A. Gates and L.M. Rocha [2014]."Structure and dynamics affect the controllability of complex systems: a Preliminary Study". Artificial Life 14: Proceedings of the Fourteenth International Conference on the Synthesis and Simulation of Living Systems: 429-430, MIT Press.

Structure and dynamics affect the controllability of complex systems: a Preliminary Study

Alexander J. Gates[1,2] and Luis M. Rocha[1,2,3]
[1]School of Informatics and Computing, Indiana University, Bloomington, IN 47406, USA
[2]Cognitive Science Program, Indiana University, Bloomington, IN 47406, USA
[3]Instituto Gulbenkian de Ciencia, Oeiras, Portugal

Complex systems are typically understood as large nonlinear multivariate systems. Their organization and behavior are commonly modeled by representations such as graphs and automata networks. Graphs, where nodes representing variables lack intrinsic dynamics, capture the *structure* or organization of complex systems. The simplest way to study multi-variate *dynamics*, is to allow network nodes to have states and update them with automata; for instance, Boolean networks (BN) are canonical models of complex systems and exhibit a wide range of dynamical behaviors [1].

The structure of networks has provided many insights into the organization of complex systems [2]. The success of this approach is its ability to capture the organization of complex systems, and how it changes in time (network evolution) without explicit dynamical rules for node variables. As the field matures, however, there is a need to move from understanding to controlling complex systems. This is particularly true in systems biology and medicine, where increasingly accurate models of biochemical regulation have been produced [3]. More than understanding the organization of biochemical regulation, we need to derive *control* strategies that allow us, for instance, to move a mutant cell to a wild-type state [4], or revert a mature cell to a pluripotent state [5]. Towards these goals, a question of central importance remains: How well does network structure represent the multivariate dynamics of the underlying complex system, especially from the point of view of control?

Network controllability aims precisely to identify a minimal set of *driver variables* (a.k.a. driver nodes) from the structural network, which can fully control system dynamics—i.e. drive system dynamics to any state-space configuration [6]. *Structural controllability* is an influential method to derive driver variables, using only structural properties of the system without any consideration of its dynamical details [7]. It has been used to suggest, for instance, that biological systems are harder to control than social systems [8]. However, applications of structural controllability have been heavily critiqued due to its stringent assumptions [9].

Here we explore the relationship between network structure and controllability through the analysis of dynamical ensembles of BN. The control problem for general BN is computationally intractable (NP-hard) [10]. Simplification techniques, such as structural controllability or those based on dynamical redundancy [4], are thus highly desirable. The BN studied here are discrete dynamical systems $\mathcal{X} = \{x_i\}$ of N Boolean variables $x_i \in \{0, 1\}$ that are updated synchronously according to deterministic logical functions. The structural network specifies all directed pairwise interactions $\{e_{ij}\}$ which indicate when variable x_i is an input for the logic of x_j. At time t, the network is in a configuration $\boldsymbol{X}^t$, which is the vector of all variable states $(x_i(t))$ at time t. The overall dynamics of temporal sequences of network configurations can be represented by the state-transition graph (STG). In this graph, each node is a configuration $\boldsymbol{X}^t$ of the BN and each directed edge is a transition from $\boldsymbol{X}^t$ to $\boldsymbol{X}^{t+1}$. Thus, the STG fully describes the 2^N possible configurations and transitions in the network's dynamical landscape.

We study the control due to a subset $\mathcal{D} \subset \mathcal{X}$ of driver variables as instantaneous perturbations to the variable's state. To capture all possible dynamically allowable trajectories due to controlled interventions on $\mathcal{D}$, we introduce the *controlled state transition graph* ($\text{CSTG}_\mathcal{D}$). The $\text{CSTG}_\mathcal{D}$ is an extension of the STG where additional edges connect a configuration to each of its $2^{|\mathcal{D}|} - 1$ perturbed counterparts. The network is fully controllable if a trajectory exists between every pair of configurations; for BN this is equivalent to requiring the $\text{CSTG}_\mathcal{D}$ to be strongly-connected.

We extend this binary notion of controllability by tallying the fraction of configurations that are controlled by driver set $\mathcal{D}$. Given a specific configuration $\boldsymbol{X}_k$, the fraction of reachable configurations $r_{\text{CSTG}_\mathcal{D}}(\boldsymbol{X}_k)$ is the number of other configurations reachable on graph $\text{CSTG}_\mathcal{D}$ via a directed path starting from $\boldsymbol{X}_k$, normalized by $2^N - 1$. The *mean fraction of reachable configurations* is then given by $\overline{R} = \langle r_{\text{CSTG}_\mathcal{D}}(\boldsymbol{X}_k)\rangle_k$, where $k = 1 \dots 2^N$. It measures the fraction of configurations which are on average reachable by controlling the variables in $\mathcal{D}$. The *mean fraction of controlled configurations* is $\overline{C} = \langle r_{\text{CSTG}_\mathcal{D}}(\boldsymbol{X}_k) - r_{\text{STG}}(\boldsymbol{X}_k)\rangle_k$. It measures the average fraction of controlled configurations by discounting those transitions which would have naturally

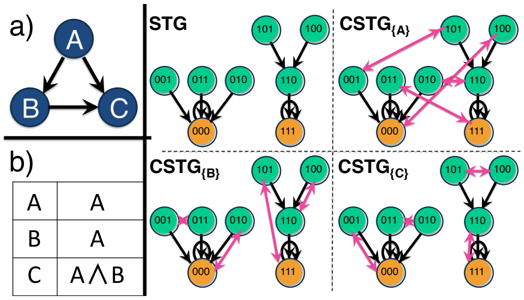

Figure 1: a) FeedForward network motif structure with dynamics given by the rules in b) give rise to the State-Transition Graph (STG). Each of the singleton sets $\{A\}$, $\{B\}$, and $\{C\}$ are used as driver variable sets to produce the Controlled STGs ($\text{CSTG}_{\mathcal{D}}$) shown.

occurred due to the uncontrolled dynamics of the network. A fully controlled network must have $\overline{R} = 1.0$, but partially controlled networks will vary in [0,1].

We start by characterizing the entire ensemble of possible BN dynamics for the $N = 3$ variable Feed-Forward network motif [11] shown in Figure 1a. For each network structure, there are $L = \sum_i 2^{k_i}$ possible transition rules, where k_i is the in-degree of variable x_i. In this case, the full ensemble consists of $2^L = 2^6 = 64$ distinct networks; the logic of one is depicted in Figure 1b. This figure also depicts its STG and the $\text{CSTG}_{\mathcal{D}}$ for various driver sets $\mathcal{D}$. If we additionally remove all variable transition functions that refer to tautologies, contradictions, and functions always controlled by a single input (fully canalizing [12]), we obtain a smaller Non-trivial ensemble. The controllability analysis for both ensembles is shown in Figure 2. Notice that structural controllability analysis predicts that variable A is capable of fully controlling the network. However, this driver variable ($\text{CSTG}_{\{A\}}$) fails to control a large majority of the possible BN. Even pairs of variables cannot fully control all networks ($\overline{R} < 1.0$); to guarantee full control for every network, all three variables need to be driven.

This simple example highlights the tenuous relationship between structure and dynamics for complex systems, and the implications for understanding and characterizing their control. This work has been extended in four significant ways [13], and will be showcased at the conference: (1) analysis of several more network motifs; (2) extend our measures of partial control to the more biologically relevant concept of attractor control (since all non-attractor configurations are transients); (3) study the space of random BN (such as Erdos-Renyi structural networks with transitions parameterized by in-degree and bias) to establish that similar results hold as the systems are scaled up; (4) study models of biochemical regulation such as the segment polarity gene-regulatory BN of *Drosophila melanogaster* [14], the control of which is known [4]. These studies show how the structure-only analysis of complex systems tends to fail to properly characterize their full or partial controllability. Additionally, we lay the groundwork for understanding which restrictions must be enforced on the transition functions of BN, such that structure may suffice for predicting controllability. Our full study and analysis has been submitted for publication [13].

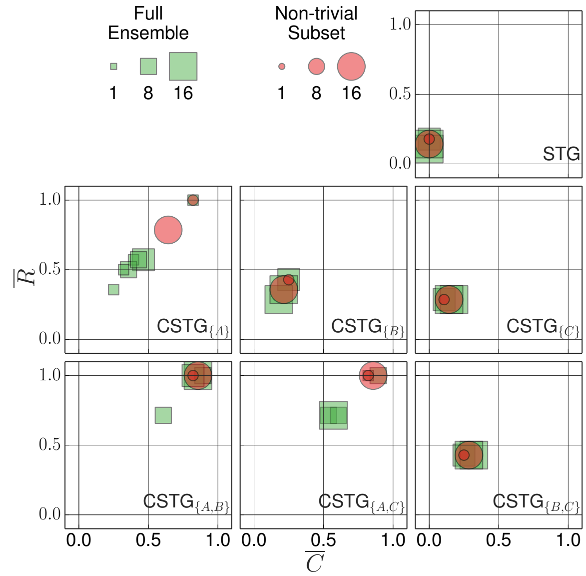

Figure 2: Measures $\overline{R}$ and $\overline{C}$ for the Full Ensemble of 64 BN (green squares and red circles) with structure given by the Feed-Forward network motif shown in Figure 1, as controlled by various combinations of driver variables. The Non-trivial subset is highlighted by red circles.

References

[1] Kauffman, S. A. (1993) *The Origins of Order: self-organization and selection in evolution*. (Oxford University Press).

[2] Boccaletti, S, Latora, V, Moreno, Y, Chavez, M, & Hwang, D.-U. (2006). *Physics Reports* **424**, 175–308.

[3] R, Albert & S.M. Assmann. (2009). *Plant Systems Biology* **35**, 41–66.

[4] Marques-Pita, M & Rocha, L. M. (2013). *PLOS ONE* **8**, e55946.

[5] Wang, J, Zhang, K, Xu, L, & Wang, E. (2011). *PNAS* **108**, 8257–8262.

[6] Valente, T. (2012). *Science* **337**, 49–53.

[7] Liu, Y.-Y, Slotine, J.-J, & Barabási, A.-L. (2011). *Nature* **473**, 167–173.

[8] Egerstedt, M. (2011). *Nature* **473**, 158–159.

[9] Cowan, N. J, Chastain, E. J, Vilhena, D. A, Freudenberg, J. S, & Bergstrom, C. T. (2012). *PLOS ONE* **7**, e38398.

[10] Akutsu, T, Hayashida, M, Ching, W.-K, & Ng, M. (2007). *J. Theor. Biol.* **244**, 670–679.

[11] Alon, U. (2007). *Nature Reviews Genetics* **8**, 450–461.

[12] Kauffman, S, Peterson, C, Samuelsson, B, & Troein, C. (2004). *PNAS* **101**, 17102–17107.

[13] Gates, A. J & Rocha, L. M. (2014) Control of complex networks requires structure and dynamics. *submitted*.

[14] Albert, R & Othmer, H. (2003). *J. Theor. Biol.* **223**, 1–18.