

# Towards a Shapley value propagation method for Boolean networks<sup>\*</sup>

Giang Pham and Paolo Milazzo

Department of Computer Science, University of Pisa

Largo B. Pontecorvo, 3, 56127, Pisa, Italy

tranhuong.giangpham@phd.unipi.it   paolo.milazzo@unipi.it

## 1 Extended abstract

Boolean networks are a common type of network often used to capture the qualitative information in biological systems, such as those involved in gene regulation processes. In a Boolean network, nodes correspond to Boolean variables (representing, for instance, activation states of different genes) and edges denote influences between nodes (representing, for instance, how the activation of some genes influences other genes). Since influence relationships can be complex, they are typically more properly represented by logic formulas that can precisely express the combinations of active nodes that cause the activation/inactivation of each other node.

Boolean networks are a common formalism for the modeling of gene regulation processes because of their simple nature and since they provide several approaches to study the dynamics of the system [1]. The network can be simulated to collect behavioral patterns of the system. Moreover, they allow an exhaustive description of all possible behaviors to be computed in terms of a transition graph (or transition system), that can then be explored to analyze behavioural patterns and attractors (i.e., steady states and sustained cyclic behaviours) [2].

In the context of the modelling of gene regulation processes through Boolean networks, it is common to consider some nodes of the network representing entities whose activation state depends on the environmental configuration as the *input* nodes of the network. Moreover, there are typically other nodes of the network that are of interest for the study, and that we can call the *target* nodes. Analysis of Boolean networks usually aims at describing how the different configurations of the input nodes impact on the activation of the target nodes, and also to evaluate the role of each other nodes in order to identify *key regulators* for the network (that could be, for instance, the target of a drug in case the modelled regulation network is responsible for a disease).

Shapley value was proposed as an unique solution for coalition games in game theory by Lloyd Shapley in 1953. Coalition games is a class of games in which players gather to form coalitions and each coalition is a subset of the full set of

---

<sup>\*</sup> Supported by project “MEDICA” (2022RNTYWZ) funded by MUR-PRIN\_2022, and by PNRR ECS00000017 - “THE - Tuscany Health Ecosystem” - Spoke 6 - CUP I53C22000780001 funded by the EC under the NextGeneration EU programme.

players. There is a function called *payoff function* which returns the *payoff value* for each coalition of players. The game can be modelled with a pair  $(N, v)$  in which  $N$  is called the carrier of the game, which is the set of all the players that participate in the game.  $v$  is the payoff function,  $v(S)$  returns the payoff value of the coalition  $S$ . The power set of  $N$  has  $2^N - 1$  non empty subsets,  $S$  belongs to this power set. The set of all the players  $N$  is called the grand coalition.

The main concern of the coalition game is to divide the payoff value of the grand coalition  $v(N)$  among  $N$  players. For this problem, Lloyd Shapley proposed an unique solution  $\phi$  that satisfies 3 axioms: symmetry; carrier; additivity [3]. Symmetry axiom states that if two players contribute the same amount to all of the coalition they join, then their value should be the same. Carrier axiom makes sure that the sum over the value of all the players in the carrier  $N$  is exactly the payoff value received by this carrier  $v(N)$ . This axiom leads to another axiom for the *null player* or *dummy player* who contributes 0 for all the coalitions, the value for this null player should be 0. The last axiom - additivity applies to 2 games  $v$  and  $w$ . For each player, the value of the game  $[v + w]$  which is the sum of two games ( $[v + w](S) = [v](S) + [w](S)$ ) is the sum of the value from each individual game  $\phi(v + w) = \phi(v) + \phi(w)$ .

The desirable properties of Shapley value makes it become popular in many fields, including systems biology. In such a context, Multi-perturbation Shapley Analysis (MSA) [4] is a method applied widely to many kinds of biological networks. Given a network, the main idea is to evaluate the contribution of the network nodes using the Shapley value. All the nodes of the network are considered as players, when some of the elements get knockout, the rest form a coalition. The payoff value for each coalition in this case is the result of the network with the remaining elements.

The Shapley value, then, can be calculated for all the elements within the network, and represents the contribution of each individual element on the network. This method is applied in many studies from feature selection in Blood Loss Severity Prediction [5] to analysing the the multivariate lesion of clinical stroke data [6] and the lesion-symptom relationship study [7], to name a few. However, one problem of MSA is that it requires to perform knockout for every individual node and every possible combination of nodes. This is time consuming and sometimes unfeasible. For example, to apply MSA to a Boolean network in order to get the payoff value (defined as a function on the target nodes) for each coalition (defined as a subset of input nodes), it is needed to simulate the network after every knockout of every node or combination of nodes. Hence, in this paper, to tackle such a difficulty, we propose a method to *propagate* the Shapley value of input nodes through the Boolean network. This would allow us to determine the contribution of each intermediate node on the target without the need of performing simulations.

In the presentation, we will discuss preliminary ideas and results about the development of such a propagation method.

## References

1. R. Barbuti, R. Gori, P. Milazzo, and L. Nasti, “A survey of gene regulatory networks modelling methods: from differential equations, to boolean and qualitative bioinspired models,” *Journal of Membrane Computing*, vol. 2, pp. 207–226, 2020.
2. J. D. Schwab, S. D. Kühlwein, N. Ikonomi, M. Kühl, and H. A. Kestler, “Concepts in boolean network modeling: What do they all mean?,” *Computational and structural biotechnology journal*, vol. 18, pp. 571–582, 2020.
3. L. S. Shapley *et al.*, “A value for n-person games,” 1953.
4. A. Keinan, B. Sandbank, C. C. Hilgetag, I. Meilijson, and E. Ruppin, “Fair attribution of functional contribution in artificial and biological networks,” *Neural computation*, vol. 16, no. 9, pp. 1887–1915, 2004.
5. A. Razi, F. Afghah, A. Belle, K. Ward, and K. Najarian, “Blood loss severity prediction using game theoretic based feature selection,” in *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 776–780, IEEE, 2014.
6. M. Zavaglia, N. D. Forkert, B. Cheng, C. Gerloff, G. Thomalla, and C. C. Hilgetag, “Technical considerations of a game-theoretical approach for lesion symptom mapping,” *BMC neuroscience*, vol. 17, pp. 1–15, 2016.
7. S. Ofir-Geva, I. Meilijson, S. Frenkel-Toledo, and N. Soroker, “Use of multi-perturbation shapley analysis in lesion studies of functional networks: The case of upper limb paresis,” *Human Brain Mapping*, vol. 44, no. 4, pp. 1320–1343, 2023.