

An efficient method for approximating attractors in large-scale asynchronous Boolean models

Trinh Van Giang¹, Kunihiko Hiraishi¹

¹Japan Advanced Institute of Science and Technology, Japan

December 16, 2020

Contents

- 1 Introduction
- 2 Preliminaries
- 3 ARBNs and GARBNs
- 4 Approximation method
- 5 Experiments
- 6 Conclusion

Contents

- 1 Introduction
- 2 Preliminaries
- 3 ARBNs and GARBNs
- 4 Approximation method
- 5 Experiments
- 6 Conclusion

Boolean networks and attractors

- **Boolean networks (BNs)** are **simple but efficient** models of gene regulatory networks. They are also interesting mathematical objects widely applied in many areas beyond systems biology [Gershenson, 2004].
- A central aim of Boolean network analysis is to find **attractors** which are important long-term behavior of BNs.
 - ▶ Analysis of attractors could provide **new insights** into system biology (e.g., the origins of cancer) [Kauffman, 1992, Huang, 2001].
 - ▶ Attractors play an important role in **the development of new drugs** [Tun et al., 2011, Biane and Delaplace, 2018].

Classification of Boolean networks

- Many different types of BNs have been proposed and widely studied [Gershenson, 2002, Gershenson, 2004, Saadatpour et al., 2010, Paulevé and Richard, 2012, Mizera et al., 2018].
- The updating scheme of BNs regulates the way that the nodes are updated through time evolution.
- **Synchronous BNs:**
 - ▶ Classical Random Boolean Networks (CRBNs)
- **Asynchronous BNs:**
 - ▶ Asynchronous Random Boolean Networks (ARBNs)
 - ▶ Generalized Asynchronous Random Boolean Networks (GARBNs)

Motivations

- In theory:
 - ▶ There are **few studies** exploring the relations among dynamics of different types of BNs [Paulevé and Richard, 2012].
- In practice:
 - ▶ The existing methods [Garg et al., 2008, Zheng et al., 2013, Mizera et al., 2018] for finding attractors of ARBNs **do not handle well large-scale networks**.
 - ▶ In [TRINH and HIRAISHI, 2020], we have proposed an efficient method called filtSAT for finding attractors of GARBNs. The **accuracy and efficiency of filtSAT** have been justified by the experiments on real biological and randomly generated networks. Especially, filtSAT can handle large-scale networks (e.g., **the T_Cell_Receptor_Signaling network with $n = 101$** [Helikar et al., 2012]).¹

¹ n denotes the number of nodes of a BN.

Contributions

- In theory:
 - ▶ We state and prove several [relations](#) in dynamics between ARBNs and GARBNs.
- In practice:
 - ▶ Based on these relations, we propose [an approximation method called ApproARBN](#) for efficiently finding attractors of large-scale ARBNs.
 - ▶ Experiments are also conducted to evaluate [the accuracy and efficiency](#) of our proposed method.
 - ▶ Experimental results consolidate the observation that the number of attractors of an ARBN is [equal to](#) that of its GARBN counterpart [in most cases](#).

Contents

- 1 Introduction
- 2 Preliminaries**
- 3 ARBNs and GARBNs
- 4 Approximation method
- 5 Experiments
- 6 Conclusion

Boolean networks

- A BN contains many nodes (genes), each node takes either value 0 (inactive) or 1 (active).
- Interactions between the nodes are expressed by **Boolean functions**.
- In this paper, BNs are implicitly considered as **random Boolean networks** (i.e., there is no restriction on Boolean functions).

Dynamics of Boolean networks

- At each time step, a BN can transit from a state to another state. This is the *state transition*.
- The dynamics of a Boolean network is captured by a *state transition graph* (STG).
- An STG is a directed graph which shows states (nodes) and state transitions (arcs).

Classical Random Boolean Networks

- CRBNs were proposed by Stuart Kauffman [Kauffman, 1969] to model gene regulatory networks in cells.
- They have the synchronous and deterministic updating scheme, i.e., **all nodes are updated synchronously**.
- The STG of a CRBN has 2^n nodes and 2^n arcs.

Asynchronous Random Boolean Networks

- Asynchronous Random Boolean Networks (ARBNs) were firstly studied by Harvey and Bossomaier [Harvey and Bossomaier, 1997].
- The updating scheme is asynchronous and non-deterministic. At each time step, **only one node is randomly selected** in order to be updated.
- The STG of an ARBN has 2^n nodes and up to $n \times 2^n$ arcs.

Generalized Asynchronous Random Boolean Networks

- GARBNs have the semi-synchronous and non-deterministic updating scheme [Gershenson, 2002].
- At each time step, they **select randomly any number of nodes to update synchronously**. This means that GARBNs can update synchronously no node, only one node, some nodes, or all the nodes.
- The STG of a GARBN has 2^n nodes and up to 2^{2n} arcs.

Classification of attractors

- We classify three types of attractors based on [Garg et al., 2008]:
 - ▶ singleton attractors
 - ▶ simple attractors
 - ▶ complex attractors
- Simple attractors can again be divided into two subclasses:
 - ▶ simple attractors (type1)
 - ▶ simple attractors (type2)

Classification of attractors (cont.)

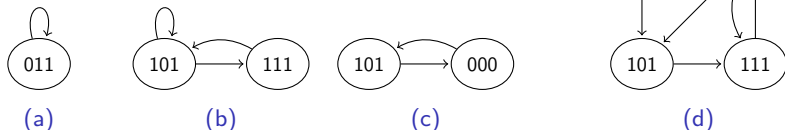


Figure: (a) Singleton attractor. (b) type1 attractor. (c) type2 attractor. (d) complex attractor.

A simple attractor is a cycle of at least two states such that each state has exactly one successor state (excluding itself) and may have a self transition.

- A type1 attractor: **any two consecutive states** of the attractor **differ in at most one** node.
- A type2 attractor: **at least two consecutive states** of the attractor **differ in more than one** node.

Example of Boolean networks

$$f_1 = x_1 \vee (\neg x_1 \wedge ((\neg x_2 \wedge x_3) \vee (x_2 \wedge \neg x_3)))$$

$$f_2 = (\neg x_1 \wedge \neg x_3) \vee (x_2 \wedge x_3) \vee (x_1 \wedge \neg x_2)$$

$$f_3 = \neg x_1 \vee (x_1 \wedge (\neg x_2 \vee (x_2 \wedge \neg x_3)))$$

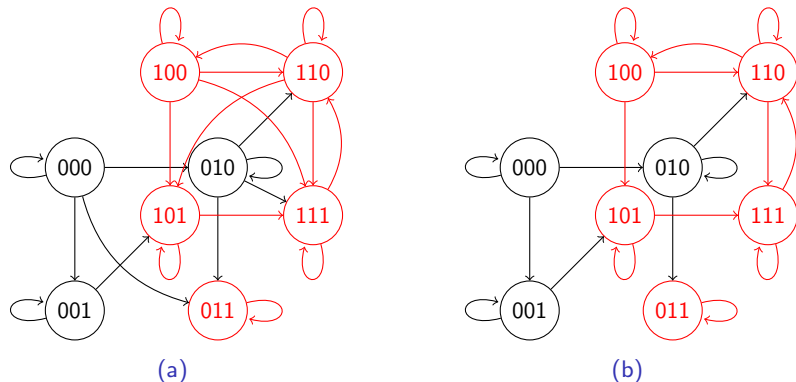


Figure: Example of BNs. (a) STG of the GARBN. (b) STG of the ARBN.

Contents

- 1 Introduction
- 2 Preliminaries
- 3 ARBNs and GARBNs**
- 4 Approximation method
- 5 Experiments
- 6 Conclusion

Relations between ARBNs and GARBNs

Proposition 1

Let \mathcal{A} be an ARBN and \mathcal{G} be its GARBN counterpart. \mathcal{A} and \mathcal{G} have the same set of singleton attractors.

Proposition 2

Let \mathcal{A} be an ARBN and \mathcal{G} be its GARBN counterpart. \mathcal{A} and \mathcal{G} have the same set of type1 attractors.

Proposition 3

Let \mathcal{A} be an ARBN and \mathcal{G} be its GARBN counterpart. Then $FR^{\mathcal{A}}(\{s\}) \subseteq FR^{\mathcal{G}}(\{s\})$ holds for any state s of \mathcal{A} , where $FR^{\mathcal{N}}(S)$ denotes the forward reachable set of the state set S of a BN \mathcal{N} .

Relations between ARBNs and GARBNs (cont.)

Theorem 1

Let \mathcal{A} be an ARBN and \mathcal{G} be its GARBN counterpart. Let $A^{\mathcal{N}}$ denote the set of attractors of a BN \mathcal{N} . There exists a mapping $m : A^{\mathcal{G}} \rightarrow A^{\mathcal{A}}$ with $m(att) \subseteq att$ for all $att \in A^{\mathcal{G}}$ and $m(att_1) \neq m(att_2)$ for all $att_1, att_2 \in A^{\mathcal{G}}, att_1 \neq att_2$.

Corollary 1

The number of attractors of an ARBN is greater than or equal to that of its GARBN counterpart.

We have reported in the paper an example for the **inequality** case of Corollary 1. However, we observed that **the equality case of Corollary 1 happens in most cases** [Gershenson, 2004, TRINH and HIRAISHI, 2020]. This suggests us to propose an efficient method for approximating the attractors of an ARBN based on the attractors of its GARBN counterpart.

Contents

- 1 Introduction
- 2 Preliminaries
- 3 ARBNs and GARBNs
- 4 Approximation method**
- 5 Experiments
- 6 Conclusion

The filtSAT method [TRINH and HIRAISHI, 2020]

- It efficiently finds attractors of a GARBN \mathcal{G} .
- It firstly finds attractors of the CRBN counterpart of \mathcal{G} .
- Then, it filters out the set of CRBN attractors by using SAT-based bounded model checking for checking the reachability in \mathcal{G} .
- The result of filtSAT includes $A_{sing}^{\mathcal{G}}$, $A_{type1}^{\mathcal{G}}$, and $F_{comp}^{\mathcal{G}}$. $F_{comp}^{\mathcal{G}}$ is a set of states of \mathcal{G} and each its state is expected to represent (i.e., belong to) a complex attractor of \mathcal{G} .
- filtSAT can handle **large networks** [TRINH and HIRAISHI, 2020].

The approximation method

- Based on filtSAT and the relations between ARBNs and GARBNs, we propose an efficient method called ApproARBN for approximating attractors of ARBNs.
- The ApproARBN method:
 - ▶ Given an ARBN \mathcal{A} and \mathcal{G} be its GARBN counterpart.
 - ▶ We firstly apply filtSAT to \mathcal{G} .
 - ▶ \mathcal{A} and \mathcal{G} share the **same sets of singleton and type1 attractors** by Propositions 1 and 2.
 - ▶ By Theorem 1, the forward reachable set of a state in $F_{comp}^{\mathcal{G}}$ in \mathcal{G} always contains an attractor of \mathcal{A} . Thus, we make a long random walk in the STG of \mathcal{A} for the set $F_{comp}^{\mathcal{G}}$. Making random walk aims at **making $F_{comp}^{\mathcal{G}}$ approach attractor states** of \mathcal{A} .
 - ▶ Finally, we obtain $F_{comp}^{\mathcal{A}}$ as the set of states of \mathcal{A} , and each state of $F_{comp}^{\mathcal{A}}$ is expected to belong to a complex attractor of \mathcal{A} .

The correctness of ApproARBN

- In this paper, we have analyzed in detail some cases in which the result of ApproARBN is **incorrect**.
 - ▶ $F_{comp}^{\mathcal{A}}$ may contain **redundant states** belonging to the same attractor of \mathcal{A} because $F_{comp}^{\mathcal{G}}$ may contain redundant states belonging to the same attractor of \mathcal{G} [TRINH and HIRAISHI, 2020].
 - ▶ $F_{comp}^{\mathcal{A}}$ also may contain **spurious states** which do not belong to any attractor of \mathcal{A} since the random walk does not guarantee always reaching attractor states.
 - ▶ Some complex attractors of \mathcal{A} may **not appear** in $F_{comp}^{\mathcal{A}}$ since the number of attractors of \mathcal{A} may be larger than that of \mathcal{G} .
- We have also provided some reasons convincing **the accuracy of ApproARBN** which will be confirmed by the conducted experiments on real biological networks.

Contents

- 1 Introduction
- 2 Preliminaries
- 3 ARBNs and GARBNs
- 4 Approximation method
- 5 Experiments**
- 6 Conclusion

Implementation

- For finding CRBN attractors, ApproARBN uses the SAT-based method by [Dubrova and Teslenko, 2011] if $n > 65$ and uses the decomposition-based method by [Yuan et al., 2019] otherwise.
- We have implemented ApproARBN in a JAVA tool.
- This tool uses the JDD library [Vahidi, 2003] for BDD manipulation and Z3 [De Moura and Bjørner, 2008] as the SAT solver.

Experiments

- All the experiments were run on a virtual machine whose environment is CPU: Intel(R) Xeon(R) Silver 4116 4x2.10GHz, Memory: 24 GB, CentOS 7 64 bit.
- The time limit was set to 10 hours.
- We applied ApproARBN, genYsis [Garg et al., 2008], and CABEAN [Mizera et al., 2018] to BNs of **32 real biological networks** whose sizes range from 19 to 101.

Experimental results: The first remark

- There are 26/32 networks in which CABEAN or genYsis succeeded to find attractors within the time limit. In all these networks (except the T_Cell_Receptor_Signaling network), the result of ApproARBN is always **correct**.
- In the T_Cell_Receptor_Signaling network, genYsis only returned the number of attractors, we thus **can not determine the correctness of ApproARBN** for this network.

Experimental results: The second remark

- Number of succeeded cases:
 - ▶ genYsis: 24/32
 - ▶ CABEAN: 22/32
 - ▶ ApproARBN: 31/32
- For the networks in which both genYsis and ApproARBN succeeded, ApproARBN **almost outperforms** genYsis.
- For the networks in which both CABEAN and ApproARBN succeeded, the running time of ApproARBN is **comparable** to the running time of CABEAN.
- Especially in the T_Cell_Receptor_Signaling network with $n = 101$, the running time of ApproARBN is only **20.84 seconds** while the running time of genYsis is **3369.89 seconds** and CABEAN failed to obtain the result within 10 hours.

Experimental results: The third remark

| name | n | ApproARBN | | | |
|---------------------------------|-----|-----------|-------|----------|----------|
| | | CRBN time | $ A $ | time | Correct? |
| Colitis_associated_colon_cancer | 70 | 701.90 | 10 | 873.26 | N/A |
| InflammatoryBowelDisease | 47 | 11944.1 | 1 | 11948.98 | N/A |

- In some networks, **most** of the running time of ApproARBN was spent for finding CRBN attractors and the time for finding CRBN attractors is long.
- The used methods for finding CRBN attractors may be **inefficient** for these networks. We can use a more efficient method or efficiently combine multiple methods.
- It is **potentially possible** since there have recently been many efficient methods (e.g., [He et al., 2018, Yuan et al., 2019]) for attractor detection in CRBNs.

Experimental results: The last remark

| name | n | # type2 | ApproARBN | | |
|--------------------------------|----|---------|-----------|----------|----------|
| | | | A | time | Correct? |
| ButanolProduction | 66 | 13312 | 8192 | 25711.04 | N/A |
| HGF_Signaling_in_Keratinocytes | 68 | 2757 | 72 | 8685.93 | Yes |

- In some networks, the running time of ApproARBN is **much long**.
- We can see that in these networks, the number of type2 attractors of the CRBN counterpart is **large**.
- Since the number of iterations of the filtering of ApproARBN is equal to this number, ApproARBN may be **inefficient** when it is large.
- Excluding redundant CRBN attractors before the filtering of ApproARBN may be a **potential improvement**.

Contents

- 1 Introduction
- 2 Preliminaries
- 3 ARBNs and GARBNs
- 4 Approximation method
- 5 Experiments
- 6 Conclusion**

Conclusion

- We have stated and proved several **relations** between dynamics of ARBNs and GARBNs.
- Based on these relations, we have developed an approximation method called ApproARBN for efficiently finding attractors of large-scale ARBNs.
- Experimental results justify the **accuracy** of ApproARBN and the **efficiency** of ApproARBN as compared to the two existing methods, genYsis and CABEAN.

Future work

- In some BNs of our experiments, **most of the running time** was spent for finding CRBN attractors. Thus, we can improve ApproARBN by **using a more efficient method or efficiently combining multiple methods** for finding CRBN attractors.
- Since ApproARBN may be **inefficient** when the number of type2 attractors of the CRBN counterpart is large, **excluding redundant CRBN attractors before the filtering of ApproARBN** may be a potential improvement.
- It is also interesting to analyze the theoretical and/or practical computational complexity of ApproARBN.

References I



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.



De Moura, L. and Bjørner, N. (2008).

Z3: An efficient SMT solver.

In *International conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340. Springer.






Dubrova, E. and Teslenko, M. (2011).

A SAT-based algorithm for finding attractors in synchronous Boolean networks.

IEEE/ACM Transactions on Computational Biology and Bioinformatics, 8(5):1393–1399.

References II

-  Garg, A., Di Cara, A., Xenarios, I., Mendoza, L., and De Micheli, G. (2008).
Synchronous versus asynchronous modeling of gene regulatory networks.
Bioinformatics, 24(17):1917–1925.
-  Gershenson, C. (2002).
Classification of random Boolean networks.
In *Proceedings of the Eighth International Conference on Artificial Life*, pages 1–8. MIT Press.
-  Gershenson, C. (2004).
Introduction to random Boolean networks.
CoRR, nlin.AO/0408006.

References III



Harvey, I. and Bossomaier, T. (1997).

Time out of joint: Attractors in asynchronous random Boolean networks.

In Proceedings of the Fourth European Conference on Artificial Life, pages 67–75. MIT Press, Cambridge.






He, Q., Xia, Z., and Lin, B. (2018).

P_UNSAT approach of attractor calculation for Boolean gene regulatory networks.

Journal of Theoretical Biology, 447:171–177.

References IV

-  Helikar, T., Kowal, B., McClenathan, S., Bruckner, M., Rowley, T., Madrahimov, A., Wicks, B., Shrestha, M., Limbu, K., and Rogers, J. A. (2012).
The cell collective: toward an open and collaborative approach to systems biology.
BMC Systems Biology, 6(1):96.
-  Huang, S. (2001).
Genomics, complexity and drug discovery: insights from Boolean network models of cellular regulation.
Pharmacogenomics, 2(3):203–222.
-  Kauffman, S. A. (1969).
Metabolic stability and epigenesis in randomly constructed genetic nets.
Journal of Theoretical Biology, 22(3):437–467.

References V



Kauffman, S. A. (1992).

The origins of order: Self-organization and selection in evolution. In *Spin Glasses and Biology*, pages 61–100. World Scientific.



Mizera, A., Pang, J., Qu, H., and Yuan, Q. (2018).

Taming asynchrony for attractor detection in large Boolean networks. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 16(1):31–42.






Paulevé, L. and Richard, A. (2012).

Static analysis of Boolean networks based on interaction graphs: A survey.

Electronic Notes in Theoretical Computer Science, 284:93–104.

References VI

-  Saadatpour, A., Albert, I., and Albert, R. (2010).
Attractor analysis of asynchronous Boolean models of signal transduction networks.
Journal of Theoretical Biology, 266(4):641–656.
-  TRINH, V. G. and HIRAISHI, K. (2020).
A study on attractors of generalized asynchronous random Boolean networks.
IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, 103(8):987–994.
-  Tun, K., Menghini, M., D'Andrea, L., Dhar, P., Tanaka, H., and Giuliani, A. (2011).
Why so few drug targets: a mathematical explanation?
Current Computer-Aided Drug Design, 7(3):206–213.

References VII



Vahidi, A. (2003).

JDD: a pure Java BDD and Z-BDD library.

<https://bitbucket.org/vahidi/jdd>.



Yuan, Q., Mizera, A., Pang, J., and Qu, H. (2019).

A new decomposition-based method for detecting attractors in synchronous Boolean networks.

Science of Computer Programming, 180:18–35.



Zheng, D., Yang, G., Li, X., Wang, Z., Liu, F., and He, L. (2013).

An efficient algorithm for computing attractors of synchronous and asynchronous Boolean networks.

PloS One, 8(4):e60593.

Thank you for your attention!