

Formal Modeling of Biological Systems

(Invited Paper)

Qinsi Wang

Computer Science Department
Carnegie Mellon University, USA
qinsiw@cs.cmu.edu

Edmund M. Clarke

Computer Science Department
Carnegie Mellon University, USA
emc@cs.cmu.edu

Abstract—As biomedical research advances into more complicated systems, there is an increasing need to model and analyze these systems to better understand them. For decades, biologists have been using diagrammatic models to describe and understand the mechanisms and dynamics behind their experimental observations. Although these models are simple to be built and understood, they can only offer a rather static picture of the corresponding biological systems, and scalability is limited. Thus, there is an increasing need to develop formalism into more dynamic forms that can capture time-dependent processes, together with increases in the models scale and complexity. In this invited review paper, we argue that the formal modeling formalisms can be applied fruitfully to biological systems, and can be complementary to the traditional mathematical descriptive modeling approaches used in systems biology. We also discuss one example: a stochastic hybrid model of the effect of estrogen at different levels in species' population in a freshwater ecosystem.

I. INTRODUCTION

Systems biology studies biological components including molecules, cells, organisms, species, and ecosystems. It aims to better understand the properties of individual parts within complex living systems as well as the dynamics of entire systems. To achieve this, given quantitative measurements of the behavior of groups of interacting elements, mathematical and computational models are constructed to reproduce and predict dynamical behaviors. For decades, graphic models have been used to describe and understand the mechanisms behind the experimental observations of biological systems. Although these models are simple to be built and understood, they can only offer a rather static picture of the corresponding biological systems, and scalability is limited. Then, models expressed mathematically (e.g. using differential equations) have occupied the leading position, where they are simply translated into computer programs simulating those models. Along with the collaboration between biologists and computer scientists becomes tighter, researchers have realized that biological systems and (distributed) computer systems share a lot of features. That is, similar to (distributed) computer systems, biological systems consist of various components, whose behaviors are partially decided by the communication with each other. This led to an increasing interest for system biologists and computer scientists in borrowing existing formal specification and analysis techniques that were designed

for computer systems and in developing biological domain-specific methods, and thus to the success of application of these techniques to biomedical systems.

As shown in Figure 1, with formal and executable models and well-founded analysis methods for them, it offers an excellent means to present knowledge about biological systems, and to reason about these systems rigorously. Moreover, traditional in-vivo and in-vitro experiments are usually expensive, and need to take an awfully long time. While, the execution of formal models, offering in silico evaluation of hypotheses, only takes comparatively little time and effort. Especially, when considering different experimental configurations, multiple wet-lab experiments need to be carried out repeatedly. Whereas, for formal models, only trivial modifications on the initial assignment of system variables and parameters are required.

In the following, a non-exhaustive literature review with regards to formal modeling formalisms that have been successfully applied to biological and biomedical systems is given. We discuss the main ideas and important features of five specification languages that have fallen on fertile ground in systems biology. (See [8], [25], [45] for reviews of more classes of formal models that have been used in systems biology.) Furthermore, this review will take the further step of demonstrating how to model the effect of estrogen at different levels in species' population in a freshwater ecosystem as a stochastic hybrid system. We hope that this short review can be an aid for both systems biologists and computer scientists to gain a better understanding of related concepts and theory that have been focused on currently in this area, so as to further promote the collaboration.

II. FORMAL MODELING OF BIOLOGICAL SYSTEMS

A. Boolean Networks

Boolean networks (BNs), as one of the most widely used formal models, were firstly introduced by Kauffman [42] in 1969, where BNs were used to model gene regulatory networks. A BN is a directed graph containing a set of nodes. Nodes are defined as Boolean variables, whose values represent the dynamic activity and behavior of the involved elements (e.g., genes or proteins). At each time step, the next value of a variable is determined by a Boolean function of its regulators. The values of all variables form a global state to be updated synchronously. In this way, the execution of a BN

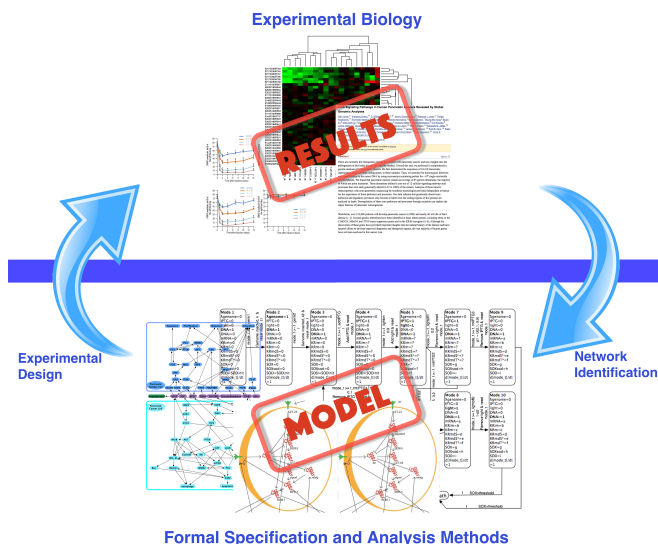


Fig. 1: Experimental biology itself is an iterative process of hypothesis-driven experimentation of a specific biological system. We can boost this process by using formal executable models, and formal analysis methods, such as model checking, to provide interesting new hypotheses for biological experimental design.

illuminates the causal and temporal relationships between the involved elements.

The main advantage of this modeling language is that, even with a strongly simplified view of biological networks, it can still capture the network structure and dynamics, and offer biologically meaningful predictions and insights. Also, using such a high abstraction, it is possible to model interactions among large numbers of elements and perform model validation and model-based prediction. Besides being applied to analyze the robustness and stability of genetic regulatory networks [4], [22], [44], BNs have also been used to study cell signaling networks and understand their impacts on distinct cell states [30], [31], [34], [39]. Moreover, BNs can be inferred directly from experimental time-series data [3], [23], [26], [52].

BNs use a coarse approximation where the status of each modeled element as either active (on) or inactive (off) by neglecting intermediate states. In real-world biological systems, some elements may have multiple states. The difference between this binary assumption and biological reality led researchers to suggest extensions of BNs, such as Qualitative Networks (QNs) [53], Gene Regulatory Networks [48], and the logical model considering the time delay mechanism THiMED [47]. In QNs, each variable can have one of a small discrete number of values. Dependencies between variables become algebraic functions instead of Boolean functions. Dynamically, a state of the model corresponds to a valuation of variables and changes in values of variables occur gradually based on these algebraic functions. QNs have been shown to be a

suitable formalism to model some biological systems [11], [19], [53]. For the logical model using THiMED, system elements are modeled by multi-valued variables. The timing details that capture relative delays between events are allowed, and implemented by truth tables. Another type of extensions of BNs copes with the inherent noise and the uncertainty in biological processes, such as Boolean networks with noise [2] and Probabilistic Boolean networks [54]. These modeling formalisms allow one to consider the uncertainty in the knowledge of signaling networks as well as stochasticity in biological systems.

B. Petri Nets

Petri nets [49] were created by Carl Adam Petri in 1962 to describe chemical processes, and then were also intensively employed in computer science to model and analyze concurrent and distributed systems. A Petri net is a graph with two types of nodes - places and transitions, which are connected by directed arcs. Places represent the resources of the system; transitions indicates the events that can change the state of the resources; and directed arcs, connecting places to transitions and transitions to places, describe which places are pre and/or post-conditions for which transitions. The data, in Petri nets, are represented as so-called tokens. The state of the system is represented by places holding tokens. Note that, one place may hold multiple tokens. Given a start configuration of a Petri net, which assigns tokens to each place, transitions change the state of the system by moving tokens along edges. For each transition, tokens are consumed from the input place through the transition and then created in the output place(s). A transition fires whenever it is enabled by the presence of some tokens in one of the places directly connected to it. In a given state of the system, there may be more than one transition that can move a token, so that the execution of a Petri net is non-deterministic.

Petri nets allow for concurrency and nondeterminism, and provide a framework that supports both qualitative analysis through the static structural topology of Petri nets, and quantitative analysis via the time evolution of the token distribution. Thus, this modeling language is more general than BNs, and holds a good balance between modeling power and analyzability. Petri nets are well-suited for modeling the concurrent behavior of biochemical networks such as genetic regulatory pathways. In detail, the places in a Petri net can represent genes, protein species and complexes; transitions represent reactions or transfer of a signal; directed arcs represent reaction substrates and products; and a transition firing is execution of a reaction where substrates are consumed and products are created. They have been used to describe the concurrent behavior of biochemical networks, including metabolic pathways and protein synthesis [16], [55].

Another good thing about Petri nets is that, there are several successful extensions forming a very versatile framework providing additional possibilities in modeling and analysis. For instance, in timed Petri nets, transitions can be timed, which allow for modeling the timing of the system as well. They

have been used to model and analyze signal transductions in an apoptosis pathway [18]. In colored Petri nets, tokens with different colors denote multiple possible values for each place, and thus allow for distinct activation levels to be assigned to resources. They have been used to analyze metabolic pathways [28]. In stochastic Petri nets, probabilities have been added to the different choices of the transitions to consider the uncertainty of biomedical systems. They have been used to analyze signaling pathways, where the number of molecules of a given type is represented by the color of a place and probabilities represent reaction rates [32], [36].

C. Rule-based Modeling

The combinatorial explosion, which emerges from the complexity of multi-protein assemblies, poses a major barrier to the development of detailed, mechanistic models of biological systems. Modeling approaches, such as differential equations, that need manually enumerating all potential species and reactions in a network are impractical. To alleviate the problem, rule-based modeling languages, such as the BioNetGen language (BNGL) [24] and Kappa [21] have been developed. To address the combinatorial complexity in biochemical systems, the key idea of the rule-based languages is to represent interacting molecules as structured objects and to use pattern-based rules to encode their interactions. So that, a rich yet concise description of signaling proteins and their interactions can be provided. In other words, rule-based modeling specifies only those components of a biological macromolecule that are directly involved in a biochemical transformation. Also, the reaction rules are defined as transformations of classes of species, avoiding the need for specifying one reaction per each possible state of a species.

Due to the similarity to the chemical reaction representation widely used in systems biology, rule-based languages have harvested a lot of attention among biologists. It has been applied in the modeling of different cell signaling pathways and networks [9], [10], [13].

Considering that these rule-based languages were designed for describing molecular level dynamics, one growing need is to extend them to span multiple biological levels of organization. ML-Rules [46] is a multi-level rule-based language, which can consider multiple biological levels by allowing objects to be able to contain collections of other objects. This embedding relationship can affect the behavior of both container and contents, and allows users to describe both inter- and intra-cellular processes. Another extension of the BNGL to enable the formal specification of not only the signaling network within a single cell, but also interactions among multiple cells is proposed in [58]. Unlike ML-Rules using continuous rate equations to capture the dynamics of intracellular reactions, this multiscale language models intracellular dynamics using BNs, which reduces the difficulty of estimating the values of hundreds of unknown parameters often involved in large models. This has been used to capture the intra- and inter- cellular dynamics involved in the pancreatic cancer microenvironment [58].

The other increasing need is to take the spatial information into consideration when carrying out the cell biological modeling. SRSim [33], as one spatial extension of the BNGL, integrates the BNGL with a three-dimensional coarse-grained simulation building upon the LAMMPS molecular dynamics simulator [51]. It allows for the formulation of reaction networks, and offers the dynamic simulation as well. SRSim has been used to model and analyze the human mitotic kinetochore [40]. Another spatial extension of the BNGL is cBNGL [35], in which structures and rules are associated with the concept of compartments and membranes. That is, cBNGL distinguishes between three-dimensional (compartment volume) and two-dimensional (surface) compartments. ML-Space [12], as a spatial variant of ML-Rules, considers compartmental dynamics, mesh-based approaches, and individuals moving in the continuous space. In ML-Space, species can be defined as individual particles that react due to collisions, or as a population of species residing in a small area [41]. It has been used to study the dynamics of lipid rafts and their role in receptor co-localization.

D. Hybrid Systems

Hybrid systems [5] are formal models that combine continuous and discrete dynamics in a piecewise manner. In detail, the state space of a hybrid system is defined by a finite set of discrete modes. In each mode, the system evolves continuously obeying processes, generally ordinary differential equations (ODEs) [20]. Transition conditions control the switch from one mode to another, which can be followed by a ‘reset’ of the involved continuous variables. In general, the temporal dynamics of a hybrid system is piecewise continuous.

By using ODEs, one of the most powerful techniques in modeling system dynamics, hybrid systems aim to bridge the gap between mathematical models and computational models by combining the two. The continuous part of hybrid systems, which are captured by differential equations, bears the closest relationship to the underlying biochemical rate laws, thus can accurately model complex biological systems. While, the discrete part of such models is the executable control mechanism that drives a hybrid system.

Hybrid systems are particularly suitable to model biological systems that exhibit clear switching characteristics over time (that is, the same system variables need to be regulated by different processes in distinct discrete states), such as the cell cycle. They have been successfully used to describe biological systems at distinct levels, including genetic regulatory networks [7], cell signaling pathways [29], the cell cycle control [17], the cardiac cell [61], bacteria-killing procedures [59], and human ventricular action potentials in tissue [14].

E. Stochastic Hybrid Models

Stochastic hybrid systems (SHSs) are a class of dynamical systems that involve the interaction of discrete, continuous, and stochastic dynamics. Due to the generality, SHSs have been widely used in systems biology, such as modeling subtilin production in *Bacillus subtilis* [38], and personalized prostate

cancer treatment [60]. To describe stochastic dynamics, uncertainties have been added to hybrid systems in various ways. A wealth of models has been promoted over the last decade.

One way expresses random initial values and stochastic dynamical coefficients using random variables, resulting in hybrid automata (HAs) with parametric uncertainty [60]. When modeling real-world biological systems using hybrid models, parametric uncertainty arises naturally. Although its cause is multifaceted, two factors are critical. First, probabilistic parameters are needed when the physics controlling the system is known, but some parameters are either not known precisely, are expected to vary because of individual differences, or may change by the end of the system's operational lifetime. Second, system uncertainty may occur when the model is constructed directly from experimental data. Due to imprecise experimental measurements, the values of system parameters may have ranges of variation with some associated likelihood of occurrence.

Another class of models integrates deterministic flows with probabilistic jumps. When state changes forced by continuous dynamics involve discrete random events, we refer to such systems as probabilistic hybrid automata (PHAs) [56]. PHAs extend HAs with discrete probability distributions. More precisely, for discrete transitions in a model, instead of making a purely (non)deterministic choice over the set of currently enabled jumps, a PHA (non)deterministically chooses among the set of recently enabled discrete probability distributions, each of which is defined over a set of transitions. Although randomness only influences the discrete dynamics of the model, PHAs are still very useful and have interesting practical applications [57]. One interesting variation of PHAs [60] allows additional randomness for both transition probabilities and resets of system variables. In other words, in terms of the additional randomness for jump probabilities, for the probabilities attached to probabilistic jumps from one mode, instead of having a discrete distribution with predefined constant probabilities, they can be expressed by equations involving random variables whose distributions can be either discrete or continuous. This extension is motivated by the fact that some transition probabilities can vary due to factors such as individual and environmental differences in real-world systems. When it comes to the randomness of variable resets, a system variable can be reset to a value obtained according to a known discrete or continuous distribution, instead of being assigned a fixed value. When continuous probabilistic events are also involved, we call them stochastic hybrid automata (SHAs) [27].

Other models replace deterministic flows with stochastic ones, such as stochastic differential equations (SDEs) [6] and stochastic hybrid programs (SHPs) [50], where the random perturbation affects the dynamics continuously. When all such ingredients have been covered, there are models such as the general stochastic hybrid systems (GSHSs) [15], [37]. In the next section, we will show how to construct a stochastic hybrid model of the effect of estrogen at different levels in species' population change in a fresh water ecosystem.

III. A STOCHASTIC HYBRID MODEL OF EFFECT OF ESTROGEN IN SPECIES' POPULATION IN ECOSYSTEM

Hormones can be potential sources of the environmental pollution. For example, estrogen could cause feminization of fish, which will lead to the population change of species in the fresh water ecosystem. To understand how different estrogen levels affect the ecosystem, based on the 2014 CMU iGem team's experiments [1], we propose a stochastic hybrid model to depict the population change of algae, fish, and bird. Distinct modes are used to consider different levels of estrogen, and within each mode, dynamics are captured by ordinary differential equations (ODEs), partial differential equations (PDEs), and stochastic differential equations (SDEs).

Our water ecosystem model follows a simple tropic pyramid structure. The algae are the food source of the fish, which in turn is the food source for the bird. The dynamics for the population of algae $X(t)$ is captured by an ODE as follows.

$$\begin{cases} \frac{dX(t)}{dt} = X(t)(p_1 - e_1Y(t) - d_1) \\ X(0) = x_0 \end{cases} \quad (1)$$

where,

- p_1 , as a constant, is the reproduction rate for algae;
- e_1 , as a constant, is the rate consumed by fish;
- d_1 , as a constant, is the natural death rate for algae; and
- x_0 is the initial population of algae.

For the fish population, we consider three different types respectively: female fish $Y_f(t)$, male fish $Y_m(t)$, and feminized male fish $Y_{m2f}(t)$. Instead of using ODEs, where all individual fish are treated as identical, PDEs are used to consider differences between individual organisms, such as age-specific fertility and death rates. The following PDEs, where the age structure of fish has been taken into consideration, describe population change of fish of different types. The dynamics for the population of female fish $Y_f(t)$ is defined as follows.

$$\begin{cases} \frac{\partial Y_f(a,t)}{\partial a} + \frac{\partial Y_f(a,t)}{\partial t} = -Y_f(a,t)(d_2(a) + e_2Z(t) + oY(t) - s_1X(t)) \\ Y_f(0,t) = \frac{1}{2}p_2 \int_{a_{fmat}}^{a_{fmax}} Y_m(a_1,t) da_1 \int_{a_{fmat}}^{a_{fmax}} Y_f(a_2,t) da_2 \\ Y_f(a,0) = y_{f0}(a) \\ Y_f(t) = \int_0^{a_{fmax}} Y_f(a,t) da \end{cases} \quad (2)$$

where,

- a_{fmat} , as a constant, is the maturity age of fish;
- a_{fmax} , as a constant, is the maximum age of fish;
- $d_2(a)$ is the natural death rate for fish. This function is defined as

$$d_2(a) = \begin{cases} 0, & a = 0 \\ d_2, & a \in (0, a_{fmax}) \\ 1, & a = a_{fmax} \end{cases}$$

- e_2 , as a constant, is the rate consumed by bird;
- o , as a constant, is the death rate caused by the over-crowding;
- s_1 , as a constant, is the surviving rate due to food consuming;

- p_2 , as a constant, is contact rate between mature male and female fish for the reproduction; and
- $y_{f0}(a)$ is the initial population and age structure of female fish.

The dynamics for the population of male fish $Y_m(t)$ is defined as follows.

$$\begin{cases} \frac{\partial Y_m(a,t)}{\partial a} + \frac{\partial Y_m(a,t)}{\partial t} = -Y_m(a,t)(d_2(a) + e_2Z(t) + oY(t) - s_1X(t)) - \int_0^{a_{fmax}} f(a)Y_m(a,t)da \\ Y_m(0,t) = \frac{1}{2}p_2 \int_{a_{fmat}}^{a_{fmax}} Y_m(a_1,t)da_1 \int_{a_{fmat}}^{a_{fmax}} Y_f(a_2,t)da_2 \\ Y_m(a,0) = y_{m0}(a) \\ Y_m(t) = \int_0^{a_{fmax}} Y_m(a,t)da \end{cases} \quad (3)$$

where, $f(a)$ is the feminized rate for male fish. As the older a fish is, the more the accumulated estrogen in its body is. As the feminized rate is positively linear to the accumulated estrogen amount in the body, we define $f(a) = \frac{a}{a_{fmax}} \cdot y_{m0}(a)$ is the initial population and age structure of male fish.

The dynamics for the population of feminized male fish $Y_{m2f}(t)$ is defined as follows.

$$\begin{cases} \frac{dY_{m2f}(t)}{dt} = Y_{m2f}(t)(s_1X(t) - d_2(a) - e_2Z(t) - oY(t)) + \int_0^{a_{fmax}} f(a)Y_m(a,t)da \\ Y_{m2f}(0) = 0 \end{cases} \quad (4)$$

Then, the total number of fish $Y(t)$ is the sum of three distinct types:

$$Y(t) = Y_f(t) + Y_m(t) + Y_{m2f}(t)$$

Last, considering the random entrance and exit of bird over the given area, the following SDE is used to capture the population dynamics of bird.

$$\begin{cases} \frac{dZ(t)}{dt} = Z(t)(p_3 - d_3 + s_2Y(t)) + \sigma_5Y(t)Z(t)dW_t \\ Z(0) = z_0 \end{cases} \quad (5)$$

where,

- p_3 , as a constant, is the reproduction rate for bird;
- d_3 , as a constant, is the natural death rate for bird;
- s_2 , as a constant, is the surviving rate due to food consuming;
- z_0 is the initial population of bird; and
- σ_5 is fluctuation rate.

If no estrogen is introduced into the environment, the ecosystem is stable and the model simulates what is essentially the predator-prey interaction. That is, initially there is a relatively high amount of fish, and relatively low amounts of bird and algae. This puts a strain on the fish population, while simultaneously making it easy for the bird to find prey due to the combination of a large food source and low competition for that food source. Thus this leads to a dip in the fish population and a peak in the bird population. The dip in the fish population also leads to a peak in the algae population, as the algae can grow without being consumed as fast due to the lack of fish. This scenario puts a strain on the bird population as there is now too much competition for a smaller food source, while simultaneously making it

easy for the fish to find food due to the combination of a large food source and low competition for that food source. Thus the population is back to the initial starting conditions, and the model continues to cycle through these scenarios ad infinitum. When various amounts of estrogen have been added, the estrogen leads to the feminization of male fish, with higher concentrations of estrogen corresponding to an increased likelihood of feminization. Feminized male fish cannot reproduce, which leads to more frequent dips in the fish population and can throw the entire ecosystem out of the equilibrium as described above. Essentially our model attempts to capture the long-term effects of estrogen, and demonstrates how sensitive a freshwater ecosystem can be to various concentrations of estrogen.

Our model involves rather complicated nonlinear integro-partial differential equations, which, moreover, often have accompanying nonlinear integro-boundary conditions, and SDEs. Such model equations are very difficult to analyze, not mentioning when hybrid dynamics are also involved. To be able to handle the probabilistic reachability analysis of this model, we are developing a platform by integrating numerical solving algorithms into our analyzers of hybrid systems - dReach [43], and of stochastic hybrid systems - SReach [60].

IV. CONCLUSION

Formal models offer system biologists a platform where they can use computational modeling and analysis tools to clarify and demystify complex systems. Models can be tested and adapted inexpensively in silico providing new insights. However, development of accurate and efficient modeling methodologies and analysis techniques are still open challenges for biological and biochemical systems. One long-term goal is that large-scale models should revolutionize biology and medicine and enable design of new therapies. To achieve this, general modeling and analysis frameworks will be needed for predictive and comprehensive models, such as whole-cell models, where various elements (e.g. genes, and proteins) and types of interactions (e.g. transport, and regulation) should be considered. Another goal is to automatically include useful and existing results from existing literature, public databases, and experimental data when building formal models for the biological systems. This will be benefited from creating a framework that will allow for creating and studying causal, explanatory models of complicated biological systems in which interactions have important causal effects. The modules included in the framework should provide functionality necessary for automation of information mining, information assembly and explanation of biological systems.

ACKNOWLEDGEMENT

This work was partially supported by ONR Award (N00014-13-1-0090), NSF CPS Breakthrough (CNS-1330014), and NSF CPS Frontier (CNS-1446725).

REFERENCES

- [1] 2014 CMU iGem. http://2014.igem.org/Team:Carnegie_Mellon. Accessed: 2016-08-06.
- [2] T. Akutsu, S. Miyano, and S. Kuhara. Inferring qualitative relations in genetic networks and metabolic pathways. *Bioinformatics*, 16(8):727–734, 2000.
- [3] T. Akutsu, S. Miyano, S. Kuhara, et al. Identification of genetic networks from a small number of gene expression patterns under the boolean network model. In *Pacific symposium on biocomputing*, volume 4, pages 17–28. Citeseer, 1999.
- [4] R. Albert and H. G. Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster. *Journal of theoretical biology*, 223(1):1–18, 2003.
- [5] R. Alur, C. Courcoubetis, T. A. Henzinger, and P.-H. Ho. Hybrid automata: An algorithmic approach to the specification and verification of hybrid systems. In *Hybrid systems*, pages 209–229. Springer, 1993.
- [6] L. Arnold. *Stochastic Differential Equations: Theory and Applications*. Wiley - Interscience, 1974.
- [7] V. Baldazzi, P. T. Monteiro, M. Page, D. Ropers, J. Geiselmann, and H. de Jong. Qualitative analysis of genetic regulatory networks in bacteria. In *Understanding the Dynamics of Biological Systems*, pages 111–130. Springer, 2011.
- [8] E. Bartocci and P. Lió. Computational modeling, formal analysis, and tools for systems biology. *PLoS Comput Biol*, 12(1):e1004591, 2016.
- [9] D. Barua, J. R. Faeder, and J. M. Haugh. Structure-based kinetic models of modular signaling protein function: focus on shp2. *Biophysical journal*, 92(7):2290–2300, 2007.
- [10] D. Barua, J. R. Faeder, and J. M. Haugh. Computational models of tandem src homology 2 domain interactions and application to phosphoinositide 3-kinase. *Journal of Biological Chemistry*, 283(12):7338–7345, 2008.
- [11] A. Beyer, P. Thomason, X. Li, J. Scott, and J. Fisher. Mechanistic insights into metabolic disturbance during type-2 diabetes and obesity using qualitative networks. In *Transactions on Computational Systems Biology XII*, pages 146–162. Springer, 2010.
- [12] A. T. Bittig, F. Haack, C. Maus, and A. M. Uhrmacher. Adapting rule-based model descriptions for simulating in continuous and hybrid space. In *Proceedings of the 9th International Conference on Computational Methods in Systems Biology*, pages 161–170. ACM, 2011.
- [13] M. L. Blinov, J. R. Faeder, B. Goldstein, and W. S. Hlavacek. A network model of early events in epidermal growth factor receptor signaling that accounts for combinatorial complexity. *Biosystems*, 83(2):136–151, 2006.
- [14] A. Bueno-Orovio, E. M. Cherry, and F. H. Fenton. Minimal model for human ventricular action potentials in tissue. *Journal of theoretical biology*, 253(3):544–560, 2008.
- [15] M. L. Bujorianu and J. Lygeros. Toward a general theory of stochastic hybrid systems. In *Stochastic Hybrid Systems*, pages 3–30. Springer, 2006.
- [16] C. Chaouiya. Petri net modelling of biological networks. *Briefings in bioinformatics*, 8(4):210–219, 2007.
- [17] K. C. Chen, L. Calzone, A. Csikasz-Nagy, F. R. Cross, B. Novak, and J. J. Tyson. Integrative analysis of cell cycle control in budding yeast. *Molecular biology of the cell*, 15(8):3841–3862, 2004.
- [18] L. Chen, G. Qi-Wei, M. Nakata, H. Matsuno, and S. Miyano. Modelling and simulation of signal transductions in an apoptosis pathway by using timed petri nets. *Journal of biosciences*, 32(1):113–127, 2007.
- [19] K. Claessen, J. Fisher, S. Ishtiaq, N. Piterman, and Q. Wang. Model-checking signal transduction networks through decreasing reachability sets. In *International Conference on Computer Aided Verification*, pages 85–100. Springer, 2013.
- [20] E. A. Coddington and N. Levinson. *Theory of ordinary differential equations*. Tata McGraw-Hill Education, 1955.
- [21] V. Danos, J. Feret, W. Fontana, R. Harmer, and J. Krivine. Rule-based modelling of cellular signalling. In *International Conference on Concurrency Theory*, pages 17–41. Springer, 2007.
- [22] H. De Jong. Modeling and simulation of genetic regulatory systems: a literature review. *Journal of computational biology*, 9(1):67–103, 2002.
- [23] P. D’haeseleer, S. Liang, and R. Somogyi. Genetic network inference: from co-expression clustering to reverse engineering. *Bioinformatics*, 16(8):707–726, 2000.
- [24] J. R. Faeder, M. L. Blinov, and W. S. Hlavacek. Rule-based modeling of biochemical systems with bionetgen. *Systems biology*, pages 113–167, 2009.
- [25] J. Fisher and T. A. Henzinger. Executable cell biology. *Nature biotechnology*, 25(11):1239–1249, 2007.
- [26] J. Fisher, A. S. Köksal, N. Piterman, and S. Woodhouse. Synthesising executable gene regulatory networks from single-cell gene expression data. In *International Conference on Computer Aided Verification*, pages 544–560. Springer, 2015.
- [27] M. Fränzle, E. M. Hahn, H. Hermanns, N. Wolovick, and L. Zhang. Measurability and safety verification for stochastic hybrid systems. In *HSCC*, pages 43–52, Apr. 2011.
- [28] H. Genrich, R. Küffner, and K. Voss. Executable petri net models for the analysis of metabolic pathways. *International Journal on Software Tools for Technology Transfer*, 3(4):394–404, 2001.
- [29] R. Ghosh and C. Tomlin. Symbolic reachable set computation of piecewise affine hybrid automata and its application to biological modelling: Delta-notch protein signalling. *Systems Biology*, 1(1):170–183, 2004.
- [30] H. Gong, Q. Wang, P. Zuliani, J. R. Faeder, M. Lotze, and E. Clarke. Symbolic model checking of signaling pathways in pancreatic cancer. In *BICoB*, page 245, 2011.
- [31] H. Gong, P. Zuliani, Q. Wang, and E. M. Clarke. Formal analysis for logical models of pancreatic cancer. In *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pages 4855–4860. IEEE, 2011.
- [32] P. J. Goss and J. Peccoud. Quantitative modeling of stochastic systems in molecular biology by using stochastic petri nets. *Proceedings of the National Academy of Sciences*, 95(12):6750–6755, 1998.
- [33] G. Gruenert, B. Ibrahim, T. Lenser, M. Lohel, T. Hinze, and P. Dittrich. Rule-based spatial modeling with diffusing, geometrically constrained molecules. *BMC bioinformatics*, 11(1):1, 2010.
- [34] S. Gupta, S. S. Bisht, R. Kukreti, S. Jain, and S. K. Brahmachari. Boolean network analysis of a neurotransmitter signaling pathway. *Journal of theoretical biology*, 244(3):463–469, 2007.
- [35] L. A. Harris, J. S. Hogg, and J. R. Faeder. Compartmental rule-based modeling of biochemical systems. In *Winter Simulation Conference*, pages 908–919. Winter Simulation Conference, 2009.
- [36] Z. Heidary, J. Ghaisari, S. Moein, M. Naderi, and Y. Gheisari. Stochastic petri net modeling of hypoxia pathway predicts a novel incoherent feed-forward loop controlling sdf-1 expression in acute kidney injury. *IEEE transactions on nanobioscience*, 15(1):19–26, 2016.
- [37] J. Hu, J. Lygeros, and S. Sastry. Towards a theory of stochastic hybrid systems. In *HSCC*, pages 160–173. Springer, 2000.
- [38] J. Hu, W.-C. Wu, and S. Sastry. Modeling subtilin production in bacillus subtilis using stochastic hybrid systems. In *International Workshop on Hybrid Systems: Computation and Control*, pages 417–431. Springer, 2004.
- [39] S. Huang and D. E. Ingber. Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks. *Experimental cell research*, 261(1):91–103, 2000.
- [40] B. Ibrahim, R. Henze, G. Gruenert, M. Egbert, J. Huwald, and P. Dittrich. Spatial rule-based modeling: a method and its application to the human mitotic kinetochore. *Cells*, 2(3):506–544, 2013.
- [41] P. Janmey, D. Fletcher, S. Gerecht, R. Levine, P. Mallick, O. McCarty, L. Munn, and C. Reinhart-King. Physical sciences and engineering advances in life sciences and oncology.
- [42] S. Kauffman. Homeostasis and differentiation in random genetic control networks. *Nature*, 224:177–178, 1969.
- [43] S. Kong, S. Gao, W. Chen, and E. Clarke. δ -reachability analysis for hybrid systems. In *International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 200–205. Springer, 2015.
- [44] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang. The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America*, 101(14):4781–4786, 2004.
- [45] D. Machado, R. S. Costa, M. Rocha, E. C. Ferreira, B. Tidor, and I. Rocha. Modeling formalisms in systems biology. *AMB express*, 1(1):1, 2011.
- [46] C. Maus, S. Rybacki, and A. M. Uhrmacher. Rule-based multi-level modeling of cell biological systems. *BMC Systems Biology*, 5(1):166, 2011.
- [47] N. Miskov-Zivanov, P. Wei, and C. S. C. Loh. THiMED: Time in hierarchical model extraction and design. In *International Conference on*

Computational Methods in Systems Biology, pages 260–263. Springer, 2014.

- [48] A. Naldi, D. Thieffry, and C. Chaouiya. Decision diagrams for the representation and analysis of logical models of genetic networks. In *International Conference on Computational Methods in Systems Biology*, pages 233–247. Springer, 2007.
- [49] C. Petri. *Kommunikation mit Automaten*. PhD thesis, Rheinisch-Westfälisches Institut f. instrumentelle Mathematik an d. Univ, 1962. Ph.D. thesis.
- [50] A. Platzer. Stochastic differential dynamic logic for stochastic hybrid programs. In *CADE*, pages 446–460. Springer, 2011.
- [51] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of computational physics*, 117(1):1–19, 1995.
- [52] K. Sachs, O. Perez, D. Pe’er, D. A. Lauffenburger, and G. P. Nolan. Causal protein-signaling networks derived from multiparameter single-cell data. *Science*, 308(5721):523–529, 2005.
- [53] M. A. Schaub, T. A. Henzinger, and J. Fisher. Qualitative networks: a symbolic approach to analyze biological signaling networks. *BMC systems biology*, 1(1):1, 2007.
- [54] I. Shmulevich, E. R. Dougherty, S. Kim, and W. Zhang. Probabilistic boolean networks: a rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2):261–274, 2002.
- [55] E. Simao, E. Remy, D. Thieffry, and C. Chaouiya. Qualitative modelling of regulated metabolic pathways: application to the tryptophan biosynthesis in e. coli. *Bioinformatics*, 21(suppl 2):ii190–ii196, 2005.
- [56] J. Sproston. Decidable model checking of probabilistic hybrid automata. In *FTRTFT*, pages 31–45. Springer, 2000.
- [57] J. Sproston. Model checking for probabilistic timed and hybrid systems. In *PhD thesis*. SCS, University of Birmingham, 2001.
- [58] Q. Wang, N. Miskov-Zivanov, B. Liu, M. Lotze, F. James, and E. M. Clarke. Formal modeling and analysis of pancreatic cancer microenvironment. In *International Conference on Computational Methods in Systems Biology*. Springer, 2015.
- [59] Q. Wang, N. Miskov-Zivanov, C. Telmer, and E. M. Clarke. Formal analysis provides parameters for guiding hyperoxidation in bacteria using phototoxic proteins. In *Proceedings of the 25th edition on Great Lakes Symposium on VLSI*, pages 315–320. ACM, 2015.
- [60] Q. Wang, P. Zuliani, S. Kong, S. Gao, and E. M. Clarke. Sreach: A probabilistic bounded delta-reachability analyzer for stochastic hybrid systems. In *International Conference on Computational Methods in Systems Biology*, pages 15–27. Springer, 2015.
- [61] P. Ye, E. Entcheva, S. Smolka, and R. Grosu. Modelling excitable cells using cycle-linear hybrid automata. *IET systems biology*, 2(1):24–32, 2008.