

Online Observability of Boolean Control Networks*

Guisen Wu Liyun Dai* Zhiming Liu
RISE & School of Computer and Information Science,
Southwest University
Chongqing, China
{wgs233,dailiyun,zhimingliu88}@swu.edu.cn

Taolue Chen
Department of Computer Science and Information Systems
Birkbeck, University of London
taolue@dcs.bbk.ac.uk

Jun Pang
Faculty of Science, Technology and Communication
University of Luxembourg
jun.pang@uni.lu

Hongyang Qu
Department of Automatic Control and Systems Engineering
University of Sheffield
h.qu@sheffield.ac.uk

Abstract—Four types of observability of Boolean control networks (BCNs) have been proposed to solve different problems. However, all of them are offline observabilities that we can't adjust the input sequence by observing the output sequence in the process of determining the initial state of BCNs. In this paper, we firstly propose the online observability that it can determine the initial state of BCNs dynamically without presupposing its initial state. The online observability accomplishes this task by deciding input sequence and observing out sequence in every time step. The node values of BCNs update at discrete time, then we use the time step to represent the discrete time. Compare with offline observabilities, the online observability can determine the initial state of some biological systems which can be checked at most once.

Index Terms—Boolean control networks, online observability, directed graph, supertree

I. INTRODUCTION

In 1960s, Nobel Prize winners Jacob and Monod found that “Any cell contains a number of ‘regulatory’ genes that act as switches and can turn one another on and off. If genes can turn one another on and off, then you can have genetic circuits.” [1], [2]. Inspired by these Boolean-type actions in genetic circuits, the Boolean networks (BNs) is firstly proposed by Kauffman [3] for modeling nonlinear and complex biological systems. Some general descriptions of the BNs and its applications to biological systems can be found in Kauffman. Since then research interests in BNs have been motivated by the large number of natural and artificial systems. These natural and artificial systems describing variables display only two distinct configurations, and hence these describing variables take only two values, i.e., $\{0, 1\}$ [8], [7], [9], [5], [6], [4].

When external regulation or perturbation is considered, BNs are naturally extended to Boolean control networks (BCNs) [10]. The study on control-theoretic problems of BCNs can date back to 2007 [11]. The work above also proves that the problem of determining the controllability of BCNs is NP-hard in the number of nodes. Furthermore, it points out that “One of the major goals of systems biology is to develop a control theory for complex biological systems.” Since then,

the study on control-theoretic problems in the areas of BNs and BCNs has drawn great attention [2], [12], [13], [14], [4]. Furthermore, controllability and observability are basic control-theoretic problems. In 2009, an algebraic framework to deal with both BNs and BCNs by using *semi-tensor product* (STP) of matrices has been developed in [2]. Moreover they give equivalent conditions for controllability of BCNs and observability of controllable BCNs. To date, there are four types of observability have been proposed.

- 1) The first type of observability proposed in 2009 [2] means that every initial state can be determined by an input sequence.
- 2) The second observability proposed in 2010 [12] stands for that for every two distinct initial states, there exists an input sequence which can distinguish them, and this observability is determined in [15].
- 3) The third observability proposed in 2011 [13] states that there is an input sequence that determines the initial state.
- 4) The fourth observability proposed in 2013 [4] is essentially the observability of linear control systems, i.e., every sufficient long input sequence can determine the initial state.

In abovementioned definitions an input is not the value of an input-node, but it represents the values of all input-nodes of the BCN on a time step. Therefore an input is a vector of the values of all input-nodes of the BCN on a time step. An input sequence consists of several inputs in sequential time steps. We will give the formal definition of four observabilities in the following pages.

However, all of four existing types of observability of BCNs are offline observabilities which means that they can't adjust the input sequence by observing the output sequence in the process of determining the initial state of BCNs. Therefore, we propose the online observability that we can determine the initial state of BCNs dynamically, that the online observability decides the input sequence in each time step by observes the out sequence. We can infer the possible states set by observe

*Corresponding author

the output of *BCN*, then we can decide the input based on the possible states set, so we call this process is a dynamic process.

Contribution: Firstly, we give the formal definition of online observability of *BCNs*. Secondly, we provide two algorithms to determine the online observability for *BCNs*. Finally, we introduce some applications of the online observability of *BCNs* and advantages of online observability compared with offline observabilities.

Structure: The remainder of this paper is organized as follows. *Section II* introduces necessary preliminaries about *BCNs*, algebraic forms of *BCNs* and the four existing kinds of observability of *BCNs*. *Section III* presents the definition of deduce function, *K* steps deterministic and online observability of *BCNs*. *Section IV* presents how to determine the online observability of *BCNs* by super tree and directed graph. *Section V* talks about some applications of the online observability of *BCNs*. We also compare the online observability with offline observabilities in this section. *Section VI* ends up with the introduction of some future works.

II. PRELIMINARIES

In this section we introduce the definition of *BCNs* and their algebraic form, and the existed four kinds of observability of *BCNs*.

A. Boolean Control Networks

A Boolean control network can be described as a directed graph together with logical equations to describe the updating rules of the nodes in the graph, the definition of *BCN* is as follows.

Definition 1: ([10]) A *BCN* consists of input-nodes, state-nodes, output-nodes, and directed edges which connect nodes. A node in *BCN* can take a logic value from $\{0, 1\}$ at a discrete time $0, 1, 2, \dots$. For one directed edge from node v_i to node v_j means that the logic value of v_j at time step $t + 1$ is affected by the value of v_i at time step t or $t + 1$.

Note that one can only know whether or not a node is affected by another node from the network graph. Different *BCNs* may have the same structure, in order to determine a *BCN* uniquely, logical equations are also needed to describe the specific updating rules of *BCNs*.

Firstly, we give a simple example to describe *BCN*.

Example 1: In Fig.1 we have a *BCN* with two input-nodes *A* and *B*, four state-nodes *C*, *D*, *E* and *F*, and two output-nodes *G*, *W*. The updating rules of the *BCN* shown in the edges of Fig.1, and they are also described as truth table Fig.2. The reason why we use truth table to describe the updating rules of the *BCN* is that, this form will be more convenient for *BCN* to convert into its algebraic form. For instance, we can know the updating rule of state-node *G* from the truth table,

$$G(t) = C(t) \wedge \neg(\neg D(t) \wedge \neg E(t) \wedge \neg F(t))$$

For convenience, we will use this example in the whole paper to explain various concepts we introduce.

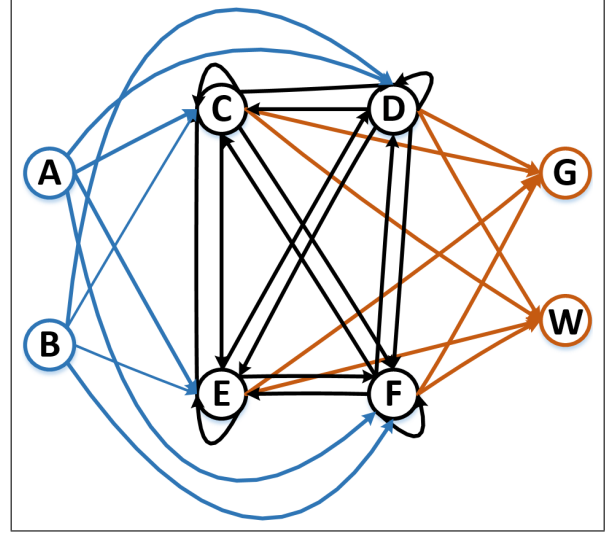


Fig. 1. A Boolean control network with two input-nodes *A* and *B*, four state-nodes *C*, *D*, *E* and *F*, and two output-nodes *G*, *W*. We use blue, black and orange, to distinguish three kinds of nodes and three kinds of edges.

Truth Table																		A(t)	B(t)
C(t)	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0		
D(t)	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0		
E(t)	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	0		
F(t)	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1		
C(t+1):	0	1	0	0	0	1	1	1	0	1	1	0	1	1	1	1	0	1	1
	1	0	0	0	0	1	1	0	1	0	0	0	1	1	1	1	1	1	0
	1	1	1	0	0	0	1	1	0	0	1	0	0	0	0	1	0	0	1
	1	0	1	1	0	1	0	0	1	1	0	1	0	0	0	0	0	0	0
D(t+1):	1	1	1	0	1	0	1	0	0	1	1	1	0	0	0	0	0	1	1
	0	1	0	1	0	1	1	0	1	0	1	0	0	0	0	1	0	1	0
	0	1	1	0	0	1	0	1	0	0	0	0	1	1	1	0	0	0	1
	1	0	0	0	1	0	1	1	0	0	0	1	1	1	0	0	0	0	0
E(t+1):	1	0	0	0	1	1	1	0	0	1	0	0	0	0	1	0	1	1	1
	0	1	0	1	1	0	0	0	1	1	0	1	1	0	1	1	1	1	0
	0	1	0	1	1	1	1	1	0	1	1	1	0	1	0	0	0	0	1
	1	1	0	1	1	0	0	0	1	1	1	0	1	0	0	0	0	0	0
F(t+1):	0	0	1	0	1	1	1	1	1	0	1	0	1	0	0	1	1	1	1
	0	1	1	0	0	0	0	1	0	0	1	1	1	1	0	0	1	0	0
	1	0	1	1	1	1	1	1	0	1	0	0	1	0	0	1	0	0	1
	1	0	1	0	1	0	1	0	1	1	1	1	0	0	0	0	0	0	0
G(t):	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0		
W(t):	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0	0	0		

Fig. 2. The truth table which describe the updating rules of the *BCN* shown in Fig.1.

B. The Algebraic Forms of BCNs

To better illustrate the concept of algebraic forms, in this paper, we investigate the following *BCN* assumes to has n state-nodes, m input-nodes and q output-nodes. Then the updating rules of the *BCN* can be described as following formulas:

$$\begin{aligned} s(t+1) &= f(i(t), s(t)) \\ o(t) &= h(s(t)) \end{aligned} \quad (1)$$

$s(t) \in \mathbb{B}^n$ are state-nodes; $i(t) \in \mathbb{B}^m$ are input-nodes; $o(t) \in \mathbb{B}^q$ are output-nodes; $f : \mathbb{B}^{n+m} \mapsto \mathbb{B}^n$ and $h : \mathbb{B}^n \mapsto \mathbb{B}^q$ are logical functions that represent the updating rules of *BCNs*. Where $\mathbb{B} : \text{the set } \{0, 1\}$; $t = 0, 1, \dots$ represents the discrete

time.

Therefore in the previously mentioned example, we have $C(t), D(t), E(t), F(t) \in s(t)$; $A(t), B(t) \in i(t)$ and $G(t), W(t) \in o(t)$; $n = 4$, $m = 2$ and $q = 2$; f and h are described in the truth table (Fig.2).

The *STP* of matrices can be used to represent the algebraic forms of *BCNs* [2], the definition of *STP* is as follows.

Definition 2 (STP): [14] Let $X \in \mathbb{R}_{m \times n}$, $Y \in \mathbb{R}_{p \times q}$ and $\alpha = \text{lcm}(n, p)$ be the least common multiple of n and p . The *STP* of X and Y is defined as

$$X \ltimes Y = (X \otimes I_{\alpha/n})(Y \otimes I_{\alpha/p}),$$

where \otimes denotes the Kronecker product.

After introduce the definition of *STP* of matrices, we introduce some related notations at first [16]:

- δ_n^i : the i -th column of the identity matrix I_n ;
- Δ_n : the set $\{\delta_n^1, \dots, \delta_n^n\}$;
- $\delta_n[i_1, \dots, i_s]$: $[\delta_n^{i_1}, \dots, \delta_n^{i_s}]$ ($i_1, \dots, i_s \in \{1, 2, \dots, n\}$) the logical matrix;
- $L_{n \times s}$: the set of $n \times s$ logical matrices.

Using *STP* of matrices, the formula (1) can be equivalently represented in the following algebraic form:

$$\begin{aligned} s(t+1) &= L \ltimes i(t) \ltimes s(t) \\ o(t) &= H \ltimes s(t) \end{aligned} \quad (2)$$

where $s(t) \in \Delta_N$, $i(t) \in \Delta_M$, and $o(t) \in \Delta_Q$ denote the states, inputs and outputs respectively the same as in (1), but they are with vector forms; $L \in L_{N \times (NM)}$; $H \in L_{Q \times N}$ denote the relation matrices; that $N = 2^n$, $M = 2^m$, and $Q = 2^q$. Since *STP* keeps most properties of the conventional product [14], the associative law, the distributive law, etc., we usually omit the symbol “ \ltimes ” hereinafter. For instance, the formula “ $s(t+1) = L \ltimes i(t) \ltimes s(t)$ ” can be written as “ $s(t+1) = Li(t)s(t)$ ”.

To construct algebraic form (2) we give a mapping $\tau : \{0, 1\} \mapsto \{\delta_2^1, \delta_2^2\}$ where $\tau(0) = \delta_2^2$, $\tau(1) = \delta_2^1$. Therefore, the logical variable $A(t)$ takes value from these two vectors, i.e., $A(t) \in \{\delta_2^1, \delta_2^2\}$. Using the *STP* of matrices, we have

$$\begin{aligned} i(t) &= i_1(t) \cdots i_m(t); \\ s(t) &= s_1(t) \cdots s_n(t); \\ o(t) &= o_1(t) \cdots o_q(t). \end{aligned}$$

And according to [17], each logical function f_p of state-nodes can be found in update rules (1). The form of f_p as:

$$f_p(i_1(t), \dots, i_m(t), s_1(t), \dots, s_n(t))$$

and there exists a structure matrix $L_p \in L_{2 \times NM}$ such that

$$\tau(f_p(i_1(t), \dots, i_m(t), s_1(t), \dots, s_n(t))) = L_p i(t) s(t) \quad (3)$$

For state-nodes s_1, \dots, s_n , we have n logical matrices L_1, \dots, L_n for them, respectively. If for each state-node s_p the logical matrix has form

$$L_p = [\delta_2^{p_1}, \dots, \delta_2^{p_{NM}}],$$

then we have that

$$L = [\delta_N^{R_1}, \dots, \delta_N^{R_{NM}}]$$

where

$$\delta_N^{R_1} = \delta_2^{1_1} \cdots \delta_2^{n_1}; \dots; \delta_N^{R_{NM}} = \delta_2^{1_{NM}} \cdots \delta_2^{n_{NM}}.$$

By this relationship we can construct the L for the algebraic forms of *BCNs*, and we can also construct the logical matrix H in the similar way.

For instance, the *BCN* whose structure is depicted in Fig.1, and whose updating rules is described as truth table in Fig.2. Hence the updating rules of this *BCN* can be represented with the algebraic form:

$$\begin{aligned} s(t+1) &= \delta_{16}[\alpha] i(t) s(t) \\ o(t) &= \delta_4[\beta] s(t) \end{aligned} \quad (4)$$

where $\alpha = \{10, 4, 11, 16, 9, 5, 1, 7, 15, 2, 3, 12, 7, 6, 8, 13, 8, 9, 15, 10, 14, 4, 3, 16, 1, 14, 12, 13, 5, 7, 2, 6, 7, 2, 3, 13, 13, 9, 5, 1, 16, 13, 6, 14, 11, 10, 4, 15, 1, 14, 7, 6, 9, 8, 11, 12, 5, 5, 13, 3, 10, 12, 16, 16\}$, $\beta = \{1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4\}$, $t \in \mathbb{N}$, $s \in \Delta_{16}$, $i \in \Delta_4$ and $o \in \Delta_4$.

C. Four Existing Observability of BCNs

In this section we introduce four existing kinds of observability of *BCNs*. Let Δ_N , Δ_M , Δ_Q be three alphabets, for all $s_0 \in \Delta_N$ and all $p \in \mathbb{Z}_+$; ∞ is the infinite natural numbers. In order to introduce four existing kinds of observability of *BCNs*, we define the mappings [16]:

$$\begin{aligned} L_{s_0}^p &: (\Delta_M)^p \mapsto (\Delta_N)^p, u_0 \cdots u_{p-1} \mapsto s_1 \cdots s_p \\ L_{s_0}^\infty &: (\Delta_M)^\infty \mapsto (\Delta_N)^\infty, u_0 u_1 \cdots \mapsto s_1 s_2 \cdots \end{aligned} \quad (5)$$

$$\begin{aligned} (HL)_{s_0}^p &: (\Delta_M)^p \mapsto (\Delta_Q)^p, u_0 \cdots u_{p-1} \mapsto o_1 \cdots o_p \\ (HL)_{s_0}^\infty &: (\Delta_M)^\infty \mapsto (\Delta_Q)^\infty, u_0 u_1 \cdots \mapsto o_1 o_2 \cdots \end{aligned} \quad (6)$$

For all $p \in \mathbb{Z}_+$, all $U = u_1 \dots u_p \in (\Delta_M)^p$, and all $1 \geq i \geq j \geq |U|$, we use $U[i, j]$ to denote the word $u_i \dots u_j$ as a input sequence. Then four existing kinds of observability of *BCNs* can be define as:

Definition 3: The first kind of observability is that, *BCN* is called observable, if for every initial state $s_0 \in \Delta_N$, there exists an input sequence $U \in (\Delta_M)^p$ for some $p \in \mathbb{Z}_+$ such that for all states $s_0 \neq s'_0 \in \Delta_N$, $Hs_0 = Hs'_0$ implies $(HL)_{s_0}^p(U) \neq (HL)_{s'_0}^p(U)$ [2].

Hence the first observability means that if a *BCN* is observable then every initial state of the *BCN* can be determined by an input sequence. But we can only use the corresponding input sequence of a state to check whether this state is the initial state of the *BCN* or not.

Definition 4: The second kind of observability is that a *BCN* is called observable if for any distinct states $s_0, s'_0 \in \Delta_N$, there exists an input sequence $U \in (\Delta_M)^p$ for some $p \in \mathbb{Z}_+$, such that $Hs_0 = Hs'_0$ implies $(HL)_{s_0}^p(U) \neq (HL)_{s'_0}^p(U)$ [12].

The second observability means that a *BCN* is called observable if for every two distinct initial states of the *BCN*, there exists an input sequence which can distinguish them.

Definition 5: The third kind of observability is that, a *BCN* is called observable, if there exists an input sequence $U \in (\Delta_M)^p$ for some $p \in \mathbb{Z}_+$, such that for any distinct states $s_0, s'_0 \in \Delta_N$, $HS_0 = HS'_0$ implies $(HL)_{s_0}^p(U) \neq (HL)_{s'_0}^p(U)$ [13].

The third observability means that a *BCN* is called observable if there is an input sequence that determines the initial state of the *BCN*.

Definition 6: The fourth kind of observability is that, *BCN* is called observable, if for any distinct states $s_0, s'_0 \in \Delta_N$, for any input sequence $U \in (\Delta_M)^\infty$, $HS_0 = HS'_0$ implies $(HL)_{s_0}^\infty(U) \neq (HL)_{s'_0}^\infty(U)$ [4].

The fourth observability means that a *BCN* is called observable if every sufficient long input sequence can determine the initial state of the *BCN*.

Then from the definitions of four existing kinds of observability, we know that [16]:

- the first one implies the second one;
- the third one implies the second one and first one;
- the fourth one implies the third one, second one and first one.

And we can't use the first one and second one to determine the initial state of *BCNs*, when we don't presuppose the initial state of *BCNs*. For example, in the first kind observability we need to assume the initial state of a *BCN*, and then check it by corresponding input sequence. If the assuming is correct, we can determine the initial state. But if the the assuming is not correct, we can't determine the initial state of the *BCN*. However we can use the third observability and fourth observability to determine the initial state of *BCNs* without presupposing the initial state of *BCNs*. But the requirements for *BCNs* are very harsh when we use the third observability and fourth observability. With these disadvantages of four existing observabilities, we proposed the online observability of *BCNs* to solve this problem.

Problem: Finding the necessary and sufficient condition of determine the initial state of *BCNs* without presupposing its initial state.

III. THE ONLINE OBSERVABILITY OF BCNS

In this paper we propose the online observability, the informal definition of it is as follows.

Definition 7: A *BCN* is called online observable, if every initial state $s_0 \in \Delta_N$ can be determined by dynamically deciding input sequence and observing output sequence at every step without presupposing the initial state of *BCN*, and this process can be accomplished in finite steps.

In this section, firstly we present the definition of deduce function, secondly we present the definition of K steps deterministic. We take them as the preparations for defining online observability. Finally, we give the formal definition of online observability of *BCNs*.

A. Deduce Function

Different with existed four types, the observability we proposed can determine the initial state online. Because in

the process of determining the initial state every input of the input sequence is decided by the output we observe at every time step. At the beginning, we can observe the output of *BCNs*, so that we can infer the possible values of state-nodes and treat them as possible states set S_0 . Then as we can know the possible states set, we need to decide the input i_0 . The input i_0 will make sure any different possible states $s_i, s_j \in S_0$ will not turn into the same state after affected by input $LS_i i_0 = LS_j i_0$. After decided input, we can observe the new output, and then we can infer the new possible states set. The cardinality number of possible states set after input will not larger than the cardinality of possible states set before input. If the cardinality number of possible states set turn into be 1 then we can determine the initial state of *BCN*. To simulate this deduction process, we give the definition of deduce function that.

Definition 8 (Deduce Function): The deduce function can be defined as $D(S, I, O)$. Based on deduction process, we have for any

$$s_i(t+1) \in D(S, I, O)$$

there exists $s_i(t)$ in S that

$$s_i(t+1) = LIs_i(t)$$

and

$$O = Hs_i(t+1).$$

where

- $S \in 2^{\Delta_N}$ is the possible states set;
- $I \in \Delta_M$ represents the input;
- $O \in \Delta_Q$ represents the output;
- $D(S, I, O) \in 2^{\Delta_N}$ is the possible states set after deduction.

From the definition of deduce, we have some equations for this function that

$$D(\emptyset, I_i, O_i) = \emptyset \quad (7)$$

Equation (7) represents that if the possible states set is an empty set \emptyset , no matter what we do we can only deduce the possible set is \emptyset too.

$$D(S_i, \varepsilon, \varepsilon) = S_i \quad (8)$$

If the possible states set is S_i and we neither input anything and nor observe the output. In this case we can only deduce that the possible states set is S_i shown in equation (8).

$$D(\Delta_N, \varepsilon, \delta_4^1) = \{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\} \quad (9)$$

Using the example mentioned before, when the possible states set $S_i = \Delta_N$, and we observe that the outputs of *BCN* is δ_4^1 before we decide input. In this case we can deduce that the possible states would be $\delta_{16}^1, \delta_{16}^2$ or δ_{16}^3 shown in equation (9).

$$D(\{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}, \delta_4^1, \varepsilon) = \{\delta_{16}^{10}, \delta_{16}^4, \delta_{16}^{11}\} \quad (10)$$

If the possible states set $S_i = \{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}$ we input δ_4^1 . Before we observe the output of BCN we can only deduce the possible states would be δ_{16}^{10} , δ_{16}^4 or δ_{16}^{11} shown in equation (10).

$$D(\{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}, \delta_4^1, \delta_4^3) = \{\delta_{16}^{10}, \delta_{16}^{11}\} \quad (11)$$

But if we observe that the output of BCN is δ_4^3 , then we can deduce that the possible state can be δ_{16}^{10} or δ_{16}^{11} shown in equation (11);

$$D(\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6\}, \delta_4^3, \varepsilon) = \{\delta_{16}^9, \delta_{16}^{13}\} \quad (12)$$

Finally if the set of states is $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6\}$ and the inputs is δ_4^3 . Before we observe the output of BCN we can deduce that the possible state values can be δ_{16}^9 or δ_{16}^{13} shown in equation (12), as the cardinality number of the possible states set decreased, we can't deduce the initial state any more.

B. K Steps Deterministic

After we defined the deduce function, we can present the definition of K steps deterministic of the states set of $BCNs$ and the range of K is the set of natural numbers. It may easier to define online observability by programming language. But we would like to define its mathematical form for preciseness of concepts. Therefore, before defining the online observability of $BCNs$, we need to define the K steps deterministic of the states set of $BCNs$ at first.

Definition 9 (K Steps Deterministic): When $K = 0$, if for a set of states S_i and $|S_i| = 1$, then S_i is K step deterministic. When $K > 0$, if for a set of states S_i ($|S_i| > 1$), there exists I_i in Δ_M implies

$$|D(S_i, I_i, \varepsilon)| = |S_i|,$$

and implies "For every O_i in Δ_Q ,

$$|D(S_i, I_i, O_i)| \neq 0$$

implies $D(S_i, I_i, O_i)$ is K' ($K' < K$) steps deterministic.', then S_i is K steps deterministic.

From the definition of K steps deterministic we know $K = 0$ means that we can determine the state without any input and observing output. Because if we know the cardinality number of possible states set is 1, then we can know the state of $BCNs$. We can only discuss the case of $K = 0$ when $|S_i| = 1$. If $K > 0$, then the definition of K steps deterministic is defined recursively, and it need to use the definition of K ($K = 0$) steps. When we talk that a states set of $BCNs$ is k_i steps deterministic we default $k_i \geq 0$.

Furthermore, if S_i is k_1 steps deterministic and $k_1 \leq k_2$, then S_i is k_2 steps deterministic. But if S_i is k_1 steps deterministic and $k_1 \geq k_2$, we can not make sure whether S_i is k_2 steps deterministic or not. Therefore you can consider the " S_i is k_i steps deterministic" as "We can determine the state of a BCN with possible states set S_i in k_i steps. And we finish this process by deciding input sequence and observing out sequence at each time step".

C. Online Observability

After the previous preparation, we present the formal definition of the online observability. The formal definition of the online observability of $BCNs$ is as follows.

Definition 10 (Online Observability of BCNs): If for every O_i in Δ_Q and $|D(\Delta_N, \varepsilon, O_i)| \neq 0$, there exists a $k \geq 0$ implies $D(\Delta_N, \varepsilon, O_i)$ is k steps deterministic, then this BCN is online observable. We even can define it simpler, if there exists $k \geq 0$ implies Δ_N is k steps deterministic, then this BCN is online observable.

The difference between the second definition and the first definition is that whether we observe the corresponding output of the initial state of BCN at first. For better performance, we use the first definition of online observability.

Because in the existed second kind of observability, we presuppose the initial state of $BCNs$, and then try to find the input sequence to distinguish it from other kinds of initial states, but the input sequence determined by the presupposed initial state may make other kinds of initial states turn into be the same state, so that other kinds of initial states can't be distinguished anymore. This problem has to be considered in the online observability of $BCNs$, so the online observability implies the existed first kind of observability, and then implies the existed second kind of observability. In the existed third kind of observability, there has to exist an input sequence that can distinguish any distinct states, but in online observability we can use different input sequences to distinguish different states set which classified by their corresponding output. So we have the existed third kind of observability implies the online observability of $BCNs$, then the existed fourth kind of observability implies the online observability.

When I learn the existed four kinds of observability of $BCNs$, I find that if we want determine the initial state of a $BCNs$ by first kind of observability, we need to guess the initial state of the BCN and then check it by its corresponding input sequence, if the initial state we guess is right, we can determine it, but if not, we need to guess again and input the corresponding input sequence untill we determine the initial state of the BCN . But if we can't repeat this process, we may can't determine the initial state of the BCN any more. Then I turn my gaze to the third observability, this kind of observability makes we can determine the initial state without presupposing the initial state. But I think if we can determine the possible states set of the BCN by observing the output at first, why can't we try to find corresponding input sequence for them? And then my teacher and I talk about this thinkness and expand it into the original idea of the online observability of $BCNs$.

From the informal definition and formal definition of online observability, we can know that the necessary and sufficient condition of determine the initial state of $BCNs$ without presuppose the initial state is the online observability of $BCNs$. By this definition we can build $BCNs$ with least output-nodes when we want to determine the initial state of $BCNs$.

IV. DETERMINING THE ONLINE OBSERVABILITY OF BCNs

In this paper we proposed two ways to determine the online observability of *BCNs*, the first one is by supertree; the second one is by directed graph. The construction process of supertree simulates deduction process mentioned before. Then we check the tree based on the definition of online observability of *BCNs* depth first or breadth first. When we used the super tree to determine the observability of *BCNs*, we need to check the existence of loops, and many nodes in the tree are repeated, which will take a lot of time overhead and space overhead. So we proposed the second way which determine the online observability by directed graph, in this way we can avoid checking the existence of loop and avoid checking repeated nodes. There are also other advantages like determining observability earlier and select the input smarter, which will reduce time and space overhead.

A. Supertree

As we mentioned before when we want to determine the initial state of *BCNs*, we can use the deduce function. Then according to the definition of online observability we will alternately observing the output and deciding the input. When the cardinal number of the states set comes into be 1 we can determine the initial state, and stop deducing the initial state of *BCNs*. According to this process, we can define the supertree for *BCNs*. For convenience, we use the states set inside the node to represent the node, and output in the edge to represent the edge.

Definition 11 (Supertree): The root node of the supertree should be Δ_N , if $(|S_i| = 1)$ then S_i would be the leave node, else if a node S_i in the $2k + 1$ ($k \geq 0$) layer of the supertree, then its son nodes would be $D(S_i, \varepsilon, O_i)$, O_i would be the edge from S_i to $D(S_i, \varepsilon, O_i)$ where $|D(S_i, \varepsilon, O_i)| > 0$, if a node S_i in the $2k + 2$ ($k \geq 0$) layer of the supertree, then its son nodes would be $D(S_i, I_i, \varepsilon)$ where $|D(S_i, \varepsilon, O_i)| > 0$, I_i would be the edge from S_i to $D(S_i, I_i, \varepsilon)$.

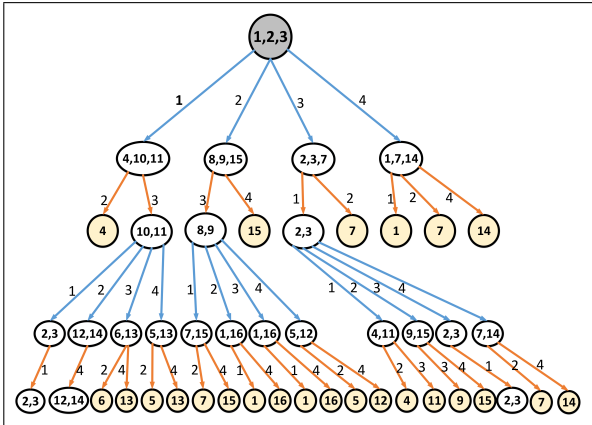


Fig. 3. Branch of the tree which represents $\{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}$. The blue edges and orange edges show the observing output processes and deciding input processes respectively; the yellow nodes are leaf nodes.

For example, the Fig.3 show branch of the tree which represents $\{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}$, the nodes represent the states sets, the blue edges represent the observing output processes, and the orange edges represent the deciding input processes. And only the yellow nodes are the leave nodes, so you can see that this branch is not completed. If we want to find all of the ways to determine the initial state, we have to build the complete tree, it will takes many additional time and space overhead. Especially when there are loops in the tree like the $\{\delta_{16}^2, \delta_{16}^3\}$ in fourth layer and the $\{\delta_{16}^2, \delta_{16}^3\}$ in fifth layer that will form a loop. In this case you can never build the complete tree, so you need to check the existence of loops and omit it. There are also some nodes take the same states set which will also take additional overhead, like that there two nodes which take $\{\delta_{16}^1, \delta_{16}^2\}$ in the fifth layer. But if we only need to find a ways to determine the initial state, when we find the leaf nodes δ_{16}^1 , δ_{16}^7 and δ_{16}^{14} in third layer by breadth-first algorithm, you can make sure that the states set $\{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}$ is 1 step deterministic.

B. Directed Graph

To improve the shortcomings of the way by supertree, we proposed the way by directed graph which takes less time and space overhead. The most difference between supertree and directed graph is that supertree is built from the root node to leaf nodes, but the directed graph is built from smaller nodes (contain less states) to larger nodes (contain more states).

The construction algorithm of directed graph is shown in the Algorithm.1, the algorithm to build nodes used in the Algorithm.1 is shown in the Algorithm.2. Some details in Algorithm.1 and Algorithm.2 are as follows:

- Sort the states inside the nodes, and then sort the n-odes: For example, the nodes $\{\delta_{16}^1, \delta_{16}^2\}$, $\{\delta_{16}^1, \delta_{16}^3\}$ and $\{\delta_{16}^2, \delta_{16}^3\}$ shown in Fig.4.
- Using other nodes to find the suitable inputs set for nodes with k states: For example, we can search inputs sets which make $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6\}$, $\{\delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$ and $\{\delta_{16}^4, \delta_{16}^7\}$ z ($z \geq 0$) steps deterministic, and then take the intersection of these sets to be the suitable inputs set of $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$.
- Check whether an input can make a nodes z steps deterministic: According to the order determined in previous steps, we check every node one by one. If for one input I_i which belongs to suitable inputs set of the states set S_i implies $|D(S_i, I_i, \varepsilon)| < |S_i|$, we can make sure the I_i is a wrong input; else if for all $O_i \in \Delta_Q$ and $|D(S_i, I_i, O_i)| > 0$, the $D(S_i, I_i, O_i)$ is z steps deterministic then I_i is a right input, then connect the node S_i to all nodes $D(S_i, I_i, O_i)$ with directed edges, the colour of directed edges represent the corresponding input; else if there exist $O_i \in \Delta_Q$ and we can not make sure whether $D(S_i, I_i, O_i)$ is z steps deterministic, we check it in the next round.

According the construction process, we have the definition of directed graph.

Algorithm 1 Algorithm to construct the directed graph of BCNs

Input: The algebraic forms of BCN

Output: The directed graph of BCN

```

1: Define a int variable  $k = 0$ , to represent the number of
   states in the nodes.
2: Define a boolean type variable  $Ob = 0$ , to represent the
   online observability of BCN.
3: while buildnode( $k$ )==1 do
4:    $k = k + 1$ 
5:   if  $k == 2$  then
6:     We make  $\Delta_M$  as the suitable inputs set for every
     node with  $k$  states.
7:   else
8:     if  $k > 2$  then
9:       We use two nodes with  $(k - 1)$  states and a node
       with two states to find the suitable inputs set for
       every node with  $k$  states.
10:    end if
11:  end if
12:  for every node with  $k$  states do
13:    for every input in the suitable input set of the nodes
    do
14:      Check whether this input can make the nodes  $z$ 
      ( $z \geq 0$ ) steps deterministic, and build edges for
      this node.
15:    end for
16:  end for
17:  if there exist one node without any edge connect it with
  other nodes then
18:     $Ob = 0$  return  $Ob$ 
19:  end if
20: end while
21:  $Ob = 1$  return  $Ob$ 

```

Definition 12 (Directed Graph): For every node S_j in the directed graph, there exists $k \geq 0$ implies S_j is k steps deterministic; for every distinct two $s_a, s_b \in S_i$ we have $Hs_a = Hs_b$; if $|S_i| = 1$ there are not edge from it to other nodes, else if there are exist one edge I' from it to one nodes then there exist p ($p \geq 1$) edges contain I' from it to nodes S_1, \dots, S_p that

$$|S_j| = |S_1| + \dots + |S_p|$$

and

$$D(S_j, I', \varepsilon) = S_1 \vee \dots \vee S_p.$$

We can also use the directed graph to determine the existed second and fourth kinds of observability. When we trying to build bottom layer and penultimate layer of the directed graph, and if there are exist some nodes in penultimate layer has no edges from it to other nodes, then this BCN is not satisfied existed second observability. When we trying to build edges for every layer, and is there exist one node whose right inputs set is not Δ_M , then this BCN is not satisfied existed fourth observability.

Algorithm 2 Algorithm to build nodes with k states

Input: The number of states in the nodes p

Output: The nodes with p states

```

1: int buildnode(int p)
2: {
3:    $p = p + 1$ 
4:   Try to build all the nodes with  $p$  states whose corre-
   sponding outputs are the same, and classify them by their
   corresponding outputs.
5:   if Failed to build then
6:     return 0;
7:   else
8:     Sort the states inside the nodes from small to large, and
     then sort the nodes based on the values inside the nodes.
9:     return 1;
10:  end if
11: }

```

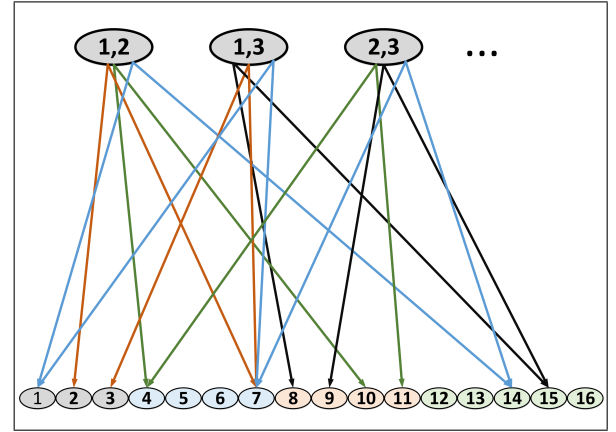


Fig. 4. Part of the directed graph which represents $\{\delta_{16}^1, \delta_{16}^2\}$, $\{\delta_{16}^1, \delta_{16}^3\}$ and $\{\delta_{16}^2, \delta_{16}^3\}$. The green, black, orange, blue edges show the inputs δ_4^1 , δ_4^2 and δ_4^3 respectively.

C. Complexity Analysis

The way use the directed graph to determine the online observability of BCNs is better than by supertree, so we analyze the complexity of it. We classify the states with their corresponding output and form the set of states with the same corresponding output: S_1, S_2, \dots, S_M .

Firstly, we need to calculate the upper bound of the number of the states in a directed graph nodes k , then :

$$k_{upb} = \text{Max}(|S_1|, |S_2|, \dots, |S_M|) \quad (13)$$

Secondly, the number of each nodes with k states:

$$k_{non} = C_{|S_i|}^k + \dots + C_{|S_p|}^k \quad (14)$$

where $|S_i|, \dots, |S_p| \geq k$.

And then the the cardinal number of suitable inputs set of each node, and the time used to check each input is a right input for a node. Finally, calculate the complexity by layer by layer. But the cardinal number of suitable inputs set of a node

depends on the number of the states in it, and the other three nodes which used to find the suitable inputs set for it. And the time used to check an input is a right input for a node of directed graph also depends on the update rules of the *BCNs*. So it is hard to give a accurate complexity of the algorithm without the complete information of *BCNs*.

We can use other three nodes like $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6\}$, $\{\delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$ and $\{\delta_{16}^4, \delta_{16}^7\}$ which is z ($z \geq 0$) steps deterministic to find suitable inputs set for $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$. Because only the input which make the subset of $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$ z steps deterministic will make the $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$ z steps deterministic, and use the this three nodes will be a convenient way to cover all the subset of $\{\delta_{16}^4, \delta_{16}^5, \delta_{16}^6, \delta_{16}^7\}$ which with cardinal number 2. By this way we can reduce the cardinal number of suitable inputs set for every nodes with more than 2 states, then reduce the time cost.

V. APPLICATIONS

If the *BCNs* we research is online observable, and we have built the directed graph for them, then we can use the online observability to determine the initial states of these *BCNs*. Using the the online observability, it will need less observation costs to determine the initial state of this *BCN*. Furthermore we can also use it to try to find the shortest path or avoid entering critical states when we determine the initial state of *BCNs*.

A. Determine the Initial State

After we build the directed graph of the *BCN*, we can use it to determine the initial state of *BCNs* as Fig.5 shows.

- Firstly, we observe the output of the *BCN* mentioned before. If we observe the output is δ_4^1 then we can infer that the possible states set should be $\{\delta_{16}^1, \delta_{16}^2, \delta_{16}^3\}$, and we record them as initial states and current states in the table.
- After that we input δ_4^1 and observe the output is δ_4^3 then we can infer that the possible states set should be $\{\delta_{16}^{10}, \delta_{16}^{11}\}$, and we record them as current states set in the corresponding position.
- Repeat the second step untill the cardinal number of the possible states set turns into 1. In that time we can determine the current state and the corresponding initial state of the *BCN*.

B. Less Observation Costs

Some biological systems, such as the immune systems which can be depicted as the *BCN* T-cell receptor kinetics model [18]. There exist 3 input-nodes, and 37 state-nodes in this model, therefore the model has 2^3 inputs and 2^{37} states. For the purpose of obtain the initial state of the *BCN*, we must select some state-nodes to be observe at first.

If we use the online observability of *BCNs* to determine the initial state of the *BCN* T-cell receptor kinetics model, it needs less observation costs. Because compare with the existing third and fourth kind of observabilities the online observability need

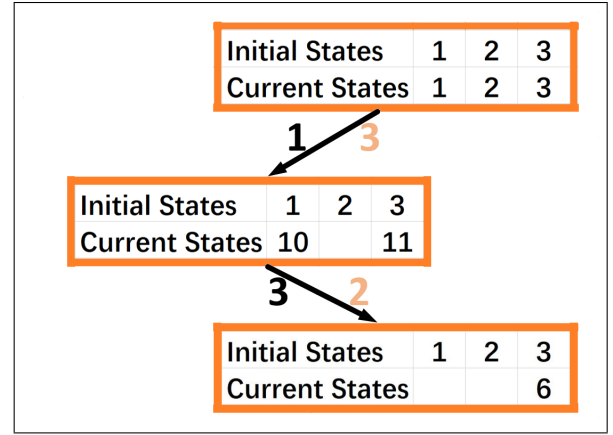


Fig. 5. The process of determining the initial state of BCNs, we change the values of current states by input and the output we observe.

weakest preconditions to determine the initial state of *BCNs* without any presupposition. In other words, the *BCNs* needs less output-nodes to determine its initial state. As we can select less state-nodes to be observe, it will takes less observation costs in this example. In addition to this, there are also some other advantages, when we use the online observability of *BCNs*.

C. Find Shortest Path

When we need to determine the initial state of a *BCN*, an important aspect that we will consider is to find the shortest path. In general, we can't find the shortest path definitely. Fortunately, we can use the directed graph to make the best decision. For the path, we introduce two functions $\text{Pe}(S, I_z)$ and $\text{Pv}(S, I_z)$ to describe its expected value and variance, respectively. For better explain our idea, we list some definitions in the following.

Some statements before definitions:

- $\{I_1, I_2, \dots, I_p\}$: the right inputs set of S ;
- $\{S_z^1, S_z^2, \dots, S_z^k\}$: the set of state sets, and its elements corresponding to the possible outputs $\{O_1, O_2, \dots, O_k\}$ after input I_z , for every I_z in $\{I_1, I_2, \dots, I_p\}$.

Definition 13 ($\text{Spe}(S)$):

$$\text{Spe}(S) = \min(\text{Pe}(S, I_1), \text{Pe}(S, I_2), \dots, \text{Pe}(S, I_p)).$$

Definition 14 ($\text{Pe}(S, I_z)$): When the $|S| = 1$, we have that $\text{Pe}(S, I_z) = 0$ for every I in Δ_M . According **Definition 13**, $\text{Spe}(S) = 0$ if $|S| = 1$. But when the $|S| > 1$, we have that

$$\text{Pe}(S, I_z) = 1 + \frac{\sum_{j=1}^k \text{Spe}(S_z^j) |S_z^j|}{|S|}$$

The function shortest path expected value $\text{Spe}(S)$ is to find the I_z from $\{I_1, I_2, \dots, I_p\}$ to calculat least $\text{Pe}(S, I_z)$ for S . From the definition of $\text{Pe}(S, I_z)$ we can know that if $|S| = 1$ then we can make sure the state of *BCNs*, and we need not input anymore to determine the state of *BCNs*. Therefore, for any input the path expected value $\text{Pe}(S, I_z)$ would be 0 and so do the shortest path expected value $\text{Spe}(S_i)$. But if $|S| > 1$

we still need more input and observe the output. Only by this way we can determine the state of $BCNs$. Therefore, we can recursively define the $\text{Pe}(S, I_z)$ and $\text{Spe}(S)$ for each input I_z in the right inputs set. The $\text{Pv}(S, I_z)$ can be defined in the similar way. Hence we omit the details in this paper.

D. Avoid Entering Critical States

In biological systems, some states of the genes may corresponding to unfavorable or dangerous situations [19]. So another important aspect that we will consider is to avoid entering critical states. We can also construct two functions $\text{Ce}(S, I_z)$ and $\text{Cv}(S, I_z)$ to describe expected value and variance of the times of entering critical states, the definition of $\text{Ce}(S, I_z)$ is as follows:

Definition 15 ($\text{Lce}(S, I_z)$):

$$\text{Lce}(S) = \min(\text{Ce}(S, I_1), \text{Ce}(S, I_2), \dots, \text{Ce}(S, I_p))$$

Definition 16 ($\text{Ce}(S, I_z)$): When the $|S| = 1$, and for every I_z in Δ_M , we have:

$$\text{Ce}(S, I_z) = |S \cap S_{cri}|$$

According **Definition 15**,

$$\text{Lce}(S) = \text{Ce}(S, I_z) = |S \cap S_{cri}|$$

But when the $|S| > 1$ we have that

$$\text{Ce}(S, I_z) = |S \cap S_{cri}| + \frac{\sum_{j=1}^k \text{Lce}(S_z^j) |S_z^j|}{|S_z|}$$

Where S_{cri} is the critical states set of the BCN we research. The definition of $\text{Ce}(S, I_z)$ has some difference with $\text{Pe}(S, I_z)$, because we need to use the critical states set S_{cri} . So that we can analyze the possibility of entering the critical states when we infer the possible states set of $BCNs$.

We can get the definitions of $\text{Cv}(S, I_z)$ in similar ways, and use all of these four functions to make the best decision we like. When we choose an input I_z with least $\text{Pe}(S, I_z)$, we may find the shortest path to determine the initial state. But the output of $BCNs$ after input I_i is uncertain, so selecting the least $\text{Pe}(S, I_z)$ may leads to a very long path. For better performance, we can also use the $\text{Pv}(S, I_z)$ to avoiding risks. The uses of $\text{Ce}(S, I_z)$ and $\text{Cv}(S, I_z)$ are similar to $\text{Pe}(S, I_z)$ and $\text{Pv}(S, I_z)$, when we trying to avoid entering critical states of $BCNs$.

In the four existing kinds of observability, we can not analyze the state of $BCNs$ dynamically, so it would be hard to find the best way we like when we determine the initial state of $BCNs$. But this problem can be solved by the online observability of $BCNs$.

VI. CONCLUSIONS

In this paper, firstly we proposed the online observability of $BCNs$ and define its mathematical form. Secondly we use the super tree and directed graph to determine the online observability. After introduced the ways to determine the

online observability we present some applications of the online observability of $BCNs$ and talk about some advantages of it.

But even we use the directed graph, it is still hard to determine the the online observability of a BCN with a large number of nodes. Therefore, in the future we will try to separate the $BCNs$, and then determine their online observability respectively. Furthermore, we also want to try to use some knowledge about formal methods to earn scalability for $BCNs$. In addition to the theoretical aspect, the realistic application is also very important. Hence we will also try to find some realistic example which can be modeled by $BCNs$. So that we can research these realistic examples well and determine the online observability their models for better performance.

REFERENCES

- [1] Waldrop and M. Mitchell, *Complexity : the emerging science at the edge of order and chaos*. Simon & Schuster, 1992.
- [2] D. Cheng and H. Qi, "Controllability and observability of boolean control networks," *Automatica*, vol. 45, no. 7, pp. 1659–1667, 2009.
- [3] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets [j]," *Journal of Theoretical Biology*, vol. 22, no. 3, pp. 437–467, 1968.
- [4] E. Fornasini and M. E. Valcher, "Observability, reconstructibility and state observers of boolean control networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 6, pp. 1390–1401, 2013.
- [5] D. G. Green, T. G. Leishman, and S. Sadedin, "The emergence of social consensus in boolean networks," in *Artificial Life, 2007. ALIFE '07. IEEE Symposium on*, 2007, pp. 402–408.
- [6] Y. Lou and Y. Hong, "Multi-agent decision in boolean networks with private information and switching interconnection," in *Chinese Control Conference*, 2010, pp. 4530–4535.
- [7] I. Shmulevich, E. R. Dougherty, and W. Zhang, "From boolean to probabilistic boolean networks as models of genetic regulatory networks," *Proceedings of the IEEE*, vol. 90, no. 11, pp. 1778–1792, 2002.
- [8] T. Akutsu, S. Miyano, and S. Kuhara, "Inferring qualitative relations in genetic networks and metabolic pathways," *Bioinformatics*, vol. 16, no. 8, pp. 727–734, 2000.
- [9] A. Faur, A. Naldi, C. Chaouiya, and D. Thieffry, "Dynamical analysis of a generic boolean model for the control of the mammalian cell cycle," *Bioinformatics*, vol. 22, no. 14, pp. e124–e131, 2006.
- [10] T. Ideker, T. Galitski, and L. Hood, "A new approach to decoding life: systems biology," *Annu Rev Genomics Hum Genet*, vol. 2, no. 1, pp. 343–372, 2001.
- [11] T. Akutsu, M. Hayashida, W. K. Ching, and M. K. Ng, "Control of boolean networks: hardness results and algorithms for tree structured networks," *Journal of Theoretical Biology*, vol. 244, no. 4, p. 670, 2007.
- [12] Y. Zhao, H. Qi, and D. Cheng, "Input-state incidence matrix of boolean control networks and its applications," *Systems & Control Letters*, vol. 59, no. 12, pp. 767–774, 2010.
- [13] D. Cheng, H. Qi, and Z. Li, "Identification of boolean control networks," *Automatica*, vol. 47, no. 4, pp. 702–710, 2011.
- [14] —, *Analysis and Control of Boolean Networks*. Springer London, 2011.
- [15] R. Li, M. Yang, and T. Chu, "Controllability and observability of boolean networks arising from biology," *Chaos An Interdisciplinary Journal of Nonlinear Science*, vol. 25, no. 2, pp. 1778–1792, 2015.
- [16] K. Zhang and L. Zhang, "Observability of boolean control networks: A unified approach based on finite automata," *IEEE Transactions on Automatic Control*, vol. 61, no. 9, pp. 2733–2738, 2016.
- [17] D. Cheng, "Semi-tensor product of matrices and its applications - a survey," *Methods & Applications of Analysis*, vol. 321, no. 2, pp. 61–79, 2003.
- [18] S. Klamt, J. Saez-Rodriguez, J. A. Lindquist, L. Simeoni, and E. D. Gilles, "A methodology for the structural and functional analysis of signaling and regulatory networks," *Bmc Bioinformatics*, vol. 7, no. 1, p. 56, 2006.
- [19] Z. Q. Li and J. L. Song, "Controllability of boolean control networks avoiding states set," *Science China(Information Sciences)*, vol. 57, no. 3, pp. 1–13, 2014.