Learning Context-sensitive Boolean Network from Steady-state Observations and Its Analysis

Huai Li[1], Jon Whitmore[1], Ed Suh[1], Mike Bittner[2], and Seungchan Kim[2]

Keywords: gene regulatory network, context-sensitive Boolean network, Markov chain simulation

1 Introduction.

Boolean network model [1] may provide useful insights for network dynamics at the coarse level. Recently Boolean network has been extended to cope with certain randomness inherent in biological system as Probabilistic Boolean network [4]. While Probabilistic Boolean network is a step forward toward a better mathematical model with capability to abstract uncertainty in biological system, it fails to describe context specific determinism of regulatory system. Context can be defined as a certain condition under which a limited number of genes are tightly regulated by each other via specific cellular mechanisms to perform a specific task [2]. This specific task can be a different developmental stage, or tissue specific function, resulting in a specific cell-type. The change of this context will result in the change in the set of genes that are highly interactive, and probably their connectivity and relationships. Different biological contexts can also correlate with different diseases or might be a reason why a certain group of patients respond to a therapy while others do not. We started to study this problem and have been developing a context-sensitive Boolean network (cBN) model that will abstract the following hypothesis; regulatory mechanism itself in cellular system is static and hard-coded in its genetic code (genomic information), but its activation and inactivation (transcriptomic information) is context sensitive.

While high-throughput gene expression profiling provide vast amount of data for cellular system, most of those measurements come from the stead state observation of the system. Suppose we infer rules from steady-state observations, would the network driven by these rules mimic behavior of biological process? For answering this question, one plausible way we can make sense out of a rule-making procedure is to see what it does in a case we understand the ground truth through the simulation of a small synthetic network driven by some artificial rules. In addition, how is the sensitivity and stability of the network to the methodology of rule formation? Although the stability of a large random Boolean network was well studied both analytically and mathematically [1,3], there has not been extensive study for their structures and the inference of model parameters based on steady-state observations, and their relevance to approximating certain biological systems behavior.

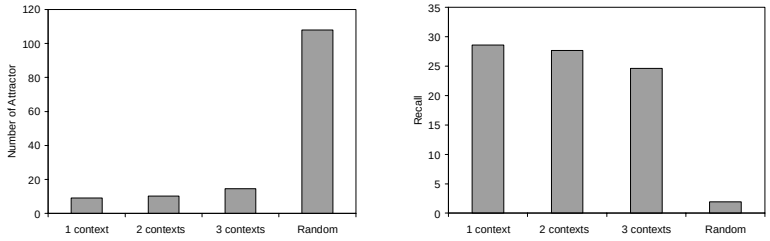
2 Results.

Fig. 1 shows the effect of varying the extent of data consistency on the dynamics of the network and the recall of rules based on simulated data. The recall is defined as the ratio of the number of relevant rules retrieved from the inferred rules to the total number of relevant rules originally constructed.

[1] Computational Biology, Translational Genomics Research Institute, 400 N. Fifth Street, Suite 1600, Phoenix, AZ 85004, USA. Email: hli@tgen.org

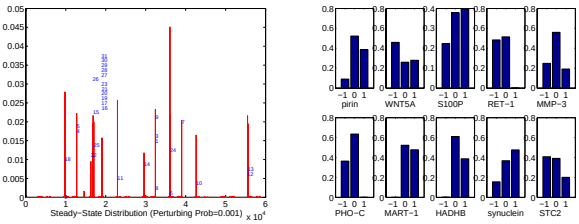
[2] Molecular Diagnostics and Target Validation Division, Translational Genomics Research Institute, 400 N. Fifth Street, Suite 1600, Phoenix, AZ 85004, USA. Email: dolchan@tgen.org

When applying cBN to microarray data, the simulation results showed that the total 31 melanoma sample states occupied 43% of the portion in the steady state distribution with perturbation probability $p = 0.001$, as shown in Fig. 2.



(a) (b)

Figure 1: Effect of varying the extent of data consistency on the recall of rules and the dynamics of the network.



(a) (b)

Figure 2: The estimated distribution after long run based on melanoma data.

3 Conclusion.

In this study, we tried to address if the rules inferred from the steady state observations and the network dynamics driven by those rules can provide us useful information by analyzing the sensitivity and stability of the network to the methodology of rule formation. We used cBN model constructed by a set of artificial rules/functions and inferred rules from steady state observations of this artificial network. By comparing various statistics estimated from the network reconstructed by the inferred rules against those estimated from the network originally constructed by the artificial rule set, even though in this very limited context, we conclude that the inference of rules from steady state observations and its analysis might be quite informative to understanding of cellular system. We also conclude that the more consistent the data is the more stable the network is and the more useful information the network can provide. When applied to microarray data, we observed the rules inferred in fact effectively drive cell states into cancer states.

References

[1] Kauffman, S.A. 1990. Requirements for Evolvability in Complex Systems: Orderly Dynamics and Frozen Components. *Physica D* **42**:135-152.

[2] Kim, S., et al. 2002. Can Markov chain mimic biological regulation? *J. Biol Systems* **10**(4):337-358.

[3] Lynch, J.F. 1993. A Criterion for Stability in Random Boolean Cellular Automat. *Ulam Quarterly* **2**:32-44.

[4] Shmulevich, I., et al. 2002. Probabilistic Boolean Networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics* **18**(2):261-74.