31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Combining Boolean Method with Delay Times for Determining Behaviors of Biological Networks

SuPing Lyu

Abstract— **Boolean functions have been used to analyze the molecular networks of cells. For example, $A \rightarrow B$ represents if A becomes active B will be activated. This method is effective for qualitatively analyzing networks but is not suitable for studies of kinetic behaviors of networks. In the present paper, a dynamic Boolean method was developed by combining Boolean operations with molecular interaction parameters (delay or response times). The Boolean operations characterize the discrete interactions among biological components. The delay times describe the quantitative kinetics. The combination of the two characterizes the discrete biological interactions of networks. For example, $A_{t_{A\to B}}B$ represents that if A becomes active B will be activated after an activation time $t_{A\to B}$. By using this dynamic logic method, we achieved the following results: we proved the general theorems to determine bistable states and oscillation behaviors of networks, we showed that time delays are essential for oscillation behaviors, we proved that single variable networks are either bistable or oscillatory, and we explained why a signal can have multiply responses from different networks. In addition, we analyzed the mitosis cycle of budding yeast cells. We showed that the mitosis cycle is not only robust against structural changes but also robust against fluctuation in kinetic parameters (e.g. delay times).**

I. INTRODUCTION

BIOLOGICAL networks are complicated. This is because there are almost infinite nuclei acids, proteins, and other biological molecules that are building components to construct various networks; these components interact with each other across all the levels and ranges. Quantifying all these components and interactions is not a very realistic goal at the present. Fortunately, it is found that many biological networks have some simple characteristics. First, many biological components take all-or-none states (*ON/OFF* or active/inactive). Secondly, many components change their values when being stimulated and return back to their previous values after their functions are completed. These two features make it possible to use simple Boolean method to study the biological networks [1,2]. However, simple Boolean method cannot be used to study the kinetic properties of networks because it does not have time components. In this paper, we introduced a dynamic logic method by combining time delays with Boolean method. This method allows us to study the kinetic properties of molecular networks without knowing the detailed values of components. This method is a discrete procedure for solving delayed differential equations (DDE).

SuPing Lyu is with Medtronic Corporate Science and Technology, 710 Medtronic Parkway, Minneapolis, MN 55432 USA (phone: 763-505-4549; fax: 763-505-4712; e-mail: suping.lyu@medtronic.com).

II. NOMENCLATURE AND ASSUMPTIONS

The terminology of graph theory is used to describe molecular networks. Protein, Gene, mRNA, or metabolic compounds are called nodes (*A, B*, etc.). Nodes can be at active states (*A, B*…) or inactive states ($\overline{A}, \overline{B}$…). The interactions among nodes are called edges. For example, $A \rightarrow B$ is an edge in which node *A* activates node *B* (Figure 1). We refer to the nodes that stimulate other nodes as "stimulators" and the nodes that respond to stimulations as "responders." Before stimulations, the responders are at "ready states (*ON* or *OFF*)." After the nodes fully respond to stimulation, they are "responding states (*OFF* or *ON*). Ready states can be either *ON* or *OFF* and so are responding states. Response time of an interaction is referred to as the interval from the time when the stimulator starts to act on its responder to the time when the responder reaches its responding state. For example, $t_{A\to B}$ is the response time for *A* to activate *B*. We refer the time within which a node is active as to active pulse width (T_A). Inactive pulse width ($T_{\overline{A}}$) is defined similarly. This paper focuses on systems in which if *A* directly interacts with *B*, T_A needs to be greater than $t_{A\to B}$ in order to activate B.

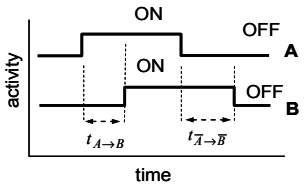

Figure 1. The states and interactions of molecular network $A \rightarrow B$. Both nodes *A* and *B* are assumed to have only ON and OFF states. Therefore, their activity curves have step shapes.

We assume network components (nodes) take discrete states such as active states and inactive states. We also assume that a component will automatically return to its ready state from its responding state within a response time after the stimulation stops. These two assumptions define the applicable ranges of the present dynamic logic method.

III. DYNAMIC LOGIC REPRESENTATIONS

A. Dynamic logic sequence and time sequence

A dynamic logic sequence of a system is defined as a sequence composed of stimulator states, response times, and responder states. For the system $A \rightarrow B$, one dynamic logic

sequence is $A_{t_{A\to B}}B$ representing A activates B within an activation time (delay) $t_{A\to B}$. Similarly, $\overline{A}_{t_{\overline{A}\to\overline{B}}}\overline{B}$ represents $\overline{A}$ deactivates B within a time $t_{\overline{A}\to\overline{B}}$. We can also write down a time sequence. For example, for the network of Figure 1 we have a time sequence $\overline{A}\overline{B}_{(t_0)}A\overline{B}_{(t_0+t_{A\to B})}AB_{(t_0+t_A)}\overline{A}B_{(t_0+t_A+t_{\overline{A}\to\overline{B}})}\overline{A}\overline{B}$. In this sequence, A and B initially are inactive. A becomes active at time t_0. B is activated then. After a period of $t_A + t_{\overline{A}\to\overline{B}}$ both return inactive (we assumed $t_A > t_{A\to B}$).

There are rules for logic sequence: (1) a dynamic logic sequence of a network contains maximum number of its node states; (2) at any given time within a sequence, a node takes only one state (*ON* or *OFF*); (3) the node states in a sequence have to meet the requirements defined by logic relationships. The time distances between stimulator states and the responder state are equal to the response times; (4) if a network system has more than one sequence, these sequences are written down in an addition format.

B. Algorithm

One of the simplest system is $A \to B$. The logic values of this network, $V(A \to B)$, are

$$V(A \to B) = A_{t_{A\to B}}B + \overline{A}_{t_{\overline{A}\to\overline{B}}}\overline{B} \tag{1}$$

Where, "+"means the system can take either one of the two values depending on the external conditions. A second simple network is a system composed of A and B with a deactivation relationship, $A-|B$. The value of this system can be obtained in a similar way. It is $V(A-|B) = A_{t_{A\to\overline{B}}}\overline{B} + \overline{A}_{t_{\overline{A}\to B}}B$. We can confirm that the logic relationship "$A-|B$" is equivalent to "$\overline{A}\to B$". Activation and deactivation are relative concept. That is,

$$(A-|B) \leftrightarrow (\overline{A}\to B) \tag{2}$$

Complicated structures can be decomposed into the above simple ones. We use product operation to represent decomposition. For example, $A\to B\to C$ can be decomposed into $(A\to B)(B\to C)$. The values are

$$V(A\to B\to C) = A_{t_{A\to B}}B_{t_{B\to C}}C + \overline{A}_{t_{\overline{A}\to\overline{B}}}\overline{B}_{t_{\overline{B}\to\overline{C}}}\overline{C} \tag{3}$$

Similarly, a double deactivation relationship has a value of $V(A \dashv\vdash B) = A_{t_{A\to\overline{B}}}\overline{B}_{t_{\overline{B}\to A}}A + \overline{A}_{t_{\overline{A}\to B}}B_{t_{B\to\overline{A}}}\overline{A}$. This system has two stable values (bistable). Similarly, a double activation network has $V(A \rightleftarrows B) = A_{t_{A\to B}}B_{t_{B\to A}}A + \overline{A}_{t_{\overline{A}\to\overline{B}}}\overline{B}_{t_{\overline{B}\to\overline{A}}}\overline{A}$.

A network with mixed deactivation and activation is interesting; it has oscillation behavior:

$$V(A \dashv\rightarrow B) = A_{t_{A\to B}}B_{t_{B\to\overline{A}}}\overline{A}_{t_{\overline{A}\to\overline{B}}}\overline{B}_{t_{\overline{B}\to A}}A... \tag{4}$$

The components of the system oscillate in the order of $AB\overline{A}\overline{B}A$ with a period equal to $T = t_{A\to B} + t_{B\to\overline{A}} + t_{\overline{A}\to\overline{B}} + t_{\overline{B}\to A}$. If delay from A to $\overline{A}$ is zero, then A and $\overline{A}$ have to co-exist and so do B and $\overline{B}$, which is impossible and oscillation does not occur. Therefore, delay time between stimulation and response is essential for oscillation.

C. Merging operation

If two directly interacting components have an activation relationship, they tend to have same states (both *ON* or *OFF* with delay times). But if they have a deactivation interaction, they tend to have different states, one takes *ON* and the other *OFF*. If two components have both activation and deactivation relationships, the system oscillates instead of being bistable. These observations can be used to simplify large systems. The procedure is referred to as merging operation; consecutive nodes with activation interactions can be merged into one group with all its nodes having a same state. For example, $A\to B\to C\to D$ can be merged into [*ABCD*], where [ ] is used to indicate that the nodes inside have activation relationships. This group of nodes is referred to as activation group node. The response times between nodes within the group can be calculated by adding all the response times between them. For examples, we have $t_{A\to D} = t_{A\to B} + t_{B\to C} + t_{C\to D}$ and $t_{\overline{A}\to\overline{D}} = t_{\overline{A}\to\overline{B}} + t_{\overline{B}\to\overline{C}} + t_{\overline{C}\to\overline{D}}$. The network $A\to B\to C\to D$ can also be merged into $A\to [BCD]$.

D. Theorems for single variable systems

A single variable network is one in which all the components (nodes) have no more than one stimulator.

Theorem 1: if a single variable network is a tree, the network has two states [3].

Theorem 2: if a single variable network is a cycle, it has two stable states as long as the number of the deactivation relationships is an even number [3].

Theorem 3: if a single variable network is a cycle and has odd number of deactivation relationships, the network does not have stable state. Instead, it oscillates with a characteristic period [3].

E. Multiple variable systems

Multiple variable systems are those in which at least one component has more than one stimulator. A simple multiply variable network contains two stimulators and one responder. The two stimulators can be either activator or suppresser. They can have either *AND* or *OR* relationships. One example is $(A\wedge C)\to B$ representing that B becomes active only after both A and C become active. Another example is $(A\wedge\overline{C})\to B$ representing that B becomes active after A becomes active and C becomes inactive. The values of each network can be calculated [3]. For example:

$$V((A\wedge C)\to B)$$
$$= (A\wedge C)_{t_{A\wedge C\to B}}B + (\overline{A}\wedge C)_{t_{\overline{A}\wedge C\to\overline{B}}}\overline{B} + (A\wedge\overline{C})_{t_{A\wedge\overline{C}\to\overline{B}}}\overline{B} + (\overline{A}\wedge\overline{C})_{t_{\overline{A}\wedge\overline{C}\to\overline{B}}}\overline{B}$$

IV. SIGNALS TRAVEL IN NETWROKS

A. Signals travel in linear chains

A linear chain has n nodes, $N_1, N_2, ...N_n$, each of which stimulates the one next (right) to it. This network is $E \to N_1 \to N_2 \to ... \to N_{n-1} \to N_n$, where E is an external signal. Before time t_0, N_1 is inactive. At t_0, E activates N_1. This

activation signal propagates into the chain. Node N_i responds to the signal after a time $t_{N_1 \to N_j} = \sum t_{N_j \to N_{j+1}}$.

If the external signal turns N_1 off at time T, N_i responds to the signal after a time $t_{\overline{N}_1 \to \overline{N_i}} = \sum t_{\overline{N}_j \to \overline{N}_{j+1}}$. If the external signal is a pulse (width T), each node has a pulse with a width of $T_i = T + \sum (t_{\overline{N}_j \to \overline{N}_{j+1}} - t_{N_j \to N_{j+1}})$. This simply indicates that the pulse width of each node is different. Each node will change the pulse width by an amount of $(t_{\overline{N}_j \to \overline{N}_{j+1}} - t_{N_j \to N_{j+1}})$ that can be either an increase or decrease. If the pulse width is decreased as the signal travels through the chain, the signal is terminated when the active pulse width of a node gets shorter than the activation response time needed for this node to activate the next one. Thus, for a same chemical signal, if it has different pulse width (T), it travels to a different depth of a same chain and results in different biological responses. If the pulse width is increased as the signal travels, the nodes of the chain can have broader active pulse width than the external signal. Then a short external stimulation can result in an elongated response. Therefore, a same type of chemical signal can have more than one response in chains that have different abilities to modify pulse width.

B. Signals travel in bistable loops

A biostable loop is composed of n nodes with the following relationship:

$$E \longrightarrow N_1 \longrightarrow N_2 \longrightarrow \cdots \longrightarrow N_{n-1} \longrightarrow N_n$$

Where, N_{i-1} activates N_i. If $E \vee N_n \to N_1$, a pulsatile activation signal can activate the entire network permanently if the pulse width of each node increases. But if the pulse width of some nodes decreases, the external signal can terminate. This is similar to the situation of signal traveling in linear structures. This simply suggests that different cycles react to a same external pulse signal differently. The cycle can be either entirely activated or be activated for a short period of time and return to inactive state. Systems where N_n and E have *AND* relationship can be analyzed in a similar way.

C. Signals travel in oscillation loops

Signal traveling in oscillation loop can be analyzed in the same way as for bistable loops. The difference is that the states of the nodes of oscillation loops change between active and inactive alternatively while it stays the same in a bistable cycle.

D. Single variable systems are either stable or oscillatory

Systems with their nodes having no more than one stimulator are either stable or oscillatory. There is no chaotic behavior. This can be proved [3].

V. GENERAL PROCEDURE

A system that has many nodes may be analyzed by using the above multiplication procedure. It also can be analyzed with a general method that is analogous to numerical methods for ordinary differential equations.

For a system composed of n nodes, N_1 to N_n, the first step is to identify state conditions of all the nodes based on the available experimental observations. For example, if it is observed that N_l and N_m jointly activate (*AND*) N_j, the active and inactive condition of N_j are

$$t_{N_l \wedge N_m \to N_j} N_j \tag{5}$$

$$t_{\overline{N}_l \wedge N_m \to \overline{N}_j} \overline{N}_j + t_{N_l \wedge \overline{N}_m \to \overline{N}_j} \overline{N}_j + t_{\overline{N}_l \wedge \overline{N}_m \to \overline{N}_j} \overline{N}_j \tag{6}$$

Where "+" in Equation 6 indicates that there are three deactivation conditions, each of them can lead to N_j to become inactive. A second step is to choose initial states for all the nodes and use them as the initial point for a time sequence of the system (initial condition). A third step is to determine the time sequence of the system based on the initial condition and the state conditions (Equations 5 & 6) of all the nodes. This is done by increasing time by a small step δt (< all the response times), checking all the nodes and determining whether their state conditions are met so they can be activated or deactivated. If any node is activated (deactivated), append the time and new state of the node to the time sequence. Update the node states and continue this operation until the sequence become stable or oscillating over time. This stable or oscillating sequence is a value of the system. A fourth step is to repeat third step with different initial conditions to find out all the other values of the system.

VI. EXAMPLE: BUDDING YEAST MITOSIS

Budding yeast cell mitosis circle is one of the most studied systems [4]. A number of proteins are involved in the cycle. Li et al. studied this system and found that the cell cycle was inherently robust against deleting certain edges. However, because quantitative response times were not available, it is not clear whether the robustness is also true against variations of response times. Here we will use the present dynamic logic method to demonstrate that this network is also robust against variations in response times.

We adopted the logic structures of cell cycle from reference [4] and plotted in Figure 2. Because there is no response time reported, we randomly set them being close to each other (e.g. the activation times of the nodes *cdh1, Mcm/SFF, cdc20,* and *cdc14* were set to 0.9 and that of the rest of nodes were set to 1. The deactivation times of the nodes *cdc20* and *swi5* were set to 0.9 and that of the rest of the nodes were set to 1). We determined the state conditions of all the nodes [3]. For example, node *SBF* is controlled by nodes *Cln3* and *Clb1,2* through a logic relationship, $C\ln 3 \wedge \overline{Clb1,2} \to SBF$. The state conditions of this node are $t_{C\ln 3 \wedge \overline{Clb1,2} \to SBF} SBF$, $t_{\overline{C\ln 3} \wedge \overline{Clb1,2} \to \overline{SBF}} \overline{SBF}$, $t_{Clb1,2 \wedge Clb1,2 \to \overline{SBF}} \overline{SBF}$, and $t_{\overline{Clb1,2} \wedge Clb1,2 \to \overline{SBF}} \overline{SBF}$. The state conditions for all the other components can be determined in the same way. Each

node has two parameters. One records its current *ON/OFF* state. The other records when it has its most recent state change.

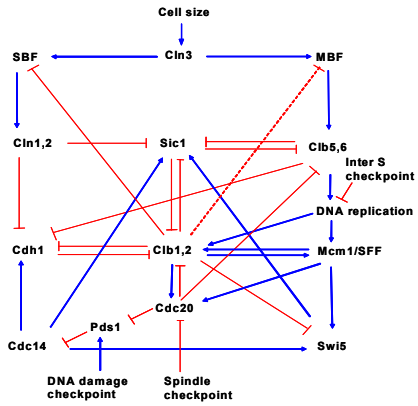

Figure 2. Simplified molecular network of yeast cell mitosis cycle adapted from reference [4]. Blue arrows indicate activation and red lines (blunt) indicate deactivation. The Detailed Boolean relationships can be found in reference [3].

Next, we set all the nodes to be active except the three nodes that represent the checkpoints (*Inter S, DNA*, and *Spindle* checkpoints). We used this set as an initial value for a time sequence of the system. Set run time to be zero.

Then, we determined the evolution of the time sequence of the system based on the initial condition (all nodes being active) and the state conditions of all the nodes. This was done by increasing time (e.g. a step of 0.07), checking all the nodes and determining whether their state conditions were met so they could be activated or deactivated. If any node was activated (or deactivated), update the state and state change time of that node and append that node to the time sequence of the system. The process was repeated until the sequence became oscillating over time. This oscillating sequence is show in Figure 3. As expected, the sequence represents a change from G_1, to S, to G_2, and to M phases. These phases were determined based on reference [4]. This calculation was repeated for 2^{13}-1 times with different initial condition (same delay times). Same mitosis cycle was obtained (although detailed time sequences might change).

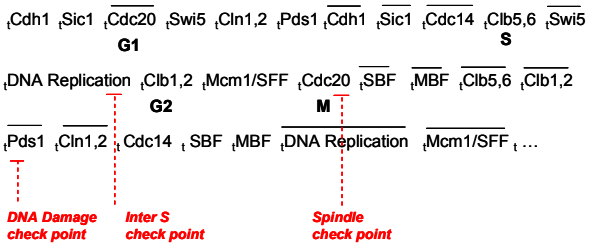

Figure 3. Calculated mitosis cycle of yeast cells. G_1, S, G_2, and M represent the four phases of mitosis [4].

We randomly changed the response times (one each time) and calculated the time sequence of the system (the initial condition was that all the nodes were active). We observed that changes in most of the response times by two orders of magnitude (e.g. *0.01* to *1*) did not disturb the cell cycle from G_1 through M (although detailed time sequences might change). A few response times had narrower ranges (0.5 – 1), such as the activation times of the nodes *Cdc20* and S*wi5* and deactivation times of the nodes *Mcm/SFF*, *Pds1*, and *Cdc20*. This new observation suggests that the mitosis cycle of yeast cells is robust against variation of interaction response times as well.

We also observed that disconnecting some edges did not disturb this cell cycle (e.g. *cdc20* —| *clb5,6* and *Mcm1/SFF* → *cdc20*). However, the cycle stops if disconnecting the edge from *DNA Replication* to *Clb1,2*. This observed sensitivity is consistent with the experimental observations: cell does not move into the next phase if its *DNA replication* is not completed. Also, it was observed the cycle was stopped if any of the three check points was active.

VII. DISCUSSIONS

The application of logic method in biological networks has been promoted by Kauffman [1] and Thomas et al [2]. The previous method was based on truth tables [2] and can be used to determine the steady state behaviors of molecular networks. The present method combines logic methods and delay times of node interactions. It allows us to determine the kinetic sequences and time sequences that characterize the steady-state and kinetic properties of the networks. Because of this, the present method can be used to analyze networks at different levels of details depending on the availability of quantitative kinetic parameters. This is an advantage of the present method over deterministic methods (e.g. ODEs or delayed DEs) because the latter require detailed kinetic parameters. It is also an advantage over the previous logic methods that are not suitable for studying kinetic properties. If a complicated network system can be decomposed into modules and motifs, these modules or motifs may be readily studied with the above methods. These advantages of the present dynamic method may help to fill in the gap between the need and availability of quantitative parameters of biological interactions. In a certain sense, the present method can be used to obtain overall pictures of molecular networks before doing detailed studies.

ACKNOWLEDGEMENTS

I like to thank Darrel Untereker for discussion and criticisms.

REFERENCES

[1] L. Glass and S. A. Kauffman. *J. Theor. Biol*., 39: 219-237, 1972.
[2] R. Thomas and R. D'Ari, *Biological Feedback* (CRC Press, Boca Raton, 1990).
[3] S. P. Lyu, arXiv 0806:1733, 2008.
[4] F. Li, Y. Lu, Q. Ouyang, and C. Tang. *Proc. Natl. Acad. Sci. USA*, 101: 4781-4786, 2004.