures how the number of admissible strings grows with the length of the string. A BN is an autonomous system, whose trajectories are uniquely determined by the initial condition. Since the number N_ℓ of distinct trajectories of length ℓ is a constant, independent of ℓ, h is necessarily zero.

The next result shows how Corollary 5 can be used to obtain more general results.

Proposition 2. *Consider the BCN*

$$X_1(k+1) = X_2(k),$$
$$\vdots$$
$$X_{n-1}(k+1) = X_n(k),$$
$$X_n(k+1) = U(k).$$

The topological entropy of this "nth-order shift-register" is $h_S = \log 2$ for every n.

Proof. Fix arbitrary $A, B \in \mathscr{S}^n$. There exists a *unique* control sequence that steers the BCN from $X(0) = A$ to $X(n) = B$, namely, $U(i)$ is bit $i + 1$ of B, $i = 0, 1, \ldots, n - 1$. In the ASSR, $[M^n]_{ij}$ is the number of distinct state trajectories with $n + 1$ symbols beginning with $\delta_{2^n}^j$ and ending with $\delta_{2^n}^i$, so we conclude that $M^n = 1_{2^n, 2^n}$, where $1_{v,w}$ denotes the $v \times w$ matrix with all entries equal to 1. The Perron root of M^n is 2^n (corresponding to the eigenvector $1_{2^n, 1}$), so $h_S = \log \lambda_M = \log 2$. $\quad\square$

4. Regulation of the mammalian cell cycle

The cell cycle is a temporal sequence of molecular events that take place in a cell, leading to its division and duplication. This is the process by which a single-cell fertilized egg develops into a mature organism, as well as the process by which hair, skin, blood cells, and some internal organs are renewed.

The cell cycle is divided into several phases. DNA replication occurs during the synthesis (or S) phase. Growth stops and cellular energy is focused on the orderly division into two daughter cells at the mitosis (or M) phase. The S and M phases are separated by two gap phases, G1 (between M and S) and G2 (between S and M). A fifth phase, called G0, corresponding to a quiescent state, can be reached from G1 in the absence of stimulation. Gap phases enable the cell to monitor its environment and internal state before committing to the S or M phase.

Mammalian cell division is tightly controlled, as it must be coordinated with the overall growth of the organism, and to address specific needs, e.g. wound healing. Faults in this control process can either kill a cell through apoptosis or cause mutations that may lead to cancer. Cell cycle coordination is achieved through extra-cellular positive and negative signals whose balance decides whether a cell will divide or remain in the G0 resting phase.

The positive signals or growth factors ultimately elicit the activation of Cyclin D (CycD) in the cell. Faure, Naldi, Chaouiya, and Thieffry (2006) developed a BCN model for the core network regulating the mammalian cell cycle. The model includes a single Boolean input corresponding to the activation/inactivation of CycD in the cell. The model also includes nine Boolean state variables $X_1(t), \ldots, X_9(t)$ representing the activity/inactivity at time t of nine different proteins: Rb, E2F, CycE, CycA, p27, Cdc20, Cdh1, UbcH10, and CycB, respectively. The BCN model is

$$X_1(t+1) = (\bar{U}(t) \wedge \bar{X}_3(t) \wedge \bar{X}_4(t) \wedge \bar{X}_9(t))$$
$$\vee (X_5(t) \wedge \bar{U}(t) \wedge \bar{X}_9(t)),$$
$$X_2(t+1) = (\bar{X}_1(t) \wedge \bar{X}_4(t) \wedge \bar{X}_9(t))$$
$$\vee (X_5(t) \wedge \bar{X}_1(t) \wedge \bar{X}_9(t)),$$
$$X_3(t+1) = X_2(t) \wedge \bar{X}_1(t),$$

$$X_4(t+1) = (X_2(t) \wedge \bar{X}_1(t) \wedge \bar{X}_6(t) \wedge (\overline{X_7(t) \wedge X_8(t)}))$$
$$\vee (X_4(t) \wedge \bar{X}_1(t) \wedge \bar{X}_6(t) \wedge (\overline{X_7(t) \wedge X_8(t)})),$$
$$X_5(t+1) = (\bar{U}(t) \wedge \bar{X}_3(t) \wedge \bar{X}_4(t) \wedge \bar{X}_9(t))$$
$$\vee (X_5(t) \wedge (\overline{X_3(t) \wedge X_4(t)}) \wedge \bar{U}(t) \wedge \bar{X}_9(t)),$$
$$X_6(t+1) = X_9(t),$$
$$X_7(t+1) = (\bar{X}_4(t) \wedge \bar{X}_9(t)) \vee X_6(t) \vee (X_5(t) \wedge \bar{X}_9(t)),$$
$$X_8(t+1) = \bar{X}_7(t)$$
$$\vee (X_7(t) \wedge X_8(t) \wedge (X_6(t) \vee X_4(t) \vee X_9(t))),$$
$$X_9(t+1) = \bar{X}_6(t) \wedge \bar{X}_7(t). \tag{21}$$

This model is based on a logical regulatory graph of the interactions between the different proteins; see Faure et al. (2006) and the references therein for the details.

Faure et al. (2006) consider the case where either $U(t) \equiv 1$ (i.e. in the presence of CycD) or $U(t) \equiv 0$, so the BCN yields two possible BNs denoted as BN_1 and BN_0, respectively. Their simulations show that BN_1 admits a globally attracting periodic trajectory composed of seven states. The sequence of state transitions along this trajectory qualitatively matches cell cycle progression. BN_0 admits a single state that is globally attracting. This state corresponds to the G0 phase.

Since $n = 9$ and $m = 1$, $L \in \mathscr{S}^{512 \times 1024}$. Set $L_i := L \ltimes \delta_2^i$, $i = 1, 2$. A calculation shows that $\det(sI - L_2) = s^{511}(s - 1)$, so Proposition 1 implies that BN_0 admits a single cycle with length 1 (i.e., an equilibrium point). Similarly, $\det(sI - L_1) = s^{505}(s^7 - 1)$, so Proposition 1 implies that BN_1 admits a single cycle with length 7. This cycle is

$$\delta_{512}^{416} \to \delta_{512}^{477} \to \delta_{512}^{469} \to \delta_{512}^{498} \to \delta_{512}^{378} \to \delta_{512}^{316} \to \delta_{512}^{284} \to \delta_{512}^{416}.$$

A calculation yields $\lambda_M = 1.8522$, so the topological entropy of this BCN is $\log 1.8522$. More work is needed in order to understand the meaning of entropy in BCNs that model biological systems.

Acknowledgments

References

Akutsu, T., Hayashida, M., Ching, W.-K., & Ng, M. K. (2007). Control of Boolean networks: hardness results and algorithms for tree structured networks. *Journal of Theoretical Biology*, *244*, 670–679.

Albert, R., & Barabasi, A.-L. (2000). Dynamics of complex systems: scaling laws for the period of Boolean networks. *Physical Review Letters*, *84*, 5660–5663.

Aldana, M. (2003). Boolean dynamics of networks with scale-free topology. *Physica D*, *185*, 45–66.

Brualdi, R. A., & Ryser, H. J. (1991). *Combinatorial matrix theory*. Cambridge Univ. Press.

Chaos, A., Aldana, M., Espinosa-Soto, C., de Leon, B. G. P., Arroyo, A. G., & Alvarez-Buylla, E. R. (2006). From genes to flower patterns and evolution: dynamic models of gene regulatory networks. *Journal of Plant Growth Regulation*, *25*, 278–289.

Cheng, D., & Qi, H. (2010). State-space analysis of Boolean networks. *IEEE Transactions on Neural Networks*, *21*, 584–594.

Cheng, D., Qi, H., & Li, Z. (2011). *Analysis and control of Boolean networks*. Springer-Verlag.

Cull, P. (1971). Linear analysis of switching nets. *Biological Cybernetics*, *8*, 31–39.

Cull, P. (1975). Control of switching nets. *Biological Cybernetics*, *19*, 137–145.

Drossel, B., Mihaljev, T., & Greil, F. (2005). Number and length of attractors in a critical Kauffman model with connectivity one. *Physical Review Letters*, *94*, 088701.

Faure, A., Naldi, A., Chaouiya, C., & Thieffry, D. (2006). Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. *Bioinformatics*, *22*, e124–e131.

Fornasini, E., & Valcher, M. E. (2013). Observability, reconstructibility and state observers of Boolean control networks. *IEEE Transactions on Automatic Control*, *58*(6), 1390–1401.

Green, D. G., Leishman, T. G., & Sadedin, S. (2007). The emergence of social consensus in Boolean networks. In *Proc. IEEE symp. artificial life, ALIFE'07, Honolulu, HI* (pp. 402–408).

Gupta, S., Bisht, S. S., Kukreti, R., Jain, S., & Brahmachari, S. K. (2007). Boolean network analysis of a neurotransmitter signaling pathway. *Journal of Theoretical Biology*, *244*, 463–469.

Hassoun, M. H. (1995). *Fundamentals of artificial neural networks*. MIT Press.

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, *22*, 437–467.

Kauffman, S. A. (1993). *Origins of order: self-organization and selection in evolution*. Oxford University Press.

Kauffman, S., Peterson, C., Samuelsson, B., & Troein, C. (2003). Random Boolean network models and the yeast transcriptional network. *Proceedings of the National Academy of Sciences United States of America*, *100*, 14796–14799.

Laschov, D., & Margaliot, M. (2011). A maximum principle for single-input Boolean control networks. *IEEE Transactions on Automatic Control*, *56*, 913–917.

Laschov, D., & Margaliot, M. (2013). Minimum-time control of Boolean networks. *SIAM Journal on Control and Optimization* (in press).

Laschov, D., Margaliot, M., & Even, G. (2013). Observability of Boolean networks: a graph-theoretic approach. *Automatica*, *49*(8), 2351–2362.

Li, S., Assmann, S. M., & Albert, R. (2006). Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. *PLoS Biology*, *4*, 1732–1748.

Li, F., Long, T., Lu, Y., Ouyang, Q., & Tang, C. (2004). The yeast cell–cycle network is robustly designed. *Proceedings of the National Academy of Sciences United States of America*, *101*, 4781–4786.

Lind, D., & Marcus, B. (1995). *An introduction to symbolic dynamics and coding*. Cambridge Univ. Press.

Szallasi, Z., & Liang, S. (1998). Modeling the normal and neoplastic cell cycle with realistic Boolean genetic networks: their application for understanding carcinogenesis and assessing therapeutic strategies. *Pacific Symposium on Biocomputing*, *3*, 66–76.

van Lint, J. H., & Wilson, R. M. (2001). *A course in combinatorics* (2nd ed.). Cambridge Univ. Press.

Williams, S. G. (2004). Introduction to symbolic dynamics. In *Symbolic dynamics and its applications* (pp. 1–11). American Mathematical Society.

Zhao, Q. (2005). A remark on scalar equations for synchronous Boolean networks with biological applications by C. Arrow, J. Heidel, J. Maloney, and J. Rogers. *IEEE Transactions on Neural Networks*, *16*, 1715–1716.

Gal Hochma received her B.Sc. in Electrical Engineering and Physics from Tel Aviv University in 2013. She is currently an M.Sc. student at the School of Electrical Engineering, Tel Aviv University. Her research interests include Boolean control networks and switched systems.

Michael Margaliot received the B.Sc. (*cum laude*) and M.Sc. degrees in Electrical Engineering from the Technion – Israel Institute of Technology – in 1992 and 1995, respectively, and the Ph.D. degree (*summa cum laude*) from Tel Aviv University in 1999. He was a post-doctoral fellow in the Department of Theoretical Mathematics at the Weizmann Institute of Science, Rehovot, Israel. In 2000, he joined the faculty of the Department of Electrical Engineering Systems, Tel Aviv University, where he is currently an Associate Professor.

Dr. Margaliot's research interests include stability theory, switched systems, optimal control theory, fuzzy modeling and control, Boolean control networks, and systems biology. He is co-author of *New Approaches to Fuzzy Modeling and Control: Design and Analysis* (World Scientific, 2000) and of *Knowledge-Based Neurocomputing: A Fuzzy Logic Approach* (Springer, 2009).

Ettore Fornasini received the laurea degrees in Electronic Engineering and in Mathematics from the University of Padova. Since November 1980 he has been full professor of Mathematical System Theory in the Faculty of Engineering, University of Padova, and taught courses in the areas of systems theory, automatic control and ecological modeling.

He spent research periods in Gainesville (Florida) at the Center for Mathematical Systems Theory, at MIT (Massachusetts), at the Technical University of Delft (Holland), at the Institut für Mathematik der Universität Innsbruck (Austria), at the University of Aveiro (Portugal).

He was an Editor of Multidimensional Systems and Signal Processing and is a reviewer for several technical journals.

His research interests include modeling and realization of nonlinear systems, dissipative systems, the polynomial approach to linear systems, structure, realization and control of multidimensional systems, modeling of 2D ecological systems, multidimensional Markov chains, 1D and 2D convolutional coding theory, and 1D and 2D positive systems.

Maria Elena Valcher (EE Master Degree 1991; Ph.D. 1995) is full professor of Automatica at the University of Padova.

She is author/co-author of about 70 papers that have appeared in international journals, 80 conference papers, two textbooks and several book chapters. Her research interests include multidimensional systems theory, polynomial matrix theory, behavior theory, convolutional coding, fault detection, delay–differential systems, positive systems and positive switched systems.

She has been involved in the Organizing Committees and in the Program Committees of several conferences. She recently was the Program Chair of the IEEE CDC 2012. She was on the Editorial Board of the IEEE Transactions on Automatic Control (1999–2002) and Systems and Control Letters (2004–2010), and she is currently on the Editorial Boards of Automatica (2006 to today), Multidimensional Systems and Signal Processing (2004 to today) and European Journal of Control (2013 to today).

She was a Member of the IEEE Control System Society BoG (appointed 2003; elected 2004–2006; elected 2010–2012). She has been a Fellow of the IEEE since January 2012. She received the 2011 IEEE CSS Distinguished Member Award.