

A Hybrid Systems Framework for Cellular Processes

Kwang-Hyun Cho^{1,*}, Karl Henrik Johansson², and Olaf Wolkenhauer³

¹College of Medicine, Seoul National University, Chongno-gu, Seoul, 110-799, Korea and Korea Bio-MAX Institute, Seoul National University, Gwanak-gu, Seoul, 151-818, Korea

²Department of Signals, Sensors and Systems, Royal Institute of Technology, 100 44 Stockholm, Sweden

³Department of Computer Science, University of Rostock, 18051 Rostock, Germany.

Abstract

With the availability of technologies that allow us to obtain stimulus-response time series data for modeling and system identification, there is going to be an increasing need for conceptual frameworks in which to formulate and test hypotheses about intra- and inter-cellular dynamics, in general and not just dependent on a particular cell line, cell type, organism or technology. While the semantics can be quite different, biologists and systems scientists use in many cases a similar language (notion of feedback, regulation, *etc.*). A more abstract system-theoretic framework for signals, systems and control could provide the biologist with an interface between the domains. Apart from recent examples to identify functional elements and describing them in engineering terms, there have been various more abstract developments to describe dynamics at the cell level in the past. This includes Rosen's (M,R)-systems. This paper presents an abstract and general compact mathematical framework of intracellular dynamics, regulation and regime switching inspired by (M,R)-theory and based on hybrid automata.

Keywords: Systems biology; Cellular processes; (M,R)-theory; Dynamical model; Hybrid automata

1. Introduction

The renewed interests in systems biology reflects the natural shift of focus occurring in modern life sciences: away from the discovery of new components and their molecular characterization, towards an understanding of functional activity, interactions and the organization of components in higher structural levels (Wolkenhauer *et al.*, 2003). As more stimulus-response, time series data become available, a need for a system-theoretic framework of cellular processes, emerges. There has been considerable progress in mathematical modeling of intracellular dynamics, employing either stochastic simulation (Paulsson, 2004; Rao *et al.*, 2002) or nonlinear differential equations to model aspects of cell signaling (Kholodenko *et al.*, 2002; Brightman and Fell, 2000; Schoeberl *et al.*, 2002; Asthagiri and Lauffenburger, 2001; Cho and Wolkenhauer, 2003; Cho *et al.*, 2003a; Cho *et al.*, 2003b; Cho *et al.*, 2003c; Tyson *et al.*, 2003; Heinrich *et al.*, 2002), gene expression or transcriptional control amongst others (see (De Jong, 2002; Smolen *et al.*, 2000) and (Rao *et al.*, 2002) for survey articles). As alternatives to the dynamic systems approach there have been also various other techniques to model dynamics, rooted in formal languages, state machines, graph theory, artificial intelligence, *etc.* (Shmulevich *et al.*, 2002; Friedman *et al.*, 2000; Maki *et al.*, 2001; Rzhetsky *et al.*, 2004; Jordan, 1996). While in some

* Corresponding author. Postal address: Korea Bio-MAX Institute, 3rd Floor, International Vaccine Institute (IVI), Seoul National University Research Park, San 4-8, Bongcheon 7-dong, Gwanak-gu, Seoul, 151-818, Korea. E-mail: ckh-sb@snu.ac.kr, Tel.: +82-2-887-2650, Fax: +82-2-887-2692.

cases parameters for models are directly identified from experimental data (Cho *et al.*, 2003a; Swameye *et al.*, 2003; Tegner *et al.*, 2003) more often a model is tuned for its simulation to match the experience of the biologist. The validation with experimental data is then somewhat indirect. In any case, mathematical modeling and simulation is gaining acceptance as a complementary tool to test or generate hypotheses in molecular and cell biology. While the semantics can be quite different, biologists and control engineers use a similar language. A more abstract system-theoretic framework for signals, systems and control could provide an interface between both conceptual domains. This would help translating common terminology, including “feedback”, “control”, “regulation”, “amplification”, “filtering”, *etc.* Apart from recent examples to identify functional elements and describing them in engineering terms (Wolf and Arkin, 2003; Iglesias and Levchenko, 2002) there have been various more abstract formalisms to describe dynamics at the cell level in the past. These include Robert Rosen’s (M,R)-systems (Rosen, 1971; Casti, 1988; Wolkenhauer, 2001), which are revisited in the present paper.

While at the molecular level many processes appear to be of random nature, at higher levels of structural organization, i.e., cells, tissue and organs, well defined principles of functional relationships emerge. The basic physical building block is the cell, which through networks of biochemical reactions realizes many fundamental processes relevant to development and disease. Biochemical reaction networks are also referred to as ‘pathways’, which are the biologist’s conceptual frameworks to organize system variables, i.e., to identify relevant genes, proteins, metabolites, and to characterize their interactions in relation to cellular functions, including cell division, cell death, cell cycle, differentiation, proliferation, *etc.* More specifically these include from gene to cell level upwards: gene expression, transcription, translation, metabolism, physiological and immunological response and control processes. Robert Rosen recognized that many of these highly specific and complex functions can be characterized by two generic aspects: a transformation process, which he called ‘metabolism’ and a regulatory component, whose function he described as ‘repair’. These two abstract components form the basis for the theory of (M,R)-systems.

Inspired by (M,R)-theory, this paper aims at developing an extended framework of intracellular dynamics, regulation and regime switching, based on hybrid automata (Lygeros *et al.*, 2003; Alur and Dill, 1994; Alur *et al.*, 2000). The theory of hybrid systems allows us to complement continuous dynamics, which form the basis for many pathway models, with other regulatory or response levels that include changes to the elementary structure of dynamics (regime change, switching, nondeterministic transitions and so forth). The proposed framework also allows for sudden structural changes which, following Rosen’s and Casti’s terminology, we call ‘mutation’. Mutation is a structural change within a gene, chromosome and protein resulting in changes to the dynamics at gene-, protein-, metabolite- or the physiological level.

2. (M,R)-Theory for Biological Cellular Processes

Considering the functionality of a cell, Rosen and Casti identified a *metabolic* component, representing basic biochemical processes and a maintenance or *repair* component, which ensures the cell’s regular functioning in response to disturbances. Historically, the original formulation of (M,R)-theory was first given by Rosen (1971) and it was further developed by Casti (1988). Following Casti’s presentation (Casti, 1988), we consider first metabolic activity, represented by the mapping

$$h : \Omega \rightarrow \Gamma$$

where Ω is the set of environmental stimuli and Γ is the set of cellular responses. We denote by $H(\Omega, \Gamma)$ the family of all these mappings, i.e., the set of all physically realizable *metabolic* processes. We consider then the following basal metabolism for normal operation:

$$\gamma^* = h^*(\omega^*)$$

where ω^* , h^* , γ^* denote the environmental input, the metabolic map, and the cellular output, respectively, when everything is working according to plan under a normal condition. For deviations to the normal function, including external and internal disturbances to the cell's chemical activity, we introduce a *repair map*:

$$\alpha_{h^*} : \Gamma \rightarrow H(\Omega, \Gamma)$$

with the boundary condition $\alpha_{h^*}(\gamma^*) = h^*$. Since the repair component itself can be subject to disturbances, we require a further regulatory element, referred to as a *replication map*:

$$\beta_{h^*} : H(\Omega, \Gamma) \rightarrow G(\Gamma, H(\Omega, \Gamma))$$

with the boundary condition $\beta_{h^*}(h^*) = \alpha_{h^*}$, where $G(\Gamma, H(\Omega, \Gamma))$ denotes the family of all repair maps. We can summarize the abstract model structure for cellular processes in terms of the following morphisms:

$$\Omega \xrightarrow{h^*} \Gamma \xrightarrow{\alpha_{h^*}} H(\Omega, \Gamma) \xrightarrow{\beta_{h^*}} G(\Gamma, H(\Omega, \Gamma))$$

with boundary conditions

$$\alpha_{h^*}(\gamma^*) = h^* \text{ and } \beta_{h^*}(h^*) = \alpha_{h^*}.$$

If the dynamics encoded by h , and its regulatory or supervisory components $\alpha_{h^*}(\cdot)$, $\beta_{h^*}(\cdot)$, are not sufficient to cope with external disturbances or fluctuations, the next level of response is a transition to an alternative dynamic regime, for which we here introduce the discrete transition map δ . The conceptual framework developed here is outlined in Figure 1.

3. Dynamical Models and (M,R)-Systems

Before we proceed to develop a concrete structure of a cellular system, we first consider an extension of the (M,R)-description in terms of mappings, into a dynamical model familiar to control engineers. This can be done by shifting our focus from the relationship between conceptual mappings to the actual processes occurring in the cell. We now regard cellular processes as a dynamical system and focus on the inputs $\omega(t) : \mathbb{R}_+ \rightarrow \Omega \subset \mathbb{R}^m$ and $\lambda(t) : \mathbb{R}_+ \rightarrow \Lambda \subset \mathbb{R}^r$, the outputs $\gamma(t) : \mathbb{R}_+ \rightarrow \Gamma \subset \mathbb{R}^p$, and states $x(t) : \mathbb{R}_+ \rightarrow X \subset \mathbb{R}^n$. Note the abuse of notation in that Ω and Γ are now subsets of Euclidean spaces, while in the (M,R)-model Ω and Γ denote signal spaces. The dynamics of the cellular system is governed by

$$\dot{x} = f(x, \omega, \sigma)$$

$$\gamma = \zeta(x)$$

where the *environment variable* $\sigma = \sigma(\lambda(t), e) \in \Sigma$ is a function of a time-varying external disturbance $\lambda(t) \in \Lambda$ and a constant internal control $e \in E$. Here Σ is the set of possible environment variables, which are depending on the cellular status and E is the set of admissible internal controls. This dynamical model describes the cellular system in the normal operational phase. As indicated by the repair map of the (M,R)-model, the cell may transit into other phases which require repair and replication. This can be covered by a state-space model, introducing the partitioning of X into two disjoint subsets, X_d and X_{ud} , which represent 'desirable' and 'undesirable' operating modes, respectively. Similarly, we define the corresponding output partitions: $\Gamma_d = \zeta(X_d)$ and $\Gamma_{ud} = \zeta(X_{ud})$. The deviation from basal metabolism in normal operating phase is assumed to occur due to external disturbances and which affect the environment variable σ . The proposed dynamical model of the cellular processes is illustrated in Figure 2.

4. Hybrid Automaton Model of the Cellular Processes

For the previous dynamical model to describe the possible state transitions taking place, we employ the semantics of hybrid automata (Lygeros *et al.*, 2003; Alur and Dill, 1994; Alur *et al.*, 2000). A hybrid automaton can capture the *nondeterminism* of a state transition which is convenient in modeling and analysis of highly uncertain cellular processes. In cellular processes, those state transitions corresponding respectively to repair, replication, and mutation are of course all triggered by the failure of a proper metabolism, but it is also nondeterministic whether a specific transition can occur for a given failure situation.

A hybrid automaton Z is a collection $Z \triangleq (Q, X, f, Init, D, A, C, R)$, where Q is a discrete state space, X is a continuous state space, $f : Q \times X \rightarrow X$ is a vector field, $Init \subset Q \times X$ is a set of initial states, $D : Q \rightarrow P(X)$ is a domain, $A \subset Q \times Q$ is a set of arcs, $C : A \rightarrow P(X)$ is a guard condition, and $R : A \times X \rightarrow P(X)$ is a reset map. We refer to $(q, x) \in Q \times X$ as the *state* of Z . Roughly speaking, hybrid automata define possible evolutions for the state. Starting from an initial value $(q_0, x_0) \in Init$, the continuous state x flows as described by the vector field $f(q_0, \cdot)$, while the discrete state q remains constant. Continuous evolution can go on as long as x remains in $D(q_0)$. If at some point x reaches a guard $C(q_0, q_1)$, for some $(q_0, q_1) \in A$, the discrete state may change to q_1 . At the same time the continuous state is reset to some value in $R(q_0, q_1, x)$. After this discrete transition, continuous evolution resumes and the whole process is repeated. It is convenient to visualize hybrid automata as directed graphs (Q, A) with vertices Q and arc A . With each vertex $q \in Q$, we associate a set of initial states $Init_q = \{x \in X \mid (q, x) \in Init\}$, a vector field $f(q, \cdot)$ and domain $D(q)$. With each arc $a \in A$, we associate a guard $C(a)$ and a reset map $R(a, \cdot)$. For a non-autonomous system, we need to further include an external input $u \in U$ in Z , denoted by Z_u , and extend the above synopsis accordingly. A trajectory or solution of a hybrid automaton is called an *execution* or *run*. The definition of an execution involves conditions on the initial state, the continuous and discrete evolution. We say that a hybrid automaton accepts an execution or not; see (Lygeros *et al.*, 2003) for more formal definitions. It is important to note that a hybrid automaton may accept many executions or none from a single initial state, i.e., a hybrid automaton can be nondeterministic or blocking. Conditions for existence and uniqueness of executions are given in (Lygeros *et al.*, 2003).

There are numerous biological processes in which a multitude of dynamic regimes are employed to ensure the functioning of the cell. These are not necessarily 'regulatory' elements but also decision processes with permanent consequences. For example, during the development of an organism, organs or tissue, cells differentiate, i.e., they adopt a specialized biochemical and/or physiological role. There are therefore two kinds of change in genome activity (Alberts *et al.*, 2002; Brown, 1999): transient or switch-like, reversible responses to external stimuli of the cell via signaling compounds that either enter the cell or act through binding to surface receptors; secondly irreversible changes of genome activity underlying differentiation, and which can be brought about by DNA rearrangements, changes in chromatin structure, and positive feedback loops. Immunoglobulins are proteins that help protect an organism against invasion by bacteria, viruses and other unwanted substances by binding to these antigens. This binding is very specific so that every antigen is recognized by only one immunoglobulin and T-cell receptor protein. DNA rearrangement is a means for the human body to produce more immunoglobulins and T-cell receptor proteins than there are genes. Changes to the chromatin structure can have an effect on gene expression by modulation of transcription or silencing larger parts of the DNA. In general, for all processes which happen in the context of higher levels of organization where collections of cells function as a whole, *intra-cellular*

dynamics are linked to *inter*-cellular coordination of the activity of genomes in different cells. This coordination involves both transient and permanent changes, and must persist over a longer period of time during development.

In the context of intracellular signaling, an example for hybrid modeling is given by the switching phenomena of ERK activities associated with the Raf-1/MEK/ERK pathway (O'Neill and Kolch, 2004; Murphy *et al.*, 2002). Intra-cellular signaling pathways enable cells to perceive changes from their extra-cellular environments and produce appropriate responses (Cho and Wolkenhauer, 2003). Pathways are networks of biochemical reactions but they are also an abstraction biologists use to organize the functioning of the cell; they are the biologist's equivalent of the control engineer's block diagram. The Raf-1/MEK/ERK signaling pathway is a mitogen-activated protein kinase (MAPK) pathway, which exists ubiquitously in most of the eukaryotic cells and is involved in various biological responses (Kolch, 2000). Figure 3, adapted from (O'Neill and Kolch, 2004) illustrates the hybrid system dynamics of the Raf-1/MEK/ERK pathway of PC12 cells. The different ERK dynamics are achieved through the combinatorial integration and activation of different Raf isoforms and crosstalk with the cAMP signaling system, which results in discrete state transitions to different cellular dynamics. PC12 cells differentiate in response to nerve growth factor (NGF), but proliferate in response to epidermal growth factor (EGF). Both growth factors utilize the Raf-1/MEK/ERK pathway. The sustained ERK activity caused by the B-raf isoform results in neuronal differentiation while the transient ERK activity caused by the activation of cAMP signaling and the inhibition of Raf-1 results in cell proliferation.

For more detailed illustration and to provide a comprehensive idea on the hybrid system modeling, we consider a non-autonomous hybrid system model of a xanthophyll cycle reaction system in plant science. Understanding the manner in which excess solar energy is dissipated by photosynthetic membranes under conditions of high light stress has been a major problem facing researchers studying photosynthesis (Demming-Adams, 1990). It has long been known that higher plants possess several defense mechanisms against excess light. Those include the xanthophyll cycle that dissipates excess light energy as heat (Demming-Adams *et al.*, 1996). The xanthophyll cycle is an interconversion process between violaxanthin (Vio) and zeaxanthin (Zea). In excessive light, the build up of a trans-thylakoid proton gradient activates the de-epoxidase which converts Vio into Zea via an intermediate antheraxanthin (Anth). The back reaction, epoxidation of Zea, is light-independent and catalyzed by an epoxidase thought to be located in the stromal side of the thylakoid membrane (Siefermann and Yamamoto, 1975).

The xanthophyll cycle reaction system is illustrated in Figure 4 where x_1 denotes the level of Vio [%], x_2 the level of Anth [%], x_3 the level of Zea [%], k_1 , k_2 the de-epoxidation rate constants, and k_3 , k_4 the epoxidation rate constants. The dynamics of the xanthophyll cycle reaction system can be modeled as follows:

$$\begin{aligned}\dot{x}_1 &= -k_1x_1 + k_4x_2 \\ \dot{x}_2 &= k_1x_1 - k_4x_2 + k_3x_3 - k_2x_2 \\ \dot{x}_3 &= k_2x_2 - k_3x_3\end{aligned}$$

There have been accumulated studies investigating the dynamics of xanthophyll cycle and the laboratory studies show that the dynamics largely depend on the light stress and the inhibitor treatment such as salicylaldehyde (SA) as an epoxidase inhibitor (Xu *et al.*, 1999). Hence the rate constants of the above equations can vary according to the external environmental conditions and the xanthophyll cycle reaction system can be therefore regarded as a hybrid system. We consider a hybrid system model of the xanthophyll cycle reaction system as follows (note that an exposition of the barley leaves is considered but detailed experimental conditions are not described here since those are out of scope of this paper):

$$Z_u = (Q, X, f, u, Init, D, A, C, R),$$

where

- $Q = \{q_1, q_2, q_3, q_4\}$;
- $X = \{x_1, x_2, x_3\}$ and $X = \mathbb{R}^3$;
- $f(q_i, x) = (-k_{i1}x_1 + k_{i4}x_2, k_{i1}x_1 - k_{i4}x_2 + k_{i3}x_3 - k_{i2}x_2, k_{i2}x_2 - k_{i3}x_3)$ for $i = 1, 2, 3, 4$;
- $u \in U = \{(u_1, u_2)\}$ with $u_1 \in \{LL, HL\}$ and $u_2 \in \{+SA, -SA\}$ where LL denotes the low light stress, HL the high light stress, $+SA$ the inhibitor treatment of SA, and $-SA$ the condition without the treatment of SA;
- $Init = Q \times \{x \in \mathbb{R}^3 \mid 12.2 \leq x_1 \leq 96.0 \wedge 4.0 \leq x_2 \leq 25.8 \wedge 0.0 \leq x_3 \leq 80.9\}$;
- $D(q_1) = \{x \in \mathbb{R}^3 \mid 33.3 \leq x_1 \leq 96.0 \wedge 4.0 \leq x_2 \leq 6.7 \wedge 0.0 \leq x_3 \leq 59.9\}$,
 $D(q_2) = \{x \in \mathbb{R}^3 \mid 33.3 \leq x_1 \leq 47.7 \wedge 6.7 \leq x_2 \leq 25.8 \wedge 31.1 \leq x_3 \leq 59.9\}$,
 $D(q_3) = \{x \in \mathbb{R}^3 \mid 12.2 \leq x_1 \leq 47.7 \wedge 6.5 \leq x_2 \leq 21.2 \wedge 31.1 \leq x_3 \leq 80.6\}$,
 $D(q_4) = \{x \in \mathbb{R}^3 \mid 12.2 \leq x_1 \leq 19.7 \wedge 5.1 \leq x_2 \leq 7.2 \wedge 75.0 \leq x_3 \leq 80.9\}$;
- $A = \{(q_i, q_j) \mid 1 \leq i, j \leq 4, i \neq j\}$;
- $C : A \rightarrow U$ with $C(q_1, q_2) = (LL, -SA)$, $C(q_1, q_3) = (LL, +SA)$, $C(q_1, q_4) = (HL, +SA)$,
 $C(q_2, q_3) = (LL, +SA)$, $C(q_2, q_1) = (HL, -SA)$, $C(q_2, q_4) = (HL, +SA)$,
 $C(q_3, q_4) = (HL, +SA)$, $C(q_3, q_1) = (HL, -SA)$, $C(q_3, q_2) = (LL, -SA)$,
 $C(q_4, q_1) = (HL, -SA)$, $C(q_4, q_2) = (LL, -SA)$, $C(q_4, q_3) = (LL, +SA)$;
- $R : A \times X \rightarrow P(X)$ with $R(q_i, q_j, x) = \{x\}$, $1 \leq i, j \leq 4$.

The respective rate constants set of different environmental conditions is as follows:

	$i = 1 (q_1)$	$i = 2 (q_2)$	$i = 3 (q_3)$	$i = 4 (q_4)$
k_{i1}	11.4	14.1	17.0	19.8
k_{i2}	218.7	430.6	330.1	141.9
k_{i3}	19.2	293.2	27.9	10.4
k_{i4}	7.2	33.2	10.1	9.0

A sample trajectory with the initial state $x(0) = (96.0, 4.0, 0.0)$ [%] is shown in Figure 5 according to the execution of the model, where it is assumed that the dark adapted barley leaves are exposed to high light during $[0, 0.1]$ (hours), exposed to low light during $[0.1, 0.2]$ (hours), SA is added during $[0.2, 0.3]$ (hours), and exposed again to high light during $[0.3, 0.4]$ (hours). It is therefore implied that $C(q_1, q_2) = (LL, -SA)$, $C(q_2, q_3) = (LL, +SA)$, and $C(q_3, q_4) = (HL, +SA)$. From Figure 5, we realize that in $q_1(D_1)$, x_3 increases rapidly due to the photoprotection mechanism under the high light stress while it decreases for the low light stress in $q_2(D_2)$; x_3 increases in $q_3(D_3)$ due to inhibition of the epoxidation reaction by SA treatment while it becomes saturated under the further high light stress in $q_4(D_4)$.

Other interesting hybrid system modeling examples of biological systems include the full reactive modeling of a multi-cellular animal in (Harel, 2003) where the *C. elegans* nematode worm is exemplified, the Delta-Notch biological cell signaling networks in (Ghosh *et al.*, 2003), and the genetic regulatory network underlying the initiation of sporulation in *B. subtilis* in (De Jong *et al.*, 2003).

5. Extended Hybrid Automaton Model

Inspired by Rosen's and Casti's model, and based on the aforementioned dynamical model, we can build an extended hybrid automaton of the hybrid dynamics denoted M . It is a minor extension of the model presented in the previous section and defined by

$$M = (Q, X, \Omega, \Gamma, f, Init, D, A, C, R),$$

where

- $Q = \{q_{normal}, q_{repaired}, q_{replicated}\};$
- $X = \mathbb{R}^n$ is an open connected set with $X = X_d \cup X_{ud}$ and $X_d \cap X_{ud} = \emptyset$;
- $\Omega \subset \mathbb{R}^m$;
- $\Gamma \subset \mathbb{R}^p$;
- $f = f_h(x, \omega, \sigma_l)$ with $x(t) \in X$, $\omega(t) \in \Omega$, and $\sigma_l = \sigma_l(\lambda(t), e)$ in which $\lambda(t) \in \Lambda$ is a time-varying external disturbance and $e \in E_k$ is a constant internal control. Here $\bigcup_{k \in K} E_k$ is a family of controls parameterized in $K \subset \mathbb{N}$ and $\Sigma = \{\sigma_l\}_{l \in L}$ is a family of environment map parameterized in $L \subset \mathbb{N}$. Moreover, $\gamma = \zeta(x)$ is an output map;
- $Init = \{(q_{normal}, x) \mid x \in X_d^0\}$ in which X_d^0 is the interior of X_d ;
- $D(q) = X_d$ for $q \in Q$;
- $A = \{(q_{normal}, q_{repaired}), (q_{repaired}, q_{repaired}), (q_{repaired}, q_{replicated}), (q_{replicated}, q_{replicated}), (q_{replicated}, q_{normal})\}$;
- $C(a) = X_{ud}$ for $a \in A$;
- $R: A \times X \times K \times L \rightarrow P(X) \times K \times L$ is a reset map. Here we extend the reset map of the hybrid automaton by including two index sets $K, L \subset \mathbb{N}$. The corresponding variables, $k \in K, l \in L$ are simply updated according to R at each discrete transition, similar to the update of the continuous state x . So, we have
 - $R(q_{normal}, q_{repaired}, x, k, l) = \{(x', k, l) \mid x' \in X\}$ (repair),
 - $R(q_{repaired}, q_{repaired}, x, k, l) = \{(x', k, l) \mid x' \in X\}$ (repair),
 - $R(q_{repaired}, q_{replicated}, x, k, l) = \{(x', k', l) \mid x' \in X, k' \in K\}$ (replication),
 - $R(q_{replicated}, q_{replicated}, x, k, l) = \{(x', k', l) \mid x' \in X, k' \in K\}$ (repair or replication) and
 - $R(q_{replicated}, q_{normal}, x, k, l) = \{(x', k', l') \mid x' \in X, k' \in K, l' \in L\}$ (mutation).

Figure 6 illustrates the hybrid automaton model and Figure 7 shows three executions accepted by the hybrid automaton of the cellular processes. The execution marked (1) illustrates repair: a discrete transition occurs due to that the continuous state enters X_{ud} . The reset map sets the new continuous state to $x \in X_d$ and a new internal control $e \in E_k$. The execution marked (2) illustrates replication: again a discrete transition occurs due to that the continuous state enters X_{ud} . Now, however, the reset map sets the new continuous state to $x \in X_{ud}$, which triggers a second transition. This time the reset map sets $x \in X_d$ but also updates $k \in K$. The latter leads to a new E_k and $e \in E_k$. The execution marked (3) illustrates mutation: Here also the second transition leads to $x \in X_{ud}$. Therefore a third transition takes place. Then the reset map sets $x \in X_d$ and updates $l \in L$, which leads to a new environment variable $\sigma_l \in \Sigma$. Note that the reset map assigns a new continuous state $x' \in X$, independent of the past state $x \in X_{ud}$.

Hence, the hybrid automaton accepts several executions and thus represents the uncertainty of the cellular dynamics. Properties such as reachable set computations, liveness, and stability can be analyzed for nondeterministic hybrid automata.

6. Concluding Remarks

With the availability of technologies that allow us to obtain stimulus-response time series data for modeling and system identification there is going to be an increasing need for conceptual frameworks in which to formulate and test hypotheses about intra- and inter-cellular dynamics, in general and not just dependent on a particular cell line, cell type, organism or technology. To this day, experimental data in molecular and cell biology are highly context-dependent but this appears to change and is going to provide control engineers with opportunities and challenges. We here presented an abstract but general, compact mathematical framework that extends (M,R)-theory to take into consideration dynamic aspects of cell signaling and gene expression and to allow for models of reversible, switching and permanent changes occurring. The hybrid automata model illustrated how highly nonlinear dynamics and nondeterminism can be captured in a formal setting. There exist several tools for the reachability calculations, stability analysis, and computer simulations of hybrid automata, which have been developed over the last decade (Lygeros *et al.*, 2003) and might be useful in the study of cellular processes. The proposed mathematical model of cellular processes could form a basis for further discussions and extensions. One direction of such an extension is the inclusion of the diverse signaling used in cellular communication. Various alternative or complementary concepts could be investigated including temporal logic, concurrency theory, stochastic automata, *etc.* Any text on modern molecular or cell biology suggests a range of problems for which established system theoretic concepts need to be extended as the complexity of these systems appear to go beyond anything that we have been familiar with in the engineering sciences. A major challenge is the generation of quantitative stimulus-response time series to enable the application of system identification techniques. However, despite of the uncertainty one faces in modeling intra-cellular dynamics, the encouraging experience is that even drastically simplified models can provide useful practical guidance for the design of experiments, helping the experimentalist to decide what to measure and why.

Acknowledgments

The preparation and revision of the xanthophyll cycle reaction system example benefited from discussion with Prof. C.-H. Lee and S.-Y. Shin. The work of K.-H. Cho was supported by a grant from the Korean Ministry of Science and Technology (Korean Systems Biology Research Grant, M10309000006-03B5000-00211). Karl Henrik Johansson acknowledges the support received by the Swedish Research Council. O. Wolkenhauer acknowledges the support received by the U.K. Department for the Environment, Food and Rural Affairs (DEFRA) as part of the *M. bovis* post-genomics programme.

References

- Alur, R., Dill, D.L., 1994. A theory of timed automata. *Theoretical Computer Science* 126, 183–235.
- Alur, R., Henzinger, T.A., Lafferriere, G., Pappas, G.J., 2000. Discrete abstractions of hybrid systems. *Proc. IEEE* 88, 971–984.
- Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., Watson, J.D., 2002. *Molecular Biology of the Cell*. 4th Edition, Garland Science.

Asthagiri, A.R., Lauffenburger, D.A., 2001. A computational study of feedback effects on signal dynamics in a mitogen-activated protein kinase (MAPK) pathway model. *Biotechnol. Prog.* 17, 227–239.

Brown, T.A., 1999. *Genomes*. βIOS Scientific Publishers.

Brightman, F.A., Fell, D.A., 2000. Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signaling in PC12 cells. *FEBS Letters* 482, 169–174.

Casti, J.L., 1988. The theory of metabolism-repair systems. *Applied Mathematics and Computation* 28, 113–154.

Cho, K.-H., Wolkenhauer, O., 2003. Analysis and modeling of signal transduction pathways in systems biology. *Biochemical Society Transactions* 31, 1503–1509.

Cho, K.-H., Shin, S.-Y., Kim, H.-Y., Wolkenhauer, O., McFerran, B., Kolch, W., 2003a. Mathematical modeling of the influence of RKIP on the ERK signaling pathway. In: Priami, C. (Ed.), *Computational Methods in Systems Biology*. Springer-Verlag, Berlin, LNCS 2602, pp. 127–141.

Cho, K.-H., Shin, S.-Y., Lee, H.-Y., Wolkenhauer, O., 2003b. Investigations into the analysis and modeling of the TNF α mediated NF- κ B signaling pathway. *Genome Research* 13, 2413–2422.

Cho, K.-H., Shin, S.-Y., Kolch, W., Wolkenhauer, O., 2003c. Experimental design in systems biology based on parameter sensitivity analysis with Monte Carlo method: A case study for the TNF α mediated NF- κ B signal transduction pathway. *SIMULATION: Transactions of the Society for Modeling and Simulation International* 79, 726–739.

De Jong, H., 2002. Modeling and simulation of genetic regulatory systems: A literature review. *J. of Computational Biology* 9, 67–103.

De Jong, H., Gouze, J.-L., Hernandez, C., Page, M., Sari, T., Geiselmann, J., 2003. Hybrid modeling and simulation of genetic regulatory networks: A qualitative approach. *Proc. 6th Int. Workshop on Hybrid Systems: Computation and Control (HSCC 2003)*. Prague, Czech Republic, Maler, O., Pnueli, A. (Eds.), LNCS 2623, Springer-Verlag: Berlin, pp. 267–282.

Demming-Adams, B., 1990. Carotenoids and photoprotection in plants: A role for the xanthophyll zeaxanthin. *Biochim. Biophys. Acta* 1020, 1–24.

Demming-Adams, B., Gilmore, A.M., Adams III, W.W., 1996. In vivo functions of carotenoids in higher plants. *FASEB J.* 10, 403–412.

Friedman, N., Linial, M., Nachman, I., Pe'er, D., 2000. Using Bayesian networks to analyze expression data. *J. of Computational Biology* 7, 585–600.

Ghosh, R., Tiwari, A., Tomlin, C., 2003. Automated symbolic reachability analysis with application to Delta-Notch signaling automata. *Proc. 6th Int. Workshop on Hybrid Systems: Computation and Control (HSCC 2003)*. Prague, Czech Republic, Maler, O., Pnueli, A. (Eds.), LNCS 2623, Springer-Verlag: Berlin, pp. 233–248.

Harel, D., 2003. A grand challenge for computing: Full reactive modeling of a multi-cellular animal. *Bulletin of the EATCS (European Association for Theoretical Computer Science)* 81, 226–235.

Heinrich, R., Neel, B.G., Rapoport, T.A., 2002. Mathematical models of protein kinase signal transduction. *Molecular Cell* 9, 957–970.

Iglesias, P.A., Levchenko, A., 2002. Modeling the cell's guidance system. *Science STKE* 148, re12 (www.stke.org/cgi/content/full/sigtrans;2002/148/re12).

Jordan, M.I. (Ed.), 1996. *Learning in Graphical Models*. Kluwer Academic Publishers: Dordrecht, The Netherlands.

Kholodenko, B.N., Klyatkin, A., Bruggeman, F.J., Sontag, E., Westerhoff, H.V., Hoek, J.B., 2002. Untangling the wires: A strategy to trace functional interactions in signaling and gene networks. *PNAS* 99, 12841–12846.

Kolch, W., 2000. Meaningful relationships: The regulation of the Ras/Raf/MEK/ERK pathway by protein interactions. *Biochem. J.* 351, 289–305.

Lygeros, J., Johansson, K.H., Simic, S.N., Zhang, J., Sastry, S., 2003. Dynamical properties of hybrid automata. *IEEE Trans. Automatic Control* 48, 2–17.

Maki, Y., Tominaga, D., Okamoto, M., Watanabe, S., Eguch, Y., 2001. Development of a system for the inference of large scale genetic networks. *Proc. of the Pacific Symposium on Biocomputing* 6, 446–458.

Murphy, L.O., Smith, S., Chen, R.H., Fingar, D.C., Blenis, J., 2002. Molecular interpretation of ERK signal duration by immediate early gene products. *Nature Cell Biology* 4, 556–564.

O'Neill, E., Kolch, W., 2004. Conferring specificity of the ubiquitous Raf/MEK signaling pathway. *British Journal of Cancer* 90, 283–288.

Paulsson, J., 2004. Summing up the noise. *Nature* 427, 415–418.

Rao, C.V., Wolf, D.M., Arkin, A.P., 2002. Control, exploitation and tolerance of intracellular noise. *Nature* 420, 231–237.

Rosen, R., 1971. Some realizations of (M,R)-systems and their interpretation. *Bulletin of Mathematical Biophysics* 33, 303–319.

Rzhetsky, A., Iossifov, I., Koike, T., Krauthammer, M., Kra, P., Morris, M., Yu, H., Duboue, P.A., Weng, W., Wilbur, W.J., Hatzivassiloglou, V., Friedman, C., 2004. GeneWays: A system for extracting, analyzing, visualizing, and integrating molecular pathway data. *J. of Biomedical Informatics* 37, 43–53.

Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., Müller, G., 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nature Biotechnology* 20, 370–375.

Shmulevich, I., Dougherty, E.R., Kim, S., Zhang, W., 2002. Probabilistic Boolean network: A rule-based uncertainty model for gene regulatory networks. *Bioinformatics* 18, 261–274.

Siefermann, D., Yamamoto, H.Y., 1975. Properties of NADPH and oxygen-dependent zeaxanthin epoxidation in isolated chloroplasts. *Arch. Biochem. Biophys.* 171, 70–77.

Smolen, P., Baxter, D.A., Byrne, J.H., 2000. Modeling transcriptional control in gene networks – Methods, recent results, and future directions. *Bulletin of Mathematical Biology* 62, 247–292.

Swameye, I., Müller, T.G., Timmer, J., Sandra, O., Klingmüller, U., 2003. Identification of nucleocytoplasmic cycling as a remote sensor in cellular signaling by databased modeling. *PNAS* 100, 1028–1033.

Tyson, J.J., Chen, K.C., Novak, B., 2003. Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology* 15, 221–231.

Tegner, J., Young, M.K.S., Hasty, J., Collins, J.J., 2003. Reverse engineering gene networks: Integrating genetic perturbations with dynamical modeling. *PNAS* 100, 5944–5949.

Wolf, D.M., Arkin, A.P., 2003. Motifs, modules and games in bacteria. *Current Opinion in Microbiology* 6, 125–134.

Wolkenhauer, O., 2001. Systems biology: The reincarnation of systems theory applied in biology? *Briefings in Bioinformatics* 2, 258–270.

Wolkenhauer, O., Kitano, H., Cho, K.-H., 2003. Systems biology: Looking at opportunities and challenges in applying systems theory to molecular and cell biology. *IEEE Control Systems Magazine* 23, 38–48.

Xu, C.C., Jeon, Y.A., Hwang, H.J., Lee, C.-H., 1999. Suppression of zeaxanthin epoxidation by chloroplast phosphatase inhibitors in rice leaves. *Plant Science* 146, 27–34.

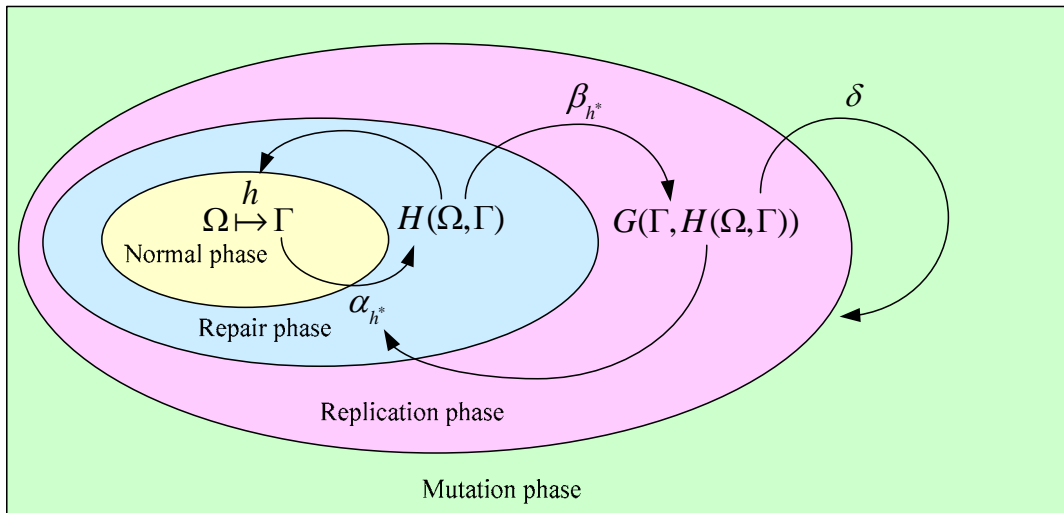


Figure 1: Conceptual modeling of cellular processes. Each transition to a different phase or an alternative dynamic regime can be modeled as a discrete transition, which further invokes a hybrid system framework.

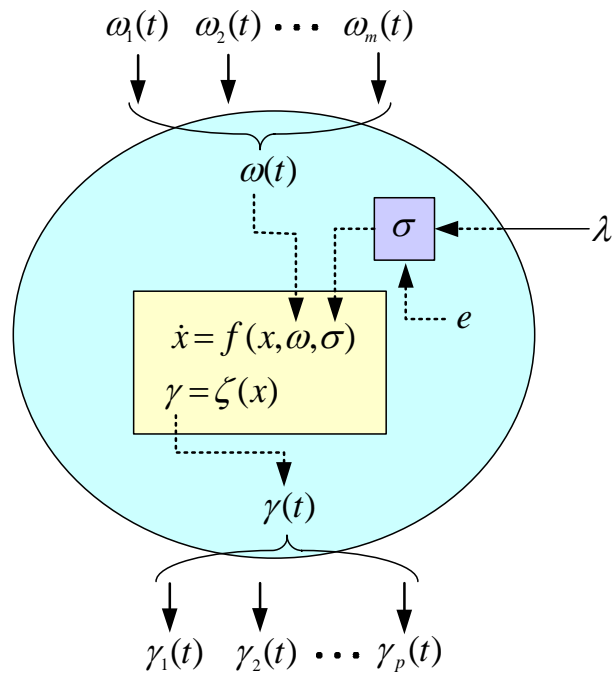


Figure 2: A dynamical model of the cellular processes. The external disturbance λ affects the environment variable σ and it is further controlled by an internal control e . Transitions into different operating phases due to the external disturbances are accounted for partitioning the state-space into the corresponding disjoint subspaces.

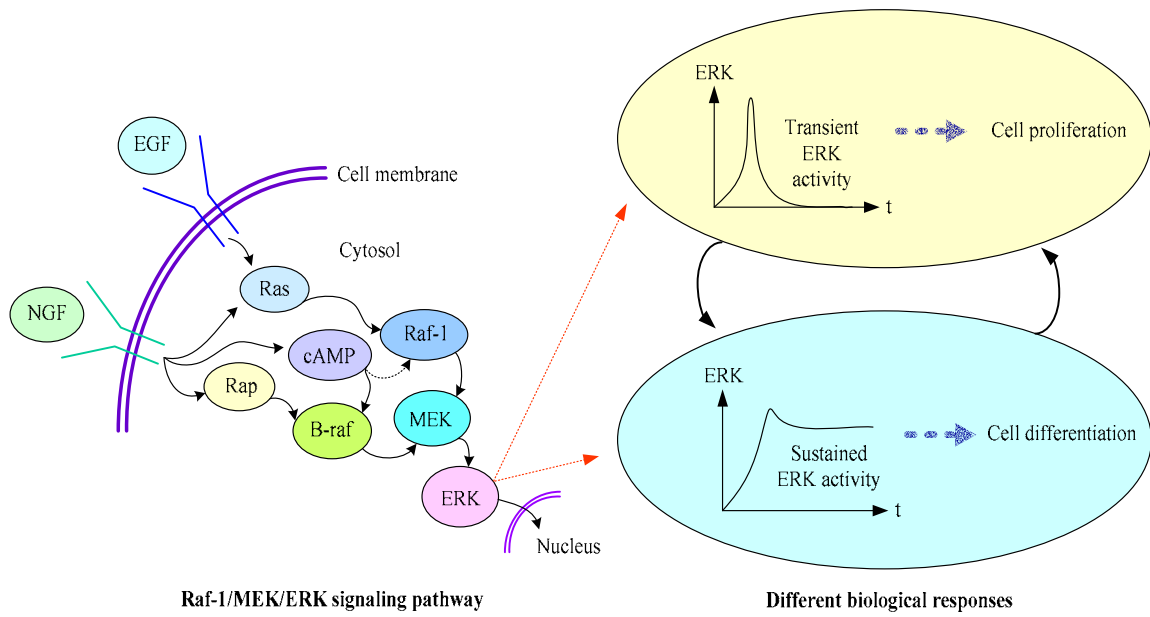


Figure 3: Hybrid system dynamics of the Raf-1/MEK/ERK cellular signaling pathway in PC12 cells, where both, the quantity and history of ERK concentrations determine discrete state transitions to different dynamics that decide upon the cell's fate.

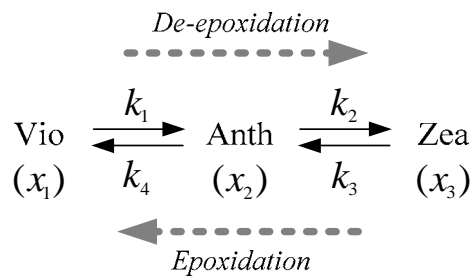


Figure 4: State diagram of the xanthophyll cycle reaction system.

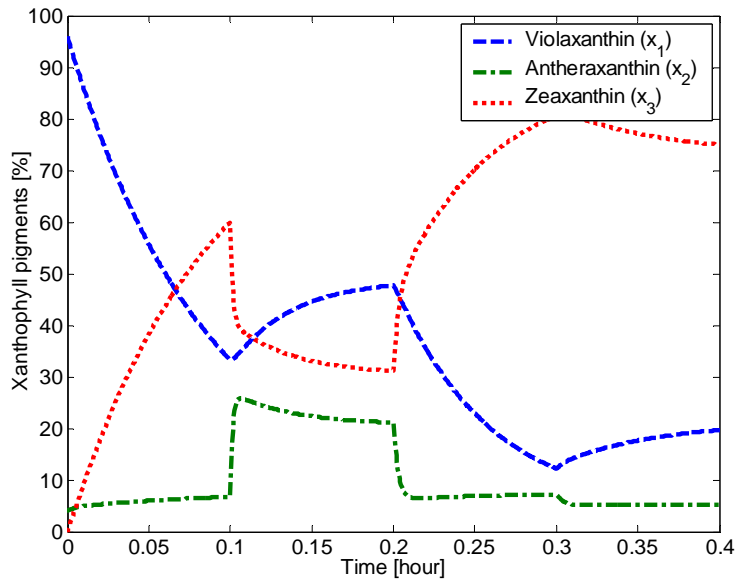


Figure 5: Sample trajectories of the xanthophyll cycle reaction system according to executions of the hybrid system model.

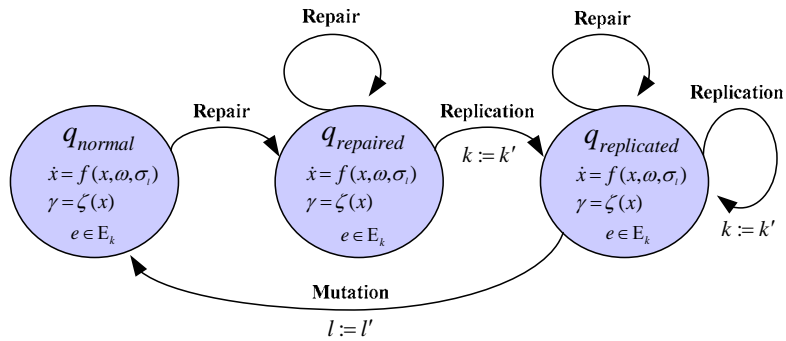


Figure 6: A hybrid automaton of cellular processes.

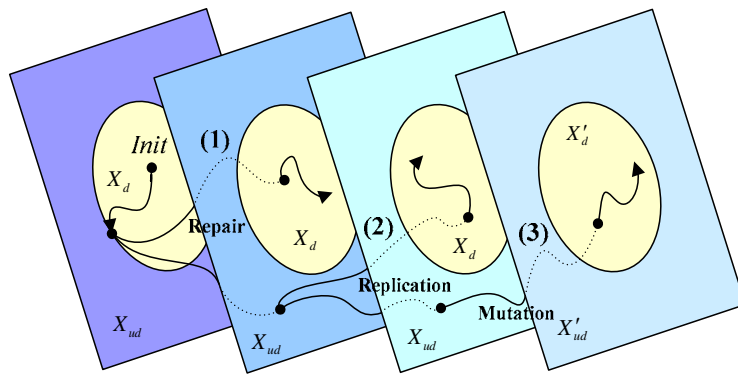


Figure 7: Examples of executions accepted by the hybrid automaton.