Inference of Boolean networks from perturbation data

Giuseppe Torrisi[1], Tessa Green[2], Chris Sander[2], Linus Schumacher[1]
[1]*University of Edinburgh, Edinburgh, United Kingdom,* [2]*Harvard Medical School, Boston, USA*

Gene regulatory networks (GRNs) identify models of genes that regulate their state via activator or inhibiting signaling propagated via a network. Recent experiments perturb the state of genes and record the effects, providing rich information on the connectivity of the GRN. The inference of GRNs is the problem of inferring interactions between genes from available measurements about genes' state. It can be mapped to a problem of link prediction in the presence of node perturbations.

To address this challenge, we develop a Bayesian model of the GRN that takes as input the states of nodes after perturbation experiments involving both node activation and inhibition, i.e., Crispr-a and Crispr-i experiments. To predict the effect of node perturbation, we define how the state of every node is controlled by the others. We model the node state in terms of a linear interaction model for Boolean variables in the presence of noise. The dynamics has been studied in literature using the dynamic cavity method for sparse random graphs and recent implementation of the dynamic cavity method using dynamic programming allows to apply to networks with realistic degree distributions, see ref. [1,2].

The network inference problem lays on a space that grows exponentially with N^2, with N the number of nodes of the network. We design an approximate algorithm that solves our Bayesian inference model and scales as N^3. Our solver uses the message-passing technique, and we design a dynamic programming method to compute the messages. Our dynamic programming method is optimised such that it takes advantage of the discrete nature of both couplings and Boolean variables. We evaluate the effectiveness of our method on synthetic networks and quantify its reconstruction error. We also apply to a real dataset performing Crispr-a and Crispr-i experiments on K562 cell line. Overall, our work provides a new tool for GRN inference that is based on disordered systems techniques and informs the design of future experiments by highlighting the importance of including both inhibition and activation perturbations for a more comprehensive understanding of gene interactions.

References

[1] G. Torrisi, R. Kühn, A. Annibale, Uncovering the non-equilibrium stationary properties in sparse Boolean networks. J. Stat. Mech. Theory Exp., 2022(5), p.053303 (2022).
[2] G. Torrisi, A. Annibale, R. Kühn, Overcoming the complexity barrier of the dynamic message-passing method in networks with fat-tailed degree distributions. Phys. Rev. E, 104(4), p.045313 (2021).