OF THE COMPLEXITY OF BOOLEAN NETWORK STATE TRAJECTORIES

T. Mäki-Marttunen[1,2], J. Kesseli[1], S. Kauffman[3], O. Yli-Harja[1] and M. Nykter[1]

[1]Department of Signal Processing, Tampere University of Technology,
[2]Department of Mathematics, Tampere University of Technology,
P.O. Box 553, FI-33101 Tampere, Finland
[3]Complex Systems Center, University of Vermont, Burlington, VT 05405
tuomo.maki-marttunen@tut.fi, kesseli@cs.tut.fi, Stuart.Kauffman@uvm.edu,
yliharja@cs.tut.fi, matti.nykter@tut.fi

ABSTRACT

We study the complexity of network dynamics in a couple of very different model classes: The traditional random Boolean networks (RBN) and Frisch-Hasslacher-Pomeau lattice gas automata (FHP). For this we formulate the FHP dynamics as a probabilistic Boolean network (PBN). We use the set complexity of successive network states to assess the complexity of the dynamics. We find that the complexity is maximised near a transition state in both types of dynamical systems.

1. INTRODUCTION

Boolean networks are one of the simplest existing dynamical systems, yet they can produce an extremely wide range of different observable types of behavior. This makes them a suitable target for complexity research: a broad diversity of dynamics that is reducible to simple building blocks. Boolean network models have been used to establish and study several fundamental properties of dynamical systems. These include, among others, characterization of attractor structure, information processing properties, dynamical regimes, and ability to store information [1]. Remarkably, many of these properties are present in multiple other classes of dynamical systems, and some of these properties have been found in real systems, such as living cells.

While Boolean networks are a highly useful model class, there are multiple limitations. Some of these become apparent when one wants to understand the complexity of processes that pertain to physical quantities such as work or energy. Additionally, standard random Boolean model does not capture spatial positioning of the nodes. To extend the analysis beyond random Boolean networks, we utilize a model class that captures spatial positioning and pertains to the quantities such as work and energy: a lattice gas model. We use both of these model classes to show that our observations about the dynamical behavior holds also for both of these systems. Finally, we will establish a connection between lattice gas model and probabilistic Boolean networks. This connection allows comparison and generalization of results between these systems.

In this work we will study complex behavior of dynamical systems during the transition towards a steady state. Specifically, we address the complexity of the trajectories during this transition. Using an information theoretical measure, set complexity, we show that the maximally complex dynamical behavior is observed during the transition period. Additionally, we show that this property is shared between Boolean networks and lattice gas model.

2. METHODS

2.1. Random Boolean networks

A Boolean network is defined as a collection of nodes $\{V_1, \ldots, V_N\}$ where at each time step t each node is assigned a Boolean value $x_i(t)$, i.e.

$$\forall t \in \mathbb{N}\, \forall i \in \{1, \ldots, N\} : x_i(t) \in \{0, 1\}.$$

Each node receives input from 0 to N nodes and the state of the node at time instant $t+1$ is a Boolean function of the states of its neighbors at time instant t:

$$x_i(t+1) = f_i(x_{I_{i1}}(t), \ldots, x_{I_{in_i}}(t)),$$

where I_{ij} ($j \in \{1, \ldots, n_i\}$) are the indices of the input nodes of node i, $n_i \in \{0, \ldots, N\}$ denoting the number of them.

By Random Boolean Networks (RBN) we mean such Boolean networks where the inputting nodes as well as the Boolean functions are picked by random. We denote the probability of the output of a Boolean function in the network being 1 by p. For simplicity, we keep the number of inputting nodes a network-wide constant: $\forall i : n_i = K$. The parameters p and K together determine the dynamics of the network through a sensitivity parameter $s = 2Kp \cdot (1-p)$. Networks with $s = 1$ are critical networks, as networks with $s > 1$ are chaotic and $s < 1$ stable [2].

RBNs should not be confused with probabilistic Boolean networks (PBN) that are such Boolean networks where each node may have a number of functions. The choice of which function to use to update the state of the node depends on a random process. In the present paper we will use PBNs but skip the formal definitions of the underlying probability spaces.

2.2. Lattice gas

Movement of gaseous particles in a 2-dimensional box can be studied using a Frisch-Hasslacher-Pomeau (FHP) lattice gas model [3]. This is a model where the spatial domain is divided into a hexagonal grid, and each hexagonal box in the grid can contain up to six particles. The particles always move into one of the six directions and can only move with velocity of one hexagon/time step. Given that two or more particles are in the same hexagon, they must have different momenta, that is, different directions of momentum.

The momentum of a particle can only change with 1) a collision with the wall and 2) a collision with another particle. A collision with a wall occurs whenever a particle enters a hexagon occupied by a wall, and a collision between particles takes place if a) exactly two particles with opposite momenta come to the same hexagon or b) exactly three particles with momenta that add up to 0 come to the same hexagon. In collisions 1 and 2b the momenta of the particles is changed in a deterministic way whereas in 2a the momenta of the particles are chosen by random between the two possibilities (see Figure 1).

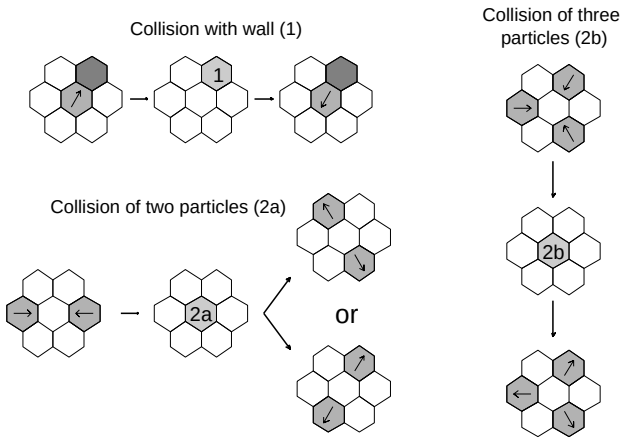
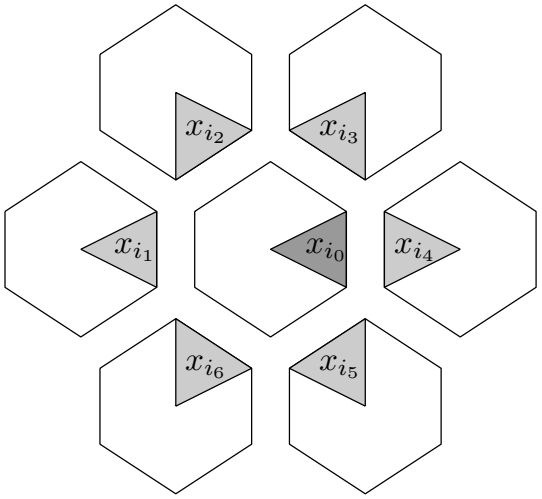


Figure 1. The collisions in FHP model.

The lattice gas automaton described above can be presented as a PBN. Each hexagon in the spatial grid contains six nodes, each of which can contain a particle moving into one of the six directions, i.e. the state of the node is '1' when there is a particle in the hexagon moving into the specified direction and '0' when there is not. The state of the node at time instant $t+1$ is determined by the state of its surrounding nodes at time instant t. The update rule is deterministic in all cases but collision 2a — in that case the outcomes of the four nodes of the possible scattering directions are random but correlate and anti-correlate with each other. The Boolean functions of a node V_{i_0} whose parent hexagon does not contain a wall are listed in Figure 2. If in contrast the parent hexagon contains a wall, then the only input node is in the opposite direction of the node (i.e. V_{i_4} would be input for V_{i_0} in Figure 2), and the Boolean function is identity.



f :

000000: 0	010000: 0	100000: 1	110000: 1
000001: 0	010001: 0	100001: 1	110001: 1
000010: 0	010010: 0/1	100010: 1	110010: 1
000011: 0	010011: 0	100011: 1	110011: 1
000100: 0	010100: 0	100100: 0	110100: 1
000101: 0	010101: 1	100101: 1	110101: 1
000110: 0	010110: 0	100110: 1	110110: 1
000111: 0	010111: 0	100111: 1	110111: 1
001000: 0	011000: 0	101000: 1	111000: 1
001001: 0/1	011001: 0	101001: 1	111001: 1
001010: 0	011010: 0	101010: 0	111010: 1
001011: 0	011011: 0	101011: 1	111011: 1
001100: 0	011100: 0	101100: 1	111100: 1
001101: 0	011101: 0	101101: 1	111101: 1
001110: 0	011110: 0	101110: 1	111110: 1
001111: 0	011111: 0	101111: 1	111111: 1

Figure 2. The input nodes of a node moving right in the Boolean network representation of the FHP model. Here, the state of the node x_{i_0} is determined as $x_{i_0}(t+1) = f(x_{i_1}(t), x_{i_2}(t), x_{i_3}(t), x_{i_4}(t), x_{i_5}(t), x_{i_6}(t))$. The values of function f are listed on the right for each input. In the type 2a collisions both values '0' and '1' are possible: in these cases the scattering direction of the collision is picked by random and the outcome of the node i_0, together with the other nodes in the hexagon, depends on this scattering direction.

2.3. Set complexity

Complexity of Boolean network dynamics can be studied by the means of the context-dependent information it carries. The dynamics of a Boolean network is represented by a set of its successive states that are read into strings. To the obtained set of strings one can apply a recently proposed all-purpose measure, *set complexity* [4], defined as:

$$S(\{x_1, ..., x_N\}) = \sum_j C(x_j) \frac{1}{N(N-1)} \sum_{j \neq k} d_{jk}(1 - d_{jk}),$$

where $C(x_j)$ is the Kolmogorov complexity — or its approximation — of string x_j. The variable d_{jk} stands for the *normalized compression distance* (NCD) of strings x_j

and x_k, defined as

$$d_{jk} = \mathrm{NCD}(x_j, x_k) = \frac{C(x_j x_k) - \min(C(x_j), C(x_k))}{\max(C(x_j), C(x_k))},$$

where $x_j x_k$ is the concatenation of strings x_j and x_k. In the present study we use the length of the LZMA encoded string as an approximation for the Kolmogorov complexity of the string.

3. RESULTS

3.1. Information content of dynamics of RBN

We consider RBNs of size $N = 1000$ with variable number of neighbors ($K = 1, 2, 3$). The state of the network is read into a string of length N, and the complexity of the dynamics is approximated by the set complexity of successive states. Figure 3 shows the set complexity of the states using a sliding window of four successive states. At the transition to a short-cycle steady state the set complexity is maximised, as seen for networks with sensitivities $s = 0.5$ and $s = 1$. In the case of $K = 3, s = 1.5$ the dynamics of the network is too chaotic for the set complexity of four states to find any pattern in — in our simulations a network of this type never returned to one of its previous states during the $T = 224$ time steps, indicating a lengthy path to a steady state, a long steady state cycle, or both. For comparison, the median steady state cycle lengths are 1 for all networks with $s = 0.5$, 16 for networks with $K = 2, s = 1$, and 10 for networks with $K = 3, s = 1$.

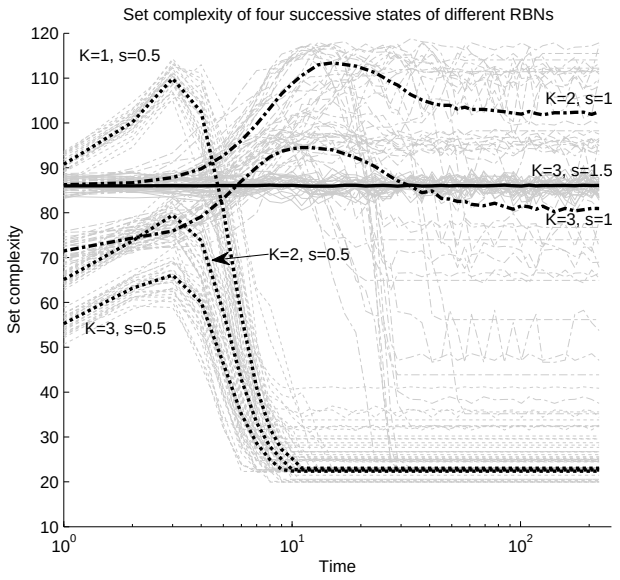


Figure 3. Set complexity of RBN dynamics. Different numbers of neighbors (K) and different sensitivities s are used. The thin curves represent samples of set complexity trajectories, as the thick curves are medians of 1000 trajectories.

Next we study the maximum values of this kind of set complexity trajectories in more detail. We consider RBNs with the same fixed numbers of neighbors ($K = 1, 2, 3$)

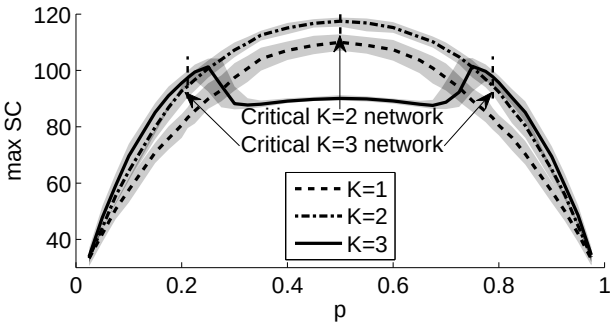


Figure 4. Set complexity maximum values of different RBNs. The curves represent medians with the area between 10% and 90% quantiles shaded. Number of samples is 130. The values of p that induce a critical BN are marked, $p = 0.5$ for $K = 2$ and $p = \frac{1}{2} \pm \frac{\sqrt{3}}{6}$ for $K = 3$ networks.

but with variable p; hence, with variable sensitivity s. Figure 4 shows the distribution (the quantiles) of the maximum set complexity value as a function of p. For networks with $K = 2$ the critical network produces the dynamics that has the greatest maximum set complexity, as for $K = 3$ that holds for slightly chaotical ($s \approx 1.1$) networks. Notice the symmetricity: the dynamical characteristics of networks with $p = q$ and $p = 1 - q$ are identical with only the role of 0's and 1's interchanged.

3.2. Information content of dynamics of the lattice gas

The lattice gases considered are set up in a $M \times M$ grid of hexagonal boxes, surrounded by impenetrable walls. In the beginning all particles are situated in three overlapping rectangle-shaped regions on the left side of the box. The first rectangle lies in the upper-left corner of the box and consists of particles that are moving lower-right, the second one occupies the whole height of the box and consists of particles moving right, and the third one lies in the lower-left corner and moves upper-right. The rectangle sizes are chosen such that it takes an exact number of time steps (16 in our simulations) before the first particles hit a wall on the opposite side of the box. The rectangles are densely packed: in each slot of the rectangle a particle exists with probability $p = 0.95$. This particular setup was chosen as a reference to a classical example in thermodynamics in which particles spread from a densely packed cluster to the whole spatial domain to obtain a maximal entropy.

In the beginning of the simulation there is a short period of time during which the particles do not collide but move in the most ordered fashion, each of the six $M \times M$ planes only being shifted in their direction of movement. In contrast, toward the end of the simulation the particles move chaotically, colliding frequently with each other as well as with the surrounding walls. Between these phases lies a period of transition from the orderly motion into the chaotical one. Figure 5 shows that the set complexity of

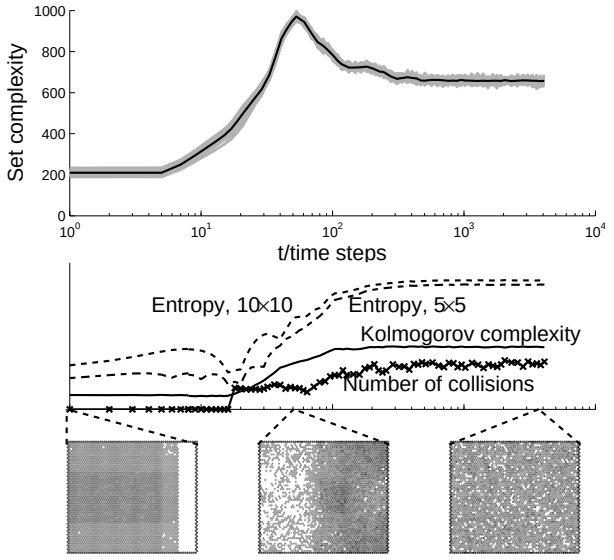


Figure 5. Simulation of a spreading lattice gas. Upper axes: The set complexity of twelve successive states of the lattice gas versus time. A number of realizations (N=100) plotted in gray, median plotted in black. Lower axes: Both entropy (dashed), Kolmogorov complexity estimate (solid) and number of collisions (solid + x) plotted versus time. The entropy is calculated using the traditional thermodynamical definition. The state space is divided into 5-by-5 (10-by-10) boxes, and the number of particles in all of these microboxes is calculated. The entropy is proportional to the logarithm of the number of all possible configurations with the said number of particles in the microboxes. For simplicity, only the interior of the spatial domain is considered when calculating the entropy, i.e. the number of particles colliding with the walls do not contribute to the entropy. Bottom figures: Lattice gas states visualized at the beginning of the simulation ($t = 0$), during the transition period ($t = 56$) and in the end ($t = 4096$).

12 successive states of the lattice gas is maximised during this transition period. In contrast, the Kolmogorov complexity of the state as well as the traditional entropy shows a mere rising trend in its trajectory. For the compression the state of the lattice gas is read into a string first column by column, last plane by plane — i.e. the distance of two nodes in the same hexagon is always a multiple of M^2 in the string.

4. CONCLUSION AND DISCUSSION

In this article, we have considered set complexity as a measure of dynamical complexity of state transitions using two models, RBNs and FHP lattice gases as case studies. Using these models, we have found that there is a point of maximal complexity in time, occurring before the system settles down to its long-term behavior. Although the generality of the phenomenon is yet to be assessed in full, it can be suggested that such maximal complexity can be found in a range of other models of dynamical systems as well.

A random Boolean network can be thought of as performing a type of a classification task, taking the initial state of the system and finding out, by computing its characteristic state transitions, to which attractor basin the input vector belongs. Such a parallel calculation can perform a huge variety of functions, and what we see in the set complexity measure can be taken to quantify this observed complexity. In accordance with this idea, critical or near-to-critical networks show the highest peaks of set complexity due to the complex dynamics emerging at the phase transition from ordered to chaotic networks. For the FHP lattice gas model, the state of the particles changes through a transient towards a steady state equilibrium. In this case, the characteristics of the steady state are not dependent on the structured initial state we have selected and statistically speaking, the end result is always the same. However, the complexity of the trajectory still reaches its maximum before equilibrium, showing the generality of our set complexity approach.

Previously, set complexity has been quantified on the attractor cycle, with critical networks showing the highest complexity in that case as well [4]. Our results agree with these previous observations, but move the focus to the transitional states before the attractors are reached. In the case of biological systems, it can be argued that such states are more significant in determining the response of the system to external inputs and variations in the state due to noise. In future work, such response should be studied explicitly, using suitable Boolean network models. The effects of different network characteristics on the complexity of calculation in the transients can then be studied together with traditional measures of criticality, to see how the network response is shaped.

5. ACKNOWLEDGMENTS

The authors would like to acknowledge the following funding: TISE graduate school, Academy of Finland projects 132877 and 213462 (Center of Excellence).

6. REFERENCES

[1] M. Aldana, S. Coppersmith, and L.P. Kadanoff, "Boolean dynamics with random couplings," *Perspectives and Problems in Nonlinear Science*, vol. 93, no. 4, pp. 23–90, 2003.

[2] P. Rämö, S. Kauffman, J. Kesseli, and O. Yli-Harja, "Measures for information propagation in Boolean networks," *Physica D: Nonlinear Phenomena*, vol. 227, no. 1, pp. 100–104, 2007.

[3] U. Frisch, B. Hasslacher, and Y. Pomeau, "Lattice-gas automata for the Navier-Stokes equation," *Physical Review Letters*, vol. 56, no. 14, pp. 1505–1508, 1986.

[4] D.J. Galas, M. Nykter, G.W. Carter, N.D. Price, and I. Shmulevich, "Biological information as set-based complexity," *IEEE Transactions on Information Theory*, vol. 56, no. 2, pp. 667–677, 2010.