

Boolean networks with variable number of inputs (K)

Metod Skarja,^a Barbara Remic, and Igor Jerman
Bion Institute, Ljubljana, Slovenia

We studied a random Boolean network model with variable number of inputs K per element. An interesting feature of this model, compared to the well-known fixed K networks, is its higher orderliness. It seems that the distribution of connectivity alone contributes to certain amount of order. In the present research, we tried to disclose some reasons for this unexpected order. We studied also the influence of various numbers of source elements (elements with no inputs) on network's dynamics. The analysis, carried out on average $K=2$ networks, revealed the correlation between the number of source elements and dynamic diversity of the network. For diversity measure, we used the number of attractors, their lengths and similarity. For quantitative measure for attractor similarity, we developed two methods, one that takes into account the size and overlapping of frozen areas and the other, which considers also active elements. As the number of source elements increases, dynamic diversity of the networks does likewise: the number of attractors increases exponentially, while the similarity of attractors diminishes linearly. The length of attractors remains approximately the same, which indicates that orderliness of the networks remains the same. We determined also the amount of order that originates from the canalizing properties of Boolean functions and the propagation of this influence through the network. This source of order can account only for one-half of frozen elements; the other half presumably freezes due to the complex dynamics of the network. Our work also shows that different ways of assigning and redirecting connections between elements may influence the results significantly. Studying such systems can also help in modeling and understanding a complex organization and ordering in biological, especially genetic, systems.

89.75.-k, 89.75.Fb

Keywords: Boolean networks, biological systems, connectivity distribution, variable K , sources of order, canalization, frozen elements, no input elements (source elements), attractor similarity, effective distribution, genetic networks, high orderliness

^a metod.skarja@bion.si

The analysis of networks with variable number K of inputs, compared to fixed K networks, revealed a considerably higher degree of order. The orderliness found was quite surprising and only half of it can be accounted for by the evident causes like the canalization of Boolean functions and the propagation of this influence through the network. The source or no-input elements act as autonomous or independent sources of the network's dynamics and enhance the dynamic diversity of the networks. This high and partly unexplainable order could characterize also living systems.

INTRODUCTION

General

Boolean networks (BNs) are mathematical models for an extensive study of statistical mechanics in complex dynamical systems with multitude of coupled variables.¹ Such systems also include biological systems on various levels of organization, like genetic networks, metabolic networks, immune system, neural networks, ecosystems, biosphere etc.

A Boolean network is composed of N elements; each of them has a randomly assigned Boolean or logical function. In discrete networks, the values of elements are binary, which means 1 or 0. Interactions between the elements are directed and defined by randomly assigned inputs among them. Once the connectivity and the Boolean functions are assigned, they remain constant for a given network. The initial values of elements are chosen randomly and are synchronously updated in time steps, according to the connectivity and Boolean functions of the network. Thus a network evolves through a succession of states. The trajectories are deterministic, which means that the states on the same trajectory always follow exactly the same succession. All trajectories that converge into the same attractor constitute a basin of attraction. Eventually all trajectories run into attractors, which can be points or cycles. An external perturbation may cause a system to leap from one trajectory to another and follows this new one. The point attractors or steady state attractors are unalterable end states, whereas the cyclic attractors are periodically repeating successions of states. Their length may vary extremely, but only short and moderate length attractors are regarded as ordered in practice, while for attractors with very long periods it can readily happen that in a real time we cannot observe any periodic patterns. The degree of order can be estimated from the number of attractors, their cycle lengths, their influence (basin of attraction) and their stability (bifurcations).

Kauffman's work

Stuart Kauffman was the first researcher that intensively studied Boolean networks^{1,2} and discovered a quite unexpected order in these predominately random constructs. He worked with a constant number of inputs (K) per element so that all elements had exactly the same K . Networks with different K had different dynamics and there was a sharp phase transition from order to disorder. For $K=2$ the network's dynamics was ordered, but any larger K resulted in increasingly disordered dynamics. In $K=2$ networks Kauffman estimated the number of attractors to be around \sqrt{N} , and also their cycle length to be

around \sqrt{N} , although recent researches show that this holds approximately only for median values around $N=100$,³ while in general it is underestimated^{4 5 6} and the rise is even faster than linear. Anyway, compared to the enormous number of all possible network states (2^{100} or approx. 10^{30} in $N=100$ network), these attractor states represent only an infinitesimal fraction of the entire state space.

The source of this incredible order were, according to Kauffman, so called *frozen elements*,¹ whose values remain constant through a given trajectory. Formation of frozen elements arises through canalizing Boolean functions^b and homogeneity bias (P).^{1 2} Kauffman pointed out that this spontaneous order is intrinsic to the specific organization of the system, and is not a consequence of far from equilibrium thermodynamics, which presumably causes self-organization and self-ordering in open thermodynamic systems. He therefore called it »order for free«. ¹ The discovery of this astonishing spontaneous order raised a question whether it is only an artifact of mathematical idealizations, and if it actually exists in real systems.

Besides this basic model, there were developed also models with a more realistic resemblance to chemical kinetics in real systems, i.e. with continuous variables and updating triggered by passing of the variables over predefined thresholds.³ It was shown there, that while some attractors may be artifacts of synchronous updating and binary idealization, the classification of dynamical attractors and several other properties in particular network remains invariant, confirming thus the broader validity of results obtained with fixed K discrete model. Asynchronous updating was investigated, too.⁷ Here, there were no cyclic attractors, but they displayed also pseudo-periodic order, based on correlation between states.

Introduction of variable K

The study of original model described above exhibits already the most important properties of it, but we think that it neglect a very important fact that in real systems K is not equal for all elements.^{8 9 10 11 12} In real systems we expect that different elements have different numbers of inputs, some many, other only a few, one, or even none. A *variable K* is therefore essential for a more realistic modeling of biological systems. An increasing number of empirical evidences^{8 9 10 11} supports the idea of widely distributed connectivity in biological systems, the distribution being scale-free with an average K that exceeds Kauffman's critical invariable $K=2$.^{8 9 11}

With our variable K network model, we wanted to emphasize the consequences of K variability. Another group of researchers carried out a similar research.¹³ Their work derives from investigations of scale-free networks, their topology, stability and growth.^{12 14 15} The authors applied scale-free connectivity distribution to Boolean networks, and compared the impact of various K distributions on network's dynamics.¹³ They proved that scale-free or power-law distribution results in more ordered dynamics than Poisson and invariable K distribution. The topology of K distribution alone (power-law distribution being the most effective) seems to be the source of order, enabling that the average connectivity exceeds $K=2$, while the dynamics still remains ordered. The

^b Boolean functions with the property, that their output is simply the value of one of their inputs; or more generally, that their output is independent of some of the inputs (see also Appendix).

effective average K is significantly smaller than the actual average K and the order originating from many elements with low effective K prevails over the disorder produced by a few elements with a high effective K . Besides having more ordered dynamics than networks with different connectivity distributions, the scale-free networks¹³ are also incredibly stable and error tolerant.^{8 14}

Specific in our research

We have studied networks with the average $K=2$, but apart from Fox and Hill¹³ we also allowed elements with 0 inputs (source elements). The distribution of inputs in our networks was skewed binomial (see Methods for more detail). We were interested in how ordered are the variable K networks, compared to the invariable K networks, and what are the sources of that order. We also studied the impact of source elements (elements without inputs) on network dynamics. We suppose that variable K networks are even more ordered than invariable K networks, because the distribution alone creates frozen components that introduce a certain amount of order. Fox and Hill¹³ independently demonstrated the impact of K distribution and supported our hypothesis. They also determined the lower limit for the size of frozen component originating from Boolean functions and propagation of their effect through the network.

Most of our work was based on numerical simulations, while the effective K distribution in a part, which depends on the canalizing properties of Boolean functions and spreading of their influence through the network, was determined analytically. Our contributions are also two originally developed methods for analyzing the similarity and the relatedness of attractors.

GENETIC INTERPRETATION OF VARIABLE K NETWORKS

In Boolean networks with variable K , each element has its own K_i indicating the number of inputs, and K_o indicating the number of outputs (they are uniformly determined by the inputs), according to the assignment of connections. Regarding that, we distinguish four kinds of elements:

- *sources* ($K_i = 0, K_o = 1$) or autonomous elements
- *connected elements* ($K_i = 1, K_o = 1$)
- *sinks* ($K_i = 1, K_o = 0$)
- *unconnected elements* ($K_i = 0, K_o = 0$)

The sources influence the network's dynamics, but the network has no feedback influence on them. If a Boolean network models a genome, these elements could be interpreted as constitutive (if 1) or silent (if 0) genes, in both cases they are non-inducible. They can also represent external factors, which regulate the expression of inducible genes. Their value depends on their initial state, because it remains constant during the computation. We assumed that sources have a strong impact on network's dynamics, so we studied them with special attention.

Connected elements influence the network's dynamics and the network influences them back, through their inputs. They could be interpreted as genes inducible by other genes that can further induce some other genes. The regulatory self-inputs are also possible.

The sinks receive influence from the network, but they have no direct feedback influence on the network. We may interpret them as inducible genes that do not induce any other genes.

Unconnected elements do not interact with the network. They are members of the network, but are not dynamically connected to it. They could be interpreted as non-inducible (constitutive or silent) genes, having no influence on other genes – but they might still be important for the organism.

METHODS OF NETWORK'S CONSTRUCTION AND COMPUTATION

General

A Boolean network is defined with three matrices. The matrix of elements (*E*-matrix) determines the number of elements and their initial values (either 1 or 0). The matrix of connections (*C*-matrix) determines the directed connections between elements i.e. inputs and outputs. It also reveals the *K* distribution and the organization of the network. The matrix of functions (*F*-matrix) determines a random Boolean function for each element. Every possible combination of input values leads to either 0 or 1 outcome value of this element. For a certain *K* there are 2^{2^K} possible Boolean functions, which means that for *K*=1, there are 4, for *K*=2, there are 16, and for *K*=3, there are already 256 possible Boolean functions. If Boolean functions are chosen randomly, there are approximately equal fractions of 0 and 1 values as their outcomes, so the homogeneity bias is 0.5. If we determine this fraction to be different, we shift the homogeneity bias towards 1 or 0, so that one value (1 or 0) prevails. In this way we can externally impose some order into the network's dynamics.

The *C*- and *F*- matrices define one particular Boolean network, while *E*- matrix of this network may vary (this represents different initial states of the same network). To see the characteristics of network's dynamics we randomly determined various initial states with the same *C*- and *F*- matrices. To see some general characteristics of a certain kind of networks we randomly select various *C*- and *F*- matrices.

Specific in our research

In our work, we constructed networks with a predetermined number, but randomly chosen positions of source elements, by randomly assigning a given number of directed connections (specifying output and input element for each) between a given numbers of elements. The distribution of inputs and outputs after this process was binomial with on average around 13% of elements without input and the same percent without output (during this process only inputs to predetermined source elements were not allowed). To eliminate unwanted elements that aroused due to that, inputs or outputs were rearranged afterwards as necessarily to meet predetermined conditions as follows.

We allowed sinks, but we did not allow any additional source elements (regarding the predetermined number) or any inputs to the pre-existing ones. In order to eliminate eventual additional source elements we rearranged inputs in a random way until all surplus source elements had at least one input each. Neither did we allow source elements without any outputs, since this would only reduce the total number of them. They were eliminated with a random rearrangement of outputs. Unconnected elements were

generally not allowed, since they actually only reduce the total number of elements. They were also eliminated with the rearrangement of inputs (they became sinks, elements with no outputs).

We rearranged the inputs and outputs in two ways: either we redirected one uniformly at random chosen connection (1RC) or we redirected a random number of connections to the element (rndRC – the number of them was determined randomly using binomial distribution). We repeated the process until the connectivity distribution corresponded to the predetermined conditions described above.^c After the rearrangement of connections, the binomial distribution became skewed with the prevailing number of one input elements, which was higher in the 1RC case, and with fewer elements with 2 and 3 inputs than in rndRC case (see Figure 1). Both distributions lie between power-law distribution on one side, and Poisson and fixed K distribution on the other one. A motivation to study two distributions came from the initial study of 1RC case due to its considerable favoring of one-input elements. Two different distributions enable us also to show that relatively small differences in final distributions yield distinct results. 1RC distribution is closer to power-law distribution and is thus expected to be more ordered.

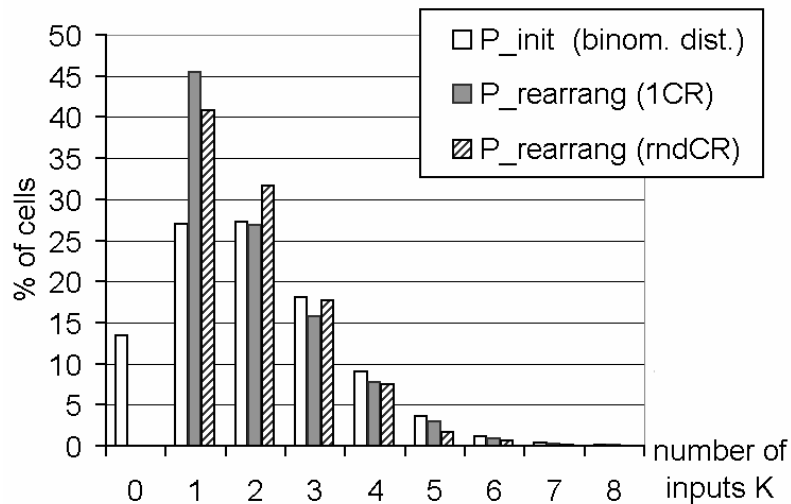


Figure 1: Average distribution of number of inputs per cell in networks with 0 source elements after: a) random assigning of connections (P_{init} - binomial distribution), b) rearrangement of the connections to avoid unwanted cells without inputs ($P_{rearrang}$ (1CR) for one connection toward such element redirected and $P_{rearrang}$ (rndCR) for random connections redirected).

In our study we also allowed selfinputs (but they appear with very low probability). The Boolean functions were not restricted in any way (contradiction and tautology, i.e. output is always 0 or 1, respectively, were also allowed). The number of inputs and outputs per

^c In general, with this iterative process only a few iterations were needed to obtain the required connectivity. An apparently more natural way to allow an initial random assignment of connections to determine the number of source elements yields in fact the networks with lower number of the latter with a too low probability.

element was not limited, but in fact, due to the nature of the distribution it rarely exceeded 7. We varied the predetermined number of source elements from 0 to 7. For each number of source elements and each case (1RC or rndRC) we constructed 25 different networks (different C - and F - matrices), each with: 1000 initial states for 0, 1, and 2 source elements in the networks, 2000 initial states for 3 and 4 source elements in the networks, and 3000 initial states for 5, 6, and 7 source elements in the networks.

Methods of results' analysis

Basic analysis of order in network dynamics usually includes the average and the median number of attractors, and their average and median length. Some authors further characterize network dynamics by analyzing the size of frozen elements fraction,¹³ or by various attractor classification methods. Attractors can be classified more precisely by determining which elements froze to a fixed value.¹⁶ The resemblance of those schemes suggests taxonomic classification of attractors.³ Regarding the unfrozen component, the attractors can be further compared and classified according to the cycle lengths of individual elements and the structure of these individual cycles.¹⁶ Both mentioned methods are useful only for the networks with a small number of elements, since they rely on a qualitative comparison of different schemes for attractors; they do not offer a quantitative measure for their similarity. This situation stimulated us to develop two quantitative methods for expressing the similarity of attractors, useful also for analyzing larger networks based on a similar classification of elements as described above.

Similarity of frozen components

The first method for analysis the similarity of attractors takes into account both the size and the degree of coincidence of frozen areas in different attractors and considers also the values of frozen elements. For this purpose, we classified elements into three groups: those frozen to 0, to 1, and active elements. This classification is analogous to the classification proposed by Weisbuch and Stauffer.¹⁶ Two attractors are considered more similar if they have more corresponding elements or sites that freeze to the same value in both attractors. Therefore, attractors in general exhibit greater similarity, if many corresponding elements in different attractors are predominantly frozen to the same value in most of them. A suitable measure for the coincidence of frozen areas in different attractors can thus be the largest frozen fraction (0 or 1) at each particular site, calculated over all attractors and then averaged over all sites (elements) in the network,

$$S = \frac{\sum_{n=1}^N (p_{n, frozen})_{\max}}{N} , \quad (\text{Eq. 1})$$

where n designates a particular element or site in the network, and $(p_{n, frozen})_{\max}$ is the maximal frozen fraction at this site calculated over all attractors. The sum runs over all sites (N) in the network.

If a particular element is frozen to the same value in all attractors, then its contribution to overall similarity (before division by N) is 1, if the half of its values is frozen to 1 and the other half to 0, then its contribution to the similarity is 0.5. Active elements diminish the largest possible fraction of frozen elements at each site, accordingly the maximal possible contribution to the similarity is therefore reduced to $[1 - \text{active fraction}]$ and the minimal possible contribution would be $[1 - \text{active fraction}]/2$. It would thus be sensible to express the similarity only above the estimated similarity, obtained in a random case. But for a better comparison with the second method presented below, we are going to use the complete quantity, as it is calculated with Eq. 1.

Overall similarity of attractors

The purpose of this method is to include a quantitative measure for the active component of each attractor and at the same time to retain the measure for the size and overlapping of the frozen component. As a quantitative measure for a particular element in an attractor, we used its average value over all the states in the attractor's cycle (therefore an average value 0.6 means that this

0	0	1	0,5	0,8	1	1	0	0	1
1	0	0,5	0	0	0,5	0	1	0	0
0	0	0	1	0,5	0	1	0	0	1
1	1	0,5	1	0,5	0	1	0	1	0,5
1	1	0,5	0,5	1	0	1	1	0	0
1	0,5	0,5	0,5	0	1	1	1	0	1
0	0	0	1	1	0	0,5	0	1	0
0	0	0	1	1	1	0,5	1	0	0
0	0	1	0	0,5	1	1	0	1	0
0	1	1	1	1	1	0	1	1	0

Figure 2: A scheme of an attractor, representing elements frozen to 0 and 1, and active elements with indicated their average values over the attractor's cycle.

particular element, when passing through one attractor cycle, had 60% values with value 1)^d. Each attractor is thus represented by a scheme with average values of particular elements over the attractor cycle (an example of such a scheme is shown in Figure 2). For comparison of different attractors, where each element can in general be frozen in some and active in other attractors, we calculated a fictitious “average” attractor, where the values of its elements were the average values over the corresponding elements in the schemes for individual attractors,

$$M_n^{avg} = \frac{1}{L} \sum_{l=1}^L M_{n,l} , \quad (2)$$

where M_n^{avg} is the value of the n -th element in the “average” attractor, $M_{n,l}$ is the value of the n -th element in the scheme for the l -th attractor and L is the number of all attractors in a given network. The similarity of corresponding elements in arbitrary two attractors l and l' can be defined as “one minus absolute difference between their values”,

^d Another possible measure for active elements were employed by Bagley and Glass.³ They counted the number of transitions between 0 and 1 of a particular element, when passing through a complete cycle.

$1 - |M_{n,l} - M_{n,l}^{\text{avg}}|$. The similarity S_l between an individual and the “average” attractor is therefore defined as the average over similarities of corresponding elements,

$$S_l = \frac{1}{N} \sum_{n=1}^N \left(1 - |M_{n,l} - M_n^{\text{avg}}| \right). \quad (3)$$

The overall similarity of attractors (S) in a given network is calculated as the average over individual attractor similarities S_l ,

$$S = \frac{1}{L} \sum_{l=1}^L S_l. \quad (4)$$

The overall similarity of attractors (Eq. 4) cannot be less than 0.5 (but it can be less between any two individual attractors or between a particular attractor and the average attractor). In an entirely random case, it would be expected that the fractions of 0 and 1 values were equal for every site in the network. The contribution of such a site to the similarity, if there were no active elements, would be 0.5, but in any other case, the similarity would be higher. Neither can active elements with their average values between 0 and 1 reduce the similarity below 0.5. The active fraction at particular site tends to increase the similarity in the case that its average value is close to the average value of this site. The similarity, obtained with this method, is always larger than that obtained with the first one (Eq. 1).

The formula (Eq. 4) does not distinguish between the cases when two average values of corresponding active elements are equal, but the successions of actual values are different or out of phase. These cases could be important in the networks with very high similarity of attractors, where there could still be differences in the structure of active component. They may also be important in the networks with larger fractions of active elements, where we would need a more precise measure to distinguish between them.

Effective distribution of inputs due to Boolean functions

Trying to get some insight into the origins of high order in K -variable networks, we analytically analyzed effective changes in the distribution of inputs that arise due to Boolean functions of particular elements and due to the propagation of their influence through the network (see Appendix for detailed description of our calculation). We first determined the fraction of elements that have fewer effective inputs k than the actual ones (K) solely due to their Boolean functions (Eq. A1 and Table AI). Considering these properties of Boolean functions, we get the first approximation for the distribution of effective inputs (Eq. A2). The influence of elements that have the number of effective inputs smaller than the actual ones is spreading through the network, so that the number of effective inputs to other elements is further diminished. Considering that, the second approximation for the distribution of effective inputs is obtained with Eqs. A3 and A4.

RESULTS

Number of attractors

The average number of different attractors for networks with different number of constitutive elements is shown in Figure 3a for both cases, when a random number of inputs is redirected during the network construction (rndRC) and when one input is redirected (1RC) (see Methods of Network's Construction and Computation). It increases approximately in an exponential manner with increasing number of source elements. In diagram, the exponential trendlines with their formulae are also shown. Such dependence is also theoretically expected, if one considers that source elements have 2^{N_c} different fixed inputs to the network. If we fix the values of source elements to be the same in all initial states, we can reduce the network to other elements with appropriate reduction of Boolean functions of elements with inputs from source elements, so that such particular network effectively behaves in the same way as before (the average number of attractors in networks with one choice of values of fixed inputs should also be slightly lower due to lower number of remaining elements).^e We see also, that in general the average number of attractors in rndRC case is higher than in 1RC case. We got similar results for the median number of attractors (see Figure 3b). The median numbers are in general lower than the average numbers of attractors. This is expected since a few networks with very large numbers of attractors less affects the median number than the average. The case with no source elements can be most directly compared to the results of other authors. In this case we got the average number of 10.1 attractors in the 1RC case and 14 in the rndRC case, while the median number of attractors was 5 and 3 respectively. For

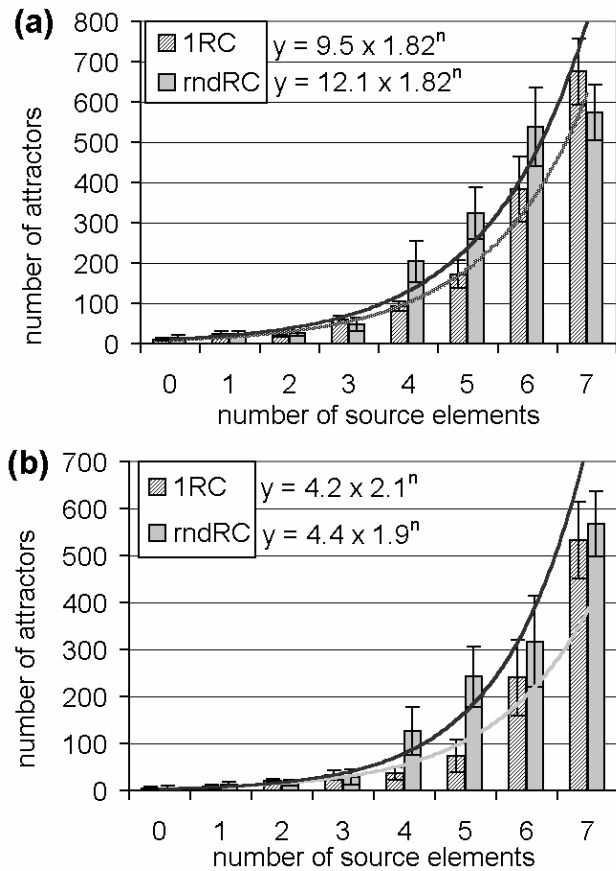


Figure 3: The average number (a) and the median number (b) of attractors found in networks with given number of source elements (elements without inputs) and with 1 (1RC) or random (rndRC) number of connection per needed element re directed at network construction. The exponential trendlines with their equations are also inserted.

^e This consideration shows also that critical point for networks with source elements remains the same (i.e. $K_{crit}=2$ for $K_{avg}=2$ and $p=0.5$, as shown generally for variable K networks for large N by Fox and Hill.¹³

comparison, in fixed $K=2$ model studied by Bagley and Glass,³ they found that their results for median are in good agreement with the $\sqrt{N/2}$ curve. For $N=100$ networks the interpolation of their results gave the median 5 attractors, while the average number was around 17 (similar result for the average number was found also by Bilke and Sjunnesson).⁴

Length of attractors

Figure 4 shows the average length and median length of attractors as a function of the number of source elements. They don't exhibit any obvious correlation to the number of source elements (this behavior is expected according to the theoretical consideration in previous subsection) and within the error limits also no difference between the networks that have one or a random number of inputs redirected. The average length of attractors over all different numbers of source elements is 8.6 for 1RC case and 10.1 for rndRC case, while the median length is 4.9 and 5.8 respectively. The interpolation of the results of Bagley and Glass³ for $N=100$ case gave a median length of 5 states (again a good agreement with $\sqrt{N/2}$ curve), while the average length was 13 states.

Distribution of networks according to the number of attractors

Table I shows the number of *different* attractors found in individual networks, arranged in columns according to the increasing number of source elements N_C . The networks in each column are arranged according to a decreasing number of attractors. As we can see, there are only a few networks with large number of attractors and many networks with just a few of them. This seems to be a common feature of Boolean networks in the most explored area around $K=2$. The lower limit for the possible number of attractors is 2^{N_C} , i.e. the number of possible different outputs from the source elements. We note also that the occurrence of networks with the highest number of attractors is subject to large fluctuations (the highest number of attractors in

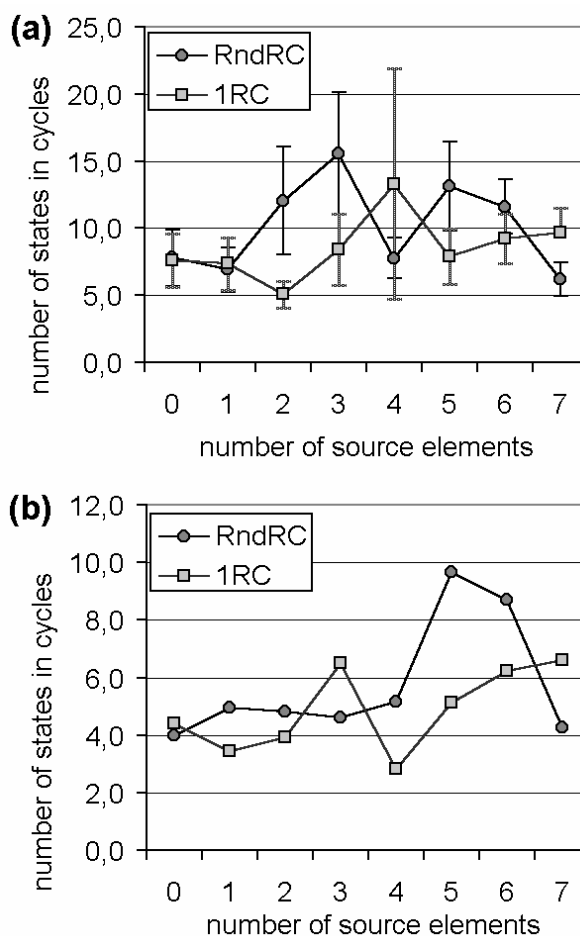


Figure 4: The average number (a) and the median number (b) of states in attractors, found in networks with given number of source elements (elements without inputs) and with 1 (1RC) or random (RndRC) number of connection per needed element redirected at network construction.

the networks with 7 source elements is lower than in the networks with 5 and 6 source elements).

Table I: The number of attractors in individual RndRC networks listed in columns and arranged from networks with higher toward the networks with lower number of attractors.

	number of source elements							
	0	1	2	3	4	5	6	7
Right: number of attractors in individual networks, sorted from networks with higher toward the networks with lower number of attractors	192	102	152	414	1153	1530	1709	1485
	55	97	90	111	671	849	1502	1326
	22	56	46	80	510	652	1295	981
	11	53	35	63	482	485	1247	904
	11	52	34	52	239	460	1150	763
	8	45	29	51	230	443	1021	733
	8	34	28	49	194	384	808	703
	6	28	26	37	190	340	609	697
	4	28	24	33	180	336	392	657
	4	26	22	32	175	329	366	624
	4	16	21	31	133	313	336	593
	3	13	21	28	131	260	320	592
	3	13	16	28	126	242	317	568
	2	12	16	26	110	221	309	532
	2	8	13	24	106	198	271	498
	2	8	11	20	85	198	262	486
	2	6	10	18	77	129	259	384
	2	4	8	16	67	127	229	359
	2	4	8	16	66	120	193	331
	2	4	8	15	66	112	192	266
2	4	8	12	49	96	191	256	
1	4	6	12	28	83	185	192	
1	2	5	9	16	77	128	138	
1	2	4	8	16	56	84	128	
1	2	4	8	16	48	64	128	
average	14	25	26	48	205	324	538	573
median	3	13	16	28	126	242	317	568

Distribution of attractors according to their lengths

We examined also the distribution of all found attractors regarding their lengths. Figure 5 shows the cumulative number of all attractors of a given length, shown for each number of source elements separately. The same attractor was counted as many times as it was found. We found the majority of point attractors, while among the cyclic ones mostly prevail the ones with even number of lengths. In all networks, there were 24% point attractors, 65% attractors with even number lengths and 11% attractors with odd number lengths.

This unequal distribution further indicates intrinsic order in Boolean networks that is not apparent at first sight. One possible explanation for this inequality is the fact that the lengths of attractors are often multiples of individual element cycle lengths or common

denominators of different individual cycle lengths¹⁶ (Weisbuch & Stauffer, 1987). If we multiply numbers among themselves, we get 75% even numbers and 25% odd numbers, which is then also the expected distribution of even and odd attractor lengths, if this were the main reason for the inequality.

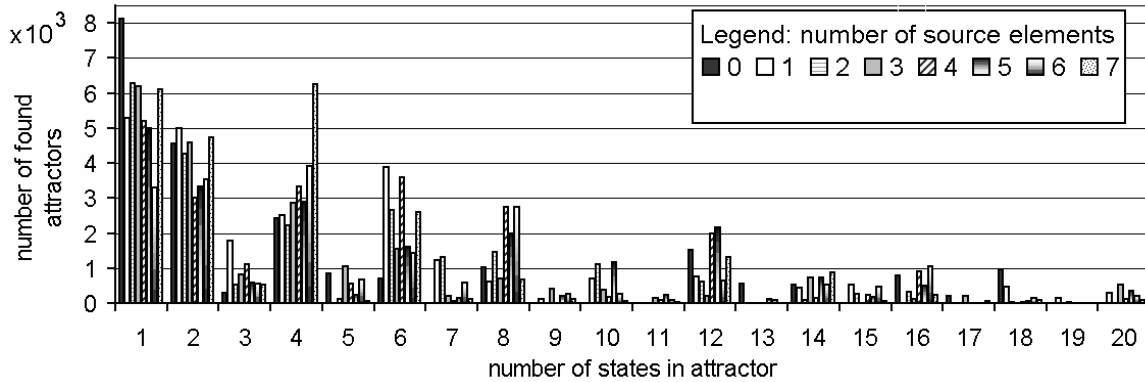


Figure 5: Number of all attractors of given length found in all examined networks, shown for each number of source elements separately. The same attractor was counted as many times as it was found.

Both for the number and for the length of attractors we consistently found lower average values than for fixed $K=2$ networks, reported in the cited literature, while the results for median values are more equalized. More detailed study including also distributions shows that regarding the cycle lengths we found 0,79% cycles with more than 100 states and 0,06% cycles with more than 1000 states (last ones in a two networks, one with 1524 states in some cycles and the other with 3124 states). Here we included all the networks into consideration since there is no dependence on the number of source elements. These results are in quite good agreement with the results that can be estimated from the work of Bagley and Glass,³ but the difference appears at still longer cycles, where they found

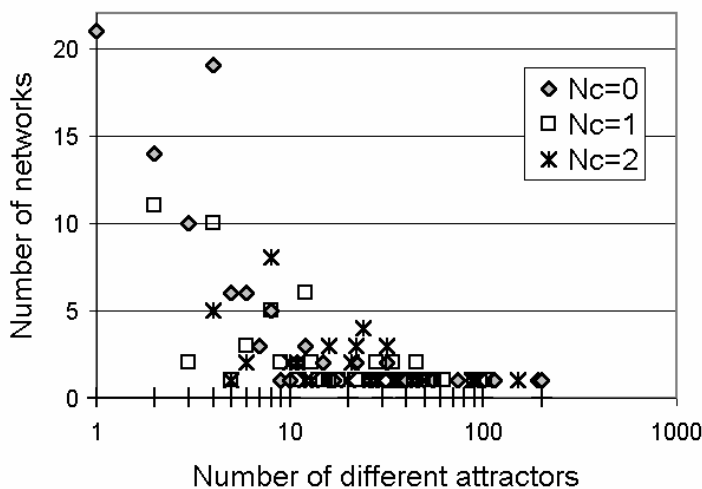


Figure 6: The number of different attractors per network, for networks with $N_c=0, 1,$ and 2 source elements.

considerable number of attractors with lengths over 3000 and even over 10^4 . Regarding the number of *different* attractors, we found fewer networks with many of them than in the above-cited work. There the number of different attractors extends up to around 800, while in our case, it extends only to around 200, and this holds for networks with 0, 1, and 2 source elements (see Figure 6^f and also Table I). Only the networks with the number of source elements 4 and above finally exceeded that number (800).

Similarity of attractors

As a first measure of attractor similarity, we compared frozen component sizes (see Figure 7a for 1RC case and in Figure 7b for rndRC case). There is no apparent correlation between the number of source elements and the average size of the frozen component. There are also considerable fluctuations in the frozen fraction size among networks with different number of source elements. Numerical values for sizes of frozen and active components are given in the last two columns of Table II.

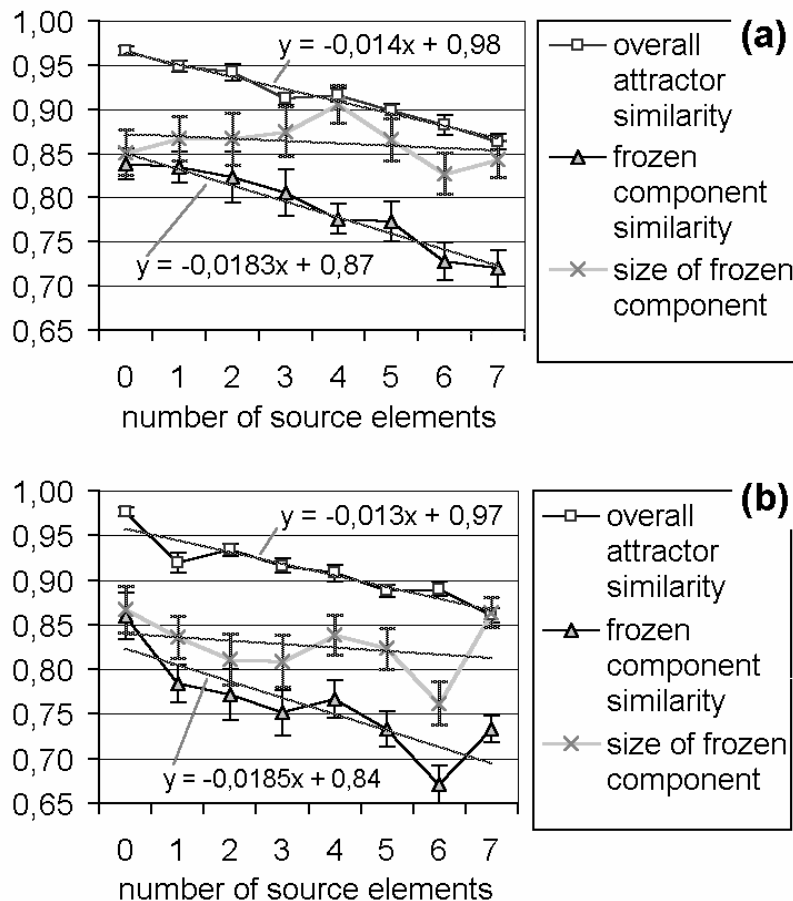


Figure 7: Size of frozen component, frozen component similarity (Eq.1) and overall attractor similarity (including also a measure for active component, Eq.4) in networks with: a) one connection per needed element at construction redirected (IRC case), and b) random number of connections redirected (rndRC case).

^fTo verify these results we extended for this analysis the number of networks, so that altogether 108 different networks without source elements and 66 networks with 1 source element were studied.

When calculating the similarity of attractors with both newly developed methods (Eq.1 and Eq.4), both similarities decrease with increasing number of source elements (see Figure 7a for 1RC case and in Figure 7b for rndRC case). To demonstrate more clearly the impact of source elements on the remaining network, we omitted the source elements from calculation of the similarity (otherwise, every source element with on average equal number of 0's and 1's would decrease the similarity for $0.5/N$). The networks with one connection redirected have in general slightly higher similarity of attractors. The curves obtained with both last methods are different in details, which indicates that contribution of the active component is not simply proportional to its size. Comparing all three curves in Figure 7b, it is apparent that the size of the frozen component is much more correlated with the similarity of frozen component (Eq.1) than with the overall similarity of attractors (Eq.2).

Table II: number of different types of elements in attractors

	FROZEN ELEMENTS					frozen + active	only active	stable froz.	SIZE OF:	
	Source elem.	Source canaliz.	Bool. + spread.	either 0 either 1	both 0 and 1				frozen comp.	active comp.
1RC	0	0	44	34	2	13	8	78	85	15
	1	2	41	32	4	15	5	74	87	13
	2	2	43	29	3	15	6	72	87	13
	3	3	38	28	7	19	1	66	87	13
	4	4	37	29	8	16	2	66	91	9
	5	4	35	26	8	19	3	60	87	13
	6	5	33	19	6	27	3	53	83	17
	7	6	32	15	8	30	2	47	84	16
RndRC	0	0	40	42	1	10	7	82	87	14
	1	2	37	29	5	21	5	66	84	17
	2	3	35	30	2	21	7	64	81	19
	3	4	35	22	5	24	7	58	81	19
	4	5	35	22	6	25	3	57	84	16
	5	6	33	19	6	31	1	51	82	18
	6	5	31	16	4	30	7	47	76	24
	7	6	32	17	10	26	2	49	86	14
Avg_1RC	3,5	3,3	38,0	26,6	5,7	19,2	3,7	64,6	86,2	13,8
Avg_RndRC	3,5	3,8	34,5	24,7	5,0	23,5	5,0	59,3	82,6	17,4
Average	3,5	3,6	36,2	25,7	5,4	21,3	4,3	61,9	84,4	15,6

NOTES: »*Source canaliz.*« are source elements canalizers (elements that effectively depend on the source elements only), »*Bool. + spread.*« are elements, that are frozen due to their Boolean functions and/or spreading of their influence«, »*either 0 either 1*« means other frozen elements, frozen in all attractors to the same value, »*both 0 and 1*« means frozen elements, that are in some attractors frozen to 1 and in some to 0 (without *source elements* and *source elements canalizers*), »*frozen + active*« are elements, that are in some attractors frozen and in some active, and »*only active*« are elements, that are active in all attractors. »*Stable froz.*« means stable frozen component, i.e. all elements with the same values in all attractors of given network (composed of »*Bool. + spread.*« and »*either 0 either 1*«).
Size of frozen and active component in last two columns shows the average size of these areas in all attractors irrespective of their position in a network.

Structure of the network and influence of Boolean functions

The results show that source elements diminish the *stable frozen component* (elements frozen due to their Boolean functions and/or spreading of their influence, and other frozen elements with the same values in all attractors), while the *frozen + active component* becomes larger, as well as the *unsteady frozen component* (i.e. frozen elements with different values in different attractors). The *active component* also becomes smaller proportionally to the growing number of source elements on the account of the growing *mixed (active/frozen) component* (see also Table II with the number of different types of elements found in our networks).

Approximately one-half of the frozen component consists of elements frozen due to Boolean functions and spreading of their influence through the network. This size of the frozen component is also in good agreement with theoretical prediction, presented in Appendix (see also Figure 8). From there it follows that the elements with a low number of inputs contribute mostly to the fractions of elements that have fewer effective inputs k than the actual ones (K) solely due to their Boolean functions (see also Table AI). Half of the elements with one input and 12.5% elements with two inputs become frozen solely due to that. This initial frozen area is further approximately doubled when the spreading of its influence is considered. The size of the frozen area emerged through this process is thus mainly dependent on the fraction of one-input elements in the original distribution (see also Figure 9). This is also seen in our results, where different frozen fractions in IRC case are in general higher than in rndRC case (see Table II).

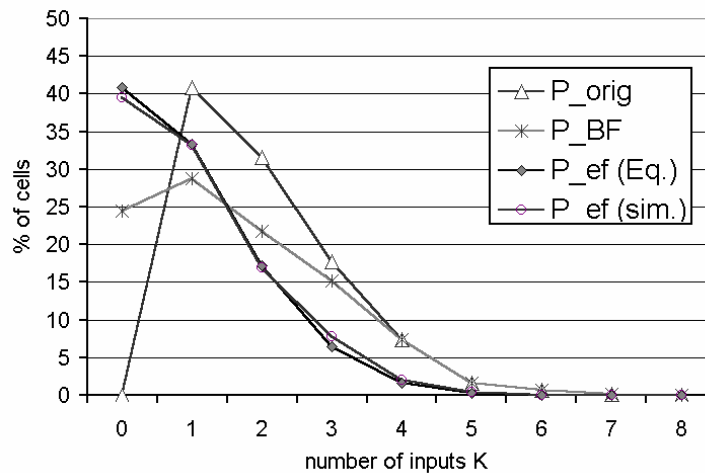


Figure 8: Effective distributions of number of inputs per cell in networks with 0 source elements and with random number of connections per needed element redirected: a) original distribution, P_{orig} , b) after taking into account properties of Boolean functions, P_{BF} , c) after taking into account also spreading of influence of Boolean functions through network, $P_{ef}(Eq.)$, d) average effective distribution due to reasons above, obtained with analysis of networks in simulations, $P_{ef}(sim.)$

We should emphasize here that we consider only Boolean functions that act as canalizers or freeze their output independently of their inputs. Therefore we do not include a class of

Boolean functions that become canalizers or freeze only in conjunction with some proper combinations of input values to them (such elements are presumably the main cause for *mixed*, i.e. *active/frozen component*, following different initial states of the same network). The growth of this component with an increasing number of source elements indicates that source elements with their various yet fixed outputs determine the behavior of some of the elements of this sort, according to the initial states. Similar growth of the *unsteady frozen component* can be attributed more to the unconditional freezing and canalizing of the elements, directly or indirectly connected with source elements. Other frozen elements, including a part of unsteady frozen component, may result also from various dynamic processes, for example different feedbacks that may extend in a complicated ways through many elements.

The effective distribution of inputs, calculated for other commonly used distributions, show the influence of different connectivity as source of order. In Figure 9 are shown original *fixed*, *Poisson* (very close to binomial), *power-law*, and our *rndRC* distribution, all for $K=2$, and effective distributions of them. We see that the most interconnected networks remain those with fixed K distribution (with 20% of no-input elements, and with 40% of one- and two- input elements each) and those with Poisson distribution (with still considerable fractions of elements with more than 2 inputs), while the least interconnected become networks with power-law distribution, where the number of elements with given number of inputs diminish approximately exponentially with increasing number of them.

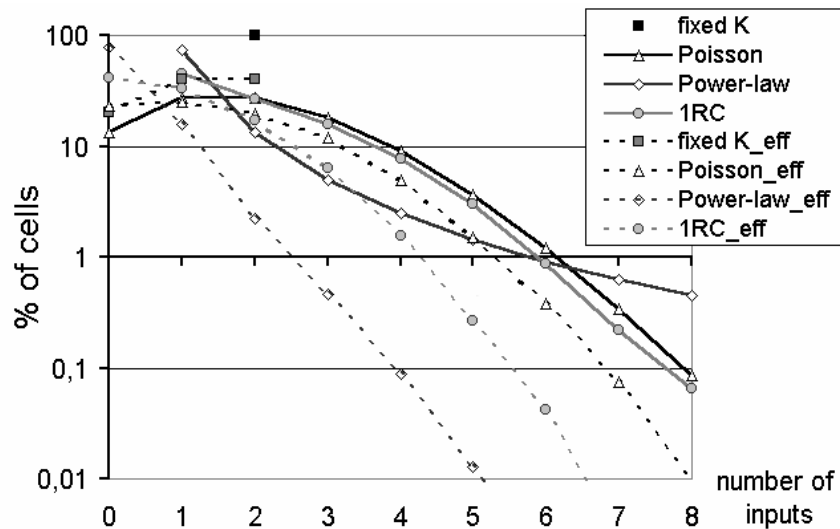


Figure 9: Fractions of cells with different number of inputs in original distributions (*fixed K*, *Poisson*, *Power-law* and our *IRC*) and corresponding effective distributions after calculation of the canalizing effect of Boolean functions and spreading of their influence (see Appendix).

DISCUSSION, CONCLUSIONS AND SOME QUESTIONS

We introduced a Boolean network model with variable number of connections per element and analyzed the degree of order in the model's dynamics. We used some standard methods, like the number of attractors, their length and distribution, the size of frozen components etc. With our newly developed methods, we also analyzed similarity of attractors. We also compared actual and effective distribution of connections. The orderliness we found was quite surprising and only the half of it can be accounted for by the evident causes like canalization of Boolean functions and propagation of this influence through the network. For the other half the causes are not so obvious, but we assume that some Boolean functions in conjunction with proper combinations of inputs and the complex dynamics itself create that additional order. This high and partly unexplainable order should be common also to living systems – at least where the structure of the system's dynamics is based on principles similar to Boolean networks. In addition, as it is seen from Kauffman's and related investigations, including our own, the cause of the high dynamic order surmounts the mere non-equilibrium thermodynamics that is considered the great architect of life. A significant amount of order is apparently an intrinsic property of living systems' complex network organization. Compared to the fixed K networks the analysis of K variable networks revealed a considerably higher degree of order, which is expected regarding the distribution of inputs, that is between the power law (more ordered) and Poisson and fixed K distributions (less ordered). Comparison of our results with fixed $K=2$ networks³ shows, that the median number of attractors and the median length of them were only slightly lower, while average number and average length were lower for around 30%. The occurrence of very long attractors and the networks with a large number of attractors are much more infrequent in our case than in fixed $K=2$ networks, too. This indicates that for low and medium numbers of attractors and their lengths both systems behave similarly, while there is more “free space” available for fixed $K=2$ networks to develop larger numbers of attractors and longer cycles. In variable K networks, there is a larger fraction of elements with low numbers of inputs. This creates larger frozen components that impose order in network's dynamics, than in the case of fixed K networks. Therefore, orderliness of K variable networks depends on both the average K and distribution of connections. K variable networks with the average $K=4.2$ and scale-free distribution are still ordered,¹³ while Kauffman's invariable $K=4$ networks are far beyond the threshold of the so-called disorder. Regarding the contribution of K distribution to the orderliness, the scale-free distribution seems to be the most effective (see also Figure 9).¹³ Our work also shows, that the (manner of) rearrangement of inputs to eliminate unwanted elements may have strong impact on the final distribution and consequently on the results obtained. It would be very interesting to determine the phase transition from order to so-called disorder in K variable networks with different distributions, and the impact of distributions on the orderliness. This would reveal the average K in networks “on the edge of chaos”, and this would apply to most biological systems, since according to many researches the majority of them seems to be poised on this edge.^{2 17} This special regime enables them to be very flexible and to adapt successfully on various environmental changes. With slight modifications of connectivity distribution, the edge of chaos would vary, so this may also represent one of the adaptational evolutionary mechanisms.

The network analysis revealed a correlation between the number of source elements and the number and similarity of attractors. As the number of source elements increases, the number of attractors does likewise almost in an exponential manner. This means that any additional autonomous influence (a source element) markedly increases the network's dynamic diversity. This fact also corresponds to a decreasing similarity of attractors, as the number of source elements increases. However, this reduction is only linear as compared to the exponential increase in the number of attractors, which indicates that the newly utilized parts of the state space remain close to the already explored ones. The source or no-input elements actually act as autonomous or independent sources of network's dynamics since the network has no feedback influence on them. The power of the source's influence most probably depends on the number of its outputs because its influence is enhanced in that way. Therefore the value of the source element dictates the network's dynamics and we assume that the perturbation (switching) of its value (for instance due to an external factor) strongly affects the network's dynamics, which consequently diverts into some other attractor. The source element is therefore able to split the dynamics of the network, since the network goes into one or another attractor, depending on the value of the source element. The smaller the number of sources, the lower is the dynamic diversity (the number and the similarity of attractors) of the network.

In genetic interpretation of the model, the source elements represent constitutive genes or external factors, which influence the expression of the genes. This means that they dictate the network's dynamics and their perturbation diverts the network into other attractor that represents different gene expression pattern. This would happen during the differentiation of the cell into a specific cell type, or during some pathological process that would disturb the normal gene expression pattern (for instance carcinogenesis). Therefore, the alteration of an external factor or a constitutive gene expression affects the expression of a large number of other genes – forming an expression module.

The number of attractors increases with the number of source elements but their distribution of lengths remains very similar. The average attractor length does not seem to be correlated to the number of source elements. Consequently, although the dynamic diversity increases with increasing number of source elements, the orderliness of networks stays the same. If the average length of attractors would also increase, those networks would become less ordered.

An interesting characteristic of attractor lengths is the prevalence of even number lengths over odd number lengths (with exception of point attractors which are the most frequent). This could be a consequence of the model design (binary, synchronous) and therefore an artifact as such (if switching is asynchronous, so that only one element switches at a time, all cycles must be of even length, as in order to return to an initial state, each element must switch an even number of times), or perhaps a form of symmetry breaking that spontaneously emerges in a complex system. It should be further investigated what are the expected frequencies of even and odd number lengths and what do they depend on (e.g. average K and connectivity distribution). If the expected frequencies would not agree with the actual ones it would mean that the symmetry breaking is also a common property of these network systems.

Further investigations of variable K networks will reveal the structure and stability of their internal order. There are some possibilities for comparing the attractors (by

comparing fractions of active elements, length of attractors, the width and the steepness of their basin of attraction) between different networks with different average K and connectivity distributions. The assumptions about the influence of source elements on network's dynamics should also be investigated. An analysis of network's dynamic homeostasis would be very important. It is essential to investigate if and how the perturbations and mutations alter the behavior of networks with different average K and different connectivity distributions. Another interesting research could analyze the manner in which different parameters (N , K , and distribution) cause changes in the networks' state space and where are the bifurcations. This could be one of tuning mechanisms, mentioned by Kauffman for adaptation on large environmental changes in evolution.

The evaluation of asynchronous and continuous models, their significance and the parallels with synchronous model dynamics is also crucial for further development of biological systems modeling. It seems that living systems like other compositions of non-linear oscillators spontaneously synchronize, so the models should include some sort of phase coherence or rhythmic patterns, but not necessarily the absolute synchrony. Regarding the modeling of genetic networks, it should be investigated, if continuous models would predict a more realistic number of different attractors than discrete Boolean networks, since many attractors in the discrete case could merge into one attractor in the continuous case.³ In the model the number of different attractors would correspond to the number of different cell types, which is approximately \sqrt{N} (N is the number of genes).

The distribution of connections raises a question about the influence the number of elements has on network's dynamics. In networks with a small number of elements, the stochastic effect of connectivity distribution is very pronounced, so the distribution can vary extremely. Moreover, since the distribution markedly affects the orderliness of the network, the degree of order should differ significantly in such different networks. To what extent the number of elements influences the orderliness of networks could represent an interesting future investigation.

An important question, which remains to be thoroughly investigated in the future, particularly in relation to natural systems, is related also to the growth of the system. Here, a way should be found to include new elements and at the same time to retain functional integrity and the main characteristics of the system, possibly on the principles of self-organization and self-ordering.

There are many interesting researches to be done with the Boolean network models in the future and those new findings are expected to give profound insights into the spontaneous orderliness of complex systems, especially the biological ones.

APPENDIX – EFFECTIVE DISTRIBUTION OF INPUTS DUE TO BOOLEAN FUNCTIONS AND THEIR CANALIZATION

In order to investigate the reasons for surprisingly high order of K -variable networks, we compared the actual and the effective distribution of inputs. First, we determined the expected number of elements that have the number of effective inputs smaller than the number of actual inputs solely due to the canalization properties of Boolean functions; i.e. the outcome of certain Boolean functions is independent of some inputs. The element

with K inputs has $n_B(K) = 2^{2^K}$ possible Boolean functions. The total number of 0's and 1's in each function equals 2^K , which is also the number of all possible combinations of inputs. Two of these functions are independent of all inputs (the function with all 1's and the function with all 0's). The number of functions that are independent of all except one input is $2K$. For 2 inputs such functions are (0011), (1100), (1010) and (0101). The first two canalize one input and the last two the other one. An example for Boolean functions with 3 inputs, which is dependent on the first input, is (11110000), while the function (10001000) depends on the second and the third input but not on the first one. A general expression for the number of Boolean functions, that effectively depend on k inputs (i.e. effective inputs) and are therefore independent of $K-k$ inputs, can be obtained in the following way. Such function is composed from (one) function for fewer number of inputs k , which is repeated in the original function 2^{K-k} times, while from the position of its elements depends, which input(s) it canalizes. In the above example for $K=3$ and $k=1$, such function is (10), which is distributed in the original function like (11..00..) and repeated in this sequence through the whole original function (the other two possible distributions are (1100....) and (1010....). Function with k inputs can be distributed in

$\binom{K}{k}$ possible ways, which is also the number of possible combinations of k inputs out of K . The number of functions, that canalize one combination of inputs, is $n_B(k)$, but we must subtract those functions, that already canalize fewer number of the m than k . This leads us to the recursive expression:

$$n_B(k, K) = \left[n_B(k) - \sum_{k'=0}^{k-1} n_B(k', k') \binom{k}{k'} \right] \binom{K}{k}, \quad \text{A1}$$

where $n_B(k, K)$ is the number of Boolean functions with k effective inputs and K actual inputs and $\binom{m}{n} = \frac{m!}{n!(m-n)!}$ is the binomial coefficient (see also Table AI for the number of Boolean functions at different values of k and K).

This expression gives us the first approximation for the effective distribution of inputs. In Boolean network with N elements and the distribution of inputs $P(K)$ we have on average $N_K = N P(K)$ elements with K inputs. The $n_B(k, K) / n_B(K)$ fraction of these elements has k effective inputs. The average number of all elements with k effective inputs is given by this sum:

$$N_B(k) = \sum_{K=k}^N \frac{n_B(k, K)}{n_B(K)} N_K = \sum_{K=k}^N P_B(k, K) N_K, \quad \text{A2}$$

and the modified distribution of effective inputs due to Boolean functions is given by $P_B(k) = N_B(k) / N$ (we call these effective inputs BF inputs).

In the next step we determine the elements that become frozen or their effective inputs are (further) reduced due to the propagation of BF frozen elements influence through the network, i.e. if they have inputs from the elements that are already frozen or become frozen through this process. Let us now calculate the total number of frozen elements (BF frozen and frozen through the spreading of BF frozen influence), $\tilde{N}(0)$. The probability,

that an element with one effective BF input becomes frozen, is $\tilde{N}(0)/N$, the probability for an element with two effective BF inputs to become frozen is $(\tilde{N}(0)/N)^2$, etc. The average frozen component due to the Boolean functions and the propagation of their influence is given by equation:

$$\tilde{N}(0) = N_B(0) + N_B(1) \frac{\tilde{N}(0)}{N} + N_B(2) \left(\frac{\tilde{N}(0)}{N} \right)^2 + N_B(3) \left(\frac{\tilde{N}(0)}{N} \right)^3 + \dots ,$$

A3

which can be solved by iteration.

For other elements, the number of effective inputs is additionally reduced if they have any effective inputs that are or become frozen. Let us calculate the total number of elements with k effective inputs $\tilde{N}(k)$. We get the average number of elements with $k > 0$ effective inputs $\tilde{N}(k)$ by subtracting the number of elements with any frozen inputs from the number of elements $N_B(k)$ with k effective BF inputs (eq. A2) and by adding the number of elements with any frozen inputs that have higher number of effective BF inputs,

$$\tilde{N}(1) = N_B(1) - N_B(1) \frac{\tilde{N}(0)}{N} + 2N_B(2) \frac{\tilde{N}(0)}{N} \frac{N-\tilde{N}(0)}{N} + 3N_B(3) \left(\frac{\tilde{N}(0)}{N} \right)^2 \frac{N-\tilde{N}(0)}{N} + \dots ,$$

A4a

$$\tilde{N}(2) = N_B(2) - N_B(2) \left(\frac{\tilde{N}(0)}{N} \right)^2 - 2N_B(2) \frac{\tilde{N}(0)}{N} \frac{N-\tilde{N}(0)}{N} + \binom{3}{2} N_B(3) \frac{\tilde{N}(0)}{N} \left(\frac{N-\tilde{N}(0)}{N} \right)^2 + \dots ,$$

A4b

$$\tilde{N}(3) = N_B(3) - N_B(3) \left(\frac{\tilde{N}(0)}{N} \right)^3 - 3N_B(3) \left(\frac{\tilde{N}(0)}{N} \right)^2 \frac{N-\tilde{N}(0)}{N} - \binom{3}{2} N_B(3) \frac{\tilde{N}(0)}{N} \left(\frac{N-\tilde{N}(0)}{N} \right)^2 + \binom{4}{3} N_B(4) \frac{\tilde{N}(0)}{N} \left(\frac{N-\tilde{N}(0)}{N} \right)^3 + \dots ,$$

A4c

⋮

The terms on the right contain factors for the probability that the estimated number of inputs is actually frozen, the probability that all other inputs are not frozen, and the number of possible combinations of such inputs.

Frozen elements may also result from other processes, for example the various feedbacks that may extend in a complicated ways through many elements. A simple example is an element with self-input only, that is frozen if its Boolean function is 00 or 11 (this possibility is already mentioned above), but also if its Boolean function determines the 1(0) output, if the input is 1(0) – we can say that it is dynamically frozen; in reversed case the element flickers. It can be demonstrated that in simple circular feedback loops with 1, 2, 3, ... elements and one effective input per element, and if Boolean functions of all elements in the loop are (10) or (01), the probability, that those elements become (dynamically) frozen is 50%, 25%, 12.5%, ..., respectively (a total probability, including all functions, that such loops freeze strongly increases, since only one frozen element freezes the whole loop). The probability for any larger feedback loops to freeze dynamically is getting progressively smaller, but the number of possible variations is increasing enormously. The case when an element with more than one effective input

becomes frozen if one of its inputs is frozen to a particular value, but it remains active if it is frozen to the other value, gives rise to the elements, that are frozen in some and active in other attractors.

Table AI: The number of all Boolean functions with K inputs $n_B(K)$ and the number of Boolean functions $n_B(k, K)$ with effective number of inputs k at the actual number of inputs K .

K	$n_B(K)$	$n_B(k, K)$					
		$k=0$	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$
1	4	2	2				
2	16	2	4	10			
3	256	2	6	30	218		
4	65536	2	8	60	872	64594	
5	4,29E+09	2	10	100	2180	322970	4.3E+09

-
- ¹ S. A. Kauffman, *Origins of Order : Self-Organization and Selection in Evolution*, (Oxford University Press, Oxford, 1993)
 - ² S. A. Kauffman, *At Home in the Universe: The Search for the Laws of Self-Organization and Complexity*, (Oxford University Press, Oxford, 1996)
 - ³ R. J. Bagley and L. Glass, "Counting and Classifying Attractors in High Dimensional Dynamical Systems," *J. Theor. Biol.* 183, 269-284 (1996)
 - ⁴ S. Bilke, F. Sjunnesson, "Stability of Kauffman Model", *Phys. Rev. E* 65, 016129 (2001)
 - ⁵ J. E. Socolar, S. A. Kauffman, "Scaling in Ordered and Critical Random Boolean Networks", *Phys. Rev. Lett.* 90, 068702 (2003)
 - ⁶ B. Samuelsson, C. Troein, "Superpolynomial Growth in the Number of Attractors in Kauffman Networks", *Phys. Rev. Lett.* 90, 098701 (2003)
 - ⁷ E. A. Di Paolo, "Rhythmic and non-rhythmic attractors in asynchronous random Boolean networks," *Biosystems* 59, 185-195 (2001)
 - ⁸ H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabasi, "The large-scale organization of Metabolic Networks," *Nature* 406, 651-654 (2000)
 - ⁹ K. W. Kohn, "Molecular Interaction Map of the Mammalian Cell Cycle Control and DNA Repair Systems," *Mol Biol of Cell* 10, 2703-2734 (1999)
 - ¹⁰ P. Uetz, L. Giot, G. Cagney, T. A. Mansfield, R. S. Judson, J. R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart, A. Qureshi-Emili, Y. Li, B. Godwin, D. Conover, T. Kalbfleisch, G. Vijayadamodar, M. Yang, M. Johnston, S. Fields, and J. M. Rothberg, "A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*," *Nature (London)* 406, 623 (2000)
 - ¹¹ S. Tavazoie, J. D. Hughes, M. J. Campbell, R. J. Cho, and G. M. Church, "Systematic determination of genetic network architecture," *Nature Genetics* 22, 281-285 (1999)
 - ¹² A. L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, Vol 286, Oct 1999

- ¹³ J. J. Fox and C. C. Hill, "From topology to dynamics in biochemical networks," *Chaos* 11, 809-815 (2001)
- ¹⁴ R. Albert, H. Jeong, and A. -L. Barabasi, "The internet's Achilles heel: Error and attack tolerance of complex networks," *Nature* 406, 378-382 (2002)
- ¹⁵ A. -L. Barabasi, R. Albert, and H. Jeong, "Mean-field theory for scale-free random networks," *Physica A* 272, 173 (1999)
- ¹⁶ G. Weisbuch and D. Stauffer, "Phase transitions in cellular random boolean nets," *J. Physique* 48, 11-28 (1987)
- ¹⁷ R. V. Solé and B. Goodwin, *Signs of Life*, (Basic Books, Member of Perseus Books Group, New York 2000)