# Simulating genes operation and interaction

Rekubratskiy V.A.
*Centre "Bioengineering" RAS*
*vrecobra@gmail.com*

Korotkova M.A.
*Moscow Physical Engineering Institute*

## Abstract

*Gene operation and interaction occur in every organism. Gene network is commonly used to describe gene interaction issues due to its complicity. Simulation of gene networks operation is an essential problem of bioinformatics. System providing tools for creating, setting up and simulating gene networks is described. Broad simulation abilities along with high extensibility are system's advantages to be noted.*

## 1. Introduction

Gene operating and interaction determines development of every organism on the Earth [1]. Though many genomes are well examined nowadays [2], gene interactions issue still remains a subject of intensive investigation. The main point is that interactions can often be revealed only through indirect indications, such as gene activity changes [3], [4]. Activity of some genes can have positive or negative influence on other genes' activity [3]. This influence can be detected during experiments. Many articles with assumption about certain gene relationships are published every year [5]-[9]. Gene network is represented by directed graph with vertices corresponding to genes and arcs corresponding to positive and negative interactions. It is commonly used to describe gene interactions of an organism [10]. Gene network is usually worked out by a biologist studying interactions issue. Unfortunately, a network based on some experimental data may contradict the data received later or by another researcher (i.e. works [4] and [7] to be compared). Several distinct gene networks are usually possible to be drawn up basing on incomprehensive data set also [11]. To find the most accurate gene network is required a network verification using some other experimental data.

Computer system providing simulation of gene interaction and operation is needed for comparing and testing gene networks. Such a system may give the possibility for automatic verification of gene network conformity with experimental data given.

## 2. Simulation system

### 2.1. Gene model

In real organisms, a gene can be active in some cells and inactive in others. The degree of activity may also vary. Each gene is responsible for producing its special protein [1]. A gene is considered active in a cell when some amount of corresponding protein is found in that cell. Amount of gene's protein can be measured in real cells and thus can be used for verifying gene network conformity. Thereby protein amount is an essential characteristic for gene model. Protein production rate depends on different factors and may vary from cell to cell and from gene to gene. The amount of other genes' proteins is one of the factors influencing the protein producing rate and this gene parameter used for gene interaction modelling. Gene networks are often simulated using Petri networks [12], [13]. Only one gene characteristic – amount of corresponding protein – is taken into account within this approach. However, a model taking some other gene parameters into consideration could express more compound dependencies of gene interactions.

The simulation system proposed takes both gene interactions and states into account. Each gene in the model is described by three parameters: product, activity and block. The first one – product – describes the amount of gene's corresponding protein. It also represents gene ability operate on other genes and thus can be defined as gene outer state. The second parameter – activity – represents gene ability to generate its product and thus can be defined as gene inner state. The speed of product generation may vary. Activity function is introduced as another gene characteristic to reflect generation speed variability. Both product and activity are real positive numbers, in contrast to the third Boolean parameter. This parameter

was introduced in the model to express the fact that some genes actually may start or stop operating at any stage of organism development process. When gene is blocked (i.e. its block parameter is true) it cannot nor generate any product.

In nature, product of every gene undergoes the degradation process, i.e. its value gradually decreases if gene activity falls to zero. This fact is also taken into account in the simulation system proposed. The speed of degradation expressed by degradation function is another gene characteristic.

## 2.2. Gene interactions model

Each interaction in the simulation system is described by a set of rules. These rules identify the conditions of the interaction arising and how gene parameters are modified through the act of interaction. Rules defining necessary conditions of interaction are expressed by inequalities and the ones defining parameters change expressed by the equations. Algebraic expressions including any combinations of real numbers and gene parameters can be used in both groups of rules. Commonly, gene product parameters alone are used in conditions of interaction reflecting the fact that product is a characteristic of gene interaction ability.

## 2.3. Simulation process

Gene network operation is simulated as step-by-step process. Check of active interactions is performed on each step. Changes to gene parameters are not applied until an interaction was executed. Instead, all changes are accumulated to be applied at the end of the step. Changes of gene product values depend on gene product generation ability (i.e. activity), so those changes are made at the end of each step also but before applying parameter changes.

## 3. Implementation methods

Simulation system was implemented using Microsoft Visual C++ as a Windows application. The application enables a user to create a gene network by inserting genes into system, drawing interactions scheme and setting up some numeric parameters. After network was created and its initial state was set up, a user can monitor overall simulation process (on network scheme and parameters table) and particular gene activity changes (on graph of function).

## 3.1. Gene network implementation

Gene network is implemented as a two-level system. The first (outer) level is used for intercommunication with a user. It enables him to set up network topology and gene parameters and then monitor gene states during simulation process. The second (inner) level is used to simulate network operation. First, all the information set on the outer level is transferred to the inner one. Simulation step is executed via the second level and then results are translated back to the first level, and so on.

## 3.2. Network editor

The system provides a graphic user interface for setting up gene network. Network topology drawing has a lot in common with working in many scalar or vector graphics editors. Arcs corresponding to gene interactions are created similar to lines drawing, genes are repositioned by mouse pointer, etc.

Extensive class system was used to implement network edition routines. All network objects are derived from common base class containing information about object's inputs and outputs, its position, id, etc. It defines an interface for operations upon network objects as well. These operations include saving, loading, setting up, visualizing and some others. Many actions and checks could be executed uniformly for vertices (genes) and arcs (interactions) due to this design. It gives rise to two related class hierarchies: one for network objects (genes, interactions) and one for operations upon objects. In terms of design patterns, a combination of "Composite" and "Visitor" patterns is used providing very flexible class architecture.

## 3.3. Extensibility

System simulating single gene network operation was initially developed. After that is was extended to perform simulation of multicellular organism development. The extension was made easily due to the system design providing for it. New network object types can be easily introduced on the outer network level, e.g. objects corresponding to some cell activity processes or to environment influence. One new object type was introduced during system expanding to provide creation of processes for cell parameters modifying. The inner network level provides uniform parameter reference system. It enabled simple introducing of new parameters, such as cell size and position, equally with existing gene parameters. The

further system extension towards more comprehensive model is still possible.

## 4. Novelty and results

Quite a few software systems for gene network simulation have been developed so far [14], [15]. Unfortunately, these systems simulate operation of a standalone network without its connection to cells life activity. In fact, there may by observed genes interactions of genes of the remote cells with different gene states in each one. These situations could hardly be described by a single gene network models.

There are only few software systems simulating multicellular organism operation that can take gene interactions into account [16]. The using of this model requires some special knowledge in programming.

The system proposed is designed to be a standalone application that does not require any other mathematical or simulation software running. It grants user a simple to use tool for drawing his own gene networks. It also provides a set of relatively simple algebraic rules for describing gene interaction details. This small set of rules, however, is shown to enable simulation of rather wide range of gene and cell interaction processes. Some processes of genetic control for Arabidopsis thaliana development, as well as for HCV infecting and developing, were successfully simulated using the designed system.

## 5. Conclusions

The simulation system and its software implementation were designed. The simulation system provides a comprehensive gene operation model and simulation of intercellular gene interactions. The software implementation enables a user to easily draw a gene network and describe intercellular gene interaction details by a set of relatively simple algebraic rules. The whole system is shown to be an extensible one and enable simulation of rather wide range of gene and cell interaction processes.

The further development is intended to make system interface simpler for a biologist providing an easier description of gene interaction processes. A more comprehensive model of organism operation is to be designed further taking advantage of system extensibility.

## 6. Support

## 7. References

[1] A.S. Konichev, G.A. Sevastyanova. "Molecular biology", "Academia" publishing center, Moscow, 2003

[2] National Center for Biotechnology Information: http://www.ncbi.nlm.nih.gov/

[3] N.A. Kolchanov, E.A. Ananko, F.A. Kolpakov et al. "Gene networks", *Mol. Biol.,* Moscow, Russia, 2000, July-August, 34(4), pp. 449-460

[4] J.L. Bowman, G.N. Drews, E.M. Meyerowitz. "Expression of the Arabidopsis floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development", *Plant Cell*, American Society of Plant Physiologists, United States, 1991, August (8), pp. 749-758

[5] T. Gedeon, E.D. Sontaq. "Oscillations in multi-stable monotone systems with slowly varying feedback", *J Differ Equ*, Elsevier, Netherlands, 2007, August 15, 239(2), pp. 273-295

[6] G. Pogorelko, O. Fursova, E. Klimov. "IDentification and Analysis of the Arabidopsis Thaliana Atfas4 Gene Whose Overexpression Results in the Development of A Fasciated Stem", *J Proteomics Bioinform*, OMICS Publishing Group, United States, 2008, October, 1(7), pp. 329-335

[7] S.V. Shestakov, A.A. Penin, M.D. Logacheva, T.A. Ezhova. "New modified scheme of genetic control of flower development", *Alive systems technology*, Moscow, Russia, 2005, V. 2, №1, pp. 37-46.

[8] C.Capelli, F. Brisighelli, F. Scarnicci, A. Blanco-Verea, M. Brion, V.L. Pascali. "Phylogenetic evidence for multiple independent duplication events at the DYS19 locus", *Forensic Sci Int Genet*, Elsevier, Netherlands, 2007, December, 1(3-4), pp. 287-290

[9] B. Zheng, X. Lu. "Using protein-semantic network metrics to evaluate functional coherence of protein groups", *AMIA Annu Symp Proc*, American Medical Informatics Association, United States, 2007, October, 11, pp. 1174

[10] N.Geard, J. Wiles. "A gene network model for developing cell lineages", *Artif. Life*, MIT Press, United States, 2005, Summer, 11(3), pp. 249-267

[11] C. Gustafson-Brown, B. Savidge, MF Yanofsky. "Regulation of the arabidopsis floral homeotic gene APETALA1", *Cell*, Cell Press, United States, 1994, January 14, 76 (1), pp. 131-143

[12] S. Grunwald, A. Speer, J. Ackermann, I. Koch. "Petri net modelling of gene regulation of the Duchenne muscular dystrophy", *Biosystems*, Elsevier Science Ireland, Ireland, 2008, May, 92(2), pp. 189-205

[13] L.J. Steggles, R. Banks, O. Shaw, A. Wipat. "Qualitatively modelling and analysing genetic regulatory networks: a Petri net approach", *Bioinformatics*, Oxford University Press, England, 2007, February, 1, 23(3), pp. 336-343

[14] "Ingeneue: gene network simulation software"
http://rusty.fhl.washington.edu/ingeneue/index.html
[15] "GeNESiS: gene network evolution simulation software"
http://genomics.iab.keio.ac.jp/genesis.html

[16] "CellModeller Project"
http://www.archiroot.org.uk/doku.php/navigation/cellmodeller