Effects of feedback and feedforward loops on dynamics of transcriptional regulatory model networks

Chikoo Oosawa[1,], Kazuhiro Takemoto[2], Michael A. Savageau[3]

[1]Department of Bioscience and Bioinformatics, Kyushu Institute of Technology
Iizuka, Fukuoka 820-8502, Japan
[2]Bioinformatics Center, Institute for Chemical Research, Kyoto University
Gokasho, Uji, Kyoto 611-0011, Japan
[3]Department of Biomedical Engineering, College of Engineering
University of California, Davis, 95616, U.S.A.

November 17, 2007

Abstract

We demonstrate the advantages of feedforward loops using a Boolean network, which is one of the discrete dynamical models for transcriptional regulatory networks. After comparing the dynamical behaviors of network embedded feedback and feedforward loops, we found that feedforward loops can provide higher temporal order (coherence) with lower entropy (randomness) in a temporal program of gene expression. In addition, complexity of the state space that increases with longer length of attractors and greater number of attractors is also reduced for networks with more feedforward loops. Feedback loops show opposite effects on dynamics of the networks. These results suggest that feedforward loops are one of the favorable local structures in biomolecular and neuronal networks.

Keywords : Boolean networks; feedback loop; feedforward loop; mutual information; entropy; transcriptional regulatory networks

1 Introduction

Recent studies of natural complex networks [17] including transcriptional regulatory networks in cells have revealed at least three stastistical properties: long-tailed global connectivity distributions having a small number of highly connected nodes; small-world properties of short path lengths between any two nodes; and highly clustered connections among adjacent nodes [2, 7, 8, 9, 18, 20, 22, 24]. The last local structures called motifs consist of a few nodes and edges among the nodes which are found to be statistically significant, and can be regarded as functional modules [18, 20, 22]. Since feedback and feedforward loops are motif structures as well as basic and ubiquitous circuits in man-made systems, one can expect that transcriptional regulatory networks also have both feedback and feedforward loops; however, only the feedforward loops prevail [18, 20, 22]. Other biological networks such as signal transduction and neuronal networks also have similar tendencies, which suggests that feedforward loops are favored in complex biological networks. In general, although massive available network data demonstrate the statistical significance, it is unclear why feedforward loops are advantageous over feedback loops in dynamical systems.

2 Model and method

2.1 Boolean network

The dynamics of the Boolean networks [6, 21] is determined by the equation

$$X_i(t+1) = B_i\left[\boldsymbol{X}(t)\right] \quad (i = 1, 2, ..., N), \qquad (1)$$

where $\boldsymbol{X}(t)$ is a binary state, either 0 or 1, of node i at time t, $B_i(\ )$ are Boolean functions [See Table 1] used to simultaneously update the state of node i, and $\boldsymbol{X}(t)$ is a binary vector that gives the states of the

Corresponding author: chikoo@bio.kyutech.ac.jp

N nodes in the network. After assigning the initial states $\boldsymbol{X}(0)$ to the nodes, the successive states of the nodes are updated by input states and their Boolean functions. The dynamical behavior of these networks is represented by a time series of binary states. The time course follows a transient phase from an initial state until a periodic pattern, called an attractor, is established [See Fig. 3].

Table 1: 4 of 16 Boolean functions with indegree K_{in} = 2. In this paper, we used Boolean functions shown below, because of the feasibility of computation and the biological meaning of the functions [14, 25].

Inputs		Output			
0	0	0	0	0	1
0	1	0	0	1	0
1	0	0	1	0	0
1	1	1	0	0	0

2.2   Numerical condition

To investigate the effects of feedback and feedforward loops on the dyamics Boolean networks, we constructed many networks with varying numbers of independent feedback or feedforward loops, where both loops consisted of three nodes and three directed edges [See Fig. 1 and Table 2]. After embedding the specified number of loops, the rest of the directed edges were assigned at random.

In total, we constructed 9 10^4 Boolean networks [See Table 2] with fixed a network size. We applied 2 10^3 initial states to each network. Four different Boolean functions [See Table 1] were used in the same frequency [See Table 2].
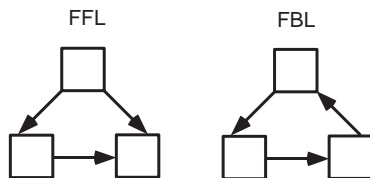
FFL   FBL

Figure 1: Two embedded structures: FFL and FBL stand for feedforward loop and feedback loop, respectively. The loops consist of 3 nodes (squares) and 3 directed edges.

Table 2: Numerical condition: All networks consist of the same amount of network resources, nodes, directed edges, and the number of Boolean functions. The difference among the generated networks lies in the style of the connections.

Size of networks N	128 nodes
Connectivity	For all nodes, $K_{in} = K_{out} = 2$
Boolean function	only AND type [See Table 1]
Types of loop	FFL, FBL [See Fig. 1]
Number of loops	0, 10, 20, 30, and 40
Number of edges	256
Number of initial states	2000 per network
Number of realizations	10^4 in each condition

2.3   Entropy and mutual information

We measured the entropy (randomness) and mutual information (coherence) of state variables to characterize the temporal (series) structure of state variables in the Boolean networks [3, 4, 5, 6] [See Fig. 2].
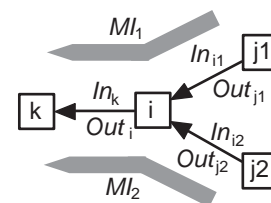

Figure 2: Flow of state variables from upstream nodes to downstream. Since the input connectivity for all nodes is 2 ($=K_{in}$), there are two pathways for mutual information in each node. The input sequence In_{i1} for the node i is the same as the output sequence of an upstream node Out_{j1}, and the output sequence of Out_i for the node i is the same as the input sequence of a downstream node In_k. When node i has multiple output connections, they have the same binary sequence because state variables in networks are subject to Eq. (1).

2.4   Complexity of state space

To characterize the complexity of state space from initial states [See Fig. 3], we use two measures [19]:

1. Sum of length of attractors: Each network may contain different number of attractors and the lengths of the attractors may also vary. The measures define the total length of the attractors in state space.

2. Basin entropy [See Appendix 5.1 for details]:

$$H_{Basin} = \sum_i p(i) \log_2 p(i) \qquad (2)$$

where, $\sum_i p(i) = 1$. The two measures indicate the complexity of state space from initial states. According to the definitions, the larger values of two characteristics sgnify higher complexity of the basin of attraction [See Fig. 3].
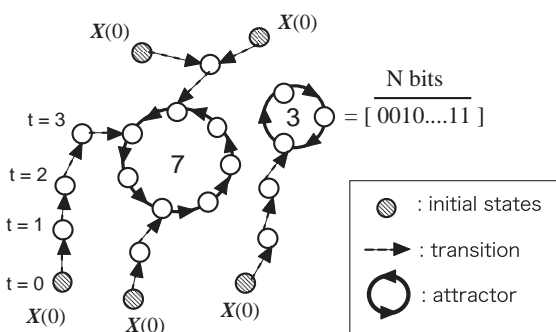

Figure 3: Example of a state space: There are different 2^N states (shown as circles) in the space for each network. The numbers 3 and 7 inside attractors indicate the lengths of the attactors. The numbers also correspond to the size of the basin of attraction.

2.5   Path length

To obtain the structual properties of propagating pathway of the state variables, we measured two properties [4, 10, 11, 12]:

1. Path length, which is the average number directed edges in the shortest path from a node to all reachable nodes.

2. Average path length, which is the average number of the path lengths for all the nodes.

3   Results

In total, we obtained 5479157 attractors from 9 10^4 networks with 1.8 10^8 intial states. The size

of entropy [See Fig. 7], mutual information [See Fig. 8], total length of attractors [See Fig. 6], and basin of entropy [See Fig. 7] are measured from the attractors. We also obtained the dependence of average path length on the number of embedded loops [See Fig. 5].
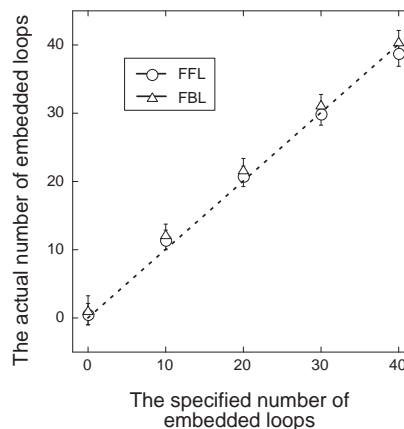

Figure 4: Relationship between the specified number of embedded loops and the actual number of embedded loops. Symbols indicate mean   SD. Dashed line is given by $y = x$.

Figure 4 shows that our successful method for embedding loops in Boolean networks. The average path lengths increases with increasing number of embedded loops [See Fig. 5 and see Appendix 5.2].   Fig-
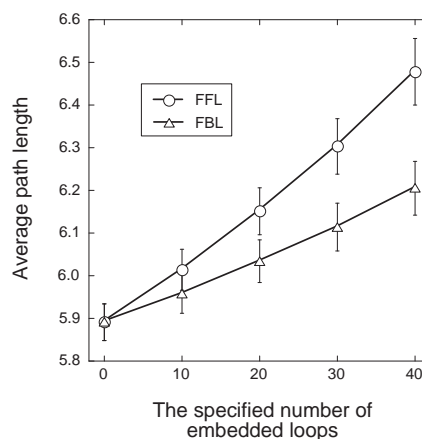

Figure 5: Relationship between the number of embedded loops and average path length. Symbols indicate mean   SD.
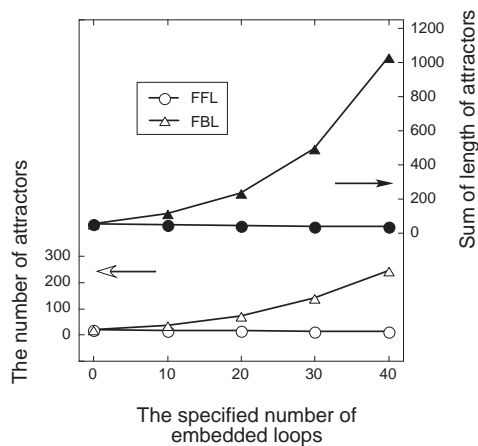
ure 6 shows the complexity of state-space structures.

Figure 6: Relationship between the number of embedded loops and sum of length of attractors (upper), and the number of attractors (bottom). Symbols indicate mean values.

Feedforward and feedback loops have opposite effects. Please note that both the number of attractors and the sum of the lengths of attractors with feedforward loops slightly decrease. Both entropies in Fig. 7 indicate opposite effects on different loop structures. Unlike Figs.
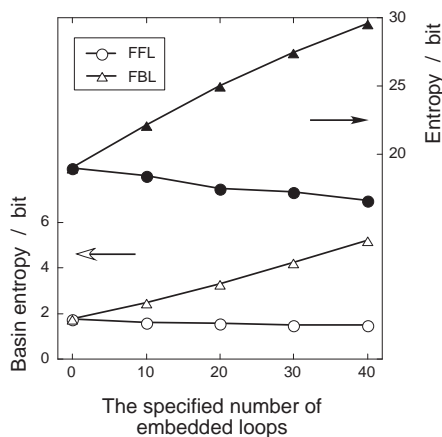

Figure 7: Relationship between the number of embedded loops and size of entropy (upper), and basin entropy (bottom). Symbols indicate mean values.

6 and 7, the size of mutual information increases with increasing number of embedding loops [Fig. 8].
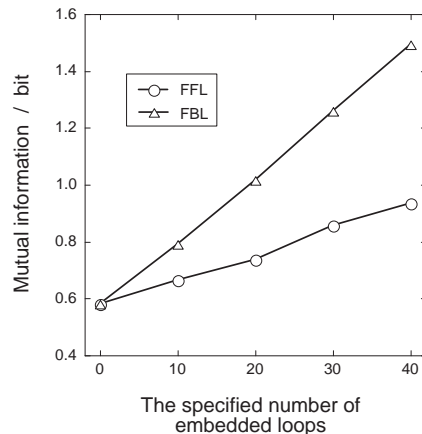

Figure 8: Relationship between the number of embedded loops and the size of mutial information. Symbols indicate mean values.

4 Summary and discussion

We examined the effects of embedding loops on dynamical and topological properties. The networks with more feedback as well as feedforward loops exhibit longer path lengths. As for the dynamical properties, the networks with more feedback loops show a larger number of attractors and greater size of entropy and mutual information, demonstrating that feedback loops increase the complexity of the state space of the networks. In other words, the feedback loops behave as pattern generators of a temporal program of gene expression or entropy generators, and the resultant mutual information (coherence) is driven by the entropy. On the other hand, the networks with more feedforward loops show a smaller number of attractors and size of entropy but larger mutual information. The effects indicate that feedforward loops play a role of stabilizing the state space as well as organizing temporal patterns with less entropy.

The structure of feedback loops resembles that of a repressilator [16], which is a synthetic genetic regulatory network consisting of three genes connected in a feedback loop. The repressilator shows self-sustained oscillations reminiscent of our results from Boolean networks with feedback loops. Other studies [13, 22, 24] demonstrate that feedforward loops can exhibit temporal and spatial order with differential equations. Our results are similar, as shown in Fig. 8.

Control parameters for the dynamics of Boolean networks are input connectivity, K_{in}, the size of network, the bias of Boolean functions, and output con-

nectivity distributions [1, 15, 6]. In this report, we change only the connection style with the same amount of network resources [See Table 2]. Nevertheless, Figs. 6 – 8 demonstrate that the internal connection style may well be regarded as a novel control parameter for the dynamics of Boolean networks.

Our results may provide a blueprint for the design of an artificial regulatory gene network [16, 23], elucidate the role of loop structures in dynamical systems, and provide some insight into the prediction of relationships between complex network structures, behaviors, and functions. Since the currently available biological network data show resultant structures after evolutionary and/or developmental processes, our constructive approach [4, 5, 6, 10, 11, 12] is one of the promising ways for disentangling natural large-scale complex networks.

Acknowledgements

References

[1] A. A. Moreira, L. A. N. Amaral, "Canalizing Kauffman Networks: Nonergodicity and Its Effect on Their Critical Behavior", *Phys. Rev. Lett.*, **94**, 218702, 2005.

[2] A.-L. Barabási, Z. N. Oltvai, "Network biology: Understanding the cell's functional organization", *Nature Gnenetics*, **5**, 101–112, 2004.

[3] C. G. Langton, "Computation at the edge of chaos: Phase transitions and emergent computation", *Physica D*, **42**, 12–37, 1990.

[4] C. Oosawa, "Roles of hubs in Boolean networks", *Proceedings of the 15th IEEE International Workshop on Nonlinear Dynamics of Electrical Systems*, 245–248, 2007.

[5] C. Oosawa, K. Takemoto, S. Matsumoto, M. A. Savageau, "Local cause of coherence in Boolean networks", *Proceedings of the 12th International Symposium on Artiflcial Life and Robotics*, 621–627, 2007.

[6] C. Oosawa, M. A. Savageau, "Effects of alternative connectivity on behavior of randomly constructed Boolean networks", *Physica D*, **170**, 143–161, 2002.

[7] C. Y. Huang, C.-T. Sun, C-.Y. Cheng, J.-L. Hsieh, "Bridge and brick motifs in complex networks", *Physica A*, **377**, 340–350, 2007.

[8] D. J. Watts, S. H. Strogatz, "Collective dynamics of 'small-world' networks", *Nature*, **393**, 440–442, 1998.

[9] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabási, "The Large-scale organization of metabolic networks", *Nature*, **407**, 651–654, 2000.

[10] K. Takemoto, C. Oosawa, "Evolving networks by merging cliques", *Physical Review E*, **72**, 046116, 2005.

[11] K. Takemoto, C. Oosawa, "Modeling for evolving biological networks with scale-free connectivity, hierarchical modularity, and disassortativity", *Mathematical Biosciences*, **208**, 454–468, 2007.

[12] K. Takemoto, C. Oosawa, T. Akutsu, "Structure of n-clique networks embedded in a complex network", *Physica A*, **380**, 665–672, 2007.

[13] K. Klemm, S. Bornholdt, "Topology of biological networks and reliability of information processing", *Proc. Natl. Acad. Sci.*, **102**, 18414–18419, 2005.

[14] L. Raeymakers, "Dynamics of Boolean Networks Controlled by Biologically Meaningful Functions", *J. Theor. Biol.*, **218**, 331–341, 2002.

[15] M. Aldana, P. Cluzel, "A natural class of robust networks", *Proc. Natl. Acad. Sci.*, **100**, 8710–8714, 2003.

[16] M. B. Elowitz, S. Leibler, "A synthetic oscillatory network of transcriptional regulators", *Nature*, **403**, 335–338, 2000.

[17] M. Newman, A.-L. Barabási, D. J. Watts, *The Structure and Dynamics of Networks*, Princeton University Press, 2006.

[18] N. Kashtan, S. Itzkovitz, R. Milo, U. Alon, "Topological generalizations of network motifs", *Phys. Rev. E*, **70**, 031909, 2004.

[19] P. Krawitz, I. Shmulevich, "Basin Entropy in Boolean Ensembles", *Phys. Rev. Lett.*, **98**, 158701, 2007.

[20] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, "Network motifs: simple building blocks of complex networks", *Science*, **298**, 824–827 2002.

[21] S. A. Kauffman, *The Origins of Order*, Oxford University Press, 1993.

[22] S. Ishihara, K. Fujimoto, T. Shibata, "Cross talking of network motifs in gene regulation that generate temporal pulses and spatial stripes", *Gene to Cells*, **10**, 1025–1038, 2005.

[23] T. S. Gardner, C. R. Cantor, J. J. Collins, "Construction of a genetic toggle switch in *Escherichia coli*", *Nature*, **403**, 339–342, 2000.

[24] U. Alon, *An Introduction to Systems Biology*, Chapman & Hall/CRC (2006).

[25] Y. Setty, A. E. Mayo, M. G. Surette, U. Alon, "Detailed map of cis-regulatory input function", *Proc. Natl. Acad. Sci.*, **100**, 7702–7707, 2003.

5 Appendix

5.1 Basin entropy

The definition of basin entropy deals with the degree of partitioning of state space by basin of attractions. For example, even though the same number of attractors may be found by starting from different initial states, the portions of the number of initial states that follow into those attractors can vary from one network to the other. Eq. (2) gives the minimum and maximum entropy when all 2000 different initial states fall into only one attractor and 2000 different attractors, respectively. By the definition, $p(i)$ in Eq. (2) is giving as

$$p(i) = \frac{a_i}{2000} \tag{3}$$

$$\sum_i a_i = 2000 \tag{4}$$

where a_i is the number of initial states that reached the i-th attractor, The number 2000 is obtained the numerical condition [See Table 2].

5.2 Average path length

Since we generated networks with random assignment of edges between nodes under fixed connectivity, $K_{in} = K_{out} = 2$, we can assume a tree-like structure of a nodes in a network (Fig. 9a). With this assumption,

it can be said that the number of downstream nodes increases exponentially, and the relationships between average path length L and network size N can be expressed as follows :

$$\frac{2(2^L - 1)}{2^1 - 1} = N - 1 \tag{5}$$

The average pathlength L can be obtained by transforming Eq. (5),

$$L = \log_2 \left(\frac{N-1}{2} + 1 \right). \tag{6}$$

When the number of embedded loops increases, the average actual number of downstream nodes decreases (Fig. 9b), leading to the reduction of the common ratio in Eq. (5). Using the reduced common ratio x where x takes $1 < x < 2$, Eq. (5) changes to

$$\frac{2(x^L - 1)}{x^1 - 1} = N - 1. \tag{7}$$

Similarly, Eq. (7) can be transformed into

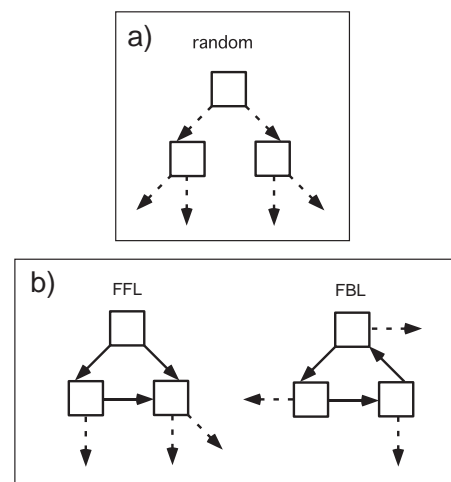$$L = \log_x \left[\frac{(x-1)(N-1)}{2} + 1 \right]. \tag{8}$$

Figure 9: Differences in the number of downstream nodes: Dashed arrows indicate output connection from a node a) and from subgraphs b). a) The number of downstream nodes increases almost exponentially with random assignment of edges. b) With embedding loops, the average number of downstream nodes decreases since the number of edges inside the loops becomes large.