Bioinformatics
doi.10.1093/bioinformatics/xxxxxx
Advance Access Publication Date: Day Month Year
Application Notes

OXFORD

Systems biology

CABEAN: A Software for the Control of Asynchronous Boolean Networks

Cui Su [1] 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 Medicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

* To whom correspondence should be addressed.

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Summary: Direct cell reprogramming, also called transdifferentiation, has great potential for tissue engineering and regenerative medicine. Boolean networks, a popular modelling framework for gene regulatory networks, make it possible to identify intervention targets for direct cell reprogramming with computational methods. In this work, we present our software, CABEAN, for the control of asynchronous Boolean networks. CABEAN identifies efficacious nodes, whose perturbations can drive the dynamics of a network from a source attractor (the initial cell type) to a target attractor (the desired cell type). CABEAN provides several control methods integrating practical constraints. Thus, it has the ability to provide a rich set of control sets, such that biologists can select suitable ones for validation based on specific experimental settings.

Availability and Implementation: The executable binary and the user guide of the software are publicly available at https://satoss.uni.lu/software/CABEAN/.

Contact: jun.pang@uni.lu

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

Direct cell reprogramming has opened up an unprecedented opportunity for treating the most devastating diseases, which are characterised by a deficiency of certain cells, or have a genetic basis (Grath and Dai, 2019). It allows us to harness abundant somatic cells, and reprogram them into the desired cell types to restore functions of the damaged tissues or organs.

Conventional experimental approaches for the identification of interventions are hindered by their long time commitment and high costs. This promotes the identification of intervention targets based on mathematical modelling of biological systems. *Boolean networks* have apparent advantages compared to other modelling frameworks. They provide a qualitative description of the biological systems. Despite the simplicity of the structure, it is capable of capturing the important dynamical properties of biological systems. In a Boolean network, biomolecules are denoted as binary-valued nodes, and the activation or inhibition regulations are described as Boolean functions. Different systems for updating the nodes, such as synchronous or asynchronous mode, lead to different dynamics. We focus on asynchronous dynamics, because this strategy can capture biological processes that happen at different classes of time scales. The steady-state behaviour of the dynamics is described as *attractors*, which are hypothesised to characterise cell phenotypes.

In the framework of Boolean networks, direct cell reprogramming is defined as the *source-target control* problem: identifying a subset of nodes, whose perturbation can drive the dynamics from the source attractor to the target attractor. Several software tools have been developed for the control of logical models. Lin *et al.* proposed a Max-SAT based automatic test pattern generate algorithm (Lin and Khatri, 2012) to identify faulty genes that cause undesired behaviours and to predict the best drug selection for curing cancer for synchronous Boolean networks under a stuck-at fault model. ActONetLib (Biane and Delaplace, 2018) is a Mathematica library designed to compute driver nodes for Boolean control networks (BCNs) based on abductive reasoning. CANA (Correia *et al.*, 2018) is a Python package integrated with several methods to study redundancy and control of synchronous Boolean networks. Murrugarra *et al.* proposed a method for the identification of intervention targets based on algebraic techniques for synchronous Boolean networks (Murrugarra

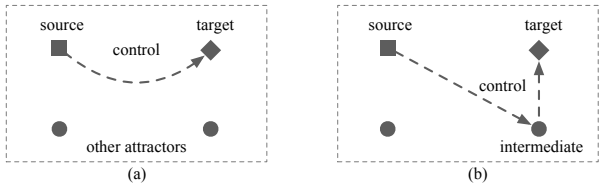

Fig. 1. (a) One-step and (b) attractor-based sequential source-target control. Squares, rhombuses and circles represent the source attractor, the target attractor and other attractors of the network, respectively.

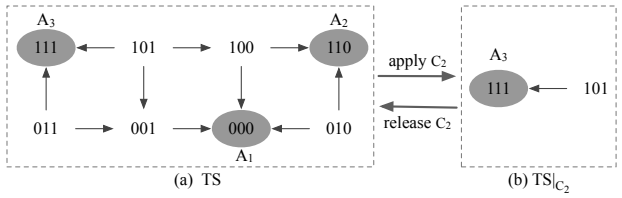

Fig. 2. (a) The transition system TS and (b) the transition system under control $\mathsf{TS}|_{C_2}$ of Example 1. We omit self-loops for all the states except for the state (101) in (a).

	Minimal One-step control	Attractor-based Sequential control
Instantaneous	Paul *et al.* (2018, 2019)	Mandon *et al.* (2019)
Temporary	Su *et al.* (2019)	Su and Pang (2020)
Permanent	Su *et al.* (2019)	Su and Pang (2020)

Table 1. Control methods that are integrated in CABEAN.

et al., 2016). However, the above mentioned methods are not directly applicable to asynchronous Boolean networks. The algorithm Kali (Poret and Guziolowski, 2018) predicts intervention targets which can reduce the reachability of attractors associated with pathological phenotypes for both synchronous and asynchronous Boolean networks. It traverses the state space of a Boolean network by performing $\max_s$ random walks of $\max_k$ steps to find all the attractors, while estimating their basins. Thus, Kali is limited to an estimation rather than an accurate computation of the attractors and their basins. BoolNet (Müssel *et al.*, 2010) is a R package, which integrates methods for reconstruction, generalisation, and attractor identification for synchronous, asynchronous, and probabilistic Boolean networks. PyBoolNet (Klarner *et al.*, 2017) is a Python package for manipulating Boolean networks, including the generation, visualisation and attractor detection. However, BoolNet and PyBoolNet do not support the computation of driver nodes that can modulate the dynamics of Boolean networks. Stable motif-based control (Zañudo and Albert, 2015) is a method for the *target control* of asynchronous Boolean networks. It identifies a subset of nodes, such that the perturbation of these nodes can drive the dynamics from any initial state to a target attractor. The main difference between the target control and the source-target control, which CABEAN is dealing with, lies in the source states: the source is a specified attractor for the source-target control, whereas the source can be any state in the state space for the target control.

In this work, we introduce our software CABEAN for the source-target control of asynchronous Boolean networks. Since it is less likely to find one control strategy that can suit diverse biological systems perfectly, CABEAN implements several methods that can modulate the dynamics in different ways. Moreover, CABEAN allows users to encode practical constraints to make the results more feasible. We have demonstrated the efficacy and efficiency of CABEAN on several real-life biological networks in our previous works (Paul *et al.*, 2018, 2019; Su *et al.*, 2019; Mandon *et al.*, 2019; Su and Pang, 2020).

2 General Features

The source-target control of Boolean networks can be achieved either in one step, called *one-step source-target control*, or in multiple steps, called *sequential source-target control*. Considering the difficulties in practical applications, we are interested in the sequential control through other attractors, called *attractor-based sequential source-target control*. Fig. 1(a) and Fig. 1(b) illustrate the one-step and attractor-based sequential control, respectively. In Fig. 1(a), we can see that one-step control drives the network from the source attractor directly to the target attractor, while attractor-based sequential control as shown in Fig. 1(b) employs other attractors as intermediates and identifies a sequence of perturbations. Attractor-based sequential control has the potential of reducing the number of required perturbations and discovering novel control paths.

Rapid advances in gene editing techniques make it possible to manipulate perturbations in various ways. Based on the application time of the perturbations, we have *instantaneous, temporary* and *permanent* controls. Instantaneous control applies perturbations instantaneously; temporary control applies perturbations for sufficient time and then releases them to retrieve the original dynamics; and permanent control applies the control for all the following time steps. Instantaneous control usually requires more perturbations, which increases the experimental costs; permanent control may lead to unforeseen consequences due to its permanent influence on the dynamics; temporary control requires the least number of perturbations and only causes a temporary shift of the dynamics, thus it is more preferable.

Combining the two categories, we can have six different control strategies as shown in Table 1. They are all integrated in CABEAN. The minimal one-step instantaneous, temporary and permanent control methods (OI, OT and OP) compute the exact and minimal control sets; while attractor-based sequential instantaneous, temporary and permanent control methods (ASI, AST and ASP) compute all the sequential control paths (including the shortest paths) with at most k perturbations. All these methods are based on the computation of strong and weak basins of attractors, which explores both the structure and dynamical properties of asynchronous Boolean networks (Paul *et al.*, 2018, 2019). The main idea is that an instantaneous control drives the network dynamics from the source state to a state in the strong basin of the target attractor, from which there only exist paths to the target attractor. Both temporary control and permanent control can make use of the spontaneous evolutions of the network dynamics by moving into the weak basin of the target attractor, from which there exist paths to the target attractor and may also exist paths to other attractors. To guarantee the inevitable reachability of the target attractor, a temporary control should drive the network dynamics to a state in the strong basin of the target attractor at the end of control, while a permanent control stirs the network from the source state to a state in the strong basin of the target attractor in the resulting transition system under control.

Practical constraints have to be taken into consideration to make the results more feasible. In practical applications, some genes may be impossible or difficult to perturb in a certain way; some attractors corresponding to undesired states, such as apoptosis or diseased states, should be avoided as the intermediate attractors for sequential control. CABEAN allows users to encode undesired perturbations and undesired intermediate attractors as preconditions. The undesired perturbations/intermediate attractors will then be avoided during the computation.

CABEAN is implemented in C and C++. It supports three types of input files, including the model file and the specification files for undesired attractors and undesired perturbations. The model file follows either BoolNet format or ISPL format of the software MCMAS (Lomuscio *et al.*, 2017). CABEAN utilises MCMAS to encode Boolean networks into the efficient data structure binary decision diagram (BDD). It explores both the network topology and dynamics to efficiently compute a set of node perturbations for solving the source-target control problem. Thus, despite the infamous state space explosion problem, CABEAN is efficient in terms of the computational time and scales well for large networks of of several hundred nodes.

Example 1. *Consider a Boolean network $G(X, F)$, where $X = \{x_1, x_2, x_3\}$, $F = \{f_1, f_2, f_3\}$, $f_1 = x_2$, $f_2 = x_1$, and $f_3 = x_2 \wedge x_3$. Its transition system* TS *is given in Fig. 2(a). This network has three attractors A_1, A_2 and A_3, indicated as the grey oval nodes. Let us assume A_1 and A_3 are the source and the target, respectively. The control $C_1 = \{x_1 = 1, x_2 = 1, x_3 = 1\}$ is a minimal* OI *control. The instantaneous application of C_1 drives the network directly from A_1 to A_3. The control $C_2 = \{x_1 = 1, x_3 = 1\}$ is a minimal* OT/OP *control. The application of C_2 fixes the values of x_1 and x_3 to* 1*. This not only drives the the network from state* (000) *to state* (101)*, but also shapes the transition system* TS *to a new one* $\mathsf{TS}|_{C_2}$ *given in Fig. 2(b), where nodes x_1 and x_3 can only take the value of* 1*. We can see that there only exists a path from state* (101) *to the target A_3 in* $\mathsf{TS}|_{C_2}$*. Hence, the temporary or permanent application of C_2 will eventually guide the network to A_3. If C_2 is applied for all the following time steps, we call it a permanent control. Otherwise, C_2 can be released once the network settles down to A_3 and we call it a temporary control. For the attractor-based sequential control, A_2 can play the role of the intermediate attractor. The path $A_1 \xrightarrow{x_1=1,x_2=1} A_2 \xrightarrow{x_3=1} A_3$ is an* ASI *control and the path $A_1 \xrightarrow{x_1=1} A_2 \xrightarrow{x_3=1} A_3$ is an* AST/ASP *control. For this example, the input and output files of different control methods can be found at the software website.*

3 Conclusion

We presented a novel software CABEAN that integrates several methods for the source-target control of asynchronous Boolean networks. These methods identify the minimal and exact control sets that ensure the inevitable reachability of the target attractor from a source attractor, such sets represent promising intervention targets for further wet-lab validation. We believe CABEAN can contribute to the study of mechanisms of biological processes and facilitate the development of direct cell reprogramming.

Acknowledgements

We thank Soumya Paul for his contribution to CABEAN, and anonymous reviewers for their useful comments.

Funding

This work was partially supported by the project SEC-PBN funded by University of Luxembourg and the ANR-FNR project AlgoReCell (INTER/ANR/15/11191283).

References

Biane, C. and Delaplace, F. (2018). Causal reasoning on Boolean control networks based on abduction: theory and application to cancer drug discovery. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **16**(5), 1574–1585.

Correia, R., Gates, A., Wang, X., and Rocha, L. (2018). CANA: A python package for quantifying control and canalization in Boolean networks. *Frontiers in Physiology*, **9**, 1046.

Grath, A. and Dai, G. (2019). Direct cell reprogramming for tissue engineering and regenerative medicine. *Journal of Biological Engineering*, **13**(1), 14.

Klarner, H., Streck, A., and Siebert, H. (2017). PyBoolNet: a python package for the generation, analysis and visualization of Boolean networks. *Bioinformatics*, **33**(5), 770–772.

Lin, P.-C. K. and Khatri, S. P. (2012). Application of Max-SAT-based ATPG to optimal cancer therapy design. *BMC Genomics*, **13**(S6), S5.

Lomuscio, A., Qu, H., and Raimondi, F. (2017). MCMAS: an open-source model checker for the verification of multi-agent systems. *Journal on Software Tools for Technology Transfer*, **19**(1), 9–30.

Mandon, H., Su, C., Haar, S., Pang, J., and Paulevé, L. (2019). Sequential reprogramming of Boolean networks made practical. In *Proc. 17th International Conference on Computational Methods in Systems Biology*, volume 11773 of *LNCS*, pages 3–19. Springer.

Murrugarra, D., Veliz-Cuba, A., Aguilar, B., and Laubenbacher, R. (2016). Identification of control targets in Boolean molecular network models via computational algebra. *BMC Systems Biology*, **10**(1), 94.

Müssel, C., Hopfensitz, M., and Kestler, H. A. (2010). BoolNet-an R package for generation, reconstruction and analysis of Boolean networks. *Bioinformatics*, **26**(10), 1378–1380.

Paul, S., Su, C., Pang, J., and Mizera, A. (2018). A decomposition-based approach towards the control of Boolean networks. In *Proc. 9th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics*, pages 11–20.

Paul, S., Su, C., Pang, J., and Mizera, A. (2019). An efficient approach towards the source-target control of Boolean networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*. accepted.

Poret, A. and Guziolowski, C. (2018). Therapeutic target discovery using Boolean network attractors: improvements of kali. *Royal Society Open Science*, **5**(2), 171852.

Su, C. and Pang, J. (2020). Sequential temporary and permanent control of Boolean networks. In *Proc. 18th International Conference on Computational Methods in Systems Biology*, LNCS. Springer. accepted.

Su, C., Paul, S., and Pang, J. (2019). Controlling large Boolean networks with temporary and permanent perturbations. In *Proc. 23rd International Symposium on Formal Methods*, volume 11800 of *LNCS*, pages 707–724. Springer.

Zañudo, J. G. and Albert, R. (2015). Cell fate reprogramming by control of intracellular network dynamics. *PLOS Computational Biology*, **11**(4), e1004193.