Recent Results on Ordering Parameters in Boolean Networks

Jürgen Klüver*  
Jörn Schmidt  
University of Duisburg-Essen, FB2,  
Campus Essen,  
Universitätsstr. 12,  
45117 Essen, Germany

The dynamics of discrete mathematical systems can be related to the values of certain ordering parameters. Experiments with the total set of possible structures of a boolean network with six elements and extended sets of two-valued boolean functions show that three simple parameters describing the topology, the bias of boolean functions, and their canalyzing potential yield necessary conditions for complex dynamics. The crucial factor determining dynamics turns out to be the topology. The impact on simulations of real systems is discussed.

1. Introduction

Simulations of real social systems by mathematical tools like cellular automata or boolean networks are widely applied by scientists as a tool to predict or control the development of complex systems. This simply means running the simulation forward through time and looking at the results [1]. Unfortunately the dynamics of mathematical models like the ones mentioned simultaneously depend on three basic conditions; namely, topological rules, transition rules, and initial conditions. The fact that topology, transition rules, and initial states are intricately interwoven constitutes a considerable obstacle for any attempts to predict the dynamics even of rather small mathematical systems from simple parameters describing the properties of the underlying structure of connections or the set of transition functions. Thus a variety of such attempts, set out under the term of “control parameters” or “ordering parameters” [2–8] did not prove as totally satisfactory. On the other hand, as we pointed out elsewhere [8, 9], the existence of meaningful ordering parameters in discrete, mathematical systems is an essential precondition of the quest for parameters used to forecast the dynamics of complex systems.

Most of the studies found in the literature, inspired by physical or biological problems, deal with random cellular automata or with their

*Electronic mail address: juergen.kluever@uni-due.de.

generalization, random boolean networks (RBN) (cf. [10] for a recent overview), which are potential universal Turing machines. Randomization seems appropriate for modeling larger systems, which those authors aim at, and it may even be necessary in order to derive hypotheses concerning large systems at all. Randomization may be applied on each of the three levels: topology, transition functions, or initial states, or on all of these levels. There are even studies with random updating schemes. These attempts usually require random inputs like scale-free topologies [11], but nevertheless yield remarkably precise, though essentially statistical results, referring, for example, to robustness and stability [11, 12], or scaling laws [13].

In this study, however, we ultimately aim at modeling the dynamics of small social groups by discrete mathematical systems. For this purpose the study has to proceed from exact topologies and precisely assigned functions, since the members of the group to be modeled are individually connected within a well defined social structure and possess individual rules of perception and action. The necessary inhomogeneity of connections and transition rules suggests the use of boolean networks, which—by explicitly including structurally inhomogeneous environments and transition functions—represent a generalization of cellular automata appropriate to map such a configuration.

The study intends to contribute to a more precise knowledge concerning the interaction between topology and transition functions.

2. Experiments

The study is essentially an experimental attempt. The experiments will be confined to examine the dynamics, in particular the appearance of high periods,\(^1\) of small boolean networks (BNs) with only six elements and two incident connections (homogeneous indegree \(k = 2\)). In the sense of the principle of “less is more,” much discussed recently [14], these small BNs may also serve as models for more expanded mathematical systems. Of course there are specific reasons for the restrictions chosen here. Restriction to \(k = 2\) or two-valued boolean functions, respectively, does not limit generality since any BN with \(k > 2\) may be substituted by an appropriately extended BN with \(k = 2\).

The study will emphasize the question of topology: Is the topology of interaction, (in a BN it is the connection structure of the underlying digraph), the crucial factor, which basically determines or limits, respectively, the dynamics of a system? However, as is well known, describing the structure of digraphs is anything but straightforward. Even in this

\(^1\)Periods of these systems with six elements are limited to \(p \leq 2^6\). We define high periods as being \(p > 2^5\). In special cases we separately consider \(p > 59\) as representing “complex dynamics.”
very restricted case of BNs with six elements and \( k = 2 \) we get \( 10^6 \) possible digraphs. With \( 16^6 \) combinations of six out of the 16 two-valued boolean functions and \( 2^6 \) possible initial states of the system, studying the complete set of approximately \( 10^{15} \) different BNs would be a rather time consuming undertaking. However, it turns out that this would be unnecessary, since there are meaningful further reductions of the possibilities to be examined.

A majority of authors have been reducing the number of structures by selecting random samples, hoping that these will be representative in a way. At least in the case of structures this strategy may be questionable. The set of \( 10^6 \) digraphs complying with the conditions of the BN(6, 2) can be partitioned into equivalence classes of isomorphic structures. The number of classes, that is, of nonisomorphic or topologically different structures, can easily be calculated as 1499. The point here is that these classes are of very different size so that any (e.g., an equally distributed) random sample taken out of the complete set of \( 10^6 \) digraphs would by far not be a representative mapping of possible structures. Hence, with respect to emphasis on structure the whole set of 1499 nonisomorphic digraphs was employed in this study. The 64 possible initial states, too, were included in total.

Restrictions had, however, to be made with regard to the number of combinations of boolean functions. Obviously, the contradiction and the tautology can be left out because they practically neutralize the elements they are assigned to, reducing the period of the BN by a factor 1/2. Out of the remaining set of 14 boolean functions 22 subsets of three functions each were selected; from any of these subsets \( 3^6 \) combinations yielding six functions were formed. Hence the bulk of the experiments reported here was based on a set of \( 1499 \times 2^6 \times 22 \times 3^6 \approx 1.6 \times 10^9 \) BNs. Updating was synchronous in all cases (cf. [15] for the relevance of other updating schemes).

One of the aims of this study is to investigate the predictive potential of ordering parameters. In this respect the discussion will be restricted to three well-known simple parameters; namely, the \( P \)-parameter [2, 3] expressing the bias of the boolean functions, a measure of the canalizing potential [5] referring to the boolean functions, and the so-called \( \nu \)-parameter concerning the underlying digraph structure [8]. The first two parameters had, of course, to be applied as the average over the set of six assigned functions.\(^2\)

The \( \nu \)-parameter measures the proportion of influence the different cells potentially exert on each other. It is easily derived from the adjacency matrices representing the structures of the digraphs the BNs are

\(^2\)We are well aware of the fact that the exact assignment of each of the boolean functions to certain vertices of the underlying digraph is an important feature of structure, see the discussion following Figure 5.
J. Klüver and J. Schmidt

Figure 1. Three different structures: (a) equally distributed, (b) centralized, (c) intermediate.

Table 1. Adjacency matrices and outdegree sequences corresponding to the structures of Figure 1.

Based on. The implications of its definition are demonstrated by the examples in Figure 1 and Table 1.

Figure 1 exhibits three different digraph structures: equally distributed interaction, extremely centralized, and “intermediate” structure.

The respective adjacency matrices with their according outdegree sequences are shown in Table 1.

The ν-parameter is defined as

\[ \nu = \frac{|(OD - OD_{\text{min}})|}{|OD_{\text{max}} - OD_{\text{min}}|}, \]

where OD is the actual outdegree sequence vector, OD_{\text{max}} and OD_{\text{min}} indicating the vectors with extremal lengths (Figures 1(a) and 1(b)).

It has to be emphasized here that ν, in contrast to other parameters describing connections in BNs, exclusively refers to the outdegree.

Kauffman [5, 6] has introduced the notion of canalyzing functions to provide for a measure of the fact that certain boolean functions do not depend upon all of their variables. Since Kauffman ([5], but c.f. [16]) defined the term binary—a function is either canalyzing or not—and since the term is not independent from other parameters of the boolean

Complex Systems, 17 (2007) 29-45
functions, like for instance $P$, it is of rather limited value. To take the property of canalizing into consideration nevertheless a slightly more sophisticated “canalyzation index” $C$ (for two-valued boolean functions) has been applied, which takes into account that boolean functions may be independent of none, of 1, of 2, or of all of the variables. The examples in Table 2 may demonstrate the definition (together with boolean function codes and $P$-parameters).

Of course, a considerable number of other parameters have been discussed in the literature. Many of those turn out to be dependent on the parameters used here. Others are basically statistical measures, which tend to wipe out structural differences and therefore were deliberately not applied here. Others again, like spectra of digraphs, need more thorough inspection to be interpretable and are left out for the sake of simplicity. For the same reason attractor basins are not presented here. Thus, within the scope of this study, the experimental parameters introduced were simply taken as the basis to relate the periods observed in the BNs, and to observe what comes out.

3. Results

Some of the results are immediately striking. Figure 2 exhibits periods observed with a typical sample of 2000 subsets of three functions, combined to 729 sets of six assigned functions each, and applied to all of the 1499 structures and 64 initial states.3

First, it is obvious that high periods ($p > 32$, or even $p > 59$) are extremely rare; only 0.04% or 0.02% of the BNs possess those high periods.4 Second, there exists a clear relation between the frequency

---

3We thank one of our students, Falko Hildebrand, for performing extensive supplementary experiments, which support the results presented here.

4The highest period observed is 63. It may be worth mentioning that certain periods (among these $p = 41, 42, 44$ up to 59, and 61) never appear.
Figure 2. Frequency of high periods \textit{versus} \( v \) in a random sample of \( \approx 1.4 \cdot 10^{11} \) BNs.

Figure 3. Frequency of high periods \textit{versus} \( v \) in a selected sample of BNs.

of BNs with high periods and the \( v \)-parameter; this will be discussed below.\(^5\)

Figure 3 shows the results obtained with the 22 subsets of boolean functions mentioned earlier. Since these subsets comprise a majority of

\(^5\)The number of BNs is related to the number of each of the structures constituting the respective value of \( v \), since these latter numbers are different for different \( v \).

functions with $P = 0.5$, the frequency of higher periods is considerably increased. The selected set of functions is therefore biased towards higher periods. It is remarkable that periods $> 59$ are more frequent than periods between 33 and 59.

The comprehensive view entailed in Figures 2 and 3, however, needs to be complemented with some more details. Of course it can be expected that only part of the 1499 structures as well as part of the combinations will lead to high periods. As it turns out, 727 of the structures—combined with appropriate combinations of boolean functions and initial states—yield instances of high periods. The remaining 772 structures are exactly those which contain at least one vertex with an outdegree of 0, that is, a sink. It is immediately evident that a sink—comparable with an assigned boolean function no. 0 or no. 15 (contradiction or tautology)—will, as it were, neutralize the respective edge. The existence of vertices with outdegree 0 influences the $v$-parameters. In Figures 2 and 3, the points with zero frequency ($y$-axis) in the lines representing high periods correspond precisely to $v$-parameters ($x$ coordinates) belonging to structures with sinks. In particular, the peculiar absence of BNs with higher periods for $v = 0.577$ is caused by sinks in each of the structures belonging to this $v$-parameter.

Hence for small BNs the occurrence of just one sink will prevent high periods, whereas in more extended BNs only a larger fraction of sinks will evoke such an effect.

Figures 2 and 3 depict the comprehensive situation for a large number of sets of boolean functions. The situation for a single set of functions, that is, the 729 combinations of a subset of three functions, is reflected in typical examples given in Figures 4 and 5. The plots exhibit essentially the same relation between structure and periods: High periods predominantly occur with $v$-parameters in the range between approximately 0.0 to 0.5, their frequency decreasing with increasing $v$.

The only exception is a single structure with $v = 0.68$. Obviously still more detailed studies are necessary in order to relate the characteristics of single structures to periods in an even more precise manner.

A simple example may emphasize that an intricate investigation will yield fruitful results: Although structures with $v = 0$ show the highest frequency of high periods there are many structures among these that do not only not allow high periods with whatever combination of functions, but, on the contrary, have extremely low periods. It turns out that these are structures which do not contain any triangle. This clearly underlines that there are other independent parameters relating structure and dynamics and that $v$-parameters though certainly a convenient measure of structural properties, are still insufficiently precise and discriminating.

---

6Boolean function codes are the decimal transforms of the binary strings given by the one-dimensional arrangement of truth tables (see also Table 1).
**Figure 4.** Frequency of high periods *versus* $v$ in a subset of all combinations of the functions no. 9, 10, and 14.

**Figure 5.** Frequency of high periods *versus* $v$ in a subset of all combinations of the functions no. 3, 5, and 6.
Recent Results on Ordering Parameters in Boolean Networks

<table>
<thead>
<tr>
<th>BF code</th>
<th>4</th>
<th>7</th>
<th>9</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>9</th>
<th>10</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P$</td>
<td>0.75</td>
<td>0.75</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.5</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>$C$</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Codes, $P$, and $C$ values of three subsets of functions applied in Figures 4, 5, and 6.

It has to be pointed out that the diagrams presented continue to constitute a comprehensive view with respect to the set of combinations of six boolean functions, neglecting that different combinations out of the applied subset of three functions may exert a very different influence on the period of the BN. Of course, the interaction between structure and boolean functions has actually to be examined for each of the particular combination of functions. A combination of functions $(6,6,6,6,6,13)$ may well result in different dynamics compared with its permutations, for example, $(6,6,6,13,6)$, if the edges that function 13 is assigned to are differently connected within the structure.

Since all of the possible 729 combinations were applied, it is guaranteed that any conceivable location of the individual functions actually occurs. However, it is not within the scope of this study to examine the interaction of structures and single functions in an overly detailed way. Instead the dynamics will be related to more comprehensive properties of the combination of functions.

Figure 6 gives an example of combinations of three boolean functions, which, compared with those in Figures 4 and 5, never lead to periods $>59$, and even rarely to periods $>32$. If we take a look at the $P$-parameters and canalyzing indices $C$ (cf. Table 3) there seems to be no salient characteristic discriminating the first from the other combinations.

Actually, it turns out that parameter values comprehensively describing the subsets of three functions, for example, the averages of $P$ or $C$ values, do not sufficiently reflect the different dynamics induced by particular subsets, or, to put it more precisely, by certain members of the set of six functions derived from a subset. It is therefore necessary to reflect on single combinations of functions within the 729 combinations. This is tentatively done with assigning mean values of $P$ and $C$ to each combination of six functions.

The first strikingly clear result is shown by Figure 7: The diagram relates the number of BNs (z-axis, log. scale) within three different ranges of periods ($p > 59$, $32 < p < 60$, and $29 < p < 33$; x-axis) to mean $P$-parameter values of the combinations of functions. The diagram is based on calculations including any of the 1499 structures, any initial states, and any of the combinations of functions derived from the 22 subsets mentioned earlier.
Figure 6. Frequency of periods >20 versus \( v \) in a subset of all combinations of the functions no. 4, 7, and 9.

Figure 7. Number of BNs (z-axis, log.) versus mean \( P \)-parameters of their functions (y-axis, left; 0.5, 0.542, and 0.583) for three ranges of periods (x-axis, \( p > 59 \), \( 32 < p < 60 \), and \( 29 < p < 33 \)).

Complex Systems, 17 (2007) 29–45
The highest periods of $p > 59$ exclusively occur with combinations of functions exhibiting a $P$-parameter of 0.5. A closer inspection, however, reveals constraints going beyond this. Combinations yielding $p > 59$ are never composed of a single function, for example, $(6,6,6,6,6,6) =$ all XOR, and never from the two noncanalyzing functions 6 and 9 (XOR and equivalence). High periods require the presence of at least one other function with $P = 0.5$, that is canalyzing ($C > 0$). These facts can be interpreted in terms of a certain inhomogeneity necessary within the set of functions, which yields high periods.

Figures 8 and 9 offer a more detailed insight into the dependency of dynamics upon $C$. The diagrams relate periods of BNs for particular structures to the mean $C$ values of single combinations of six functions. The diagrams are typical examples within the set of structures and combinations of functions that were examined. They demonstrate that high periods $p > 59$ can appear with combinations possessing mean $C$ values between 1/3 and 1, the range depending on the particular structure.

Not surprising, the range of appropriate $C$ values broadens for lower periods. This is shown for periods above 30 in Figure 8, which are exhibited for $1/3 < C < 5/3$. It may be noted that combinations with $C = 0$ yield only very short periods. A closer inspection of data again lends support to the inhomogeneity assumption.

Comparing the notion of canalyzing functions as originally defined by Kauffman (canalyzing vs. noncanalyzing) with the canalyzing index $C$ proposed here it can be stated that $C$ by virtue of its higher discriminating power is a more appropriate parameter, although the ranges of $C$, within which particular periods may be induced, are rather wide. Furthermore, the results prove that the $P$-parameter and canalyzing index $C$—although derived from the same, simple binary representation of a function—reflect different characteristics of the functions, which differently express themselves in the dynamics.

4. Conclusions

The experimental results presented in section 3 are based on a very small mathematical system. The advantage of this attempt is that it is possible to derive rather accurate relations between the three basic components determining the dynamics of the system, namely the topological structure, the transformation rules or functions, respectively, and the initial

---

7Structures are characterized here by a graph code. The graph code is derived by transforming the rows of the two-dimensional adjacency matrix into binary strings, consecutively read as a set of six decimal numbers (two digits each). Within any equivalence class of isomorphic graphs the maximum (or minimum) of these codes is unique for the whole class. The graphs can be immediately constructed from this code.
Figure 8. Periods (max., mean, most frequent) \textit{versus} canalyzing index $C$ for a single structure and all combinations from functions no. 1, 5, and 6.

Figure 9. Periods (max., mean, most frequent) \textit{versus} canalyzing index $C$ for a single structure and all combinations from functions no. 6, 9, and 10.
states. And it can be no doubt that studies that will even more deeply delve into detailed characterizations of structure, functions, and their interplay will lead to very precise predictions of the dynamics of individual systems. The question, however, remains: How far can results referring to such small systems be applied to larger mathematical systems or even to real; for example, social, systems?

Experiments which we have carried out in various contexts support our assumption that the very parameters determining the dynamics of small systems will also limit the potential dynamics of larger systems. That means that the parameters—at least those referred to in this study—predominantly act as constraints or necessary conditions, respectively, restricting the space of possibilities for the systems.

Obviously, with increasing system size, the predictions derived from certain parameters will be “softer” in the sense, that they provide statistical answers, offering only probabilities and leaving room for exceptions. That, of course, will also depend on the questions and definitions. For example, if you imagine a boolean network (BN) with 1000 elements, the question of which set of parameters will allow complex dynamics—complex defined in the sense of Wolfram’s classification—can only be answered with a certain probability. But we are convinced that, if you would define “complex” restrictively as having periods

\[ p > 2^{1000/2} = 2^{999} \]

the statement will hold that you will not get such complexity unless you assign exclusively functions with \( P \)-parameters of 0.5.

Bearing these precautions in mind we propose some tentative conclusions concerning discrete mathematical systems in general as far as they are comparable with BNs.

1. Among the configuration of parameters or variables, respectively, defining a certain system only a tiny fraction will lead to complex dynamics (defined by periods near the maximal possible period of the system). Typically this fraction lies below 1%. Put in other terms: Most system configurations will exhibit simple dynamics like short periods or point attractors. Researchers experimenting with large systems are well aware of this fact.

2. Any kind of parameter of the type discussed here represents specific constraints in the sense of necessary conditions for complex dynamics. At present we are not aware of any parameter to be interpreted as a sufficient condition. Instead it can be expected that additional structural parameters can be defined which will exclude ever more structures from the set of structures with the potential of yielding complex dynamics.

3. The topology of a system, here represented by graph structures, is the decisive factor determining the potential for complex dynamics or higher periods, respectively. The \( \nu \)-parameter is a rather coarse variable for describing the structure. Nevertheless, there is a clear relation between
\(\nu\)-parameter values and the frequency or probability, albeit very small, of developing complex dynamics. In our example BNs this probability increases from approximately \(\nu = 0.6\) to \(\nu = 0.0\). We assume that this range will also apply to larger systems. The results strongly suggest that the distribution of outdegrees within the graph structure is a crucial factor determining the possible range of dynamics.

4. There exist excluding structural conditions which reduce or even prevent complexity. Among those are the occurrence of sinks and the absence of cyclic interaction patterns.

5. Provided that the structure permits complex dynamics, certain properties of the set of assigned functions will further reduce the probability for complexity. In our experiments only combinations of functions with a \(P\)-parameter of 0.5, that is, an equal number of zeros and ones, yield the highest periods, without exceptions.

6. As to the functions, a further condition may be based on the canalizing index \(C\). There seems to be a range of mean \(C\) values for the combinations of functions which is necessary for complex dynamics. This range is \(1/3 < C < 1\) in our case.

7. The influence of different initial states on the dynamics has not been discussed here because the number of possible initial states is very limited and longer periods comprise a considerable fraction of all states. A period of 63, for example, contains all states but one. This will be different for large systems. In particular the dynamics of cellular automata are very sensitive regarding initial states because of the highly localized interaction, an observation well known to players of the “Game of Life.”

8. The preceding conclusions provide a more or less comprehensive view. A more detailed inspection of those combinations of functions yielding the highest periods reveals a kind of “principle of minor inhomogeneity”: As mentioned earlier neither sets of six equal functions nor sets of functions which are too “similar,” with respect to possessing the same \(C\) parameters, show high periods. It seems to be a necessary condition that at least one of the functions in a set of six must have deviant parameters. But it must be noted that this does not hold for \(P\). Having an equal probability to transform to either possible state is an indispensable condition, at least for small systems. In other words: Complex systems have to be nearly homogeneous, but not totally.

9. The “principle of inhomogeneity” can also be observed concerning the topology. Although structures with an outdegree sequence of \((2,2,2,2,2,2)\), indicating maximal distribution of interactions, exhibit the highest probability to yield high periods, structures with the most homogeneous distribution of arcs (like the example in Figure 1(a)) induce simple dynamics.

To summarize these conclusions and view them with reference to modeling complex systems we can propose the following conjectures.
Systems will exhibit complex dynamics only if the ordering parameters characterizing structure, transformation rules, and possibly also initial states are within specific ranges or, in other words, lie in a certain domain of the parameter space. Obviously a single parameter cannot provide a sufficiently comprehensive characterization. For small BNs we have offered a set of three such parameters. Further studies might show whether those parameters will be applicable for other and larger systems and whether a three-dimensional parameter space is sufficient.

From our study of small systems it turns out that structure or topology, respectively, is the predominant factor referring to dynamics. We may assume that this holds for social systems, too, since social structures are in a sense global within a social system whereas transformation rules are far more open to change and local variations.

Parameter values within the ranges found characteristic for complex dynamics by no means guarantee that a system with these values can actually develop more than simple dynamics. Complex dynamics occur with very low probability compared with simple dynamics. Appropriate parameter values provide therefore necessary, but by far not sufficient conditions for complex dynamics. Thus far sufficient conditions have not been proposed, and maybe do not exist.

The observations above suggest a translation with respect to real systems. The range of parameters with potential for complexity, interpreted in terms of social sciences, describes a rather democratic system with almost—but not totally—equally distributed potential interactions and opportunities to use these. It is not surprising that structures and functions which will be associated with centralized, directive social systems, are characterized by parameters outside the range. It is obvious that systems with high equality, but not total conformity as the “principle of small inhomogeneity” tells us, suffer from the disadvantage of potentially complex dynamics—at the “edge of chaos.” However, this very complexity makes those systems so capable of adapting to changing environments and conditions.

We have shown that a detailed investigation of ordering parameters seems to be a promising way to obtain a more thorough understanding of potential universal mathematical systems and, moreover, that the results may have an impact on the discussion of the dynamics of real systems. As far as ordering parameters represent logical conditions determining or limiting the dynamical properties of mathematical systems they have also to be interpreted as conditions limiting the potential dynamics, or even determining the probability of certain developments or adaptations in real systems. This could give us a Kantian Stance (Kauffman): Measuring mathematical characteristics that all real systems must have—regardless of their empirical contents—as a mathematical a priori.
References


Complex Systems, 17 (2007) 29–45