Qualitative Simulation of Large and Complex Genetic Regulatory Systems

Hidde de Jong [1] and **Michel Page** [1,2]

Abstract. Modeling and simulation techniques developed within qualitative reasoning might be profitably used for the analysis of genetic regulatory systems. A major problem with current qualitative simulation techniques is their lack of upscalability. We describe a method that is able to deal with large and complex systems, and discuss its performance in simulation experiments with random regulatory networks.

1 Introduction

In the last few years, biologists have completed the sequencing of the entire genome of model organisms like *S. cerevisiae* and *E. coli*, and the human genome is expected to follow without much delay. The analysis of these huge amounts of data involves such tasks as the prediction of folding structures of proteins and the identification of genes and regulatory signals. It is clear, however, that the structural analysis of sequence data needs to be complemented with a functional analysis to elucidate the role of genes in controlling fundamental biological processes.

One of the central problems to be addressed is the analysis of *genetic regulatory systems* controlling the spatiotemporal expression of genes in the organism. The structure of these regulatory systems can be represented as a network of interactions between genes, proteins, metabolites, and other small molecules. The study of genetic regulatory networks will contribute to our understanding of complex processes like the development of a multicellular organism.

Computer tools are indispensable for the analysis of genetic regulatory systems, as these usually involve many genes connected through regulatory cascades and feedback loops. Currently, only a few regulatory networks are well-understood on the molecular level, and quantitative knowledge about the interactions is seldom available. This has stimulated an interest in modeling and simulation techniques developed within qualitative reasoning (QR), most notably QSIM [9] and QPT [2]. QR methods have been applied to the regulation of tryptophan synthesis [7] and λ phage growth [6] in *E. coli*, and to the regulation of the transcription factor families AP-1 and NF-κB in different classes of animals [19].

A major problem is the lack of upscalability of these approaches. As a consequence of the weak nature of qualitative constraints, and the difficulty to identify implicit constraints, behavior trees and envisionments quickly grow out of bounds. This causes the range of application of the methods to be limited to regulatory systems of modest size and complexity. Systems of even a few genes related by positive and negative feedback loops cannot be handled, unless these systems have been so well-studied already that behavior predictions can be tightly constrained.

In this paper we will show that it is possible to qualitatively analyse genetic regulatory networks of larger size and complexity. In order to achieve this, we describe the systems by a class of piece-wise linear differential equations (PLDEs) putting strong constraints on possible trajectories in the phase space. Simulation is carried out by an algorithm tailored to this class of models. The method has been implemented in Java and used for the simulation of regulatory networks of currently up to 18 genes involved in complex feedback loops.

In the next two sections, we will introduce the class of PLDEs by which genetic regulatory systems can be described and review its mathematical properties. The subsequent sections introduce the qualitative simulation algorithm and present the results of simulation studies.

2 Modeling genetic regulatory systems

In Fig. 1(a) a simple example of a regulatory network is shown, involving three genes and their mutual interactions. A *regulatory interaction* is here defined as a relation between a regulated gene and one or more regulating genes. The regulating genes code for proteins that control the expression of the regulated gene, by functioning as a transcription factor or otherwise. An interaction has more than one regulating gene, if the corresponding proteins fulfill their regulatory function cooperatively, as when two proteins form a heterodimer. In the figure, the expression of gene 2 is controlled through one interaction involving two regulating genes, and gene 3 through two interactions involving one regulating gene. A gene *positively (negatively)* regulates another gene, if the protein coded for by the former tends to activate (inhibit) the expression of the latter.

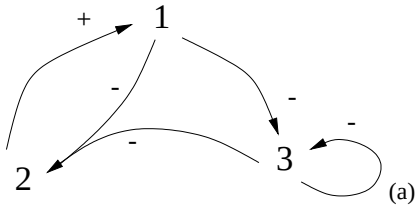
Gene regulation is often modeled by differential equations of the form

$$\dot{x}_i = f_i(\boldsymbol{x}) - \gamma_i x_i,\ x_i \geq 0,\ 1 \leq i \leq n, \tag{1}$$

where $\boldsymbol{x}$ is a vector of cellular concentrations of gene products (mRNAs or proteins), γ_i the decay rate of x_i, and f_i a usually highly nonlinear function [3, 11, 18]. The rate of expression of gene i is dependent upon the concentrations $\boldsymbol{x}$, possibly including the concentration of the product of gene i. The term $-\gamma_i x_i$ states that x_i degrades at a rate proportional to the concentration itself. Eqs. (1) are called the *state equations* of the regulatory system.

The functions f_i in (1) can be further specified as a sum of interaction terms corresponding to the interactions in the regulatory

[1] Institut National de Recherche en Informatique et en Automatique (INRIA), Unité de Recherche Rhône-Alpes, 655 avenue de l'Europe, 38330 Montbonnot Saint-Martin, France, email: Hidde.de-Jong@inrialpes.fr
[2] Université Pierre Mendès France, Grenoble, France, email: Michel.Page@inrialpes.fr



(a)

$$\dot{x}_1 = \kappa_{12} h^+(x_2, \theta_{12}, m) - \gamma_1 x_1$$
$$\dot{x}_2 = \kappa_{213} h^-(x_1, \theta_{21}, m) h^-(x_3, \theta_{23}, m) - \gamma_2 x_2$$
$$\dot{x}_3 = \kappa_{31} h^-(x_1, \theta_{31}, m) + \kappa_{33} h^-(x_3, \theta_{33}, m) - \gamma_3 x_3 \quad \text{(b)}$$

$$0 < \theta_{21} < \theta_{31} < max_1$$
$$0 < \theta_{12} < max_2$$
$$0 < \theta_{23} < \theta_{33} < max_3 \quad \text{(c)}$$

$$0 < \kappa_{12}/\gamma_1 < \theta_{21}$$
$$\theta_{12} < \kappa_{213}/\gamma_2 < max_2$$
$$0 < \kappa_{31}/\gamma_3 < \theta_{23}$$

$$0 < \kappa_{33}/\gamma_3 < \theta_{23}$$
$$\theta_{33} < (\kappa_{31} + \kappa_{33})/\gamma_3 < max_3 \quad \text{(d)}$$

Figure 1. (a) Example regulatory network modeled by the state equations in (b) and the threshold and nullcline inequalities in (c)-(d). A regulatory network is a directed graph of genes (vertices) and interactions (directed edges), where interactions involving several regulating genes are represented by directed hyperedges. The symbols '$+$' and '$-$' denote activating and inhibitory relations, respectively.

network. More precisely, for each interaction involving a regulated gene i and a set of regulating genes J, the sum contains a term $\kappa_{iJ} \prod_{j \in J} r(x_j)$, where $r(x_j)$ is a *regulation function* and κ_{iJ} a *rate constant* determining the maximum expression level of i under the influence of J.

A regulation function $r(x_j)$ accounts for the variation in expression level of gene i with the concentration x_j of the product of gene j. A regulation function often found in the literature is the Hill curve:

$$h^+(x_j, \theta_{ij}, m) = x_j^m / (x_j^m + \theta_{ij}^m), \tag{2}$$

where θ_{ij} denotes the *threshold* for the influence of j on i, and $m > 1$ a parameter determining the steepness of the function around θ_{ij}. The function ranges from 0 to 1, and increases as $x_j \to \infty$, so that j positively regulates i. In order to express that j negatively regulates i, the regulation function $h^+(x_j, \theta_{ij}, m)$ must be replaced by $h^-(x_j, \theta_{ij}, m) = 1 - h^+(x_j, \theta_{ij}, m)$. In Fig. 1(b) the state equations for the example network are shown.

Due to the nonlinear character of the functions f_i, analytical solution of the state equations (1) is not possible. The nonlinear terms can be eliminated by replacing the continuous Hill function by the discontinuous step function:

$$s^+(x_j, \theta_{ij}) = \begin{cases} 1, & x_j > \theta_{ij}, \\ 0, & x_j < \theta_{ij}. \end{cases} \tag{3}$$

The resulting equations are piecewise-linear differential equations (PLDEs) of the form

$$\dot{x}_i = b_i(\boldsymbol{x}) - \gamma_i x_i, \; x_i \geq 0, \; 1 \leq i \leq n, \tag{4}$$

where b_i is a piecewise-constant function. In particular, b_i is a sum of products of step functions weighted by a rate constant. The approximation of a continuous sigmoid by a discontinuous step function has been justified on the ground of the switch-like character displayed by genes whose expression is regulated by steep sigmoid curves [3, 4, 18]. In what follows, we will assume that genetic regulatory systems are modeled by PLDEs of the form (4).

3 Mathematical analysis

Eqs. (4) have been well-studied in mathematical biology [3, 4, 5, 10, 11, 12, 13, 14, 16, 17, 18]. Consider an n-dimensional (hyper)box of the phase space defined as follows:

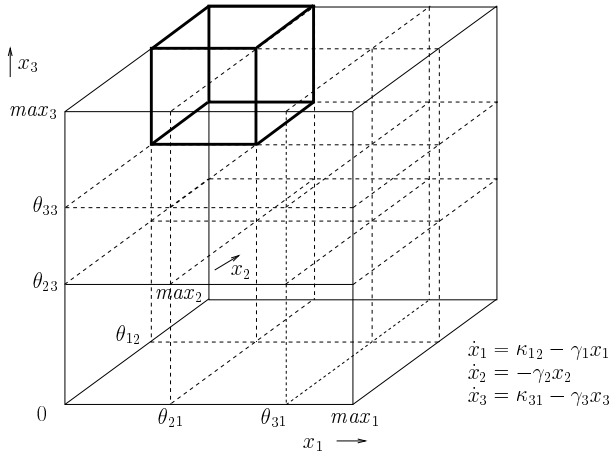


Figure 2. The phase space box of the model in Fig. 1, divided into 18 volumes by the threshold planes. The state equations for the volume in bold defined by $0 \leq x_1 < \theta_{21}$, $\theta_{12} < x_2 \leq max_2$, and $\theta_{33} < x_3 \leq max_3$ are shown in the lower right corner.

$$0 \leq x_i \leq max_i = \max_{\boldsymbol{x} \geq \boldsymbol{0}} b_i(\boldsymbol{x})/\gamma_i, \; 1 \leq i \leq n. \tag{5}$$

It can be shown that all trajectories starting inside the n-box will remain in it, while trajectories starting outside will enter the box at some time or approach it asymptotically as $t \to \infty$. We assume that $\theta_{ji} < max_i$ for all genes j regulated by gene i. The $n-1$-dimensional threshold (hyper)planes $x_i = \theta_{ji}$ divide the n-box into *volumes*. The volumes of the n-box are determined by the *threshold inequalities*

$$0 < \sigma_i^{(1)} < \ldots < \sigma_i^{(p_i)} < max_i, \tag{6}$$

obtained by ordering and renaming the p_i thresholds θ_{ji} of gene i. Since the step function is not defined at its threshold, Eqs. (4) are not defined in the threshold planes separating the volumes. Fig. 2 displays the phase space box corresponding to the example network.

In each volume of the n-box, Eqs. (4) reduce to *volume state equations* with a constant production term μ_i composed of rate parameters

in b_i:

$$\dot{x}_i = \mu_i - \gamma_i x_i,\ x_i \geq 0,\ 1 \leq i \leq n. \tag{7}$$

Notice that Eqs. (7) are linear and orthogonal. Fig. 2 gives an example of the state equations corresponding to the volume $0 \leq x_1 < \theta_{21}$, $\theta_{12} < x_2 \leq max_2$, and $\theta_{33} < x_3 \leq max_3$. It can be easily shown that within a volume all trajectories evolve towards a single, stable *focal state* $\boldsymbol{\mu/\gamma}$, which lies at the intersection of the *nullcline* (hyper)planes $x_i = \mu_i/\gamma_i$ defined by $\dot{x}_i = 0$. As the nullclines are assumed not to coincide with the threshold planes, the focal state will be located at some distance from the threshold planes.

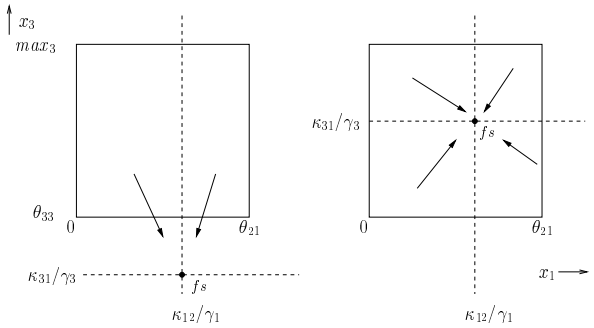


Figure 3. The focal state of the volume in Fig. 2 projected on the x_1-x_3 plane. Depending on whether $\kappa_{31}/\gamma_3 > \theta_{33}$ or $\kappa_{31}/\gamma_3 < \theta_{33}$, the focal state lies inside or outside the volume.

The focal state of a volume may lie inside or outside that volume. Whether the focal state lies inside or outside the volume is determined by the *nullcline inequalities*, which locate the nullclines $x_i = \mu_i/\gamma_i$ between two subsequent thresholds of x_i:

$$\sigma_i^{(l_i)} < \mu_i/\gamma_i < \sigma_i^{(l_i+1)},\ 1 \leq l_i < p_i, \tag{8}$$

with the special cases $0 < \mu_i/\gamma_i < \sigma_i^{(1)}$ and $\sigma_i^{(p_i)} < \mu_i/\gamma_i < max_i$. If for every i, μ_i/γ_i lies between the threshold boundaries of the volume, then the focal state lies inside the volume. If not, it lies outside the volume (Fig. 3). The nullcline inequalities for the example regulatory system are shown in Fig. 1(d). Notice that several nullcline inequalities have been specified for x_3, as a consequence of the fact that μ_3 changes between different volumes. More generally, the set of possible nullclines in the ith dimension is given by $\{b_i(\boldsymbol{x})/\gamma_i \mid \boldsymbol{0} \leq \boldsymbol{x} \leq \boldsymbol{max}\}$.

If the focal state lies outside the volume, the trajectories will tend towards one or several of the threshold planes bounding the volume. Since (4) is not defined at the thresholds, special attention should be given to the behavior of the system as it approaches the threshold planes. Following [14], the behavior of the piecewise-linear equations (4) at the threshold planes is defined as the behavior of the original nonlinear equations (1) in the limit $m \to \infty$ (see also [13]). This is motivated by the observation that, as m goes to ∞, the sigmoid function (2) approaches the step function (3).

Given this definition, two different things can happen when a trajectory approaches a threshold plane $x_i = \theta_{ji}$. First, the trajectory may be continued by a trajectory in the neighbouring volume moving towards a different focal state determined by the volume state equations of the new volume. In this case a *transition* from the volume to its neighbouring volume takes place and the threshold plane is *transparent*. Second, if the focal state of the neighbouring volume is such that trajectories in that volume also approach the threshold plane $x_i = \theta_{ji}$, no transition between the volumes is possible and the threshold plane is *non-transparent*.

The global behavior of the PLDEs may be quite complex and is not well understood. Continuations of trajectories in several volumes may give rise to (oscillations towards) additional steady states located at the intersection of threshold planes, cycles, limit cycles, or even chaotic oscillations (for $n \geq 4$) [5, 10, 12, 13, 14, 17]. Numerical simulation studies have shown that, in many cases, the global behavior of the piecewise-linear systems (4) and nonlinear systems (1) with steep sigmoids exhibit the same qualitative properties [4, 13, 18].

4 Qualitative simulation method

Our method performs a qualitative simulation of regulatory systems described by PLDEs (4). The basic idea underlying the method is to determine, in an iterative way, all volumes that are reachable from an initial volume through successive volume transitions. For each volume that has been found reachable, the position of the focal state, and hence the possible transitions to new volumes, are calculated.

Consider a volume defined in the ith dimension by two consecutive thresholds $\sigma_i^{(l_i)}$ and $\sigma_i^{(l_i+1)}$, $1 \leq l_i < p_i$.[3] The inequalities

$$\sigma_i^{(l_i)} < x_i < \sigma_i^{(l_i+1)} \tag{9}$$

form the *qualitative value* of x_i, denoted by qv_i. In addition to the qualitative value for x_i, we have a qualitative value $\dot{qv}_i$ for $\dot{x}_i$, being one of the following three inequalities

$$\dot{x}_i > 0,\ \dot{x}_i < 0,\ \text{or } \dot{x}_i \lesseqgtr 0. \tag{10}$$

If the nullcline plane for x_i lies outside the volume, i.e., $\mu_i/\gamma_i < \sigma_i^{(l_i)}$ or $\mu_i/\gamma_i > \sigma_i^{(l_i+1)}$, the qualitative value will be $\dot{x}_i < 0$ or $\dot{x}_i > 0$, respectively, everywhere in the volume. If the nullcline runs through the volume, i.e., $\sigma_i^{(l_i)} < \mu_i/\gamma_i < \sigma_i^{(l_i+1)}$, it holds that $\dot{x}_i < 0$ on one side of the nullcline plane, $\dot{x}_i > 0$ on the other side, and $\dot{x}_i = 0$ in the nullcline plane. The qualitative value of $\dot{x}_i$ in the volume is then written as $\dot{x}_i \lesseqgtr 0$.

Given a volume v with a vector $\boldsymbol{qv}$ of qualitative values for $\boldsymbol{x}$, $\dot{\boldsymbol{qv}}$ can be easily inferred from the equations and inequalities (5)-(8) by means of basic algebraic rules. As a consequence of the orthogonality of the volume state equations, this can be done separately in each dimension, thus requiring only $\mathcal{O}(n)$ inferences. For the volume emphasized in Fig. 2, we find the vector $[\dot{x}_1 \lesseqgtr 0, \dot{x}_2 < 0, \dot{x}_3 < 0]$.

The vector $\dot{\boldsymbol{qv}}$ of qualitative values for $\dot{\boldsymbol{x}}$ expresses the position of the focal state with respect to v, and hence allows one to determine the possible transitions from v to neighbouring volumes. A volume v' defined by $\boldsymbol{qv}$' is a *candidate successor volume* of v, if there is exactly one i such that $qv_i' = succ(qv_i)$, where $succ$ is defined in Table 1, and $qv_j' = qv_j$ for all $j \neq i$, $1 \leq j \leq n$. That is, only one variable has changed its qualitative value in the candidate successor volume, implying that concentrations of gene products are assumed to never reach their thresholds simultaneously.

For a candidate successor v' to be an actual successor of v, the threshold plane separating v and v' must be transparent. This implies

[3] The procedure can be easily generalized to the cases $0 < \mu_i/\gamma_i < \sigma_i^{(1)}$ and $\sigma_i^{(p_i)} < \mu_i/\gamma_i < max_i$.

qv_i and $\dot{qv}_i$	$succ(qv_i, \dot{qv}_i)$
$\sigma_i^{(l_i)} < x_i < \sigma_i^{(l_i+1)},\ \dot{x}_i > 0$	$\sigma_i^{(l_i+1)} < x_i < \sigma_i^{(l_i+2)}$
$\sigma_i^{(l_i)} < x_i < \sigma_i^{(l_i+1)},\ \dot{x}_i < 0$	$\sigma_i^{(l_i-1)} < x_i < \sigma_i^{(l_i)}$
$\sigma_i^{(l_i)} < x_i < \sigma_i^{(l_i+1)},\ \dot{x}_i \lesseqgtr 0$	$-$

Table 1. The function $succ$ mapping qualitative values for x_i and $\dot{x}_i$ to a successor qualitative value for x_i, $1 < l_i < p_i - 1$. The successor relations are motivated by basic continuity restrictions, as in [9].

that, for the x_i changing its qualitative value in the transition, the qualitative values of its derivative should not be opposite. That is, if $qv_i' \neq qv_i$, then not $\dot{qv}_i' > 0\ (< 0)$ and $\dot{qv}_i < 0\ (> 0)$, $1 \leq i \leq n$.

The simulation algorithm iteratively generates, in a depth-first manner, all volumes that are reachable from an initial volume v_{init} defined by qualitative values $\boldsymbol{qv}_{init}$.

push($stack, v_{init}$)
determine $\boldsymbol{\dot{qv}}_{init}$
while not $stack$ is empty **do**
    current volume $v \leftarrow$ pop($stack$)
    declare v to be reachable
    generate candidate successor volumes of v
    for all candidate successors v' **do**
        determine $\boldsymbol{\dot{qv}}'$
        if v' is actual successor
        then if not v' is reachable **and not** v' on stack
            then push($stack, v'$)

The volumes and their reachable successors form a directed *transition graph*. The graph may contain volumes without successors and volume cycles, which will be together referred to as *attractors*. If a volume has no successors, it either contains a steady state or all outgoing trajectories approach non-transparent threshold plane(s). In the worst case, the algorithm will generate $\mathcal{O}((p+1)^n)$ reachable volumes, where p is the maximum number of genes influenced by a single gene.

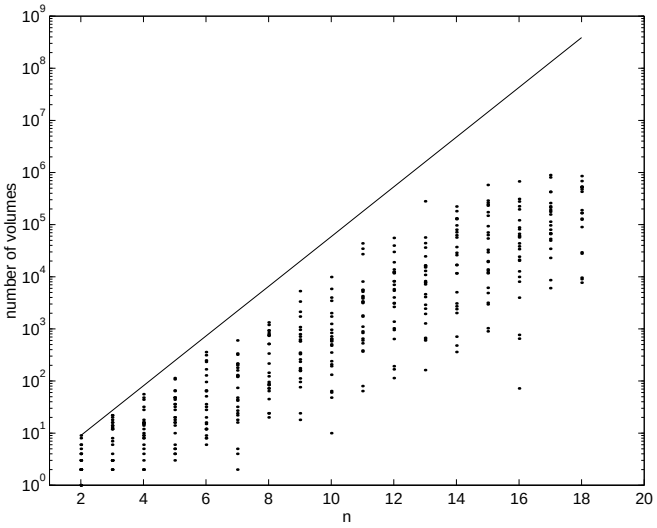


Figure 4. The number of reachable volumes from an initial volume for models with $k = 2$ and $n = 2, \ldots, 18$. Each dot in the plot represents a simulation.

5 Experimental results

The simulation algorithm predicts the attractors that may be reached from an initial volume following a sequence of volume transitions. From a biological point of view, this means that possible functional states of the regulatory system are identified, given certain initial gene expression levels [8, 18]. As shown in the previous section, the number of reachable volumes theoretically grows in an exponential fashion. This compromises the objective to deal with larger-scale regulatory systems. In order to test whether the average-case behavior is more favorable, we have performed a series of computer experiments.

The experiments have been carried out with an implementation of the simulation method in Java 1.2. The program reads and parses input files with the equations and inequalities specifying the model of the system (state equations, threshold and nullcline inequalities) as well as the initial volume. The core of the program consists of an inequality reasoner for the determination of $\boldsymbol{\dot{qv}}'$ in the main loop of the algorithm. We have developed a version of Simmons' [15] Quantity Lattice, adapted to the particularities of the class of PLDEs we are dealing with. The output produced by the program consists of a tabular representation of the volume transition graph, a list of attractors, and run-time statistics. The simulations reported below were run on a SUN Ultra 10 workstation with 128 Mb of RAM.

In order to study the upscaling properties of the methods in a systematic way, we have carried out experiments with random regulatory networks. For each of the n genes in a network, k inputs were randomly chosen among the other genes. Next, the functions f_i were randomly selected from the set of all possible functions with k inputs. Further, a random order between the thresholds of the regulation functions was generated, as well as a lower and upper threshold bound for the nullcline terms μ_i/γ_i. Each of the models thus obtained was simulated from a randomly-selected initial volume in the phase space.

The results of experiments with $2 \leq n \leq 18$ and $k = 2$ are shown in Fig. 4. For each n, 25 simulations were carried out, each with a different model and initial volume. The number of volumes reachable from the initial volume is displayed as a function of n. The most important observation to be made is that the average-case behavior is much more favorable than the worst-case behavior, shown as the drawn line in the figure (notice the logarithmic scale of the y-axis). For $n = 12$, the median number of volumes reachable is 5418, about 1% of the total number of volumes in the phase space.

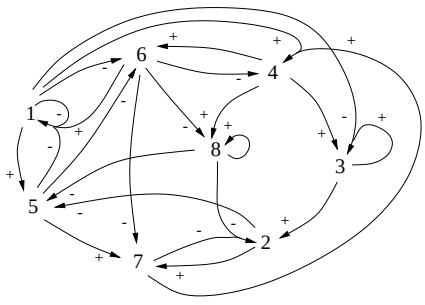


Figure 5. Example regulatory network for $n = 8$ and $k = 3$.

Fig. 4 shows a large spread in the simulation results. Simulations for $n = 16$ give results varying from 72 to 675,216 reachable volumes, although most of the time around 10^5 volumes are generated. A number of factors contribute to these differences, in particular the distance of the initial volume to the reachable attractors and the num-

ber and the size of the attractors. The number of reachable attractors also strongly varies between simulations. For $n = 14$, the median number of reachable attractors is 2, with 7 simulations yielding more than 100 attractors, mostly cycles.

Although no simulation took more than half an hour to complete, for $n > 18$ over one third of the simulations ended with a memory overflow due to the large number of volumes generated. This seldom if ever happened for lower n. We have also carried out simulations for $k = 3$ and $k = 4$, that is, for more densely connected networks. In these cases, the number of reachable volumes theoretically grows as $\mathcal{O}(4^n)$ and $\mathcal{O}(5^n)$. As for $k = 2$, the average-case behavior tends to be more favorable. However, for $n > 13$ simulations started to become intractable with the current implementation.

Fig. 5 shows an example of a network with $n = 8$ and $k = 3$ and a large number of positive and negative feedback loops. The model is defined by a total of 120 equations and inequalities. Simulated from a random initial volume, 3892 volumes turn out to be reachable. The trajectories either end in the single volume with a steady state or in one of the 14 cycles.

6 Discussion

The method for qualitative simulation of genetic regulatory systems presented in this paper has been shown capable of dealing with networks of larger size and complexity than possible with existing QR methods. We have modeled regulatory systems by a class of differential equations putting strong constraints on the local behavior in the phase space, in combination with a simulation algorithm adapted to these equations. Currently we are able to deal with networks of up to 18 genes with 2 to 4 regulators per gene and complicated feedback structures. The simulation studies described here present one of the first attempts to systematically investigate upscaling of QR methods in the context of a realistic application. The simulation method has been tailored to one class of models, but the principles underlying our approach might be applicable to other problems as well.

Adaptation to a specific class of models is the principal respect in which the approach presented in this paper differs from well-known QR methods like QPT and QSIM [2, 9]. The expressivity and generality of the formalism have been traded for the capability to deal with larger and more complex systems. For instance, the description of the state of a regulatory system is achieved on a higher level of abstraction. The basic element in our formalism is a volume, defined by a vector of qualitative values $\sigma_i^{(l_i)} < x_i < \sigma_i^{(l_i+1)}$. In QSIM one would have to distinguish individual states inside a volume as well, such as boundary states defined by $x_i = \sigma_i^{(l_i)}$ or $x_i = \sigma_i^{(l_i+1)}$, and nullcline states defined by $\dot{x}_i = 0$. The method presented here thus abstracts from trajectories inside a volume, which among other things allows a more compact representation of the behavior of the system.

Qualitative methods for the analysis of genetic regulatory systems have been developed in mathematical biology as well, the best-known example being Boolean networks [8]. Simulation of Boolean networks rests on the assumption that a gene is either active or inactive, and that genes change their activation state synchronously. Translated to the formalism of this paper, this means that there is only one threshold per gene and that thresholds are reached simultaneously. For many purposes, these assumptions are too strong. The use of random networks to study the upscaling properties of the method has been stimulated by Kauffman's [8] simulation studies with Boolean networks. The observation that trajectories remain localized in a small part of the phase space agrees with the results obtained for Boolean networks.

Thomas and colleagues [18] have proposed a generalized logical method that permits multivalued activation states and asynchronic transitions. In fact, Snoussi [16] has demonstrated that their formalism can be seen as an abstraction of a special case of (4). Although simulation is possible in the generalized logical method, the emphasis is on the identification of steady states, including steady states located on the threshold planes [13, 14, 17]. The use of logical equations abstracting from differential equations makes it difficult to integrate (semi-)quantitative information [1]. With the advent of cDNA microarrays and other new measurement technologies, (semi-)quantitative gene expression data is becoming available in large amounts.

The simulation method presented in this paper forms the core of a system currently under development called the *Genetic Network Analyzer (GNA)*. The system will be used to address a problem of high biological relevance, namely the validation of hypothesized regulatory networks by means of expression data.

ACKNOWLEDGEMENTS

The authors would like to thank O. Bernard, J. Cohen, H. Geiselmann, S. Maza, F. Rechenmann, D. Thieffry, and I. Vatcheva.

REFERENCES

[1] D. Berleant and B. Kuipers, 'Qualitative and quantitative simulation: Bridging the gap', *Artif. Intell.*, **95**, 215–256, (1997).
[2] K.D. Forbus, 'Qualitative process theory', *Artif. Intell.*, **24**, 85–168, (1984).
[3] L. Glass, 'Classification of biological networks by their qualitative dynamics', *J. Theor. Biol.*, **54**, 85–107, (1975).
[4] L. Glass and S.A. Kauffman, 'The logical analysis of continuous nonlinear biochemical control networks', *J. Theor. Biol.*, **39**, 103–129, (1973).
[5] L. Glass and J.S. Pasternack, 'Stable oscillations in mathematical models of biological control systems', *J. Math. Biol.*, **6**, 207–223, (1978).
[6] K.R. Heidtke and S. Schulze-Kremer, 'Design and implementation of a qualitative simulation model of λ phage infection', *Bioinformatics*, **14**(1), 81–91, (1998).
[7] P.D. Karp, 'Design methods for scientific hypothesis formation and their application to molecular biology', *Machine Learning*, **12**, 89–116, (1993).
[8] S.A. Kauffman, *The Origins of Order*, Oxford University Press, 1993.
[9] B. Kuipers, *Qualitative Reasoning*, MIT Press, 1994.
[10] J.E. Lewis and L. Glass, 'Steady states, limit cycles, and chaos in models of complex biological networks', *Int. J. Bifurcation Chaos*, **1**(2), 477–483, (1991).
[11] T. Mestl, E. Plahte, and S.W. Omholt, 'A mathematical framework for describing and analysing gene regulatory networks', *J. Theor. Biol.*, **176**, 291–300, (1995).
[12] T. Mestl, E. Plahte, and S.W. Omholt, 'Periodic solutions in systems of piecewise-linear differential equations', *Dyn. Stab. Syst.*, **10**(2), 179–193, (1995).
[13] E. Plahte, T. Mestl, and S.W. Omholt, 'Global analysis of steady points for systems of differential equations with sigmoid interactions', *Dyn. Stab. Syst.*, **9**(4), 275–291, (1994).
[14] E. Plahte, T. Mestl, and S.W. Omholt, 'A methodological basis for description and analysis of systems with complex switch-like interactions', *J. Math. Biol.*, **36**, 321–348, (1998).
[15] R. Simmons, ''Common sense' arithmetic reasoning', in *Proc. AAAI-86*, pp. 528–532. Morgan Kaufmann, (1986).
[16] E.H. Snoussi, 'Qualitative dynamics of piecewise-linear differential equations', *Dyn. Stab. Syst.*, **4**(3-4), 189–207, (1989).
[17] E.H. Snoussi and R. Thomas, 'Logical identification of all steady states', *Bull. Math. Biol.*, **55**(5), 973–991, (1993).
[18] R. Thomas and R. d'Ari, *Biological Feedback*, CRC Press, 1990.
[19] R.B. Trelease, R.A. Henderson, and J.B. Park, 'A qualitative process system for modeling NF-κB and AP-1 gene regulation in immune cell biology research', *Artif. Intell. Med.*, **17**, 303–321, (1999).