

## **SUMMARY**

### **INTRODUCTION**

Systemic biology is an area of biology that deals with a systemic view on biological systems and is trying to understand their general characteristics and laws that govern them. Its methodology bases on abstract models of these complex systems. This kind of theoretical approach is a necessary complement to the present reductionistic paradigm and limited experimental methods that do not give satisfactory results. Therefore we have to accept theoretical biology as a concept of thinking, that will help us understand collective dynamics of complex networks and enable us to predict their behavior at different initial states and changes. One of most successful abstract models are Boolean networks that enable us to study in a simple way the changes in network dynamics caused by alterations of different parameters.

### **BOOLEAN NETWORKS**

A Boolean network is composed of a given number of binary elements ( $N$ ) that are connected with a given number of directed connections. Each element has assigned a Boolean function that determines the value of that element, according to the permutation of its input values. The number of inputs per element ( $K$ ) can be equal for all elements or it can vary from element to element. The connections and the Boolean functions remain constant in a certain network, but the initial states from which we compute the network vary. The values of the elements are synchronously updated during computation, according to their Boolean functions and their input values so the network is following a succession of discrete states.

A succession of states is deterministic and can be represented with a curve in a state space. A state space is a mathematical tool for representation of complex system's behavior. Each point in a state space represents a state of the system that is uniformly determined with the values of all its variables. The curves that describe a temporal succession of states are called trajectories. Trajectories run into attractors that are the end-states of network's dynamics. They are either point attractors, which means that end-states are unalterable, or cyclic, which means that the network cycles in a periodic succession of states. All trajectories that run into the same attractor constitute a basin of attraction. A perturbation causes the system to leap from one trajectory to another, which can run into the same attractor as previous one or into some other attractor. The length and the number of attractors determine the orderliness of network's dynamics.

S. A. Kauffman was the pioneer of studying Boolean networks and he discovered what he called "order for free", that at certain parameter values emerges in these networks. His networks had a constant number of inputs per element and at two inputs per element ( $K=2$ ) he noticed in their dynamics a strong phase transition from disorder to order. He estimated the number and the length of attractors in  $K=2$  networks to be  $\sqrt{N}$ , which is an extremely small part of the state space that has  $2^N$  of all possible states (for  $N=100$  networks this means  $2^{100}$  or approx.  $10^{30}$  possible states). The sources of this order had to be the frozen elements, because their values remain constant during the computation of the network.

This orderliness of Boolean networks raised a question whether it is only a consequence of mathematical idealizations of the model, or it also exists in real systems. The researchers tried to create more realistic models and study their dynamics.

Asynchronous Boolean networks abolished the synchronous updating of values and continuous Boolean networks the binary idealization of variables. The models showed that at the same parameters (N and K) they both have a similar degree of orderliness. A direct comparison of dynamics, with attractor counting and their length comparison is not possible though, but the measures that are specific for the type of network give a similar estimation of their dynamic orderliness.

Different number of inputs per element is essential for more realistic modeling, because in natural systems the number of inputs is not equal for all elements. A number of empirical evidences confirm this. Our model is one of the first models with unequal K. A similar model that in some important details differs from ours was made only by one other group of researchers. Our results agree with their findings that heterogeneity in distribution of connections introduces order in a network. Besides determining the orderliness of networks our investigation also includes the study of dynamic diversity of networks, their homeostasis and the alterations in their state space due to changes in network's matrices.

## **GENETIC INTERPRETATION OF THE MODEL**

I interpreted our model as a gene network but with smaller modifications it can also represent a metabolic or an ecological network. In genetic interpretation of the model

- no-input elements represent genes that are under exclusive control of external factors
- connected elements represent genes that are regulated by other genes (and also external factors) and influence the expression of other genes
- no-output elements represent regulated genes that do not influence the expression of other genes
- unconnected elements represent non-regulated genes that do not influence the expression of other genes, but they are important for the organism

## **METHODS OF NETWORK CONSTRUCTION AND DATA ANALYSIS**

A Boolean network is defined with three matrices, namely the matrix of elements, the matrix of connections and the matrix of functions. Our networks had 100 elements and 200 connections, distributed according to two slightly modified binomial distributions. I predetermined the number of no-input elements and then randomly assigned the connections to elements. I redirected the inputs in two different ways (redirection of one input or redirection of a random number of inputs) to change eventual additional no-input or unconnected elements. To analyze the characteristics of a certain network, I computed the network from different random initial states. To analyze the characteristics of a certain type of networks I constructed different random matrices of connections and functions and also computed them from random initial states.

With the random initial states method a number of attractors was found for each network. I perturbed the attractors and counted how often the network returned into the perturbed attractor, and how often it ended in other attractors. A perturbation of attractor means that the values of a given number of randomly chosen elements are switched from 0 to 1 or vice versa. I perturbed different numbers of connected elements and at once all no-input elements. With perturbations of attractors I have also found new attractors, especially in networks with large numbers of found attractors. I perturbed these new attractors and often found some new attractors again. I repeated the process, until I found no new attractors. For comparison of both methods efficiency I repeated the search for attractors with an increased number of random initial states in some networks.

I was also interested what alterations of the state space causes a small change in the matrix of the network, so I added an input to a no-input element in networks with one no-input element.

I estimated the orderliness of networks with the number and the length of attractors, their similarity and the size of attractor basins. We determined the shift of effective input distributions analytically, with computing the probabilities of frozing of elements. For perturbations of attractors I calculated the relative frequency of returning into attractor to determine the homeostasis and by comparing the frequencies in pairs of attractors the perturbational reciprocity between them. I also calculated the correlation between the homeostasis curves of a network at different modes of perturbation and between the homeostasis curve and the size of attractor basins. For additions of an input to a no-input element I calculated the relative frequency of returning into attractor, the summated relative frequency of old attractors, the portion of new attractors according to all attractors and the summated relative frequency of new attractors that were found from that attractor.

## **RESULTS**

The average and the median number of attractors increase approximately exponentially with increasing number of no-input elements, and the similarity of attractors decreases accordingly. Networks without no-input elements are most readily comparable with networks of other authors. Networks with unequal  $K$  are more ordered than networks with constant  $K$ . The distribution of the number of attractors in networks with different numbers of no-input elements remains very similar. There are a lot of networks with a small number of attractors and a few networks with a large number of attractors. The average and the median length of attractors show no correlation to the number of no-input elements. The distribution of lengths of attractors is very similar in all networks. There are a lot of attractors with short cycles and a few attractors with long and very long cycles.

No-input elements reduce the size of frozen and active component, and enlarge the size of mixed component. Approximately the half of the frozen component froze due to canalization of Boolean functions, the other half froze as a consequence of network's dynamics. The comparison of different distributions when considering effective inputs showed that scale-free distribution has the most heterogene distribution so the networks with this distribution are the most ordered.

The homeostasis of networks with perturbations of connected elements is very high. It decreases with increasing number of perturbed elements and at 50 perturbed element approaches the estimated sizes of attractor basins. The pattern of network's homeostasis is very similar for different numbers of perturbed elements, but the level of curves and their amplitudes are getting lower with increasing number of perturbed elements. The homeostasis of network in an attractor with perturbations of connected elements is mostly in correlation to its basin size, but that does not apply for perturbations of no-input elements. The homeostasis with perturbations of no-input elements is very low and the pattern of homeostasis is totally different from the one with perturbation of connected elements. Also it is not in correlation to attractor basin sizes.

The perturbational reciprocity of attractors showed groups of attractors with a very high reciprocity, which are concentrated around the diagonal of perturbational table. This means that when perturbed only attractors with similar basin sizes fall into each other. Groups of attractors with high reciprocity probably also show the structure of the state space, in which these attractors form a local cluster.

With the addition of an input to a no-input element the state space of the network slightly changed. The sizes of attractor basins changed and some new attractors formed. The changes were mostly small, but occasionally some large changes appeared, in which case individual attractors or groups of attractors disappeared.

## **CONCLUSIONS**

In networks with unequal  $K$  there is a large portion of elements with a small number of inputs and this elements have higher probability to become frozen, than elements with a large number of inputs. The dynamic orderliness of these networks therefore depends on the number of inputs and their distribution. The distribution that introduces the largest amount of order is the scale-free distribution. No-input elements increase the dynamic diversity of the networks and have a strong influence on their homeostasis. Perturbations of no-input elements mostly divert the network into other attractor. This is expected since in these networks a large portion of attractors is defined with combinations of no-input elements values, so a changed combination usually defines one other attractor.

Perturbations of attractors are also a very efficient method for exploring the local state space, because this way we find new attractors that with the random initial state method we do not. Besides the number of attractors this method reveals also their relative location and relations among them. High perturbational reciprocity also indicates the structure of the state space, because this probably means that these basins of attractors are closer together or that the edges of their basins are lower, than with other attractors.

The addition of input into a network changed the state space of the network. The changes were mostly small, but occasionally there were some large changes appeared in which case some attractors appeared or disappeared. The incidence of large changes is accidental and we cannot predict them, they probably happen when the added input causes major changes in dynamic structure of the network.

## **SYSTEMIC BIOLOGY**

Attractors in biological networks are for instance different gene expression patterns or associations in succession of ecosystems. In genetic interpretation of a model no-input elements represent genes that are regulated exclusively by external factors. These genes have a strong influence on dynamic diversity of gene network. Their perturbations divert the network into other attractor or gene expression pattern. Perturbing a large number of connected elements, i.e. genes that are regulated by other genes, can also divert the network into other attractor. The process of gene network redirection happens during cell differentiation, during adaptation to external changes, during normal physiological processes that change expression of genes, or during pathological disturbances of normal gene expression pattern.

Mutations of genes cause changes in matrices (connections and/or functions) of the network and consequently the changes of state space. The old attractors change and some new can appear that are new patterns of gene expression and never appear in a normal tissue. Extremely high mutability and consequent fast alterations of state space have malignantly transformed cells, which enables them to be extremely adaptable and strongly competitive.

The strategy of continuous more or less rapid changing of the state space on all levels of organization is one of the laws of biological evolution. The alteration speed is limited with minimal biological information stability. Alterations of the state space are mostly small, similarly like the changes in evolution, but occasionally some large alterations in the state space occur and those can be the source of evolution novelties.

Another law of biosphere evolution determines the direction of development towards increasing complexity and appearance of emergencies. It seems that this direction is determined with physical, as with mathematical laws, because the more complex organization is physically (smaller production of entropy) and mathematically (more complex state space) favorable. Both characteristics are obviously also evolutionally advantageable.

Recent investigations of metabolic networks revealed that the distribution of interactions in these networks is scale-free and that networks with different complexity have a constant diameter. The common core of all metabolic networks indicates their common evolutionary origin. The scale-free distribution of interactions is thought to be characteristic also for other networks in organisms. The gene networks are therefore also thought to have the scale-free distribution and a constant network diameter. This would mean that the more complex the organism is, the more connected is its gene network, because their mutual concordance is needed to maintain the multicellular level of organization.

We can imagine the entire network of all possible molecular interactions of organism as a temporal and spatial superposition of all molecular interaction potentials. The actual molecular networks are temporally and spatially defined manifestations of parts of potential network. The state space of an actual network comprise only a part of potential network's state space, which could be also called the hyperspace of an organism on molecular level. The changes in network's matrices that are consequences of mutations or biochemical reactions with foreign

molecules manifest only if they are a part of actual networks. They continually change depending on the presence of molecules, their concentration and activity.

Influencing the network through no-input elements is an efficient way of directing the networks into different functional patterns. I assume that the introduction of a feedback connection (input) to a perturbed no-input element would increase the homeostasis of the network. This principle could be used for instance for cancer treatment or remediation of ecosystems. Models of complex networks will enable the prediction of effects, caused by their modifications. This would greatly improve the effectiveness of gene therapy, enable the prediction of consequences caused by human degradation of ecosystems and their protection, enhance the efficiency of biotechnological production of metabolites, enable the control of bacterial gene expression, a model of human metabolic network would be incredibly useful for testing the effects of possible drugs and by that achieving their maximal efficiency and minimal side effects.

Modern reductionistic technologies are less efficient than expected. The problem lies in their basic idea that a cell can be redirected in a desired direction simply by inhibition or inactivation of a single gene or protein. It is becoming obvious that an individual element is not important, but the network of interactions between them. The development of efficient molecular network models will be an extraordinary break through that will be the fundament for a new generation of technologies, founded on the idea of directing the network's collective dynamics.

## **FUTURE RESEARCH**

Future studies of networks with unequal  $K$  should reveal the correlation between the changes of parameters ( $N$ ,  $K$  and the distribution of connections) and their orderliness. Studies of network dynamics should be repeated in networks with larger number of elements ( $N=10\ 000$ ) and extrapolate the dynamics to the thermodynamic limit ( $N \rightarrow \infty$ ). The evaluation of continuous models would be necessary for further development of biological modeling. They are very important from the aspect of computation in biological systems. From empirical researches it would be for further development of gene network models very important to analyze the topological differences in gene networks of prokaryotes and eukaryotes and of unicellular and multicellular organisms. It would be also very interesting to model the evolutionary growth of a network. An important study would examine how damages of the network influence the network's homeostasis. It is possible that because of the deletion of elements and connections a sudden decrease in homeostasis appears, so at this point even small perturbations could cause the crash of the system. These kind of critical systems would be degraded ecosystems, injuries on different levels of an organism (metabolic, genetic, physiological), different kinds of stress, ...

Why would be these abstract models important for science? Because they reveal the general properties of complex dynamical networks and mathematical foundations of order in nature. With thorough studying of these aspects of order that are very important also in biological systems is now, as an important field of theoretical biology, beginning to deal the emerging systemic biology.