Introduction to NeuraBASE: A Novel Neuronal Network Approach to Express Network Motifs

Robert Hercus
Choong-Ming Chin
Kim-Fong Ho

neuraMATIX

Introduction

Since the discovery by Alon et al. [1-3], network motifs are now at the forefront of unravelling complex networks, ranging from biological issues to managing the traffic on the Internet. Network motifs are used, in general, as a technique to detect recurring patterns featured in networks. This is based on the assumption that information flows in distinct networks and, hence, common causal associations between nodal networks can be surmised statistically.

In essence, the motif discovery algorithm [1-4] begins with a list of variable-information of a particular network and, their connections to other network variables. An analysis is subsequently made on the most prevalent causal relationships in the network based on the frequency of occurrences. These are then compared with a randomised network with the same number of variables and directed edges. Finally, the algorithm lists out statistically significant patterns of associations that occur more frequently than they would at random.

In the paper by Alon et al. [2], it was reported that most of the networks analysed demonstrate identical motifs, even though they belong to disparate network families. For example, the said literature discovered that gene regulation in genetics, neuronal connectivity network and electronic circuit systems share two identical network motifs, which involved less than four variables. This observation implies that some similarities exist in these three network architectures.

Besides the MFinder tool [1], researchers have also developed other network motif discovery tools. However, MAVisto [6] was found to be as computationally costly with long run-time periods as the MFinder tool uses the same network search for the same subgraph sizes. FANMOD [4-5], on the other hand, has shorter run-times and incorporates the NAUTY algorithm [7] which tests for graph isomorphism. However, it is limited to a subgraph size of 8. Kavosh [8] performs in the same way as the FANMOD, except it is not limited by the subgraph size. The size of Kavosh's subgraph queries is only limited by a computer's processing power. Grochow [9] incorporates the symmetry breaking technique, which enables the unique counting of subgraphs, while MODA [10] is similar to Grochow except for the addition of an expansion tree technique that builds patterns that make subgraphs. NeMoFinder [11] is capable of detecting meso-scale motifs up to size 12 but it is limited to undirected networks. Common difficulties faced in the development of an efficient motif discovery algorithm include challenges in graph isomorphism and the exponential increase in the number of network motifs as the size of the network increases [16]. The graph isomorphism problem is a NP-complete problem. A motif discovery algorithm needs large computational power and an efficient search paradigm to avoid redundant searching and long run-times.

In response to the current development of using network motifs to explain the complexity of network structures, the aim of this paper is to provide an alternative method, which expresses the network motifs in a systematic way. The aim of this paper is not to compare the proposed approach to any existing motif discovery tools in terms of run-time performance and solving graph isomorphism problems. The proposed method uses the NeuraBASE neuronal network as described in the published patent [15]. In this method, directional associations from the parent nodes to the child nodes occupy a central position in encapsulating the causal relationship between the variables concerned.

The key features of this proposed method are:

- Propagated child nodes can be used to represent the causal relationship between network variables. As described in Section 3 (NeuraBASE Neuronal Network Approach), this can be accomplished notwithstanding the complexity of interactions by the variables.
- Based on the directional association from parent nodes to the child nodes, a time-lag factor is present in the neuronal network node building.

- Simultaneous interactions of network variables require a sufficiently dynamic, neuronal network to produce nodes that can accommodate these real-time interactions.

Finding Network Motifs

The central idea of finding network motifs is to extract the shared global statistical features that are present in a multitude of complex network systems. Examples of such networks are biochemistry (transcriptional gene regulation), ecology (food webs) and engineering (electronic circuits, the World Wide Web).

To do so, researchers have developed an algorithm called the MFinder [1], which is capable of detecting network motifs that are most frequently recurring and statistically significant by comparing them with randomised network models.

Theoretically, the algorithm begins with a full network system of N nodes where interactions between the nodes are represented by directed edges as shown in **Figure 1A** and **Figure 1B**. In this network setup, the algorithm scans for all possible n-code motifs (usually n=3 or 4 in the MFinder). Subsequently, the number of occurrences for each motif will be recorded.

As the network contains numerous types of network motifs, a comparison can be made to an ensemble of randomised networks having the same number of nodes and the same number of incoming and outgoing edges as in a live network. The algorithm then reports statistically significant network motifs, i.e. motifs that are more prevalent in the original network as compared with the randomised networks. Patterns that are functionally important but rarely occur in the network will not be reported by the algorithm.

Real Network	**Randomised Networks**

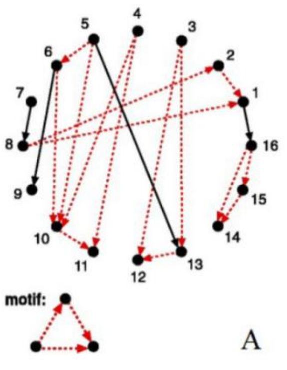
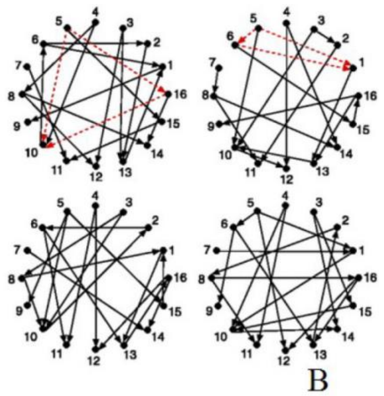

	

Figure. 1A. *A real directed network of N = 16 nodes and 19 edges with a network motif to be detected (as shown in lower left-hand corner)*	***Figure. 1B.*** *This figure depicts an ensemble of randomised networks in relation to Fig. 1A each having N = 16 nodes and 19 edges*

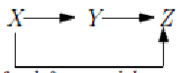
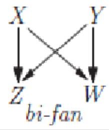
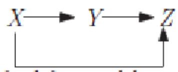
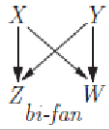
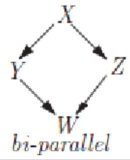
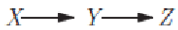
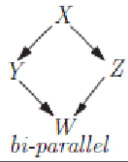
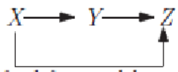
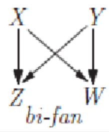
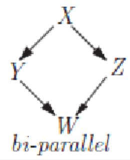
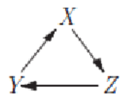
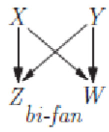
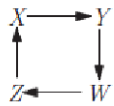
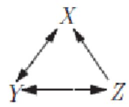
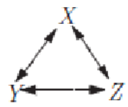
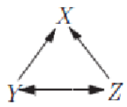
Network	Common Network Motifs		
Gene regulation (transcription)	 *feed-forward loop*	 *bi-fan*	
Neurons	 *feed-forward loop*	 *bi-fan*	 *bi-parallel*
Food Webs	 *three-chain*	 *bi-parallel*	
Electronic Circuits (forward logic chips)	 *feed-forward loop*	 *bi-fan*	 *bi-parallel*
Electronic circuits (digital fractional multipliers)	 *three-node feedback loop*	 *bi-fan*	 *four-node feedback loop*
World Wide Web	 *feed forward with two mutual dyads*	 *fully connected triad*	 *uplinked mutual dyad*

Table 1

Table 1 depicts some of the most common network motifs found in biological and technological networks. For further information on the sources of the networks analysed, see Alon et al. [2]. In **Table 1**, gene regulation, neural networks and electronic circuits (forward logic chips) share a common feed-forward loop and bi-fan motifs in their respective networks.

One of the possible functions of a feed-forward loop motif is the activation of the output signal Z through the direct activation of either the input signal X or through an intermediate signal Y. The signal Y or Z can be activated if the input signal X is persistent enough. The output or intermediate signal can be deactivated when the input goes off.

For the bi-fan motif, different combinations of disjoint input signals X and Y will give different possible output signals Z and W. The only independent motif, which is not shared by other networks, is the three-chain motif found only in food webs network. From the structure of the motif, it can be assumed that predators X do not usually hunt for the same food as their prey Y.

For the bi-parallel motif, the motif can also be found in gene regulation, neural networks, food webs and forward logic chip networks. In the network of food webs, it can be assumed that different species of preys Y, Z of a given predator X may compete for the same food W.

Similarly, if two neurons Y and Z are activated by the same neuron X, both these neurons are most likely needed to activate a subsequent neuron W. The generated motifs for the World Wide Web do not seem to be shared with other network motifs. They are much more complex, with network nodes having feedback functions. Such motifs may reflect a design aimed at short paths between related hyperlink pages, and so, they are distinct from other types of biological or technological networks.

This suggests that network motifs can define broad classes of networks, each with a specific type of structure. The motifs themselves reflect the underlying processes that generate each type of network.

NeuraBASE Neuronal Network Approach

This section presents a new method to express the network motifs given in **Table 1** based on a neuronal network approach described by Hercus [15].

Instead of a built-in unsupervised learning approach, a neuronal network node was built based on the causal relationships between the variables inherited from the network motifs. In this case, the propagated child nodes provide the necessary information of all unique causal relationships between the variables in each motif to aid in detecting the network motif.

Network motifs such as the bi-fan, bi-parallel, or the fully connected triad assume that some of the variables have simultaneous interactions with other variables. In this neuronal network approach, a time-lag concept is introduced, where variables interact with other variables, one step at a time. This approach allows complex networks to be understood in a more systematic way.

In the context of network motifs, the method of forming neuronal network nodes is based on linking two parent nodes at Level 1 of the network motif (variables from the motif), to form an association at Level 2, in the form of a child node (causal relationship between two variables). By doing so, this new node will encapsulate the information of both its parent nodes, which will in turn link it with other nodes if the node has a causal relationship with other variables.

Figure 2 depicts how the neuronal network works in a simple two-variable network motif A ↔ B. In this figure, the central point of this method is the directional associations from the parent nodes (representing variables in a network motif) to the child node. Here, the child node is viewed as a representation of the causal relationship between its parent variables. Instead of both parent nodes associating directly onto the child node, as in conventional approaches, this neuronal network strategy works one step at a time, hence, a time lag is present in forming child nodes.

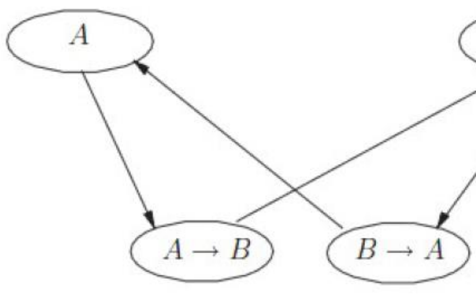


Figure 2. *An example to express the causal relationship of a two-variable network motif A ↔ B. The parent nodes A and B are at Level 1, while the child nodes A → B, and B → A are at Level 2 of the neuronal network setup*

As shown in **Figure 2**, since there is a causal network relationship from A to B, the node $A \to B$ is formed by first linking the variable A to the child node with a directional association followed by another directional association to node B.

For a causal relationship from B to A, the linking process is based on linking a directional association from node B to the child node, followed by a directional association to node A, which then forms the child node $B \to A$.

In **Figure 3** below, this neuronal network setup allows new nodes to grow by linking the node $A \to B$ to $B \to A$, where B is an intermediate link to form $A \to B \to A$. In addition to using A as an intermediate link, node $B \to A$ can also be linked to $A \to B$, to form $B \to A \to B$.

New child nodes can be propagated at Level 4 of the network, say $A \to B \to A \to B$ with $B \to A$ as an intermediate link between nodes $A \to B \to A$ and $B \to A \to B$. A similar argument can be applied when growing the new node $B \to A \to B \to A$. As the flow of information for these new child nodes is repetitive, this neuronal network setup does not propagate them, instead the focus is on finding the unique flow of information between the variables of the network motif. This is to ensure that the neuronal network structure will only have a finite number of child nodes for any finite number of network motif variables.

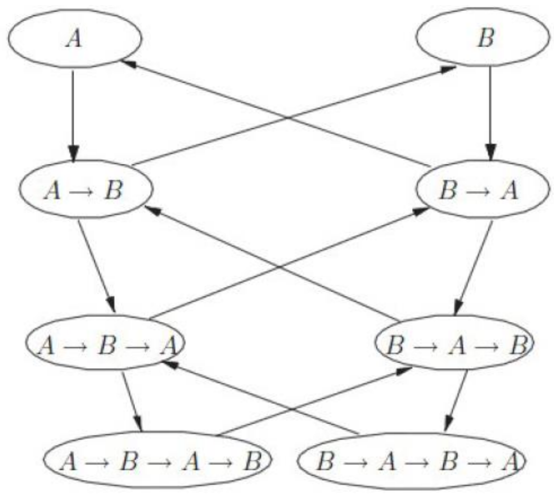


Figure 3. *An example of expressing causal relationships of a two-variable network motif $A \leftrightarrow B$ within a four-generation neuronal network setup. Nodes at Levels 3 and 4 of the neuronal network have repetitive flow of information that can be encapsulated by Level 2 nodes.*

Table 2 and **Table 3** address the nodal build-up on common network motifs as reported by Alon et al. [2]. From these tables, it can be inferred that for any particular network having a statistically significant 3 or 4-node network motif, such a network must also fulfil all the neuronal network information as expressed by the propagated child nodes for each of the corresponding motifs. As opposed to the focus on the simultaneous "firing" of information from variables to variables, this proposed neuronal network is based entirely on the time-lag principle, which links causal relationships between the variables on a step-by-step basis. Using this methodology, the flow of information can be expressed without further complicating the overall complexity of the networks concerned.

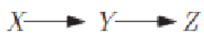
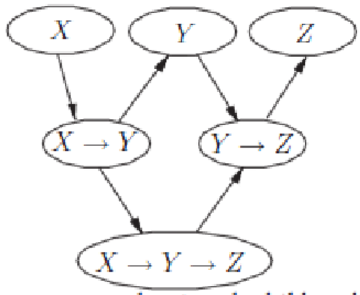
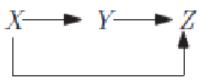
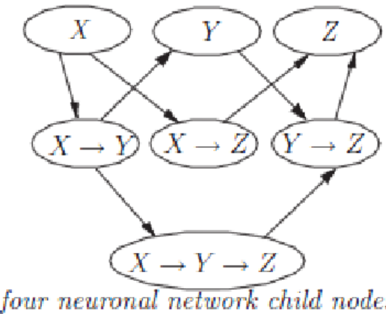
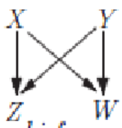
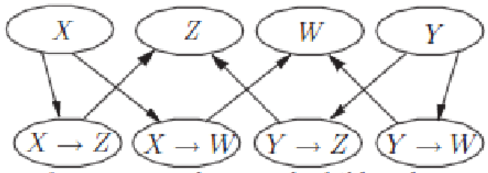
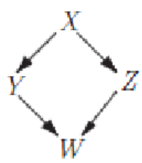
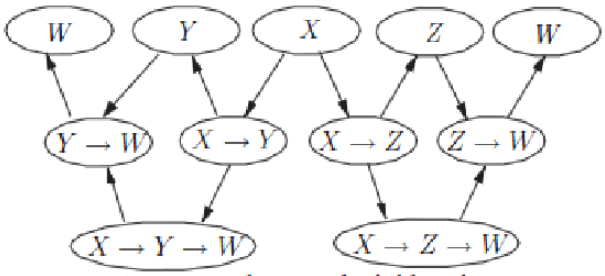
Network Motif	Neuronal Network Setup
 *three-chain*	 *three neuronal network child nodes to represent the three-chain motif*
 *feed-forward loop*	 *four neuronal network child nodes to represent the feed-forward loop motif*
 *bi-fan*	 *four neuronal network child nodes to represent the bi-fan motif*
 *bi-parallel*	 *six neuronal network child nodes to represent the bi-parallel motif*

Table 2. *A neuronal network setup for a class of network motifs*

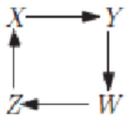
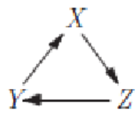
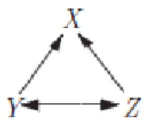
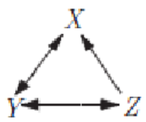
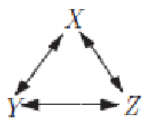
Network Motif	Neuronal Network	
	Parent Nodes	Child Nodes
 *four-node feedback loop*	X, Y	$X \rightarrow Y$
	Y, W	$Y \rightarrow W$
	W, Z	$W \rightarrow Z$
	Z, X	$Z \rightarrow X$
	$X \rightarrow Y, Y \rightarrow W$	$X \rightarrow Y \rightarrow W$
	$Y \rightarrow W, W \rightarrow Z$	$Y \rightarrow W \rightarrow Z$
	$W \rightarrow Z, Z \rightarrow X$	$W \rightarrow Z \rightarrow X$
	$Z \rightarrow X, X \rightarrow Y$	$Z \rightarrow X \rightarrow Y$
	$X \rightarrow Y \rightarrow W, Y \rightarrow W \rightarrow Z$	$X \rightarrow Y \rightarrow W \rightarrow Z$
	$Y \rightarrow W \rightarrow Z, W \rightarrow Z \rightarrow X$	$Y \rightarrow W \rightarrow Z \rightarrow X$
	$W \rightarrow Z \rightarrow X, Z \rightarrow X \rightarrow Y$	$W \rightarrow Z \rightarrow X \rightarrow Y$
	$Z \rightarrow X \rightarrow Y, X \rightarrow Y \rightarrow W$	$Z \rightarrow X \rightarrow Y \rightarrow W$
 *three-node feedback loop*	X, Z	$X \rightarrow Z$
	Y, Z	$Z \rightarrow Y$
	X, Y	$Y \rightarrow X$
	$X \rightarrow Z, Z \rightarrow Y$	$X \rightarrow Z \rightarrow Y$
	$Z \rightarrow Y, Y \rightarrow X$	$Z \rightarrow Y \rightarrow X$
	$Y \rightarrow X, X \rightarrow Z$	$Y \rightarrow X \rightarrow Z$
 *uplinked mutual dyad*	X, Y	$Y \rightarrow X$
	X, Z	$Z \rightarrow X$
	Y, Z	$Y \rightarrow Z$ and $Z \rightarrow Y$
	$Z \rightarrow Y, Y \rightarrow X$	$Z \rightarrow Y \rightarrow X$
	$Y \rightarrow Z, Z \rightarrow X$	$Y \rightarrow Z \rightarrow X$
 *feed forward with two mutual dyads*	X, Y	$X \rightarrow Y$ and $Y \rightarrow X$
	X, Z	$Z \rightarrow X$
	Y, Z	$Y \rightarrow Z$ and $Z \rightarrow Y$
	$Z \rightarrow X, X \rightarrow Y$	$Z \rightarrow X \rightarrow Y$
	$Z \rightarrow Y, Y \rightarrow X$	$Z \rightarrow Y \rightarrow X$
	$Y \rightarrow Z, Z \rightarrow X$	$Y \rightarrow Z \rightarrow X$
	$X \rightarrow Y, Y \rightarrow Z$	$X \rightarrow Y \rightarrow Z$
 *fully connected triad*	X, Y	$X \rightarrow Y$ and $Y \rightarrow X$
	X, Z	$X \rightarrow Z$ and $Z \rightarrow X$
	Y, Z	$Y \rightarrow Z$ and $Z \rightarrow Y$
	$X \rightarrow Y, Y \rightarrow Z$	$X \rightarrow Y \rightarrow Z$
	$Y \rightarrow X, X \rightarrow Z$	$Y \rightarrow X \rightarrow Z$
	$X \rightarrow Z, Z \rightarrow Y$	$X \rightarrow Z \rightarrow Y$
	$Z \rightarrow X, X \rightarrow Y$	$Z \rightarrow X \rightarrow Y$
	$Y \rightarrow Z, Z \rightarrow X$	$Y \rightarrow Z \rightarrow X$
	$Z \rightarrow Y, Y \rightarrow X$	$Z \rightarrow Y \rightarrow X$

Table 3. *A neuronal network setup highlighting only the parent and child nodes for a class of network motifs*

This approach does not constitute any loss of information should simultaneous interactions between variables are needed. For example, given a network motif $Y \leftarrow X \rightarrow Z$, where variable X is seen to fire information to both Y and Z, in the neuronal network, the nodes $X \rightarrow Y$ and $X \rightarrow Z$ can be further linked to form $X \rightarrow \{Y, Z\}$ to denote simultaneous causal relationship as shown in **Figure 4**.

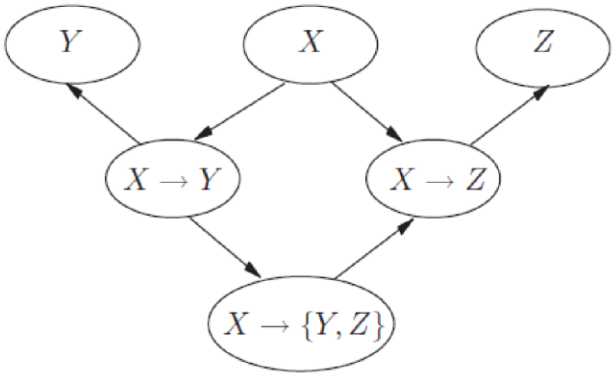


Figure 4. *A neuronal network setup corresponding to the motif $Y \leftarrow X \rightarrow Z$ to show simultaneous connection. The node $X \rightarrow \{ Y, Z \}$ at Level 3 shows that X has an influence on both Y and Z.*

Conclusion

As the number of new network motifs increases, the neuronal network's approach in propagating child nodes which represent the causal relationships of the motif variables, and its ability to include simultaneous interactions, can be potentially exploited to detect other complex motifs in any network structures.

References

[1] Alon U, Kashtan N, Itzkovitz S, Milo R (2002) Mfinder Tool Guide. Department of Molecular Cell Biology and Computer Science & Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel

[2] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U (2002) Network motifs: simple building blocks of complex networks, Science 298(5594) 824–827. DOI: 10.1126/science.298.5594.824

[3] Milo R, Itzkovitz S, Kashtan N, et al (2004) Superfamilies of evolved and designed networks, Science 303(5663):1538 – 1542. DOI: 10.1126/science.1089167

[4] Schwobbermeyer H (2008) Analysis of Biological Networks, NJ John Wiley & Sons, Inc

[5] Wernicke S, Rasche F (2006) FANMOD: a tool for fast network motif detection, Bioinformatics 22(9):1152-1153. doi: 10.1093/bioinformatics/btl038

[6] Schreiber F, Schwobbermeyer F (2005) MAVisto: A tool for the exploration of network motifs, Bioinformatics Applications Note 21(17) 3572-3574. DOI: 10.1093/bioinformatics/bti556

[7] Ribeiro P, Silva F, Kaiser M (2009) Strategies for network motifs discovery, Fifth IEEE International Conference on e-Science '09 80-87. DOI: 10.1109/e-Science.2009.20

[8] Kashani Z, Ahrabian H, Elahi E, Nowzari-Dalini A, Ansari E, Asadi S, Mohammadi S, Schreiber F, Masoudi-Nejad A (2009) Kavosh: A new algorithm for finding network motifs, BMC Bioinformatics, 10(318). DOI: 10.1186/1471-2105-10-318

[9] Grochow J, Kellis M (2007) Network motif discovery using subgraph enumeration and symmetry-breaking, Research in Computational Molecular Biology, 4456 92-106. DOI: 10.1007/978-3-540-71681-5_7

[10] Omidi S, Schreiber F, Masoudi-Nejad A (2009) MODA: An efficient algorithm for network motif discovery in biological networks, Genes Genetic System, 84(5) 385-395. DOI: 10.1.1.178.3656

[11] Chen J, Hsu M, Lee L, Ng SK (2006) NeMofinder: dissecting genome-wide protein-protein interactions with meso-scale network motifs, Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining KDD 06, 17(2) 115. DOI: 10.1145/1150402.1150418

[12] Ripley BD (1996) Pattern Recognition and Neural Network, Cambridge University Press

[13] Kohonen T (1997) Self-organizing maps, Springer-Verlag New York, Inc., Secaucus, NJ, USA

[14] Navrat P, Benuskova M, Kapustik LI, Uger M (2002) Artificial Intelligence, STU Publishing, Bratislava

[15] Hercus RG (2009) Neural Network with Learning and Expression Capability, U.S. Patent 20090119236,

[16] Wong E, Brittany B, Quader S, Huang C-H (2011) Biological network motif detection: principles and practice, Briefing in Bioinformatics doi: 10.1093/bib/bbr033