**Boolean networks as a way to explore dynamics in ecological assemblages under environmental change**

Paulo R. Guimarães Jr & Carlos J. Melián

1. **Introduction**

A key challenge in ecology lies in comprehending the resilience and vulnerability of ecological assemblages. These assemblages serve as quintessential examples of complex adaptive systems, marked by the persistent diversity and individuality of their constituent elements. Within such systems, a selective process operates, favoring the differential replication or increase of a fraction of the elements. Moreover, localized interactions among these elements define the intricate nature of complex adaptive systems. In species-rich ecological assemblages (Figure 1A), the sustained diversity of individuals and species forms the foundational material upon which natural selection exerts its influence. The diversity and selective pressures, in turn, are ultimately shaped by ecological interactions among individuals of various species. Consequently, these interactions emerge as pivotal factors for comprehending the dynamic patterns and enduring presence of ecological assemblages over time.

Traditionally, the established approach for investigating the dynamics of ecological assemblages involves the utilization of sets of differential equations. In the examination of ecological systems, the customary state variables used to portray these assemblages are the abundances (or biomasses) of each species. The set of equations articulates how the abundances of co-occurring species evolve over time, influenced by both the environment, the intrinsic features of each species, and the ecological interactions between species. The study of differential equations has benefited from centuries of robust theoretical development, making substantial contributions to our understanding of ecological systems. This approach facilitates the characterization of trajectories, outcomes, feasibility, and stability of ecological systems (Figure 1B-E). Nevertheless, challenges arise when employing differential equations to explore the dynamics of species-rich ecological systems.

Three challenges arise when applying differential equations to species-rich assemblages. Firstly, monitoring the abundances of each species in a species-rich ecological assemblage in nature proves to be an impractical task. Obtaining a one-time estimate of the abundances for a single species is already challenging and requires an extensive amount of fieldwork; this difficulty amplifies exponentially when dealing with dozens of species. Secondly, many observed responses in ecological systems, driven by ecological interactions, appear qualitative in nature. Systems experience abrupt shifts in the effects of ecological interactions, leading to sudden changes in system features that are often difficult to reproduce using smooth functions. Finally, the number of parameters involved in the set of equations describing the ecological system rapidly escalates with the number of species (S). If we assume that each species interact with *k* other species, the total number of parameters in the set of equations increases explosively as a function of *PS*+*QSk*, where *PS* represents the parameters controlling the environmental and intrinsic effects of each species (e.g., parameters governing intraspecific competition), and *QSk* denotes the total number of parameters controlling the effects of ecological interactions (see Figure 1F). The daunting task of estimating these parameters complicates the characterization of species-rich empirical systems.

In this paper, we delve into the utilization of an approach based on Boolean networks to address these challenges. While Boolean Networks are widely used in other scientific disciplines, their application in ecology is still in its early stages. We commence by elucidating Boolean Networks, along with the associated challenges in applying these methodologies to species-rich systems. Subsequently, drawing inspiration from statistical mechanics thinking, we propose alternatives to tackle two specific challenges inherent in applying Boolean Network approaches to species-rich ecological assemblages—specifically, the formulation of truth tables for each species and the characterization of system states. Finally, we integrate the statistical mechanics approach to Boolean networks with dynamic interaction landscapes to investigate how various forms of environmental changes may reorganize ecological assemblages.

*Figure 1’s idea: (A) an ecological network, (B) a plot describing abundances trajectories of some species, (C) fixed points for abundances, (d) stability of fixed points, (e) feasibility, (f) explosion of the number of parameters*

1. **Boolean Networks**

*Basic features*

Boolean networks are models developed to describe the dynamics of complex systems. Boolean networks have found applications in various fields, including biology (such as gene regulatory networks), computer science, physics, and social sciences, providing a simplified yet effective representation of dynamic processes in complex systems. Similar to other modeling approaches, we begin by describing the underlying network of interactions. In this context, the elements of a system are designated as nodes, and the interactions between these elements are portrayed as links, constituting a network. Each node is linked to a binary variable, indicating that it can exist in one of two states: either 0 or 1 (off or on). Links are affiliated with Boolean functions, which dictate the state of each node based on the states of the interacting nodes.

The two states of a node in a Boolean network can be interpreted in multiple ways. The most fundamental interpretation for ecological assemblages is to associate "off" with extinct species and "on" with extant species. However, this is not the only possible interpretation. "Off" could also represent "low abundance/biomass," while "on" indicates high abundance/biomass. Alternatively, "off" and "on" can describe two distinct prevalent phenotypes with different roles at the community level. In this context, Boolean networks offer a way to circumvent the challenge of tracking species abundances when studying the ecological dynamics of species-rich systems. Given the difficulty of obtaining reliable estimates of abundances for multiple species over time in a single site, particularly in species-rich systems, a potential solution is to adjust the scale of description. This entails characterizing the system in a coarse but reliable manner, using binary descriptors such as "absence" and "presence" or "low" and "high." The concept of changing the scale of description from continuous abundances to coarse, binary descriptors is commonly employed in ecology for various problems, including metapopulation/metacommunity dynamics and describing patterns of species occurrence at broader scales.

The Boolean Network approach also allows for overcoming the challenges of modeling abrupt transitions and the seemingly qualitative changes driven by ecological interactions. This circumvention is partly achieved through a simpler, coarser description of the state of each species ("on"/"off") rather than a quantitative state variable such as abundance. Another crucial aspect is the use of logical rules that streamline the description of non-linear effects of ecological interactions. For instance, if the presence of species *j* is sufficient for the presence of species *i*, regardless of other species interacting with *i*, a Boolean rule can be coded stating that if species *j* is "on," then species *i* is "on." Alternatively, if the presence of a species is shaped by compensatory effects of its partners, a Boolean rule can be coded indicating that species *i* is "on" if the interacting species *j* AND/OR species *k* are "on". Now, we illustrate the Boolean Networks approach using empirical examples as references.

*Illustrating Boolean Network approach using ecological systems*

First, let's consider the pairwise dynamics of the Canada Lynx (*Lynx canadensis*) and the Snowshoe Hare (*Lepus americanus*). Empirical evidence reveals rise-and-fall cycles in the abundances of both species. This system can be modeled as a Boolean network, where the Canada Lynx is "on" if the Snowshoe Hare is "on," and the Canada Lynx is "off" if the Snowshoe Hare is "on." Conversely, the Snowshoe Hare is "on" if the Canada Lynx is "off," and the Snowshoe Hare is "off" if the Canada Lynx is "on." These rules allow for the qualitative capture of the rise-and-fall cycles, described here as sequential on-off repeating patterns (see Figure 2A-B). The aforementioned rules can be organized into tables for each species. The set comprising all possible combinations of input states and the corresponding output states, based on the Boolean functions associated with each species, is referred to as the truth table for that species. This truth table comprehensively enumerates all possible combinations of binary states for the input nodes and displays the resulting states for the output nodes.

Now, let us delve into a more intricate example: a cascading effect within the kelp forests of the Pacific North Coast. Within these systems, a trophic cascade involves sea otters, sea urchins, and kelp forests. Sea otters (*Enhydra lutris*) prey on sea urchin species, such as *Mesocentrotus franciscanus* and *Strongylocentrotus purpuratus*, creating a top-down effect. Similarly, when sea urchins graze in large numbers, it results in the destruction of kelps (Laminariales). This trophic cascade triggers effects that propagate throughout the ecological network, leading to bottom-up effects that may influence various ecological features. These effects encompass ecosystem matter cycles, predatory behavior of bald eagles (*Haliaeetus leucocephalus*), and potentially the extinction of a marine megaherbivore, the Steller’s Sea cow (*Hydrodamalis gigas*) (Bullen et al. 2021, Estes PNAS’s paper).

One can model this system in its simplified form as a network comprising five nodes: hunters, sea otters, sea urchins, other sea otter resources (e.g., crabs), and kelp forest. The system’s dynamics can be represented with the following rules: (1) sea otters are "on" only if hunters are "off" and (crabs and/or sea urchins) are "on"; (2) sea urchins are "on" only if sea otters are "off"; (3) crabs are "on" only if kelp forests are "on"; (4) kelps are "on" only if sea urchins are "off." In this system, activating hunters "on" generates a trophic cascade, qualitatively reproducing the collapse of kelp forests and leading to the alternative state of the system: a barren seabed with a high density of sea urchins.

*Challenges in applying the Boolean Network approach*

Creating logical rules associated with the states of interacting species allows for the generation of various ecological effects in Boolean networks. While Boolean networks offer a natural way to describe qualitative changes in the system, the challenge lies in parameterizing the model. Specifically, it is difficult to specify the state of a given species *i* based on the combination of the states of each of its interacting partners. For instance, if species *i* interacts with *k* other species in the community, one needs to specify the outcomes of *2k* combinations for species *i* in its truth table. Assuming all *S* species have *k* interactions, we will have *S2k*, which increases rapidly with both *S* and, particularly, with *k*. Thus, the Boolean network framework, by itself, does not circumvent the issues of a large number of parameters that one needs to specify for simulations.

The best solution would undoubtedly be to incorporate information about the natural history of the system. However, even decades of detailed study of natural history may not fully elucidate the potential of ecological interactions to control the presence/absence of interacting partners in the community. In this situation, one alternative is to assign random truth tables for each species (the Random Boolean Network framework). However, the amount of information that needs to be stored for systems with large species richness (*S*) and a large number of interactions per species (*k*) is prohibitive even if randomly generated. Moreover, Boolean networks are known to generate very complex dynamics, even for moderate *S* and low *k*, often more challenging to characterize than steady states in systems of differential equations. Finally, to explore changes in ecological assemblages caused by environmental change, we need to incorporate ways to reorganize the topology of the underlying network. This requires finding solutions to circumvent these challenges when applying the Boolean networks framework to species-rich ecological systems. One possible alternative is to shift to statistical descriptions of the truth tables and outcomes at the community level of the Boolean dynamics. This option will be explored next. Then, we will address the reorganization of network topology through dynamic interaction landscapes.

*Statistical Mechanics thinking applied to Boolean Networks*

We first describe the macroscopic state of the system. Classically, the state of a Boolean network is described by a vector **Vt** that describe the state of each node at the time step *t*. Therefore, **Vt** describes the microstate in which the system is at the time step *t*. These microstates may show a variety of patterns that are often difficult to describe in detail, especially for species-rich systems (Figure 3). To come up with simple

We define three metrics that can summarize microstates that share similar features (i.e., macrostate). First, we define *Nt*as the proportion of nodes in the state “on” at the time step *t*. Second, we define , which is defined as . So, measures the differences between time steps in the composition of species in the Boolean networks. Third, we define , which is defined as , in which () is the state of the node *i* at the time *t*(*t-1*). So, measures the differences between time steps in the composition of species in the Boolean networks. These three macrostate descriptors capture several fundamental dynamical properties of Boolean Network dynamics (Figure 4).

*Figure 3. Plots illustrating the microstates of Boolean networks*

*Figure 4. Plots combining for combinations of N, , . (A) N* vs *, (B) N* vs *, (C)* vs .

We now turn our attention to the problem of truth tables. Instead of defining a single truth table for each species, we describe classes of species. To do so, we first define two types of links:

1. if species *j* is “on” favors species *i* to be “on” and if species *j* is “off” favors species *i* to be “off”, e.g., *j* is a resource or facilitate the presence of species *i*.
2. if species *j* is “on” favors species *i* to be “off”, e.g., *j* is a natural enemy of species *i,* and if species *j* is “off” favors species *i* to be “on”, e.g., *j* is a natural enemy of species *i*, such as a competitor, parasite or predator.

Then, we define five ways the effects can be computed to each species:

1. Additive
2. Additive
3. Multiplicative
4. Multiplicative
5. **Dynamic Interaction Landscapes**