

Scalable Control of Asynchronous Boolean Networks

Cui Su¹, Soumya Paul², and Jun Pang^{1,2}

¹ Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Esch-sur-Alzette, Luxembourg

² Faculty of Science, Technology and Communication, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Abstract. We summarise our recent research results on developing efficient and scalable control methods for gene regulatory networks modelled as asynchronous Boolean networks. Our methods compute a minimal subset of nodes of a given Boolean network, such that (different types of) perturbations of these nodes, in one step or a sequence of steps, can drive the network (from an initial state) to a target steady state.

1 Introduction

The ground-breaking discovery of cell reprogramming has overturned the conventional thinking that cell differentiation was irreversible. With cell reprogramming techniques, it is possible to reprogram cell fates in many different ways, such as trans-differentiation, de-differentiation, retro-differentiation, etc. This has opened up a great opportunity for regenerative medicine to treat the most devastating diseases, such as Parkinson's disease, Alzheimer's disease, etc.

A big obstacle for the application of *in vivo* cell reprogramming is the effective identification of target genes. Numerical-experimental methods are infeasible [13], due to the combinatorial complexity of target genes and high experimental costs. This gives rise to the need of computational modelling of gene regulatory networks (GRNs), which makes it possible to analyse GRNs with formal reasoning and tools. Among various modelling frameworks, Boolean networks (BNs) have distinct advantages: being simple and yet able to capture important properties of nonlinear dynamical GRNs [2]. In BNs, genes are modelled as binary nodes, being either 'expressed' or 'not expressed'. Activation/inhibition regulations between genes are described by Boolean functions. The updating of nodes can be either *synchronous* or *asynchronous*. The steady states of GRNs are described as *attractors*, to one of which the system eventually settles down.

Asynchronous BNs are considered more realistic than synchronous BNs, because only the asynchronous updating allows for the biological processes occurring on different time-scales [7]. Owing to the non-determinism of asynchronous BNs, the control methods for synchronous BNs [3, 15] are not applicable to asynchronous BNs. Recently, a stable motif based method was developed to control asynchronous BNs [14]. However, this method does not guarantee the minimality of the control sets, which may result in unnecessary experimental costs.

		Instantaneous control	Temporary control	Permanent control
Source-target control	One-step control	[8, 9]	[11]	
	Sequential control	[4, 5]		
Target control	One-step control	[1]		

Table 1. Different control strategies for asynchronous BNs.

The limitations of the existing methods motivate us to develop scalable, efficient and practical methods to control asynchronous BNs.

2 Control Problems in Boolean Networks

Attractors of a BN characterise cell phenotypes, which are biologically observable states [2]. Only the control of attractors is meaningful. Thus, the control objective for GRNs, in the context of asynchronous BNs, can be described as: finding a subset of nodes, called driver nodes, such that the control of these nodes can drive the network (from a source attractor) to a target attractor. If the source state is known, we call it **source-target control**; otherwise, we call it **target control**.

The control (perturbation) of a node means to change the expression of the node to either ‘1’ or ‘0’. Based on the amount of time that the control is applied, we distinguish the following three types of controls:

- (a) **instantaneous control** – the control is applied instantaneously;
- (b) **temporary control** – the control is applied for finite steps and then released;
- (c) **permanent control** – the control is applied for all the following time steps.

Thanks to the rapid advances in cell reprogramming, these three controls can be realised in biological experiments with different bimolecular tools. In particular, for the source-target control, based on the number of control steps, we also have:

- (a) **one-step control** – perturbations are applied simultaneously for one step;
- (b) **sequential control** – perturbations are applied in a sequence of steps.

3 Results

The major challenge in the control of asynchronous BNs lies in the infamous state space explosion problem: the state space grows exponentially with respect to the number of nodes of a BN. It prohibits the *efficiency*, *scalability* and *minimality* of the control methods. To cope with this problem, we employ the ‘divide and conquer’ strategy to explore both the structural and dynamical properties of a BN. As shown in Table 1, we have developed efficient methods to solve the minimal one-step source-target control with instantaneous, temporary and permanent perturbations [8, 9, 11], the minimal sequential source-target control

with instantaneous perturbations [4, 5], as well as the target control with instantaneous perturbations [1]. Among these methods, sequential source-target control identifies a sequence of intermediate states and the associated perturbations. At each control step, we apply a set of perturbations, wait until the network reaches the intermediate state and then apply another set of perturbations. Based on the status of intermediate states, we have developed a **general sequential control** [5], where any state (transient states or steady states) can act as intermediate states, and an **attractor-based sequential control** [4], where only steady states can play the role of intermediate states.

Our methods are implemented as part of the software ASSA-PBN [6]. We have evaluated our methods on a variety of real-life biological networks.³ The results show that our methods are quite efficient in terms of the computation time and scale well for large networks. Moreover, our methods identify the minimal control sets for different strategies, which can reduce the experimental costs to a great extent. For the source-target control, both sequential control methods [4, 5] require less perturbations than the one-step control [8, 9]. The attractor-based sequential control [4] is considered more practical than the general sequential control [5], since it uses biologically observable states as intermediate states and thus only requires partial observability of the system. Considering different types of perturbations, the temporary perturbation is preferable than the instantaneous and permanent perturbations due to: (1) temporary control requires the least number of perturbations, which not only leads to less experimental costs, but also makes the experiments easier to conduct; and (2) temporary perturbations are eventually released and thus less invasive. In this way, we can avoid some unknown side effects compared to permanent perturbations.

Furthermore, in order to make the results more practical, experimental constraints need to be incorporated. We have made the following improvements to integrate practical and experimental concerns.

- Our methods can avoid certain perturbations (perturbing a gene from ‘expressed’ to ‘not expressed’ and/or the reverse direction) during the computation. In this case, we can avoid perturbing genes that are essential for cell survival and genes that are hard or expensive to perturb.
- For attractor-based sequential source-target control, our method can avoid undesired attractors as intermediate states, such as apoptosis.
- Considering the minimal solution may not be the best solution, an upper bound of the number of perturbations can be set as a prerequisite. Then our methods can compute all the solutions within that upper bound.

4 Discussion and Future Work

Due to the diversity of GRNs, it is less likely to find one strategy that suits different networks. Furthermore, the cost and success rate of different cell reprogramming techniques vary a lot. Taking these into consideration, we can

³ We refer details on their evaluation to our previous works [1, 4, 5, 8, 9, 11].

compute a bunch of reprogramming paths with different control methods. Afterwards, biologists can make a choice from the provided solutions based on specific biological systems and experimental settings.

Given the strengths of temporary perturbations, currently we are working on sequential source-target control and target control with temporary perturbations. In future, we plan to extend our work to the control of PBNs [10, 12].

References

1. Baudin, A., Paul, S., Su, C., Pang, J.: Controlling large Boolean networks with single-step perturbations. *Bioinformatics* (2019)
2. Kauffman, S.A.: Homeostasis and differentiation in random genetic control networks. *Nature* **224**, 177–178 (1969)
3. Kim, J., Park, S., Cho, K.: Discovery of a kernel for controlling biomolecular regulatory networks. *Scientific Reports* **3**(2223) (2013)
4. Mandon, H., Su, C., Haar, S., Pang, J., Paulevé, L.: Sequential reprogramming of Boolean networks made practical. In: *Proc. 17th International Conference on Computational Methods in Systems Biology*. LNCS, Springer-Verlag (2019), accepted
5. Mandon, H., Su, C., Pang, J., Paul, S., Haar, S., Paulevé, L.: Algorithms for the sequential reprogramming of Boolean networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2019), accepted
6. Mizera, A., Pang, J., Su, C., Yuan, Q.: ASSA-PBN: A toolbox for probabilistic Boolean networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **15**(4), 1203–1216 (2018)
7. Papin, J.A., Hunter, T., Palsson, B.O., Subramaniam, S.: Reconstruction of cellular signalling networks and analysis of their properties. *Nature Reviews Molecular Cell Biology* **6**(2), 99 (2005)
8. Paul, S., Su, C., Pang, J., Mizera, A.: A decomposition-based approach towards the control of Boolean networks. In: *Proc. 9th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*. pp. 11–20. ACM Press (2018)
9. Paul, S., Su, C., Pang, J., Mizera, A.: An efficient approach towards the source-target control of Boolean networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (2019), accepted
10. Shmulevich, I., Dougherty, E.R.: *Probabilistic Boolean networks: the modeling and control of gene regulatory networks*. SIAM Press (2010)
11. Su, C., Paul, S., Pang, J.: Controlling large Boolean networks with temporary and permanent perturbations. In: *Proc. 23rd International Symposium on Formal Methods*. LNCS, Springer-Verlag (2019), accepted
12. Trairatphisan, P., Mizera, A., Pang, J., Tantar, A.A., Sauter, T.: optPBN: An optimisation toolbox for probabilistic Boolean networks. *PLOS ONE* **9**(7), e98001 (2014)
13. Wang, L.Z., Su, R.Q., Huang, Z.G., Wang, X., Wang, W.X., Grebogi, C., Lai, Y.C.: A geometrical approach to control and controllability of nonlinear dynamical networks. *Nature Communications* **7**, 11323 (2016)
14. Zañudo, J.G.T., Albert, R.: Cell fate reprogramming by control of intracellular network dynamics. *PLoS Computational Biology* **11**(4), e1004193 (2015)
15. Zhao, Y., Kim, J., Filippone, M.: Aggregation algorithm towards large-scale Boolean network analysis. *IEEE Transactions on Automatic Control* **58**(8), 1976–1985 (2013)