Quantifying Important Genes in a Gene Regulatory Network with Applications to Network Compression

Eunji Kim, Ivan Ivanov, Edward R. Dougherty

It has been long recognized that canalizing genes possess broad regulatory power, and can enforce corrective actions on cellular processes for the purpose of biological robustness to maintain a constant phenotype despite genetic mutations or environmental perturbations. The notion of genes that can constrain a biological system to a specific behavior was originally proposed in [1]. Waddington suggested the existence of genes that can produce reliable developmental effects against genetic mutations or environmental changes during evolution [1], [2]. Lehner studied Waddingtons intuition and stated that canalizing genes are hub genes that provide robustness when faced with environmental, stochastic and genetic perturbations [3]. The term canalizing gene has been used in [4] to refer to genes that possess broad regulatory power, and their action sweeps across a wide swath of processes for which the full set of affected genes are not highly correlated under normal conditions. Such genes are crucial in a complex system so it can buffer itself from the effects of random alterations or operating errors. Canalizing genes are analogous to master switches that set in motion a cascade of regulatory events that have huge impacts on downstream genes for the sake of driving the system to a desired condition. Therefore, analyzing canalizational properties of the network and preserving canalizing genes are important to maintain homeostasis or to derive intervention strategies for beneficially altering cell dynamics. Probabilistic Boolean networks (PBNs) form a widely accepted mathematical model for cellular systems and gene regulatory networks [5]. One of their important applications is to design intervention strategies that beneficially alter cell dynamics through studying long-run network behavior. Because the dynamics of a PBN are represented by an ergodic and irreducible finite Markov chain, the model possesses a steady-state distribution (SSD) reflecting its long-run dynamics. However, even relatively small networks can pose serious difficulties in assessing the dynamics, considering that a Boolean network of .... genes has 2n states and the transition probability matrix has size 2n2n. Given the exponential dependence of the state space on the number of nodes, there is a need for network reduction mappings that produce models that are more tractable. The major focus of this talk is the preservation of the canalizing properties of genes in the original network under a wide class of compression mappings. For this purpose, we first introduce, [6], the important quantification of a gene Canalizing Power (CP) and then examine what happens to the network CP vector when genes are consecutively deleted. It is hypothesized that deleting a gene with the smallest canalizing power may help to preserve the canalizational properties of the original network under network reduction. In addition, we point out to an important observation related to this hypothesis; namely, that genes in some networks

retain their canalization properties after network compression, while there is a class of networks that do not possess this property. Naturally, this hypothesis leads to definitions of reducible and irreducible networks. Thus, one can formulate a related classification problem that aims to find relevant network features that can separate reducible from irreducible networks.

References

[1] C. H. Waddington, Canalization of development and the inheritance of acquired characters, Nature, vol. 150, no. 3811, pp. 563–565, 1942.

[2] A. Wagner, Robustness and evolvability in living systems. 2005.

[3] B. Lehner, Genes Confer Similar Robustness to Environmental, Stochastic, and Genetic Perturbations in Yeast, PLOS ONE, vol. 5, no. 2, p. e9035, Feb. 2010.

[4] D. C. Martins, U. M. Braga-Neto, R. F. Hashimoto, M. L. Bittner, and E. R. Dougherty, Intrinsically Multivariate Predictive Genes, IEEE J. Sel. Top. Signal Process., vol. 2, no. 3, pp. 424–439, 2008.

[5] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang, Probabilistic Boolean networks: a rule-based uncertainty model for gene regulatory networks, Bioinformatics, vol. 18, no. 2, pp. 261–274, Feb. 2002.

[6] E. Kim, I. Ivanov, and E. R. Dougherty, Quantifying the notions of canalizing and master genes in a gene regulatory networka Boolean network modeling perspective, Bioinformatics, vol. 35, no. 4, pp. 643–649, Jul. 2018.