State Analysis of Three-node Network Motifs using Boolean Model

Philipe Chan[1], Rlan Manero[1], Matthew Te[1*] and Angelyn Lao[2]
[1] *Senior High School, De La Salle University, 2401 Taft Avenue, Manila, 0922 Philippines*
[2] *Mathematics & Statistics Department, De La Salle University, 2401 Taft Avenue, Manila, 0922 Philippines*
**Corresponding Author: matthew_christian_te@dlsu.edu.ph*

Abstract: A Gene Regulatory Network (GRN) is a biological system that represents interactions between molecular regulators and their potential targets inferred from gene expression data. GRN can be complex and complicated, which often foregoes the more valuable micro-dynamic behavior occurring between small-node interactions within the network. For this reason, we opt to study three-node boolean network motifs which are simpler and may be optimal indicators of the whole-network behavior. Eleven biologically relevant feedforward and feedback loop GRN motifs are chosen and represented as Boolean network models. Various state analyses such as drawing the state transition graph (STG), locating the shortest path presented as transient time (TT) graphs, and attractor identification are performed. Our results provide a better understanding of the dynamical state of the three-node network motifs.

Key Words: gene regulatory networks; Boolean modeling; three-node network motifs; feedforward loops; feedback loops

1. INTRODUCTION

Gene Regulatory Networks (GRNs) are networks of causal interactions between molecular regulators and gene sets that are depicted as directed graphs inferred from gene expression data coming from mRNA and proteins (Karlebach & Shamir, 2008). These networks, when compounded, form a genomic regulatory sequence that are representative of logic maps of various regulatory functions that are of equivalent to biochemical cascades such as reactions, transformations, upregulations, and downregulations (Davidson & Levin, n.d.). This is why it is important to study GRNs as they provide an inward understanding of the mechanisms which occur in a biological system such as the nature of participating genes, interaction patterns between genes, and other factors which influence gene interactions. In general, GRN's provide unique insights into the overall dynamic behaviour of a network.

The usage of boolean Networks in studying various biological networks has been widely used to predict and locate key components within the biological network or to gain a further understanding of the underlying process (Garg et al., 2009). GRNs contain over-represented subgraphs called network motifs, which help in the analysis and identification of the patterns present in the interactions of the nodes (Nakhleh,2015; Moreno, 2019). Studying the topological property of these motifs and analyzing their dynamics help in understanding their functionality. Many studies have performed various forms of establishing and analyzing boolean models (Huang et al., 2013; Fumia & Martins, n.d.; Barman & Kwon, 2018).

In this study, we focus on three-node boolean networks that are based on the feedforward and feedback network motifs. Simulations are performed on the state transition graph of the identified three-node network motifs using GINsim. We also identify the shortest path for the initial state to reach the stable state, and the attractors of the boolean networks.

2. METHODOLOGY

Eleven biologically-relevant network motifs were generated using GINsim, an open source Java software capable of performing regulatory simulation and analysis. It also allows the user to create specific regulatory models with multileveled logical functions, perform synchronous and asynchronous simulation, mutation-testing, perturbations, and more.

In a network, a state is defined as a unique configuration of information withheld by all nodes; but for boolean models, states are expressed as binary switches of 0 (OFF) and 1 (ON). The finitude of states in a discrete boolean network (2N) implies that trajectories are bound to reoccur after a specified time interval; given that this network has deterministic dynamics, the trajectory will evolve towards a steady/stable state called an attractor (Wuensche, 2016). Weisstein (2002) formally defines an attractor as a set of states (points in the phase space), unchanging under the dynamics, towards which neighbouring states in a given basin of attraction approach in the course of dynamic evolution.

State Transition Graphs (STG) are generated to depict the trajectory of the network which consists of the set of states along the evolution path as the network converges toward the attractor state from the foremost state or initial state.

Finally, the model simulation measures the transient time needed by a node to reach the attractor in the shortest possible path. Transient time graphs are plotted as step functions along a graph separated by 1-second timepoints in GNU Plot. The shortest path determines the shortest distance between two vertices in a directed graph (a graph where all the edges are directed from one vertex to another) where there are weights imputed to each edge (Sheng and Gao, 2016). Shortest path finding can be used to search for optimal sequences (shortest possible state transitions) of node states to arrive at a goal state or to expedite a transition process needed to reach a goal state (Ahuja et al., 1995).

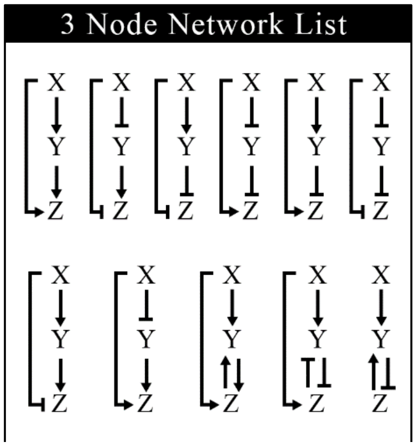


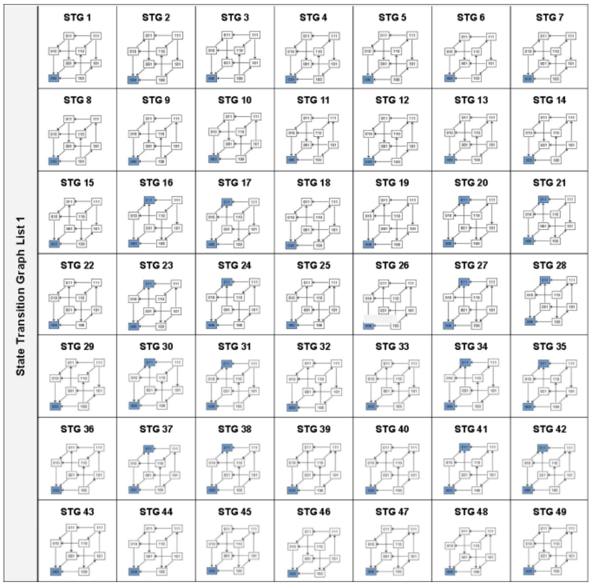
Fig. 1. List of 3-node network motifs

Table 1. Active Interactions Used

Active Interactions							
	AI 1.1	**AI 1.2**	**AI 1.3**	**AI 1.4**	**AI 1.5**	**AI 1.6**	**AI 1.7**
	G1 : G0	G1 : G2	G1 : G0, G2	G1 : G0-G2	G1 : G0, G0-G2	G1 : G2, G0-G2	G1 : G0, G2, G0-G2
	AI 2.1	**AI 2.2**	**AI 2.3**	**AI 2.4**	**AI 2.5**	**AI 2.6**	**AI 2.7**
	G2 : G0	G2 : G1	G2 : G0-G1	G2 : G0-G1	G2 : G0, G0-G1	G2 : G1, G0-G1	G2 : G0, G1, G0-G1

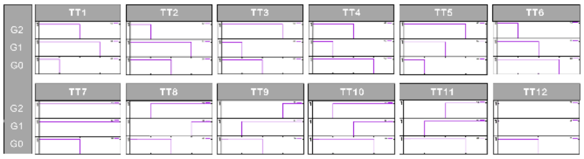
Table 1 lists all the notations and the corresponding commands of active interactions (AI) used to analyze the three-node network motifs (comprised of the nodes G0, G1, G2—also labelled as gene X, gene Y, and gene Z, respectively). In the upper half of the table, the notations pertain to the activations of node G1 and are denoted by AI 1.X, where X = {1, 2, …, 7}. Similarly for the second half of the table that is denoted by AI 2.X. Take AI 1.7 for example, node G1 is the recipient of the interaction of any of the following node(s) combinations: G0, G2, and G0-G2. For command G0-G2 (with hyphens), this requires both nodes G0 and G2 to be activated or turned-on in order to interact with G1. Henceforth, for AI 1.7, G1 will receive an interaction if either G0, G1, or both G0 and G1 are turned-on. When AI 1 and AI 2 are taken by combination, each form the series of choices depending on the type of network motif.

Table 2. List of the Generated State Transition Graphs



In Table 2, we listed all the possible state transition graphs (STGs) that could be generated for the proposed network motifs listed in Figure 1. The STGs are composed of eight nodes with their states being represented by three-digit binary numbers (0 or 1) pertaining to the states of G0, G1, and G2. Although the states of the nodes may vary based on the network motif, the STGs generated have the same number of nodes and edges regardless of the network motifs and AIs. The blue-colored nodes are the stable state nodes of the STGs. Each STG may have either one or two stable states. For stable state nodes, the nodes adjacent to them will always point toward them. The directions of the edges of the STGs may also vary depending on the network motif and AIs utilized. This leads to the 49 unique STGs.

Table 3. List of All the Transient Time Graphs Generated.



All the possible transient time (TT) graphs that could be generated in the analysis performed are shown in Table 3. The TT graph is the graphical representation of the shortest paths from the initial state node to the stable state node of a STG. The TT graph is plotted as the state of the nodes (y-axis) with respect to time (x-axis). The plots of the states of the nodes are separated into three rows: G0 is on the 3rd row, G1 is on the 2nd row, and G2 is on the top row. Take TT1 as example, at the initial time step, it can be seen that all the nodes start at state “1”. In the next time step, G1 and G2 still remain at their previous state values, while G0 shifts to state “0”. In the succeeding time point, G2 ends up shifting to state “0”. Finally, in the last time step, G1 shifts to state “0” and all nodes of the STG end up at the stable state values of “000”.

3. RESULTS AND DISCUSSION

In network motifs 1-8, AIs 1.2 – 1.7 are not possible due to the absence of an edge connecting G2 to G1. Whereas for network motif 11, it only involves AI 2.2 as its possible AI due to the absence of an edge connecting G0 to G2. Network motifs 1-8 all have identical STGs and stable states with the same combination of AIs (i.e. STGs 8-14). This results in the network motif being robust and interchangeable due to their similarities. Next, network motifs 9-10; for the same AIs chosen, have the same STGs. Moreover, network motifs 9-10 are the only network motifs in the study to have all 49 STGs as their possible results — having differing STGs for every combination of AIs. This makes network motifs 9-10 relatively flexible in terms of various choices of AI combinations that allow plethora of paths to reach a specific state. Another interesting thing to note is that the STGs of all the other network motifs can be found within the 49 STGs of network motifs 9-10. This puts network motifs 9-10 as a point of interest on their possible connections to other network motifs.

Table 4. List of Transient Time Graphs for Each Network Motif and Its Corresponding Active Interactions

Motif, AI		Active Interaction																				
		Motif 1-8							Motif 9-10													
		1.1							1.1							1.2						
	TT	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	TT1	x			x	x			x			x	x			x			x	x		
2	TT2	x	x	x					x	x	x					x	x	x				
3	TT3								x	x	x	x	x	x	x	x	x	x	x	x	x	x
4	TT4									x		x		x			x		x		x	
5	TT5	x	x	x	x	x	x	x	x	x	x	x	x	x	x							
6	TT6															x	x	x				
7	TT7																x	x			x	x
8	TT8																	x				x
9	TT9																					
10	TT10																					
11	TT11																					
12	TT12	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Motif, AI		Active Interaction																				
		Motif 9-10																				
		1.3							1.4							1.5						
	TT	1	2	3	4	5	6	7	1	2	3	4	5	6	7	1	2	3	4	5	6	7
1	TT1	x			x	x			x			x	x			x			x	x		
2	TT2	x	x	x					x	x	x					x	x	x				
3	TT3	x	x	x	x	x	x	x														
4	TT4		x		x		x															
5	TT5								x	x	x	x	x	x	x	x	x	x	x	x	x	x
6	TT6								x	x	x											
7	TT7		x	x			x	x														
8	TT8			x				x														
9	TT9		x	x			x	x														
10	TT10																					
11	TT11						x	x														
12	TT12	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Motif, AI		Active Interaction																				
		Motif 9-10														Motif 11						
		1.6							1.7							1	2	3	4	5	6	7
	TT	1	2	3	4	5	6	7	1	2	3	4	5	6	7	2.2						
1	TT1	x			x	x			x			x	x									
2	TT2	x	x	x					x	x	x											
3	TT3															x	x	x				
4	TT4															x	x	x				
5	TT5															x			x	x		
6	TT6	x	x	x																		
7	TT7		x	x			x	x		x	x			x	x		x	x			x	x
8	TT8			x				x			x				x							
9	TT9									x	x			x	x			x				x
10	TT10			x				x			x				x							
11	TT11													x	x			x				x
12	TT12	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 4 shows all the TT generated for each network motif (see Figure 1) with the possible AI combinations (listed in Table 1). In Table 4, one may observe that the TTs for network motifs 1-10 are generated in recurring patterns. For example, TT 1 occurs only in STGs with AI 2.1, 2.4, and 2.5. As for TT 2, it occurs only for STGs with AI 2.1, 2.2, and 2.3, and so on and so forth. The simple reason why network motif 11 did not follow the same patterns as network motifs 1-10 have for TT is because network motif 11 is the only network motif that excludes the arc connecting G0 to G2 (check Figure 1). Another pattern recognized is the STGs with AI 1.2, 1.3, 1.6, and 1.7 has TT 7-11. In these AI combinations, G2 interacts with G1 and is independent from G0. This observation suggests that the presence of G2, not bound to G0 (check Table 1), is responsible for having stable state at "011" (TT 7-11 is the shortest path toward 011). Lastly, TT12 is the only TT graph wherein it is generated in all the network motifs regardless of the AI combinations. This means that TT12 is the most ubiquitous shortest path present in the network motifs chosen for this study. It can be taken or explained as the most accessible path, regardless of what network motif is being considered or analyzed.

4. CONCLUSIONS

From a qualitative perspective, network motifs represent underlying functional and regulatory mechanisms which directly influence system-wide dynamics of a GRN. For this reason, opting to analyze networks motifs using stability analysis, state transition graphs, and shortest path determination (plotted against transient time graphs) proved efficient in narrowing system complexity.

In this study, eleven feedforward and feedback loops were analyzed. Networks motifs 1-8 exhibited similar state transition graphs—all possessing the same shortest path pattern for at least 1 STG—implying that the same shortest path can be used interchangeably across motifs.

Network motifs 9-10 are relatively flexible in terms of various choices of AI combinations that allow plethora of paths to reach a specific state. As compared to motifs 1-8, motifs 9-10 have more profuse and richer dynamics because they outnumber the former in terms of STG's generated by a factor of 7.

Lastly, network motif 11 produced 2-4 shortest paths leading to stable state "000" and 1-3 shortest paths leading to stable state "011". Interestingly, the absence of an arc connecting genes X and Z produces a jarring pattern to the trend exhibited by the aforementioned motifs by restricting active interactions that would otherwise converge to a greater number of STGs.

Although the study is limited to analyzing selected three-node network motifs, further work can be initiated to understand, in dynamic terms, the behavior of other biologically relevant multimodal motifs.

5. REFERENCES

Ahuja, R. K., Orlin, J. B., & Tiwari, A. (2000). A greedy genetic algorithm for the quadratic assignment problem. *Computers & Operations Research*, *27* (10), 917-934.

Barman, S., & Kwon, Y. K. (2018). A boolean network inference from time-series gene expression data using a genetic algorithm. *Bioinformatics (Oxford, England),34* (17), i927–i933.

Davidson, E., & Levin, M. (2005). Gene regulatory networks. *Proceedings of the National Academy of Sciences*, *102* (14), 4935-4935.

Fumia, H., & Martins, M. (n.d.). *Boolean network model for cancer pathways: Predicting carcinogenesis and targeted therapy outcomes.* Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23922675

Garg, A., Mohanram, K., Di Cara, A., De Micheli, G., & Xenarios, I. (2009, June). Modeling stochasticity and robustness in gene regulatory networks. *Bioinformatics (Oxford, England), 25* (12), i101–109.

Huang, X., Chen, L., Chim, H., Chan, L. L. H., Zhao, Z., & Yan, H. (2013). Boolean genetic network model for the control of C. elegans early embryonic cell cycles. *Biomedical Engineering Online, 12* (1).

Karlebach, G., & Shamir, R. (2008, October). Modelling and analysis of gene regulatory networks. Nature Reviews. *Molecular Cell Biology, 9* (10), 770–780.

Moreno, L. (2019). Graphlets and motifs in biological networks. *Encyclopedia of Bioinformatics and Computational Biology, 2*, 814 – 820.

Nakhleh, L. (2015). *Network motifs bioinformatics: Sequence analysis*. Retrieved from https://www.cs.rice.edu/~nakhleh/COMP571/SlidesSpring2015/NetworkMotifs.pdf?fbclid=IwAR0iB2YrCRzYB95mxnJn6txmObOj9YNeKhvtqviK3BrkFo5J3FYIZzde3

Sheng, Y., & Gao, Y. (2016). Shortest path problem of uncertain random network. *Computers & Industrial Engineering, 99*, 97-105.

Weisstein, E. W. (2002). *Lorenz attractor*. Retrieved from http://mathworld.wolfram.com/LorenzAttractor.html

Wuensche, A. (2016). *Exploring discrete dynamics.* Luniver Press.