

Simone Bassis
Anna Esposito
Francesco Carlo Morabito
Editors

SMART INNOVATION,
SYSTEMS AND TECHNOLOGIES ■ 26



Recent Advances of Neural Network Models and Applications

Proceedings of the 23rd Workshop of the
Italian Neural Networks Society (SIREN),
May 23–25, Vietri sul Mare, Salerno, Italy

SIREN

 Springer

Smart Innovation, Systems and Technologies

Volume 26

Series editors

Robert J. Howlett, KES International, Shoreham-by-Sea, UK
e-mail: rjhowlett@kesinternational.org

Lakhmi C. Jain, University of Canberra, Canberra, Australia
e-mail: Lakhmi.jain@unisa.edu.au

For further volumes:
<http://www.springer.com/series/8767>

marco.villani@unimore.it

Identifying Emergent Dynamical Structures in Network Models

Marco Villani^{1,2}, Stefano Benedettini¹, Andrea Roli^{1,3}, David Lane^{1,4},
Irene Poli^{1,5}, and Roberto Serra^{1,2}

¹ European Centre for Living Technology, Ca' Minich, S. Marco 2940, 30124 Venezia, Italy

² Dept. of Physics, Informatics and Mathematics,
University of Modena e Reggio Emilia, v. Campi 213b, 41125 Modena, Italy
{marco.villani, roberto.serra}@unimore.it

³ DISI Alma Mater Studiorum University of Bologna Campus of Cesena,
via Venezia 52, I-47521 Cesena, Italy
andrea.roli@unibo.it

⁴ Dept. of Communication and Economics, University of Modena e Reggio Emilia,
v. Allegri 9, 41121 Reggio emilia, Italy
lane@unimore.it

⁵ Department of Environmental Sciences, Informatics and Statistics,
University Ca' Foscari, Venice, Italy
irene.poli@unive.it

Abstract. The identification of emergent structures in dynamical systems is a major challenge in complex systems science. In particular, the formation of intermediate-level dynamical structures is of particular interest for what concerns biological as well as artificial network models. In this work, we present a new technique aimed at identifying clusters of nodes in a network that behave in a coherent and coordinated way and that loosely interact with the remainder of the system. This method is based on an extension of a measure introduced for detecting clusters in biological neural networks. Even if our results are still preliminary, we have evidence for showing that our approach is able to identify these “emerging things” in some artificial network models and that it is way more powerful than usual measures based on statistical correlation. This method will make it possible to identify mesolevel dynamical structures in network models in general, from biological to social networks.

Keywords: Dynamical systems, emergent dynamical structures, cluster index, boolean networks, emergent properties.

1 Introduction

Emergent phenomena are among the most intriguing ones in natural as well as in artificial systems. Indeed, it can be argued [1] that neural networks represent an attempt at shaping the emergent properties of a set of models in order to perform some required tasks. An intriguing aspect is the "sandwiched" nature of most emergent

phenomena: while past researches were almost exclusively focused on bottom-up emergence in two-level systems (like e.g. Benard-Marangoni convection cells emerging from the interaction of the water molecules [2]) it is becoming increasingly clear that in the most interesting cases the new entities and levels do emerge between pre-existing ones. The paradigmatic example may be that of organs and tissues in multicellular organisms: both the lower (cellular) level and the upper one (organism) predate the appearance of the intermediate structures. Other examples come from the physical world (e.g. mesolevel structures in climate) and social systems (e.g. various factions within political parties). What is more interesting in the present case is that also in artificial systems, like neural networks, one observes the formation of intermediate-level circuits between the single neurons and the global properties. It goes without saying that some neural architectures have been devised precisely to stimulate the formation of these mesolevel structures, but here we are concerned with structures that come into being by spontaneous processes without being explicitly designed from the outside (although a certain type of design may ease or prevent the formation of these spontaneous structures).

A central question is then that of identifying the emerging "things": these may be either static entities or dynamical patterns, or some mixture of the two. In dynamical networks, static emergent structures take the form of topological features, like e.g. motifs in genetic networks or communities in a broader context. There is an extensive literature on community detection, so we will concentrate here on a different type of mesolevel structures, namely those that are created by the dynamical interactions in the network. Nodes may work together although they are not directly linked, since the dynamical laws may give rise to different parts working together. If the topology were regular, these nodes might be identified by visual inspection, but in the case of irregular topologies this approach seems hopeless.

In this paper we present a first step towards the development of formalized methods to identify these mesolevel "things": since they may have a topological as well as a dynamical nature, we refer to them as mesolevel dynamical structures (MDS). The task of identifying MDSs is a formidable one, so we will show here the outline of a promising approach and some preliminary results, while remarking that there are still more open questions than answers. However, the interest of the task motivates in our opinion the opportunity to report our preliminary results.

In order to escape "bird's eye" detection methods, we will consider different subsets of the network, looking for those whose nodes appear to be well coordinated among themselves and have a weaker interaction with the rest of the nodes. For each subset of nodes we will measure its so-called cluster index, a measure based on information theory that had been proposed by Tononi and Edelman [3]. After a suitable normalization procedure (see the following for the details) we rank the various subsets in order to identify those that are good candidates for the role of partially independent "organs" (note that they not necessarily exist in any network).

The cluster index has been defined so far for quasi-static systems, and we will discuss its extension to nonlinear dynamical systems. We will also show the result of the application of this ranking method to some model systems, including some synthetic dynamical networks and some genetic regulatory networks proposed by

examining the biological literature. The method draws our attention on subsets that are functionally correlated and that represent an interesting hypothesis about possible MDSs. In the end we will also comment on the fact that our method, although not yet fully developed, already outperforms usual correlation techniques.

2 Some Useful Definitions

For the sake of definiteness, let us consider a system U , our "universe" that is a network of N nodes that can change in discrete time, taking one of a finite number l of discrete values (in the examples we will choose $l=2$ for simplicity). The value of node i at time $t+1$, $x_i(t+1)$, will depend upon the values of a fixed set of input nodes at time t , possibly including the i -th (self-loops are not prohibited). In several cases, networks start with a random state and change according to the evolution rules so the initial state may bear no relationship to the system itself. Since we are interested in finding out some properties of the networks themselves, we will consider their behaviors after an adequate relaxation time. For the time being we will also ignore external influences on some nodes, although these might be easily included.

The entropy of a single node is estimated from a long time series by taking frequencies f_v of observed values in time as proxies for probabilities, and is defined as

$$H_i = -\sum_{v=1}^m f_v \log f_v \quad (1)$$

where the sum is taken over all the possible values a node can take.

If the system is deterministic and is found in a fixed point attractor, $H_i=0$ for every node, since each node takes its value with frequency one. In order to apply entropy-based methods, Edelman and Tononi considered a system subject to gaussian noise around an equilibrium point. In our case it is however appropriate to deal with a richer time behavior since nonlinear networks can have several different attractors, each attractor contributing to the behavior of the system (though in different times). So our "long data series" will be composed by several repetitions of a single attractor, followed by repetitions of another one, etc. (ignoring the short transients between two attractors). The number of times a single attractor is represented in the data series should be weighted in some way: there are possible several different strategies, depending on the nature of the system we are analyzing. In case of noisy systems a possibility is that of estimating the weights of the attractors by measuring the persistence time of the systems in each of them [4]; deterministic systems might be analyzed by weighting attractors with their basins of attraction. For simplicity in the following we opt for this second choice.

Now let us look for interesting sets of nodes (clusters, from now on). A good cluster should be composed by nodes (i) that possess high integration among themselves and (ii) that are more loosely coupled to other nodes of the system. The measure will define, called the cluster index, is not a Boolean one, but it provides a measure of "clusterness" that can be used to rank various candidate clusters (i.e., emergent intermediate-level sets of coordinated nodes).

3 Measuring the Cluster Index

Following Edelman and Tononi [3], we define the cluster index $C(S)$ of a set of k nodes S , as the ratio of a measure of their integration $I(S)$ to a measure of the mutual information $M(S|U-S)$ of that cluster with the rest of the system.

The integration is defined as follows: let $H(S)$ be the entropy (computed with time averages) of the elements of S . This means that each element is a vector of k nodes, and that the entropies are computed by counting the frequencies of the k -dimensional vectors. Then:

$$I(S) = \sum_{j \in S} H(x_j) - H(S) \quad (2)$$

The first term is the sum of the single-node entropies, the last one is computed using vectors of length k , so I measures the deviation from statistical independence of the k elements in S^l . The mutual information of S to the rest of the world $U-S$ is also defined by generalizing the usual notion of mutual information between nodes to k dimensional vectors

$$M(S; U-S) \equiv H(S) - H(S|U-S) = H(S) + H(U-S) - H(S, U-S) \quad (3)$$

where, as usual, $H(A|B)$ is the conditional entropy and $H(A,B)$ the joint entropy.

Finally, the cluster index $C(S)$ is defined by

$$C(S) = \frac{I(S)}{M(S; U-S)} \quad (4)$$

The cluster index vanishes if $I=0$, $M \neq 0$, and is not defined whenever $M=0$. For this reason, the approach based upon cluster indices does not work properly when the mutual information of S with the rest of the system vanishes; these cases, in which S is statistically independent from the rest of the system – a significant property because they signal particularly strong structures – can be diagnosed in advance.

In this way, for every subsystem S we will get a measure of its quality as a cluster. In order to identify potential MDSs it is necessary to compare the indices of various candidate clusters. It is straightforward to compare clusters of the same size using $C(S)$, but unfortunately C scales with the size of the subsystem, so that a loosely connected subsystem may have a larger index than a more coherent, smaller one. In order to deal with these cases we need to normalize the clusters with respect to their size. The analysis may turn out quite cumbersome, but in most cases we found it sufficient to use a simple prescription, used by Tononi and Edelman in their original paper, which results in the calculation process outlined in the following.

The first step is to define a “null system”, i.e., a non-clustered homogeneous system, from which we sample a series. This system provide us with a null hypothesis

¹ $H(S)$ is estimated from the same time series used to calculate the frequencies f_i of eq. (1). So, to compute $H(S)$ we calculate the frequencies f_v^S of the observed values of S seen as a whole.

and allows us to calculate a set of normalization constants, one for each subsystem size. For each subsystem size, we compute average integration $\langle I_h \rangle$ and mutual information $\langle M_h \rangle$ (subscript h stands for “homogeneous”); we can then normalize the cluster index value of any subsystem S in its universe U using the appropriate normalization constants dependent on the size of S :

$$C'(S) = \frac{I(S)}{\langle I_h \rangle} \bigg/ \frac{M(S; U - S)}{\langle M_h \rangle} \quad (5)$$

We apply this normalization to both the cluster indices in the analyzed system and in the null system.

The definition of “null system” is critical: it could be problem-specific, but we prefer a simple solution which is fairly general: given a series of Boolean vectors, we compute the frequency of ones b and generate a new random Boolean series where each bit has the same probability b of being one. This random null hypothesis is easy to calculate, related to the original data and parameter-free; moreover we believe it satisfies the requirements set by Tononi of homogeneity and cluster-freeness.

The second step involves the computation of a statistical significance index, called T_c , that is used to rank the clusters in the analyzed system. The T_c of a cluster S is:

$$T_c(S) = \frac{C'(S) - \langle C'_h \rangle}{\sigma(C'_h)} \quad (6)$$

where $\langle C'_h \rangle$ and $\sigma(C'_h)$ are respectively the average and the standard deviation of the population of normalized cluster indices with the same size of S from the null system [5].

4 A Controlled Case Study

As a first step, we show the results of the application of our method on simple cases in which the systems analyzed have clusters by construction. These experiments make it possible to assess the effectiveness of the approach on controlled case studies, in which the expected outcome of the analysis is known a priori. Since our method aims at finding clusters of nodes which work together --- independently of their connections --- on the basis of sample trajectories of the system, we directly generated trajectories in which some groups of values behave coherently, i.e., they are the clusters to be detected. The trajectories are sequences of binary vectors of length n , $[x_1(t), x_2(t), \dots, x_n(t)]$. At each time step t , the values of the first c vector positions are generated according to the following procedure: $x_1(t)$, the leader, is a randomly chosen value in $\{0,1\}$; the values from position 2 to c , the followers, are a noisy copy of the leader, i.e., $x_i(t) = x_1(t)$ with probability $1-p$ and $x_i(t) = \sim x_1(t)$ otherwise, being p the noise rate. Values $x_{c+1}(t), \dots, x_n(t)$ are randomly chosen in $\{0,1\}$. This way, the first block of the vector is composed of strongly correlated values and it should be clearly distinguished from the rest of the positions. Besides series with one cluster, with the

same procedure we also generated trajectories with two independent clusters of size c_1 and c_2 , respectively. In this case, the clusters can be found in positions $1, \dots, c_1$ and c_1+1, \dots, c_1+c_2 , where leaders are x_1 and x_{c_1+1} . The trajectories were generated with p in $\{0, 0.01, 0.1\}$.

We applied our method based on the calculation of the normalized cluster index and we sorted the clusters as a function of the significance index T_c . In all the cases, the score based on T_c returns correctly the clusters in the first positions of the ranking. As an example, in Figure 1a we show the results of a representative case with two clusters with $c_1=c_2=4$ in a trajectory with 15 vector positions and $p=0.01$. The figure shows the ten T_c highest values and the corresponding cluster size. The bars are sorted in not increasing order of T_c . The highest peaks correspond indeed to the two clusters created in the trajectories. Each row of the matrix represents a cluster: white cells are the vector positions included in the cluster and they are ranked, from the top, by not increasing values of T_c . We can see that the first two clusters detected are indeed the ones corresponding to the positions 5, ..., 8 and 1, ..., 4 (small differences in T_c between the two clusters are due to noise).

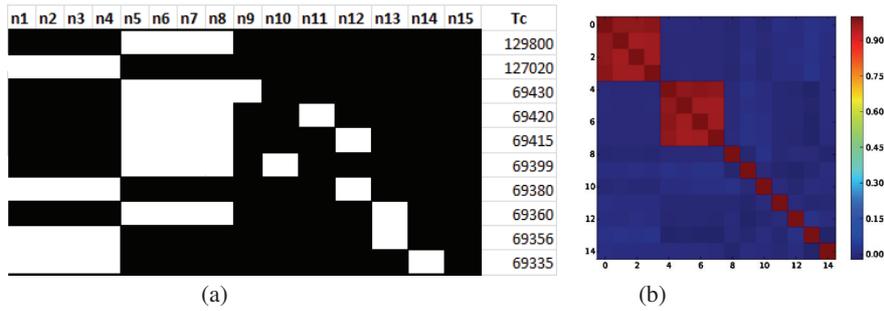


Fig. 1. (a) Matrix illustrating the elements of the clusters and the corresponding T_c values. The first two clusters are the ones introduced in the trajectory. (b) The heatmap shows the correlation values between pairs of vector positions in the trajectory.

The detected clusters are composed of correlated values, therefore we expect to find them also by simply calculating the correlation between every pair of positions. The correlation is computed by taking the value of the Pearson correlation coefficient for each pair of variables; in the case of binary variables, this quantity is usually called the phi coefficient. Since we are interested indifferently in both positive and negative correlations, we take the absolute value of the phi coefficients. Results can be plotted as heatmaps, with correlation values associated to colors from blue (lowest) to red (highest). An example is given in Figure 1b. As we can observe, the blocks composing the clusters are clearly detected and this result holds for all the trajectories we analyzed. This result is not surprising, as the vector positions composing a cluster are indeed strongly correlated (the only variance is introduced by noise). One might then object that the correlation measure is sufficient to detect clusters. In fact, this argument is only valid in some simple cases and does not extend to the general case. The reason is that correlation is a pairwise measure, while the cluster index accounts

for multiple relations. These first tests enable us to state that our method based on the cluster index can be effectively used to capture multiple correlations among variables. In the next section, we will show that this approach can be particularly powerful in detecting clusters of nodes in networks.

5 Cluster Indices Applied to Network Models

The case study we are going to examine consists of three synchronous deterministic Boolean Networks (BNs) – the BN being a very interesting case of complex systems [6] [7], also applied to relevant biological data [8] [9] [10] and processes [11] [12]. The aim of this case study is to check whether CI analysis is capable of recognizing special topological cases, such as causally (in)dependent subnetworks and oscillators, where the causal relationships are more than binary. Note that in all the following cases the phi analysis is ineffective (doesn't relate any variable, having values different from zero only on the diagonal of the matrix).

The first example is a BN made of two independent sub-networks (RBN1 - Figure 2a); in this case we expect the analysis to identify the two subsystems. The second example (RBN2 - Figure 2b) is a BN made of an oscillator (nodes 0 and 1) and one of the subnetworks from the previous example, node 2 has no feedback connections. In the last example we simply merge the networks from the previous examples (RBN3 system). Figures 3 show the top 5 most relevant clusters according to T_c . CI analysis is able to correctly identify the two subnetworks in the first example (first and second rows). The analysis clusters together 5 of 6 nodes of RBN2: those already clustered in RBN1, plus nodes 1 and 2 (which negates each other - figure 2b) and the node that compute the XOR of the signal coming from the two just mentioned groups. Indeed, all these nodes are needed in order to correctly reconstruct the RBN2 series.

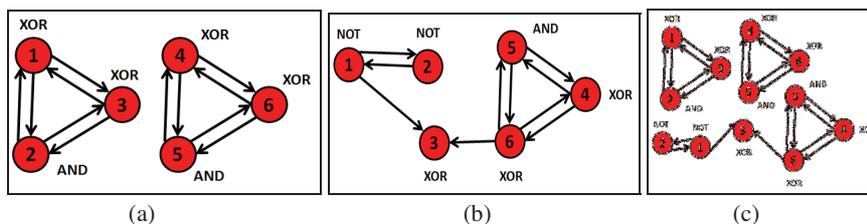


Fig. 2. (a) independent Boolean networks (RBN1); (b) interdependent networks (RBN2); (c) A system composed by both the previous networks (RBN3). Beside each boolean node there is the boolean function the node is realizing.

In the third example the top two clusters correspond respectively to the 5 nodes already recognized in RBN2 and to the whole RBN2 system, while the third and fourth rows correspond to the independent subgraphs of RBN1: all MDSs are therefore correctly identified.

We would like to point out that CI analysis does not require any knowledge about system topology or dynamics. This information is normally unavailable in real cases; on the other hand, our methodology just needs a data series.

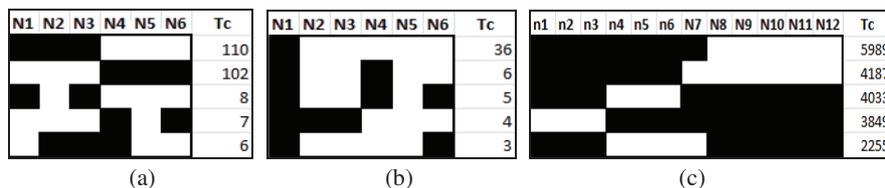


Fig. 3. Matrix illustrating the elements of the clusters and the corresponding T_c values, for (a) RBN1, (b) RBN2 and (c) RBN3 systems

As a final note, it is important to point out that covariance analysis is inadequate in this scenario as it is not able to identify any cluster. We took the same series we applied CI analysis upon and computed the correlation matrix between the node variables; the correlation indices between nodes are uniformly low in magnitude. The inadequacy of this method can be explained by the fact that correlation only takes into account binary linear interactions between variables as opposed to *CI*, which does not necessitate these hypotheses. Experiments performed using asynchronous update yielded essentially the same results with respect to both *CI* and correlation analyses.

6 Evolved Network: Some Examples

We have shown that our method makes it possible to discover clusters of coordinated nodes in a network. We may then raise the question as to what extent this technique can be used to extract meaningful information about clusters in networks subject to evolution. This would provide insights both into biological and artificial evolutionary processes involving networks. The subject of evolution of clusters is strictly linked to the evolution of modularity [13][14] but we conjecture that clusters form only when certain conditions are verified in the evolutionary process: in particular, we expect that clusters are not needed if the environmental conditions do not require organized subsystems (devoted to specific tasks).

To test this conjecture, we studied cases in which BNs with random topology are evolved for maximizing the minimal distance between attractors and for two classification tasks [15]. These tasks are static and not intrinsically modular; therefore, we expect not to find clusters in these evolved networks. The outcome of our analysis is that all these tasks can be optimally solved by BNs possessing only two complementary attractors. It can be easily shown that in homogeneous cases (systems without clusters) the cluster index scales linearly with the number of nodes of the cluster. Take a subsystem S and compute $I(S)$; all $H(X_i)$ are equal to 1 (I observes exactly two equally probable symbols on every node); moreover, $H(S)=H(X)=H(X\setminus S)=1$ because on any subsystem I again observes only two symbols with equal probability. To sum it up:

$$C(S) = \frac{I(S)}{M(S; X - S)} = \frac{N-1}{1} = N-1 \quad [7]$$

where N is the number of nodes in S .

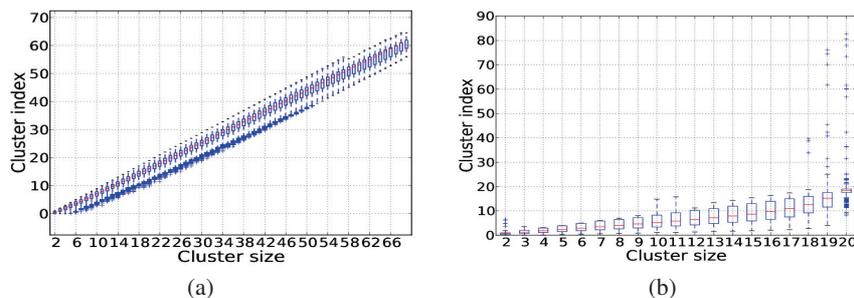


Fig. 4. Distribution of maximum of CI for each cluster dimensions for evolved system, where (a) the task is the maximization of the minimal distance between attractors (systems with 70 nodes) and (b) the task is the Density Classification Problem (DCP), a simple counting problem [16] and a paradigmatic example of a problem hardly solvable for decentralized systems (the results regard networks with 21 nodes). Essentially, it requires that a binary dynamical system recognize whether an initial binary string contains more 0s or more 1s, by reaching a fixed attractor composed respectively by only 0s or 1s.

In figure 4 you can indeed observe this kind of behavior (note that only the averages have some meaning, because of no T_c has significant value – so, the few exceptions to the general behavior on the right side of figure 4b can be discarded. More details are available in [15]).

These are just preliminary experiments and we are currently studying cases in which the formation of clusters is indeed expected. Note however that there are data of evolved systems having well defined clusters: indeed, biological evolution is affecting living systems since 3.8 billion years.

In particular we are analyzing the gene regulatory network shaping the developmental process of *Arabidopsis thaliana*, a system composed by 15 genes and 10 different asymptotical behaviors [17]: our tool was able to group together the three genes core of the system (the first two clusters resulting from T_c ranking): in this case we are identifying clusters having T_c values very significant (see [18] for details).

7 Conclusions

A central question in distributed dynamical system is that of identifying the emerging "things": these may be either static entities or dynamical patterns, or some mixture of the two (neural networks representing an attempt at shaping the emergent properties of a set of models in order to perform some required tasks). In this paper we present a first step towards the development of formalized methods – a research initially started within studies on the brain activities [3] - to identify these mesolevel organizations

(MDSs in the work), which may have a topological as well as a dynamical nature. As examples of application we used time series of simple artificial systems and more complex data coming from Boolean Networks and biological gene regulatory systems (*A.thaliana*). So, the analysis performed by our system is able to identify several interesting mesolevel dynamical structures, and we think it could suggest interesting new ways in dealing with artificial and biological systems.

Acknowledgments. This article has been partially funded by the UE projects “MD – Emergence by Design”, Pr.ref. 284625 and “INSITE - The Innovation Society, Sustainability, and ICT” Pr.ref. 271574, under the 7th FWP - FET programme.

References

1. Serra, R., Zanarini, G.: Complex Systems and Cognitive Processes - A Combinatorial Approach. Springer (1990)
2. Haken, H.: Synergetics. Springer, Heidelberg (2004)
3. Tononi, G., McIntosh, A.R., Russell, D.P., Edelman, G.M.: Functional Clustering: Identifying Strongly Interactive Brain Regions in Neuroimaging Data. *Neuroimage* 7 (1998)
4. Villani, M., Serra, R.: On the dynamical properties of a model of cell differentiation. *EURASIP Journal on Bioinformatics and Systems Biology* 2013, 4 (2013)
5. Benedettini, S.: Identifying mesolevel dynamical structures ECLT (European Center for Living Technologies) technical report, Venice (2013)
6. Kauffman, S.A.: The Origins of Order. Oxford University Press, Oxford (1993)
7. Kauffman, S.A.: At Home in the Universe. Oxford University Press, Oxford (1995)
8. Serra, R., Villani, M., Semeria, A.: Genetic network models and statistical properties of gene expression data in knock-out experiments. *Journal of Theoretical Biology* 227, 149–157 (2004)
9. Shmulevich, I., Kauffman, S.A., Aldana, M.: Eukaryotic cells are dynamically ordered or critical but not chaotic. *Proc. Natl. Acad. Sci.* 102, 13439–13444 (2005)
10. Villani, M., Serra, R., Graudenzi, A., Kauffman, S.A.: Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data. *J. Theor. Biol.* 249, 449–460 (2007)
11. Serra, R., Villani, M., Barbieri, B., Kauffman, S.A., Colacci, A.: On the dynamics of random boolean networks subject to noise: attractors, ergodic sets and cell types. *Journal of Theoretical Biology* 265, 185–193 (2010)
12. Villani, M., Barbieri, A., Serra, R.A.: Dynamical Model of Genetic Networks for Cell Differentiation. *PLoS ONE* 6(3), e17703 (2011), doi:10.1371/journal.pone.0017703
13. Espinosa-Soto, C., Wagner, A.: Specialization Can Drive the Evolution of Modularity. *PLoS Comput. Biol.* 6(3) (2010)
14. Clune, J., Mouret, J.-B., Lipson, H.: The evolutionary origins of modularity. *Proceedings of the Royal Society B* 280, 20122863 (2013)
15. Benedettini, S., Villani, M., Roli, A., Serra, R., Manfroni, M., Gagliardi, A., Pinciroli, C., Birattari, M.: Dynamical regimes and learning properties of evolved Boolean networks. *Neurocomputing* 99, 111–123 (2013)

16. Packard, N.: Adaptation toward the edge of chaos. In: Kelso, J., Mandell, A., Shlesinger, M. (eds.) *Dynamic Patterns in Complex Systems*. World Scientific, Singapore (1988)
17. Chaos, A., Aldana, M., Espinosa-Soto, C., Ponce de Leon, B.G., Garay Arroyo, A., Alvarez-Buylla, E.R.: From Genes to Flower Patterns and Evolution: Dynamic Models of Gene Regulatory Networks. *J. Plant Growth Regul.* 25, 278–289 (2006)
18. Villani, M., Filisetti, A., Benedettini, S., Roli, A., Lane, D., Serra, R.: The detection of intermediate-level emergent structures and patterns. In: *Proceeding of ECAL 2013, the 12th European Conference on Artificial Life*. MIT Press (2013) ISBN: 9780262317092