Computable Information Content and Boolean Networks Dynamics

C. Bonanno[1]
G. Menconi[2]
V. Benci[1]
P. Cerrai[3]

[1] *Dipartimento di Matematica Applicata, Università di Pisa, Italy*
[2] *Dipartimento di Matematica, Università di Bologna, Italy*
[3] *Dipartimento di Matematica, Università di Pisa, Italy*

We propose the application of information content to analyze the time and space evolution of some boolean networks, and to give a classification of their state as either solid, jelly, liquid, or gaseous, following the ideas of Kauffman. To these aims both individual nodes and the global network are studied by means of data compression. Some experiments are presented using the compression algorithm *CASToRe*, developed along Lempel–Ziv algorithms.

1. Introduction

In [1], Stuart Kauffman thinks of a metabolic network as a *boolean network*: each node of the network has an internal state (0 or 1). At every time step, the node changes its state according to boolean functions depending on the states of some randomly chosen nodes of the network (called *neighbors*). The boolean function and the neighbors, in principle, are different for each node. A system well fitted by this model is given by the network of genes, where each gene is either active or inactive (state 0 or 1) and its state depends on the state of other different genes, that can be close or not.

Kauffman also presented a classification of these networks, based on the length of their periodic orbits. Using extensive experimental simulations, he divided the networks into three different regimes: *ordered*, *chaotic*, and *transition-to-chaos regime*. The motivations for a network to be in one of the regimes are to be found in two features: the number of neighbors and the driving rules, that is, the connections and the boolean function. Few connections guarantee a strong order, while a highly connected network is usually chaotic. Moreover, simulations suggest that some boolean functions lead to order, while with other rules chaotic dynamics prevail.

Denoting the number of nodes in the network by n and the number of neighbors chosen for each node by k (a fixed value, equal for all

Complex Systems, **16** (2005) 155–174;

https://doi.org/10.25088/ComplexSystems.16.2.155

nodes), Kauffman exploited computer simulations of several different such (n, k)-networks and observed the results. For $k = 2$, the state cycles are approximately as long as the square root of n. These networks tend to rapidly become ordered. On the other hand, if $k = n$, the state cycles have length of the order of the square root of the number of states, that is, the square root of 2^n. These networks are extremely chaotic. Nevertheless, it is also possible to obtain chaotic networks for $k \ll n$, such as $k = 4$ or $k = 5$. A sort of "tuning" from chaos to order is also observed when the boolean function is modified. With this in mind, Kauffman introduced a parameter p measuring the probability of having 1 and 0 as outputs of the boolean functions. If the two states are balanced, namely $p = 0.5$, the network is chaotic, but when p decreases to 0 the behavior of the network becomes ordered.

Whereas the research of Kauffman has given rise to a huge amount of experimental investigations, to our knowledge a theoretical analysis of Kauffman networks has been carried out only to some extent, and only in some cases, namely for $k = 1$ in [2] and $k = n$ in [3]. Moreover, in [4] and [5], a theoretical analysis is accomplished about the transitions from order to chaos depending on k (at $k = 2$) and on p.

In this paper, we consider a class of boolean networks that are very similar to Kauffman networks, but preserve the spatial closeness between nodes. In this sense, the networks we analyze are also similar to cellular automata [6], another field well-studied from the experimental point of view. We use the tools of algorithmic information theory, and in particular compression algorithms, to give a *local* and *global* classification of the networks in terms of the *entropy* of the nodes. That is, we first analyze the orbit of each node (the sequence in time of its states), and then compare the behavior of the different nodes to determine the "global" characteristics of the network. In particular, we consider the number of neighbors fixed and equal for all nodes, and the feature of the networks that we study is the distribution of 1 and 0 as outputs of the boolean functions. Varying the distribution, we obtain a classification of the networks similar to Kauffman's method. However, the main aim of this paper is to introduce the method and not to present extensive experimental simulations.

An application of algorithmic information theory to cellular automata can also be found in [7], where the authors introduce a classification parameter based on the local rules of the automata. The difference with our approach is that we are interested in the "phenomenological" aspect, that is, we want to study the effects of the local rule only looking for some distinguishing feature of the rule, in particular the probability parameter p introduced earlier. Some ideas of the application of compression algorithms to cellular automata are also introduced in [8], chapter 10.

In section 2, we introduce the class of boolean networks studied and the definitions and tools from algorithmic information theory. In section 3, we perform the local analysis of some different networks. These results are used in section 4 to obtain a classification of the networks in different regimes. We shall distinguish two distinct regular regimes, where the "less complex" may be associated to Kauffman's idea of an ordered network, and two "chaotic" ones. Conclusions can be found in section 5.

2. Definitions and tools

2.1 Networks

We define a *network* to be a square lattice of side N, hence with $N \times N$ nodes, and assume periodic boundary conditions. Each node may be either in state 0 or 1. At the beginning, a boolean function is assigned at random to each node and each function is chosen such that the neighbors of each node are the eight closest ones. A parameter p measures the probability of a node to turn into the state 1; that is, a fixed number of randomly chosen inputs (among the 2^8 possible inputs) have 1 as output. At every time step, the states of the nodes are simultaneously updated following these rules. Hence, for every fixed *simulation time* T, a symbolic binary sequence $\omega_1^T \in \{0,1\}^T$ is associated to each node. We shall call this the *orbit* of the node, and it describes all the states of the node from the beginning up to the simulation time. We have chosen to fix only the parameter p and do not control the boolean functions assigned to the nodes. Hence our results should be intended to hold for "almost all" networks with a fixed p.

Next, we define the notion of *entropy* of a symbolic string, which will be used to classify the orbits of a network. We study the entropies of the orbits when p varies in $(0, 0.5]$, for different size N and simulation times T.

2.2 Compression entropy

Let us consider a finite alphabet $\mathcal{A}$ and the set $\mathcal{A}^*$ of finite strings on $\mathcal{A}$, that is, $\mathcal{A}^* = \cup_{n=1}^{\infty} \mathcal{A}^n$. Denote by $\mathcal{A}^{\mathbb{N}}$ the set of infinite strings whose symbols belong to $\mathcal{A}$. A stochastic process with realizations in the space $\mathcal{A}^{\mathbb{N}}$ is called an *information source*. If $\sigma \in \mathcal{A}^{\mathbb{N}}$, then the first n symbols in σ are denoted by $\sigma_1^n \in \mathcal{A}^n$.

One of the most intriguing problems about symbolic strings is the notion of randomness of a string. Whereas *Shannon entropy* [9] gives a notion of chaoticity for information sources, it is interesting to have a notion of randomness of a string, independent of the properties of its information source. A way to introduce such a notion is to calculate the

https://doi.org/10.25088/ComplexSystems.16.2.155

amount of information necessary to exactly reproduce the string, that is, the so-called *information content* of the string.

The most intuitive definition of information content was introduced independently by Chaitin [10] and Kolmogorov [11]. Given a *universal Turing machine* U, which can be considered as a personal computer, we say that a binary string P is a program for a given symbolic string $s \in \mathcal{A}^*$ if $U(P) = s$. The algorithmic information content (AIC) of a string s is defined as the length of the shortest binary program P that gives s as its output, namely

$$\mathrm{AIC}(s, U) = \min\{|P| : U(P) = s\}.$$

We have the following theorem due to Kolmogorov.

Theorem 1 ([11]). If U and U' are universal Turing machines then

$$|\mathrm{AIC}(s, U) - \mathrm{AIC}(s, U')| \leq K\left(U, U'\right)$$

where $K(U, U')$ is a constant that depends only on U and U' but not on s.

Theorem 1 implies that the AIC of s with respect to U depends only on s up to a fixed constant, then its asymptotic behavior does not depend on the choice of U. For this reason we will not specify the choice of the machine in the notation, but assuming a fixed machine U we will write $\mathrm{AIC}(s) = \mathrm{AIC}(s, U)$. For further properties and applications of the AIC see [12].

The shortest program that gives a string as its output is a sort of encoding of the string. The information that is necessary to reconstruct the string is contained in the program. Unfortunately, a very deep statement, in some sense equivalent to the Turing halting problem or to the Gödel incompleteness theorem, states that the AIC is not computable by any algorithm.

For simulations it is then necessary to introduce a notion of computable information content (CIC). Let us suppose to have some recursive lossless (i.e., reversible) coding procedure $C : \mathcal{A}^* \to \{0, 1\}^*$ (e.g., the data compression algorithms that are in any personal computer). Since the coded string contains all the information that is necessary to reconstruct the original string, we can consider the length of the coded string as an approximate measure of the quantity of information that is contained in the original string.

Definition 1 (CIC). The *computable information content* of a string $s \in \mathcal{A}^*$, relative to a given reversible recursive coding procedure C, is $I_C(s) = |C(s)|$, that is, the binary length of the encoded string $C(s)$.

Of course not all the coding procedures are equivalent and give the same performances, so some care is necessary in the definition of CIC. For this reason we introduce the notion of *optimality*.

Definition 2 (Optimality). A reversible coding algorithm C is *optimal* if for almost all[1] infinite strings $\sigma \in \mathcal{A}^{\mathbb{N}}$

$$\lim_{n\to+\infty} \frac{I_C(\sigma_1^n)}{n} = H \tag{1}$$

where H is the Shannon entropy of the information source.

An example of optimal codings are the well-known Lempel–Ziv compression schemes [13].

For finite symbolic strings, we simply consider the compression ratio of the whole string as a measure of chaoticity of the string.

Definition 3 (C-entropy). Given an optimal coding C, we call *compression entropy* $K_C(s)$ (C-entropy) of a string $s \in \mathcal{A}^n$ the ratio

$$K_C(s) := \frac{I_C(s)}{n}.$$

If a symbolic string $\sigma \in \mathcal{A}^{\mathbb{N}}$ is constant or periodic, then the information content of the substrings σ_1^n is expected to be of the order of the logarithm of the length n, hence the C-entropy $K_C(\sigma_1^n)$ is expected to be very small for n big enough. Conversely, if an infinite string is patternless, then the information content of its substrings is expected to grow linearly with their length n, and their C-entropy is, for n big enough, strictly positive and close to the logarithm of the cardinality of the alphabet (for binary strings, such a limit is 1).

When applying this notion to the symbolic orbits of the nodes, we assume that the simulation time T is big enough so that the C-entropy of the orbits $K_C(\omega_1^T)$ is close to its "asymptotic" value.

2.2.1 The algorithm CASToRe

We have created and implemented a particular compression algorithm called CASToRe, which is a modification of the well-known LZ78 algorithm. Its theoretical advantages with respect to LZ78 are shown in [14] and [15]: it is a sensitive measure of the information content of low entropy sequences. That is why we called it **CASToRe**: **C**ompression **A**lgorithm, **S**ensitive **To** **Re**gularity.

As proved in Theorem 4.1 in [15], the information I_C of a constant sequence $s \in \{0,1\}^n$ is $4+2\log(n+1)[\log(\log(n+1))-1]$, if the algorithm C is CASToRe. The theory predicts that the best possible information for a constant sequence of length n is AIC(s) = log(n)+const. In [14], it is shown that the LZ78 encodes a constant n-digits long sequence to a string with length about const + $n^{1/2}$ bits; so, we cannot expect that LZ78 is able to distinguish a sequence whose information grows like n^{α}

[1] With respect to a probability measure on $\mathcal{A}^{\mathbb{N}}$.

https://doi.org/10.25088/ComplexSystems.16.2.155

($\alpha < 1/2$) from a constant or periodic one. This motivates the choice of using CASToRe. Moreover, even if not formally proved, the optimality of CASToRe has been tested with success on some chaotic dynamical systems in [16]. After writing this paper, we have found that a slight modification of CASToRe has been proved to be optimal [17].

Now we briefly describe the internal running of CASToRe.

Similar to the algorithm LZ77, CASToRe is based on an adaptive dictionary. One of the basic differences in the coding procedure is that LZ77 splits the input strings into overlapping words, while CASToRe (as does LZ78) parses the input string into nonoverlapping words.

At the beginning of the encoding procedure, the dictionary contains only the alphabet. In order to explain the principle of encoding, let us consider a step h within the encoding process, when the dictionary already contains h words $\{e_1, \ldots, e_h\}$.

The new word is defined as a pair (*prefix pointer*,*suffix pointer*). The two pointers refer to two (not necessarily different) words ρ_p and ρ_s chosen among the ones contained in the current dictionary as follows. First, the algorithm reads the input stream starting from the current position, looking for the longest word ρ_p matching the stream. Then, we look for the longest word ρ_s such that the joint word $\rho_p\rho_s$ matches the stream. The new word e_{h+1} that will be added to the dictionary is then $e_{h+1} = \rho_p\rho_s$. The "new" current position is at the end of the word ρ_s in the input stream.

The output file contains an ordered sequence of the binary encoding of the pairs (i_p, i_s) such that i_p and i_s are the dictionary index numbers corresponding to the prefix word ρ_p and to the suffix word ρ_s, respectively. The pair (i_p, i_s) is referred to the new encoded word e_{h+1} and has its own index number i_{h+1}.

2.3 Remarks

Being deterministic and with a finite number of configuration states, all the networks generate periodic orbits. Nevertheless, the periodic patterns of the orbits may be as long as the maximum number of configurations, that is, 2^{N^2}. Hence, if $N \geq 5$, the extent of the period may be of the order of hundreds of millions of symbols. An orbit may be considered chaotic as far as the simulation time T is shorter than its period.

In this sense, we may exploit a finer analysis than Kauffman's, who used the length of periodic orbits to classify the networks. Indeed, by means of the compression algorithm, we can easily identify periodic orbits with increasing period length and differently complicated periodic patterns by varying the simulation time and comparing the results, and we can use this analysis to give a definition of the state of the network, as shown in section 4. Moreover, further details on the disorder of the

periodic orbits may be inferred from the analysis of the parsing of the symbolic string generated by the compression (see section 3).

3. Dynamics of individual nodes

The first part of the analysis on the defined networks concerns the behavior of the symbolic orbits of single nodes with increasing values of the parameter p. As a $N \times N$ network evolved for T time steps, we stored the N^2 symbolic orbits and compressed them using CASToRe. In the following, we simply write $I(s)$ to denote the CIC $I_C(s)$ of a finite string s with compression algorithm C = CASToRe.

The first approach is to evaluate the C-entropy $I(\omega_1^T)/T$ of the orbits and to make a distinction among nodes according to this value. We picture the results on a grayscale network (C-entropy map), where each pixel corresponds to a node; and, with increasing C-entropy from 0 to 1, the node becomes lighter.

An important feature of the C-entropy map of Figure 1 is the color variability that underlines the differences in the chaoticity of the nodes. This feature depends on the dimension N of the network and on the value of p, which, as stated previously, tunes the disorder in the network.

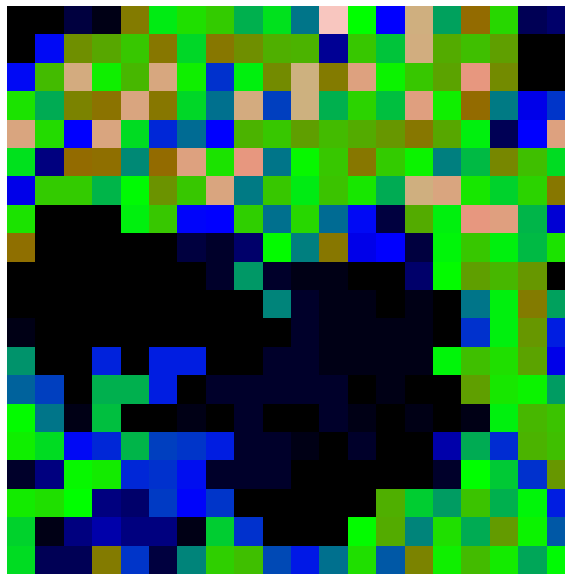


Figure 1. An example of a C-entropy map of a network with $N = 20$, evolution time $T = 10^6$, and $p = 0.14$.

https://doi.org/10.25088/ComplexSystems.16.2.155

We analyzed networks of dimension $N = 10$, 15, and 20, and the variability feature is evident for $N = 20$ but almost absent for $N = 10$.

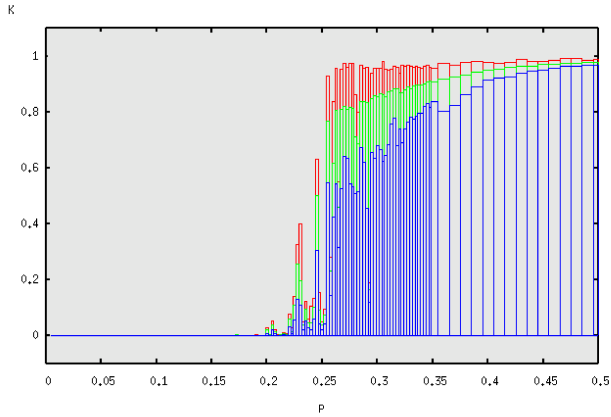
We performed a two-fold analysis of the C-entropy maps. First, we studied the evolution of the C-entropy of the nodes with time. Second, we considered the variability of the C-entropy map as the parameter p increases. The latter point is part of section 4.

Given a network $N \times N$, for small values of the parameter p the orbits of the nodes are regular and their C-entropy is already low for a short simulation time T. Hence the C-entropy map is homogeneously colored with all nodes almost black. As p increases to 0.5, the C-entropy of the nodes remains high for bigger and bigger simulation times and in the C-entropy maps light nodes frequently arise. In particular, when p approaches 0.5 all nodes are light up to the maximum simulation time T_M of our experimental simulations (e.g., for $N = 10$ we have $T_M = 3 \times 10^6$, for $N = 15$ we found light nodes up to $T_M = 3 \times 10^7$).

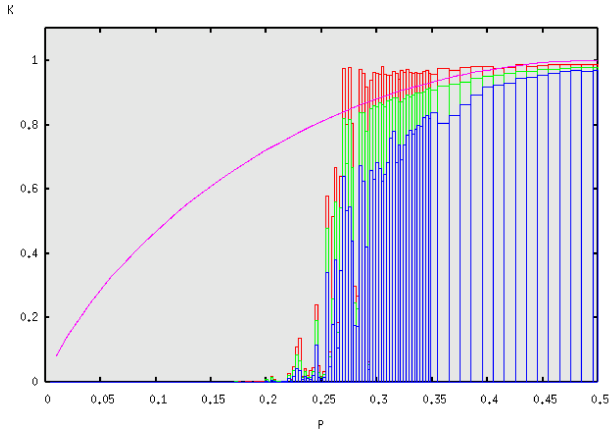
We fixed the dimension N of the networks and studied how the maximum, minimum, and mean values of the C-entropy of the nodes change when p varies from 0 to 0.5. In Figures 2(a) and 2(b) the results are shown for $N = 10$ with $T = 10^6$ and $T = 3 \times 10^6$, respectively. We remark that these results have to be considered as "statistical," indeed given the side N and the parameter p we randomly generated a network with those characteristics. The figures show the results for these randomly generated networks. From the figures one can infer that there is a first transition from homogeneous regular to arising irregular dynamics of the nodes around some p^* ($p^* \sim 0.2$ for $N = 10$). Indeed for p bigger than p^*, after some fluctuations, all nodes have positive C-entropy, the maximum values are insensitive on p and almost equal to 1, whereas the minimum values monotonically increase to 1. Concerning the mean values, it is noticeable that after p^* they are close to the Shannon entropy function relative to uncorrelated events; that is, $H(p) = -p \log_2(p) - (1 - p) \log_2(1 - p)$ (see Figure 2(b)). We could say that "on average" the orbits of the nodes are as chaotic as a binary symbolic string generated by a stochastic process of independent random variables, distributed on $\{0, 1\}$ with $P_1 = p$ and $P_0 = 1 - p$. Further investigations in this direction are shown in the final part of this section.

When comparing the two pictures of Figure 2, notice that the values of the C-entropy of the nodes decrease as simulation time increases. This feature is evident in the figure for values of p lower than 0.3, since for bigger values of p the same feature would appear for bigger simulation times. This remark points out the fact that all nodes are eventually periodic and this regularity becomes evident for a simulation time that depends on the parameter p.

The same behavior of the values of the C-entropy can be found varying the parameter p for networks with $N = 20$. In Figure 3 we show the analog of Figure 2(b) with simulation time $T = 10^6$. Again we



(a)



(b)

Figure 2. Behavior of maximum, minimum, and mean values of the C-entropy for networks with $N = 10$ for $p \in (0, 0.5]$: (a) the results on C-entropy maps with $T = 10^6$; (b) for $T = 3 \times 10^6$ together with the function $H(p)$ (see text).

can identify a first transition from homogenous regularity to variable irregularity around $p^* = 0.14$, with a small hole immediately after. Then, again, there is a slow increase in the minimum values and a jump to 1 for the maximum values. The mean C-entropy, where nontrivial, again increases close to the function $H(p)$. The decrease of the value of p^* with the dimension of the network increasing is intuitively clear. Indeed bigger networks are expected to have a greater variability, hence irregular dynamics for the nodes are expected to appear already for relatively small values of the probability parameter p.

https://doi.org/10.25088/ComplexSystems.16.2.155

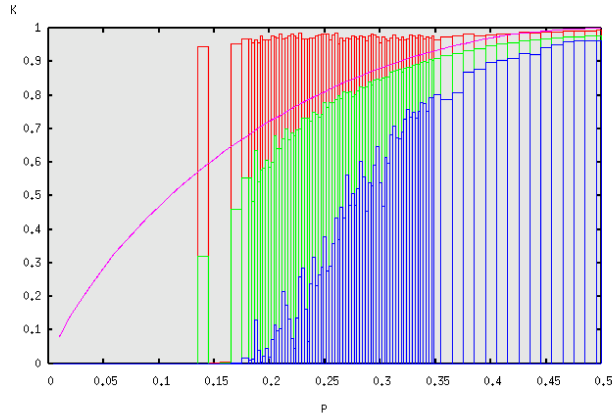


Figure 3. Behavior of maximum, minimum, and mean values of the C-entropy for networks with $N = 20$ for $p \in (0, 0.5]$ and simulation time $T = 10^6$, together with the function $H(p)$.

The difference between the minimum and the maximum values of the C-entropy in Figures 2 and 3 is an indication of the variability in the chaoticity of the nodes of a given network. It is evident that networks with $N = 20$ display a higher variability. However this feature is analyzed in more detail in section 4, where further transitions can be identified in the behavior of the C-entropy maps by varying p.

We finish the local analysis of the networks with an interesting remark about the individual symbolic orbits of the nodes. When compressing a symbolic orbit, CASToRe creates an adaptive dictionary of the words it uses for the compression. Hence it is possible to analyze the statistical features of this dictionary. Although features like the number of words in the dictionary and their frequency in the symbolic orbit are usually studied, here we concentrate only on the length of the words along the symbolic orbit. In particular, for a T-symbols long orbit we plot the length of the new word created by CASToRe *versus* the initial position of this word in the symbolic orbit. Due to the internal structure of CASToRe, the length of the new words is given by the sum of the lengths of two already known words; hence, for instance, in the case of a constant symbolic orbit, the length of a new word is always twice the length of the last word. A doubling in the length of new words is eventually found also for periodic symbolic orbits with nontrivial periodic patterns.

Now, recalling that the orbits of the nodes in any $N \times N$ network are eventually periodic, the length of the words in the dictionary is an indicator either of the length of the period (for orbits with a period much shorter than the simulation time) or of the "complexity" of the periodic pattern (for orbits with a period of the order of the simulation time).

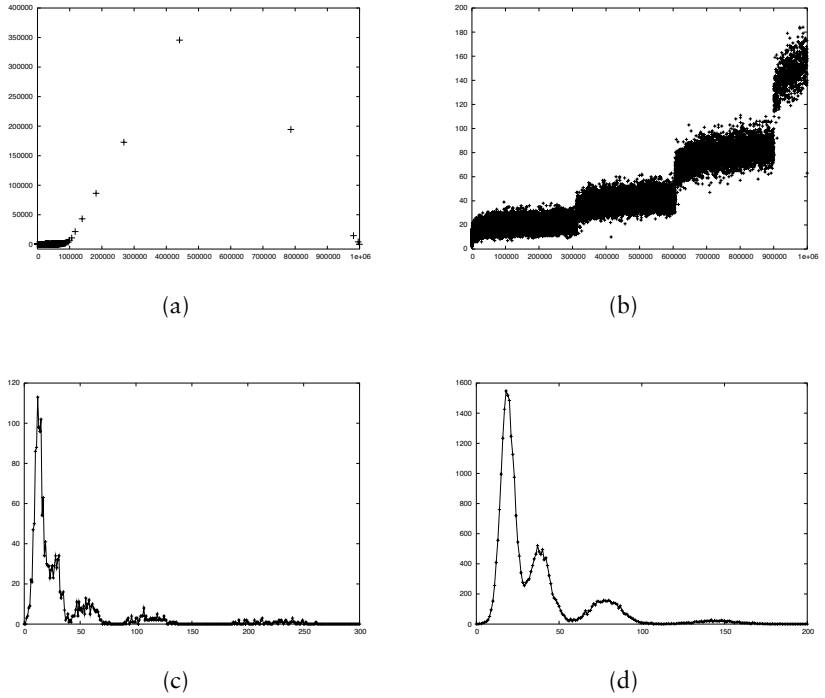


Figure 4. (a) and (b) show two word length plots, the length of the new words created by CASToRe *versus* their initial position in the symbolic orbit. The plots refer to two nodes from networks with $N = 10$, simulation time $T = 10^6$, and parameters $p = 0.2$ and $p = 0.245$, respectively. (c) and (d) show the distribution of the length of the words for cases in (a) and (b), respectively. In (c) the distribution is shown only for small lengths, where the periodic behavior is not already evident. For bigger lengths the doubling starts.

In Figure 4, we show the behavior for two nodes chosen from networks with $N = 10$. In 4(a) and 4(b) we show the word length plots of the two nodes. In case 4(a), the chosen node has a small period with respect to the simulation time $T = 10^6$, indeed the doubling in the word lengths is evident in the picture after CASToRe has compressed the first 10^5 symbols. At most the period length (simply obtained by the length L of the word when doubling first happens) is $L = 10800$, moreover it is a divisor of L. The distribution of the lengths of the words for this node is shown in 4(c), only for small lengths, since for bigger lengths the doubling starts and hence the statistics are trivial. The analysis of 4(c) already reveals the internal rules of the compression algorithm: the pairing of words groups them into "clouds" of doubling mean length. Furthermore, as the algorithm proceeds in the compression of the sym-

https://doi.org/10.25088/ComplexSystems.16.2.155

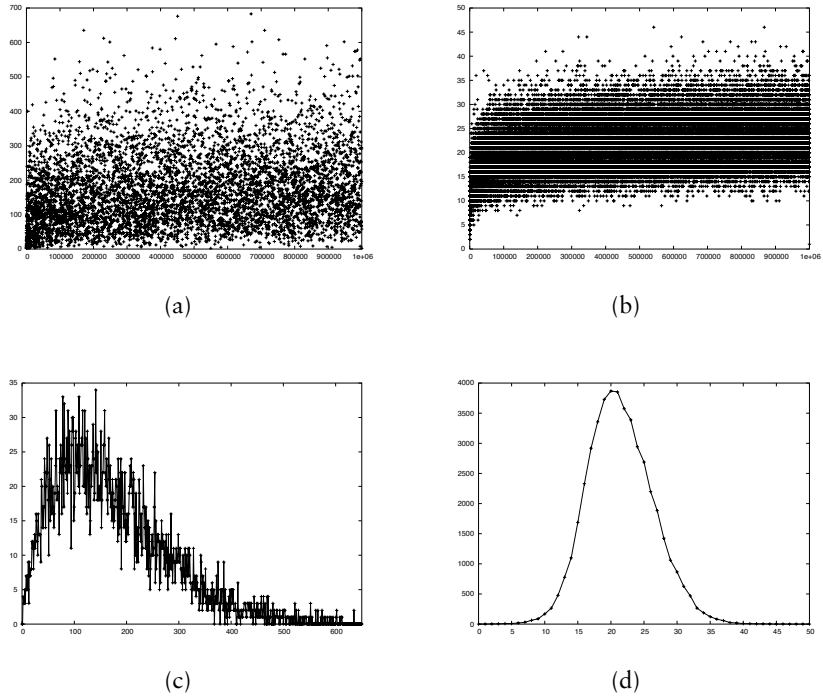


Figure 5. (a) and (b) show two word length plots, the length of the new words created by CASToRe *versus* their initial position in the symbolic orbit. On the left ((a) and (c)) is node 13 and on the right ((b) and (d)) node 11 of a network with $N = 20$, simulation time $T = 10^6$, and $p = 0.14$. (c) and (d) show the distribution of the length of the words for cases in (a) and (b), respectively.

bolic orbit, the words in the clouds decrease until there is one single word, when the periodic pattern is found.

The grouping feature is more evident in cases 4(b) and 4(d), where the chosen node has a period still smaller than the simulation time. In 4(b), the grouping occurring in case 4(a) at the beginning of the compression is enlarged to the whole picture, and big steps in the word lengths occur. However in this case there is no doubling, and the period can only be estimated using the grouping in the words by 3×10^5, the width of the steps. The distribution of the lengths in 4(d) follows the shape of case 4(c) in all the lengths.

In Figure 5 examples of chaotic nodes from a network with $N = 20$, $p = 0.14$, and simulation time $T = 10^6$ are shown (see also Figure 1). The enumeration of the nodes begins from 0 in the upper left corner and proceeds horizontally downwards. Nodes 13 and 11 are shown in Figure 5, on the left and on the right, respectively. Looking at the

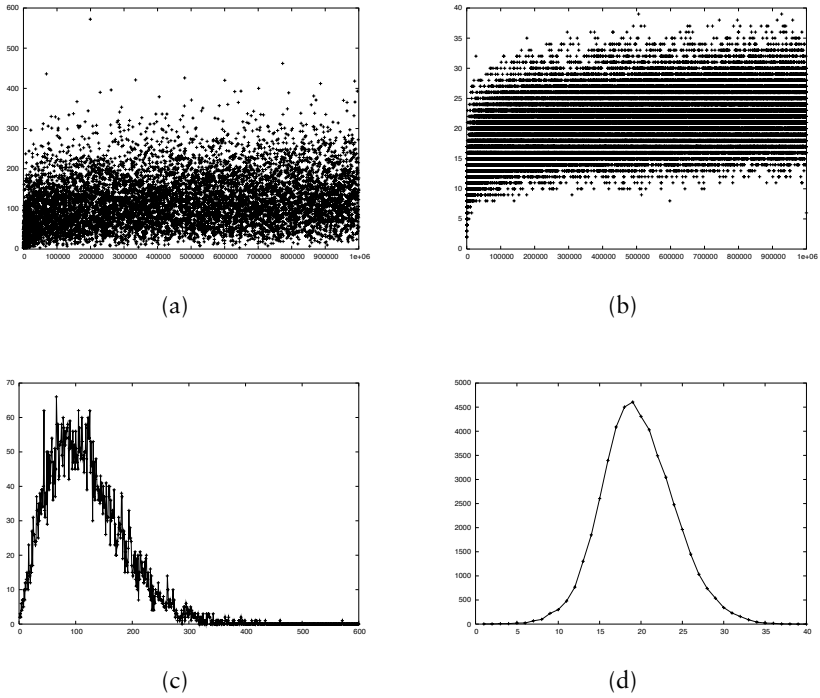


Figure 6. (a) and (b) show two word length plots for randomly generated binary symbolic orbits. Orbit σ (on the left) is generated with $P_0 = 0.02$ and $P_1 = 0.98$. Orbit τ (on the right) with $P_0 = P_1 = 0.5$. (c) and (d) show the distribution of the length of the words for cases in (a) and (b), respectively.

C-entropy map of Figure 1, node 11 is light and node 13 is dark, hence they have very different C-entropy values. This difference is evident in the word length plots 5(a) and 5(b). In 5(b), for node 11, we find small words with very similar lengths, whereas in 5(a) we find longer words with different lengths. Looking at the distributions, 5(c) and 5(d), this difference is even more recognizable.

Again, we remark that the statistics of the lengths strongly depend on the internal rules of the compression algorithm; Figures 5(a) and 5(b) should be viewed with this in mind. Figure 6 shows the analogs of Figure 5 for two random binary symbolic orbits σ and τ, generated with probabilities $P_0 = 0.02$ and $P_1 = 0.98$ on the left, and $P_0 = P_1 = 0.5$ on the right. The Shannon entropy for orbit σ is $H(0.02) = 0.141441$, while for orbit τ it is $H(0.5) = 1$; hence, orbit σ may be read as "less random" than orbit τ. A perfect analogy may be established between node 11 and orbit τ, and between node 13 and orbit σ. From this it turns out that the difference between nodes 11 and 13 is in the "complexity"

https://doi.org/10.25088/ComplexSystems.16.2.155

of the periodic pattern of the corresponding symbolic orbits. Again, this result enhances the similarities between the orbits of the nodes and the symbolic orbits generated by stochastic processes of independent random variables.

4. Global dynamics and classification of networks

In section 3 we studied the values of the C-entropy of the nodes of a given network varying the parameter p (the probability of transition to state 1, see section 2) and used this analysis to experimentally establish the presence of a first transition from homogeneously regular networks to irregular ones. We can call this first approach "local" since we did not pay attention to the relative behavior of the nodes, but simply looked at the presence of chaotic nodes. We now turn to the "global" analysis of the network, studying by means of algorithmic information theory the variability of the C-entropy in a given network. The aim is to obtain an experimental classification of networks similar to that considered by Kauffman. Analyzing the dynamical behavior of the networks, Kauffman concentrated his attention on the lengths of the periods of the orbits of the nodes. In particular, he defined a network to be: in a solid state if all the nodes are fixed or have very small periods (on the order of five time steps); in a liquid state if there are nodes with a longer period (on the order of hundreds of time steps) on a background of fixed nodes; or in a gaseous state if all nodes have long periods. Using the techniques of algorithmic information theory, we give a similar classification of nodes, choosing the C-entropy of the orbits as our indicator.

We shall make the following heuristic characterizations of networks. Let us fix a simulation time T.

- A network is in a *solid* state when the orbits of all the nodes are periodic with very short periodic patterns; the network appears homogeneous in the C-entropy map.
- A network is in a *jelly* state when all the orbits of the nodes are periodic, hence the C-entropy map is homogeneous, but periods are not very short.
- A network is in a *liquid* state when its C-entropy map at time T is heterogeneous.
- A network is in a *gaseous* state when its C-entropy map at time T is homogeneously chaotic.

This characterization needs two main ingredients: the definition of homogeneity of a network and a classification of the periodic patterns.

In order to characterize the homogeneity of a network, we will refine the analysis presented in Figure 2, where only the variation in the C-entropy of the nodes of a network was considered. In particular we use

algorithmic information theory to quantify the ratio of regular nodes in a network. As an example, we expect networks whose C-entropy map is such as in Figure 1 to be classified in the liquid state. Our approach is to study the "chaoticity" of the C-entropy map of a network, using a method to extract the C-entropy of a bidimensional lattice. Later on, we apply this technique to a suitably defined "space-time cut" of the complete time evolution of the network, in order to study the periodic patterns.

For the sake of simplicity, assume the network has size $N \times N$ where $N = 2^l$. Following the results in [18], we preprocessed the lattice by scanning it *via* the Peano–Hilbert plane-filling curve. For each $l \geq 1$, the curve covers the network visiting each node; hence a symbolic string of length N^2 corresponds to the network. In [18] it is proved that the compression ratio achieved by encoding the given symbolic string is lower than the one obtained *via* any other two-dimensional compression on the lattice, asymptotically when N tends to infinity. Therefore, we exploited this reliable bidimensional compression by means of CASToRe.

First, we applied this method to the two-dimensional color picture representing the C-entropy map and obtained a symbolic string v. Then, the C-entropy of v is the C-entropy of the C-entropy map obtained after a simulation time T. We chose an eight-symbol alphabet to codify the C-entropy values into symbols, then normalized to have C-entropy values in $[0, 1]$.

We show the results on 16×16 networks in Figure 7. Following the qualitative variation identified by the maximum, minimum, and mean C-entropy values, the C-entropy of v varies with parameter p. Starting

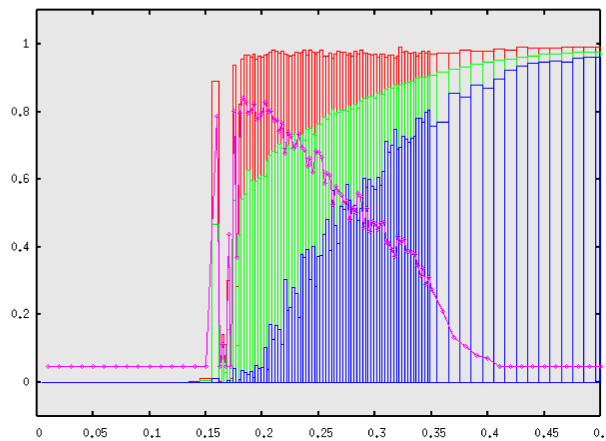


Figure 7. Behavior of maximum, minimum, and mean values of the C-entropy for networks with $N = 16$ for $p \in (0, 0.5]$ and simulation time $T = 5 \times 10^5$, together with the corresponding C-entropy values of v (see text).

https://doi.org/10.25088/ComplexSystems.16.2.155

from values very close to zero, it begins to oscillate when $p \sim 0.15$, stabilizing around its maximum at $p \sim 0.2$. Later, it softly decreases to zero, which is approximately reached around $p = 0.4$. This behavior points out for the 16×16 network a transition to the liquid state at the parameter value $p_{\mathrm{JL}} \sim 0.15$. The network remains in the liquid state until $p \sim 0.4$. Around this latter value, denoted by p_{LG}, another state transition occurs and from this point the network is in a gaseous state. Note that the transition to the liquid state corresponds to the transition already found in section 3 occurring at the values called p^*. Moreover, we remark that this transition is much more visible for networks with an increasing number of nodes (compare Figures 2 and 3), as one should expect from statistical mechanics where transitions are observed in the thermodynamic limit.

However, this first analysis is not sufficient to distinguish between the solid and jelly states in the regime $p < p_{\mathrm{JL}}$. Indeed, the C-entropy map denotes homogeneously periodic behavior of the orbits of the nodes, but it does not contain information about the lengths of the periods. The transition point p_{SJ} from solid to jelly corresponds to the kind of transitions identified by Kauffman.

In order to find the transition point p_{SJ}, we have considered a "space-time" approach to the orbits of the nodes. Given the picture of the network at the simulation time T, that is, the bidimensional lattice of cells in state 0 or 1, we have used the Peano–Hilbert curve to convert the picture into a binary string $s_T \in \{0,1\}^{N^2}$. Making the same for the picture of the network at times $T-1, T-2, \ldots, T-d+1$, we have constructed d binary strings $s_T, s_{T-1}, \ldots, s_{T-d+1}$ each of length N^2. By concatenating these d strings, we obtain a binary string $s = (s_T s_{T-1} \ldots s_{T-d+1}) \in \{0,1\}^{dN^2}$.

The C-entropy of the string s gives an indication of the lengths of the periodic patterns of the orbits of the nodes, indeed the higher or lower regularity shown by this "space-time scan" of the network is a marker for the global presence of periodic patterns with respectively low or high complexity, in particular short or long patterns.

Figure 8 shows the C-entropy variation for string s as p varies in $(0, 0.5]$ and $d = 100$, compared with the C-entropy of the corresponding strings ν (dashed line), when the network is 16×16. We have exploited simulations for 500 values of parameter p, with a simulation time $T = 5 \times 10^5$. To enhance the main trends, the plot shows smoothed curves, where for each point a local weighted average is drawn. From the picture, it is straightforward that the C-entropy of s adds further knowledge about the state of the network only when $p < 0.15$, since for larger values it nearly coincides with the Shannon entropy function $H(p)$ (dotted line).

Whereas the C-entropy of ν is negligible for $p \in (0, 0.15]$, the C-entropy of s shows an abrupt increase from 0.1 on (see Figure 9). This

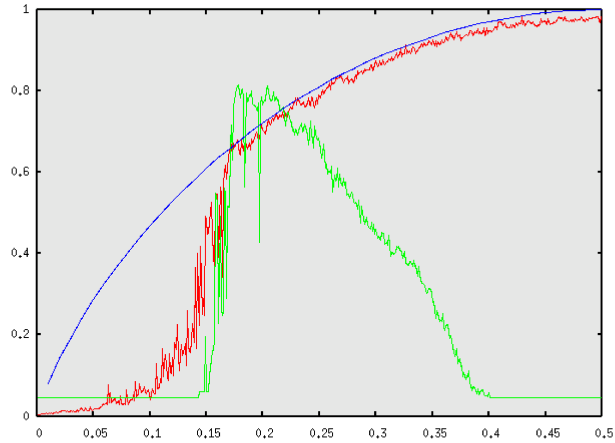


Figure 8. Networks with $N = 16$ and $T = 5 \times 10^5$. C-entropy values for string s (solid line) with $d = 100$ and string ν (dashed line) for 500 values of parameter $p \in (0, 0.5]$. The dotted line is the Shannon entropy function $H(p)$. See text for details.

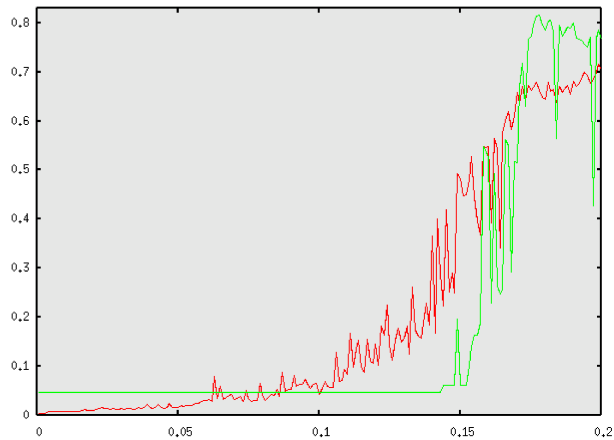


Figure 9. Networks with $N = 16$ and $T = 5 \times 10^5$. C-entropy values for string s (solid line) with $d = 100$ and string ν (dashed line) for 500 values of parameter $p \in (0, 0.2]$.

is a clear consequence of the presence of short periodic patterns with nontrivial complexity. For lower values of p, the C-entropy of s is also negligible. Hence, we can identify the transition point from solid to jelly as $p_{SJ} \sim 0.1$.

https://doi.org/10.25088/ComplexSystems.16.2.155

5. Conclusions

In this paper we addressed the question of classifying boolean networks into different classes, following the ideas of Kauffman. He looked at the periods of orbit for a network as a distinguishing feature, and some analytic results confirmed his experimental investigations. We studied the problem by applying numerical tools from algorithmic information theory, in particular a compression algorithm, identifying the classification characteristic to look for in the information content of the orbits of the nodes. Our approach led to a two-fold analysis.

First, we studied the single orbits of the nodes, showing that, as a parameter p (representing the *a priori* probability of activating a node) varies, the networks display a transition from a homogeneously regular behavior (the orbits of the nodes have small information content with respect to their length) to an irregular behavior (large information content for all orbits).

Second, we performed a "global" analysis of the network in order to make a distinction between a liquid and a gaseous state in the irregular regime and clarify what happens in the regular regime. To this end, we switched to a space-time approach, looking at the information content of some pictures suitably defined and associated to the network. On one hand, we derived a transition from the liquid to the gaseous state from the picture given by the entropy of the single nodes (C-entropy map). On the other hand, two different regular states (solid and jelly) may be defined by analyzing the information content of a space-time scan of the evolution of the network. These transitions appear clearer as the number of nodes increases, as one should expect from statistical mechanics.

In section 3, we studied the behavior of the compression algorithm on the orbits of the nodes. The results show that the irregularity of the periodic patterns of the orbits is very similar to that of the symbolic sequences obtained as realizations of a $(1-p, p)$ Bernoulli process. This is also confirmed in section 4 by the fact that the C-entropy of the space-time scan coincides with the Shannon entropy function for liquid and gaseous networks.

Finally, we remark that in this paper we restricted our attention to a "hybrid" class, namely Kauffman networks with local interactions, with the aim of introducing the method rather than to present extensive experimental simulations. However this approach can be repeated for any class of networks, from cellular automata to neural networks.

References

[1] S. A. Kauffman, *At Home in the Universe: The Search for Laws of Self-Organization and Complexity* (Oxford University Press, New York, 1995).

[2] H. Flyvbjerg and N. J. Kjaer, "Exact Solution of Kauffman's Model with Connectivity One," *Journal of Physics A: Mathematical and General*, **21** (1988) 1695–1718.

[3] B. Derrida and H. Flyvbjerg, "The Random Map Model: A Disordered Model with Deterministic Dynamics," *Journal de Physique*, **48** (1987) 971–978.

[4] B. Derrida and Y. Pomeau, "Random Networks of Automata: A Simple Annealed Approximation," *Europhysics Letters*, **1** (1986) 45–49.

[5] B. Derrida and G. Weisbuch, "Evolution of Overlaps Between Configurations in Random Boolean Networks," *Journal de Physique*, **47** (1986) 1297–1303.

[6] S. Wolfram, *Theory and Application of Cellular Automata* (World Scientific, Singapore, 1986).

[7] J. C. Dubacq, B. Durand, and E. Formenti, "Kolmogorov Complexity and Cellular Automata Classification," *Theoretical Computer Science*, **259** (2001) 271–285.

[8] S. Wolfram, *A New Kind of Science* (Wolfram Media, Inc., Champaign, IL, 2002).

[9] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Systems Technical Journal*, **27** (1948) 379–423, 623–656.

[10] G. J. Chaitin, *Information, Randomness, and Incompleteness. Papers on Algorithmic Information Theory* (World Scientific, Singapore, 1987).

[11] A. N. Kolmogorov, "Three Approaches to the Quantitative Definition of Information," *Problems Inform. Transmission*, **1** (1965) 1–7.

[12] M. Li and P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications* (Springer-Verlag, New York, 1993).

[13] J. Ziv and A. Lempel, "A Universal Algorithm for Sequential Data Compression," *IEEE Transactions on Information Theory*, **23** (1977) 337–342; "Compression of Individual Sequences via Variable-rate Coding," *IEEE Transactions on Information Theory*, **24** (1978) 530–536.

[14] F. Argenti, V. Benci, P. Cerrai, A. Cordelli, S. Galatolo, and G. Menconi, "Information and Dynamical Systems: A Concrete Measurement on Sporadic Dynamics," *Chaos, Solitons, and Fractals*, **13** (2002) 461–469.

[15] C. Bonanno, and G. Menconi, "Computational Information for the Logistic Map at the Chaos Threshold," *Discrete and Continuous Dynamical Systems, Series B*, **2** (2002) 415–431.

[16] V. Benci, C. Bonanno, S. Galatolo, G. Menconi, and M. Virgilio, "Dynamical Systems and Computable Information," *Discrete and Continuous Dynamical Systems, Series B*, **4** (2004) 935–960.

https://doi.org/10.25088/ComplexSystems.16.2.155

[17] M. Degli Esposti, C. Farinelli, and A. Tolomelli, "Compression Estimates for a Class of Dictionary Based Compressors," *Journal of Computational and Applied Mathematics*, **1** (2005) 177–198.

[18] A. Lempel and J. Ziv, "Compression of Two-dimensional Data," *IEEE Transactions on Information Theory*, **32** (1986) 2–8.