Intuitive Modeling of Dynamic Systems with Petri Nets and Fuzzy Logic

Lukas Windhager*, Ralf Zimmer

Institut für Informatik, Ludwig-Maximilians-Universität München,
Amalienstrasse 17, 80333 München, Germany
Lukas.Windhager@bio.ifi.lmu.de

Abstract: Current approaches in modeling dynamic biological systems often lack comprehensibility, especially for users without mathematical background. We propose a new approach to overcome such limitations by combining the graphical representation provided by the use of Petri nets with the modeling of dynamics by powerful yet intuitive fuzzy logic based systems. The mathematical functions and formulations typically used to describe or quantify dynamic changes of systems are replaced by if-then rules, which are both easy to read and formulate. Precise values of kinetic constants or concentrations are substituted by more natural fuzzy representations of entities. We will show that our new approach allows a semi-quantitative modeling of biological systems like signal transduction pathways or metabolic processes while not being limited to those cases.

1 Introduction

To gain insight into a biological system, computational models are built based on current knowledge and hypotheses. The behavior of these models is investigated under different constraints and compared to experimental observations, known facts or other data to verify or falsify the current model. Many of the currently availabe approaches for modeling biological systems are based on ordinary differential equations (ODEs), Bayesian or boolean networks, different types of Petri nets (PNs), combinations thereof as well as other, less common techniques like signal-flow diagrams and system dynamics models. See [GFG+06, MPLD04, OSV+05] for some reviews concerning computational modeling. ODE based modeling of dynamic changes in systems is probably the most widespread method. Entities of the modeled system (proteins, metabolites, etc.) are described by state variables which typically correspond to the concentrations or amounts of those entities at a given time. The change of these variables over time is hereby described by a set of differential equations which involve not only the state variables but also several kinetic constants. ODE based modeling was applied for example for the analysis of yeast cell cycle [CCNG+00], E. coli carbohydrate uptake [KBG07], dynamics of yeast pheromone pathway [KK04] or the modeling of the EGF receptor induced MAP kinase

*Corresponding author.

cascade [SEJGM02]. Some widely used graph-based approaches to systems biology modeling are based on Petri nets (see [Cha07] for a recent review and [Mur89] for an extensive introduction to Petri net theory). Generally, Petri nets are graphical representations of (biological) entities like proteins, genes and metabolites as well as (biological) processes like enzymatic reactions, transport, degradation, etc.. There are several different types of Petri net modeling techniques in use, ranging from the basic type (see [RLM96]) to more involved and extended types like hybrid functional Petri nets (HFPN; [MTA+03]). HFPNs extend the definition of basic Petri nets by introducing additional arc types (inhibitory and test arcs), a more sophisticated definition of tokens and the use of arbitrary functions instead of fixed arc-weights. These functions are typically similar to ODEs, incorporating concentrations of neighboring places and pre-defined kinetic constants. See [GKV01] for an executable Petri net model of glycolysis and citric acid cycle, [LZLP06] for a colored Petri net model of the EGF receptor induced MAP kinase cascade or [LGN+07] for a timed Petri net model of the apoptosis pathway.

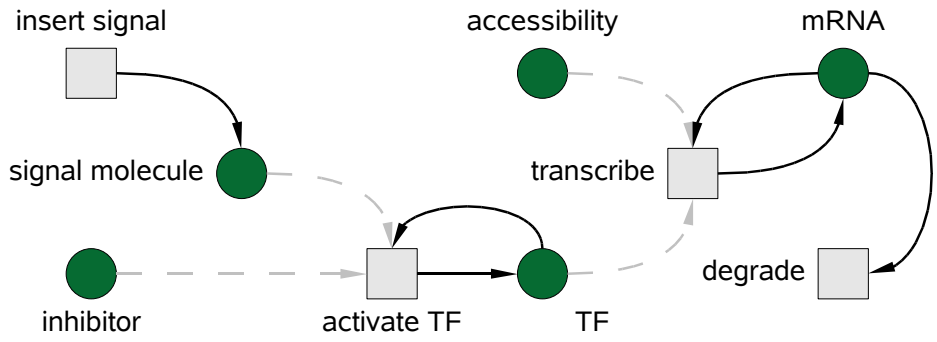


Figure 1: A Petri net representation of a small biological system. *Places* (filled circles) visualize arbitrary system entities or properties like proteins, metabolites, etc.. In our framework, their current states are described using weighted fuzzy sets replacing the commonly used tokens. *Transitions* (grey squares) visualize arbitrary (biological) processes like enzymatic reactions, transport, etc.. *Arcs* visualize dependencies of places and transitions (test arcs, dashed arrows) or define which and how places are affected whenever a transition fires (input and output arcs, solid arrows). In our framework, these effects are defined using fuzzy logic systems instead of the commonly used weights or other mathematical functions.

In this article, we introduce and motivate a new modeling approach (termed **PNFL**, Petri Nets with Fuzzy Logic) which provides a powerful and intuitive tool for investigating biological processes and systems. PNFL provides an environment where hypotheses in biological systems can be formulated, visualized and simulated in a quite intuitive and natural way and overcomes limitations of ODE-based modeling by:

1. Replacing mathematical formulations of dynamics by natural language based rule systems to facilitate comprehensibility.
2. Omitting use, definition and estimation of exact parameter values through a fuzzier, thus natural, definition of typically qualitative knowledge about entities and processes.

3. Allowing for incorporation of entities and their concentrations as well as other, arbitrary properties of entities or systems by a uniform framework based on fuzzy logic.
4. Using Petri nets as graphical frameworks for development and simulation of user-defined systems to provide a clear visualization and distinction of entities and processes.

The main innovation of our PNFL approach is the use of elements from fuzzy logic theory to describe biological systems: *Fuzzy sets* describe arbitrary entities or properties of a system; *Fuzzy logic systems* define the dynamics of biological processes and dependencies between entities. Petri nets are used as a scaffold for the fuzzy logic based definitions of biological entities and processes (figure 1).

2 Fuzzy Logic Based Modeling

The real world has an approximate and inexact nature and sets of objects in this world are usually characterized by inexact boundaries. For example, defining the "set of highly concentrated metabolites" as "the set of metabolites present at a level of more than $1 * 10^6$ molecules per mol" is unsatisfactory as this strict border is probably artificial. It is difficult to argue, that a metabolite present at $1.01 * 10^6$ molecules per mol is "highly concentrated" while it would not be "highly concentrated" at $0.99 * 10^6$ molecules per mol. In order to capture the inexact nature of our surrounding world, Lotfi A. Zadeh introduced the notion of fuzzy sets and extended the two-valued $\{0,1\}$ logic to the interval [0,1], allowing a gradual transition from falsehood to truth [Zad65, Zad96]. Fuzzy sets also allow the representation of imprecise, subjective knowledge and linguistic information. Elements are not seen as being either part of a set or not but instead they are defined as being similar to elements *described* by a set. The similarity is quantified by assigning a value between 0 (dissimilar) to 1 (equal). A fuzzy set, defined over a universe of discourse U, is characterized by its *membership function* $FS : U \rightarrow [0, 1]$. The membership function defines the similarity of an item to the fuzzy set. The universe of discourse U contains all elements that could possibly be part of the set, e.g. a set describing "high concentrations" may be defined over $[0, \infty]$ (all possible concentrations). For an extensive introduction to fuzzy logic see [Men95, Lee90a, Lee90b].

As different fuzzy sets may describe elements of the same (biological) concept, for example the concept "concentration of a protein P", we subsume fuzzy sets to *fuzzy concepts*, which correspond to the real-world concepts. Fuzzy concepts are defined as tuples $< FS_1, ..., FS_n >$, where all fuzzy sets FS_i are defined over the same universe of discourse. The fuzzy sets combined to a fuzzy concept usually have differently shaped membership functions as they describe different aspects of the underlying (biological) concept. An exemplary fuzzy concept *concentration* may include fuzzy sets *low*, *medium*, *high* and *saturated*, each describing a different "level" of concentration (figure 2).

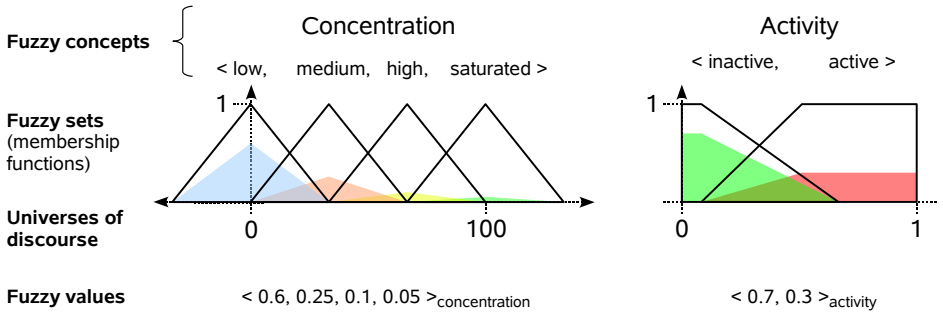


Figure 2: The current state of an entity with respect to a (biological) concept (e.g. its current concentration) can be described by the membership function values of fuzzy sets, which belong to an according fuzzy concept. A *fuzzy value* is a tuple $< w_1, ..., w_n >_{FC}$ specified with respect to a fuzzy concept $FC =< FS_1, ..., FS_n >$. The membership function values $w_i \in [0, 1]$ are called *weights* and describe the current state of an entity with respect to the fuzzy concept. Colored areas visualize the currently assigned weights.

2.1 How Fuzzy Values Represent Concentrations and Other Properties

Concentrations or amounts of proteins, RNA, metabolites, etc. are typically represented as positive real numbers. Such real numbers can in turn be represented as fuzzy values with respect to a fuzzy concept. To create suitable fuzzy concepts, several modeling decisions have to be made:

1. **Define the universe of discourse.** Basically, the whole set of real numbers could be used but it is also possible to define an explicit range, for example when the concentration of an entity is bounded by some equilibrium constraints.
2. **Define the number of fuzzy sets.** A higher number of fuzzy sets allows more detailed representations of states and the associated dynamics. On the other hand the size of rule tables in fuzzy logic systems increases, allowing a higher number of different outcomes.
3. **Define the shape and position of membership functions.** Arbitrary membership function shapes can be defined although symmetric triangular, trapezoidal or gaussian shapes should suffice for most applications. Position, shape and spread of fuzzy sets can be freely defined according to modeling requirements.

It is part of the modeling decision to utilize the same fuzzy concept for only one, some or all entities of a system. If the concentration of an entity is known it can be transferred (fuzzyfied) to a fuzzy value simply by computing the according membership function values for each fuzzy set. If it is not known but only some rough guesses are available, weights can be assigned directly to fuzzy sets. For example, if the concentration of an entity is only known to be “quite low” a suitable fuzzy value may look like $< 0.8, 0.2, 0.0, 0.0 >_{concentration}$.

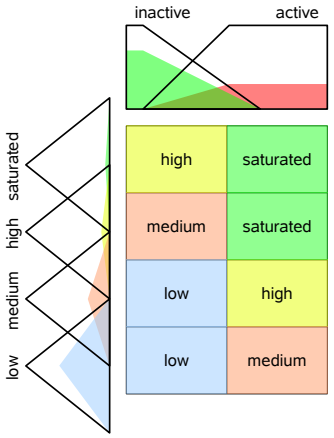
2.2 How Fuzzy Logic Systems Replace Differential Equations

Dynamic processes within a system are induced and influenced by the current state of the system and its entities and in turn influence and change them. If the current states of entities are defined by fuzzy values, processes have to be modeled by functions that operate on weighted fuzzy sets. These functions can be defined using natural language terms and without use of mathematical formulas.

A fuzzy logic system (FLS) consist of a set of rules mapping (weighted) fuzzy sets of several places (premises) to a set of output fuzzy sets (conclusions), thereby defining new weights for them. Fuzzy logic systems are specified as sets of natural language based rules. Single rules are defined as IF-THEN sentences, where several fuzzy sets, connected by AND-operators, in the IF-clause (premises) are mapped to a single concluding fuzzy set (conclusion).
Fuzzy logic theory offers several set theoretic operations to evaluate a fuzzy logic system. We decided for the frequently used and very intuitive sum-product logic ([Men95]):

1. **Inference of the weight of single conclusions depending on their premises.** Weights of premises are multiplied to infer the weight of a conclusion (product-inference).
2. **Combination of those conclusions referring to the same property.** Weights of conclusions with identical fuzzy sets are summed (sum-composition).

Generally (and intuitively) it holds that the higher the confidence of the premises (the higher they are weighted), the more confident is the conclusion (the higher it is weighted) (figure 3).



Rule set:

1) IF concentration.low	AND activity.inactive	THEN concentration.low
2) IF concentration.medium	AND activity.inactive	THEN concentration.low
3) IF concentration.high	AND activity.inactive	THEN concentration.medium
4) IF concentration.saturated	AND activity.inactive	THEN concentration.high
5) IF concentration.low	AND activity.active	THEN concentration.medium
6) IF concentration.medium	AND activity.active	THEN concentration.high
7) IF concentration.high	AND activity.active	THEN concentration.saturated
8) IF concentration.saturated	AND activity.active	THEN concentration.saturated

Sum-Product Logic (e.g. rules 3 and 5):

0.1 (*high*) x 0.7 (*inactive*) = 0.07 (*medium*)
0.6 (*low*) x 0.3 (*active*) = 0.18 (*medium*)

0.25 (*medium*)

Resulting Fuzzy Value:

$< 0.595, 0.250, 0.110, 0.045 >_{concentration}$

Figure 3: An example of a fuzzy logic system which uses the fuzzy sets and values described in figure 2 as premises to calculate weights of a fuzzy value of type *concentration* (conclusion). The rule set can be represented as a table (left). The premises (top and left, with membership functions and visualized weights) are mapped to the conclusions (center, without visualized membership functions or weights).

3 On the Use of Fuzzy Values and Fuzzy Logic Systems in Petri Nets

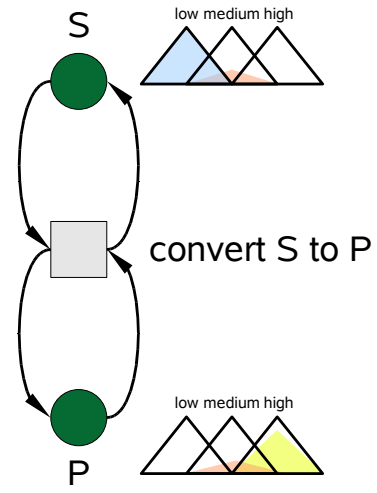
Fuzzy values are used to describe the current state of an entity with respect to a fuzzy concept. An arbitrary number of different fuzzy values can be used to describe each entity. All fuzzy values describing a single entity form a set of *fuzzy tokens* on the respective place in the Petri net model of the system. The set of fuzzy tokens represents those properties (concepts) an entity could *possibly* exhibit while the current weight assignment reflects the properties an entity *currently* exhibits. Fuzzy logic systems define the dynamics of a system. One or several of them serve as inscriptions of arcs. Whenever a transition fires, fuzzy tokens of adjacent places are consumed and a new set of fuzzy tokens is created by the FLS's of incident arcs. We distinguish three types of arcs which correspond to input, output and test arcs as defined for hybrid functional Petri Nets ([MTA$^+$03]). Input and output arcs consume and produce tokens whenever the incident transition fires while test arcs do not affect tokens. Test arcs symbolize a functional dependency of processes and entities, they allow the usage of fuzzy tokens of incident places as premises of fuzzy logic systems without consuming them.

4 Results and Conclusion

The adaption of fuzzy sets for representing states and properties and fuzzy logic based reasoning for describing processes can be used to model biological systems. Fuzzy sets capture the typically inexact, qualitative knowledge about biological entities and are well suited to represent limited knowledge, inexact measurements as well as error prone data. Due to the fact that they can stand for arbitrary properties, it is possible to uniformly represent all types of external and internal factors influencing a system. Fuzzy sets can be designed freely by a user according to his needs. Fuzzy logic systems allow the formulation of biological processes using simple yet powerful rule systems, which can be formulated using natural language. Therefore, hypotheses concerning the behavior of entities or influences between entities can be translated directly into executable systems (application 1, figures 4 and 5). The representation using Petri nets clearly visualizes entities, processes and dependencies within a biological system. A Petri net and fuzzy logic based system can easily be outlined in a pen-and-paper style by creating drafts of entities and their dependencies and describing the desired properties and effects of dependencies and influences in natural language.

The extension of fuzzy sets, fuzzy concepts and fuzzy values to represent arbitrary (non-quantifiable) properties or states of entities is straightforward. In fact, no changes of the definitions of these terms are necessary. Properties which are not per se quantifiable, like the current state of a cell in the cellcycle, may be described similar to concentrations using several fuzzy sets. Such fuzzy sets, for example belonging to the fuzzy concept *cellcycle state*, are then weighted to define the current state of an entity and represented as a fuzzy value. Although the described entity (the "cell cyle state") has no inherent reference to a real value, the universe of discourse of these fuzzy sets can still be defined as arbitrary range within the set of real numbers for the sake of uniformity. Modeling the state of a

Application 1: Minimal model of a Higgins-Sel'kov oscillator



Rule set defining concentration of *P*:

IF S.low		THEN P.low
IF !S.low	AND P.low	THEN P.medium
IF !S.low	AND !P.low	THEN P.high

Rule set defining concentration of *S*:

IF	P.high	THEN S.low
IF S.low	AND !P.high	THEN S.medium
IF !S.low	AND !P.high	THEN S.high

ODE model:

$dS / dt = v_0 - k_1 S P^2$

$dP / dt = k_1 S P^2 - k_2 P$

where $v_0 = 1$, $k_1 = 1$, $k_2 = 1.00001$,
$S(0) = 2$ and $P(0) = 1$.

Figure 4: A minimal model of an oscillator similar to the Higgins-Sel'kov Oscillator (ODE model taken from [KHK+05]). The underlying process and the ODE model (figure 5) can be described by few sentences: (1) S increases P; (2) P increases P strongly; (3) If P reaches a high level, S decreases strongly; (4) If S reaches a low level, P decreases strongly; and directly converted to a set of rules. The stated six rules suffice to create an oscillating behavior qualitatively similar to the ODE model. If a fuzzy set is negated, its current weight w is replaced by $(1 - w)$ during the execution of a FLS.

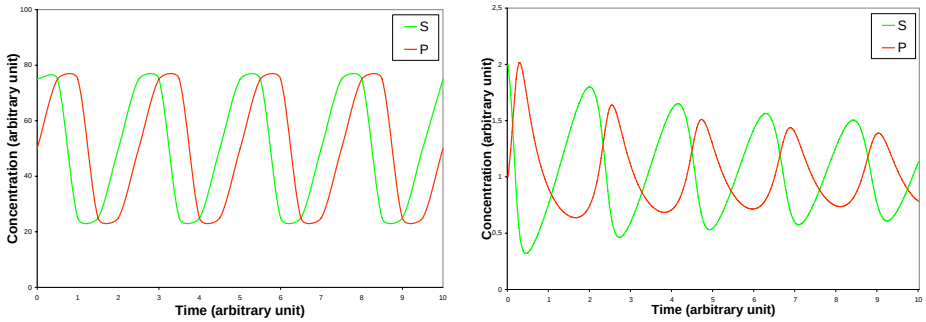


Figure 5: Time courses of the minimal fuzzy logic based model (left) and the ODE based model (right) of the Higgins-Sel'kov Oscillator described in figure 4. The PNFL model qualitatively reflects the oscillating behavior of S and P. A more involved PNFL model with extended fuzzy logic systems and two additional transitions modeling input of S and output of P is able to reproduce the dampening of oscillations as observed in the ODE model (data not shown).

Application 2: Hierarchical modeling of oscillating behavior

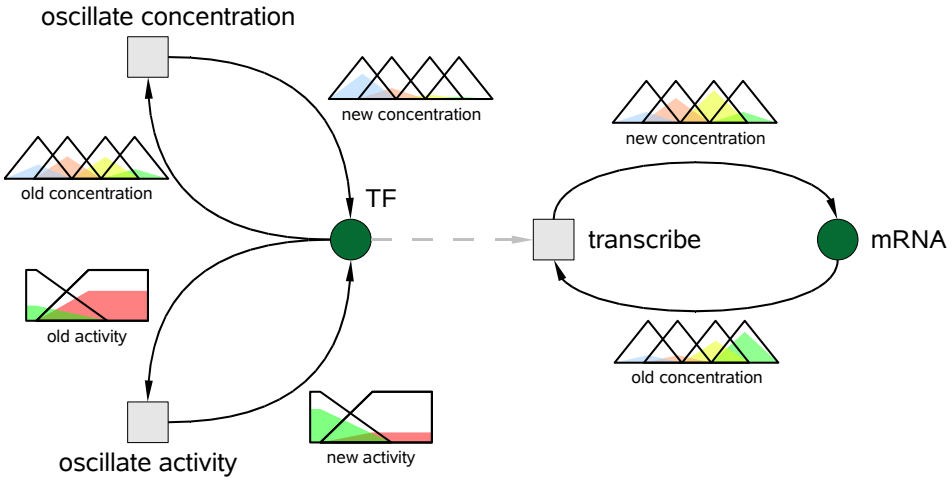


Figure 6: The concentration and activity of a transcription factor *TF* are conrolled by two transitions and exhibit an oscillating behavior. The underlying biological processes are not explicitly modeled but are described using appropriate rule systems to reduce the size of the model. Concentration and activity in turn influence the current concentration of mRNA molecules via the fuzzy logic system described in figure 3.

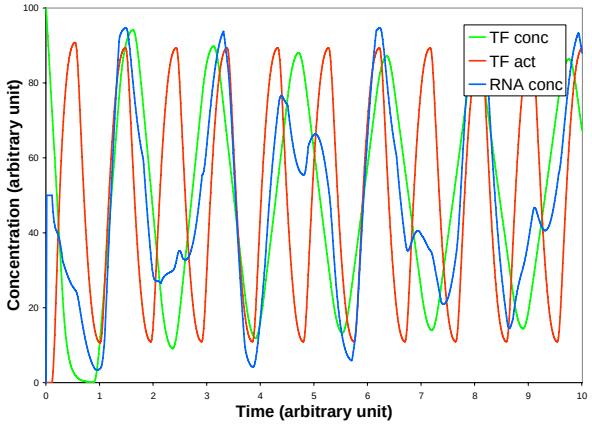


Figure 7: A time course of the dynamical behavior of the system described in figure 6. The oscillations of *TF*s concentration and activity differ in frequency and induce a quite irregular behavior of mRNA concentration.

system and its properties by the same framework as used for quantifiable entities is one of the main advantages of our fuzzy logic based approach. A uniform representation of quantifiable entities and other, more abstract properties is possible while dynamical changes of those parts of a system can be performed using the same technique, namely fuzzy logic systems. Their rule-based description allows modeling of complex behavior and is more powerful than a simple description of dependencies as *activating* or *inhibiting*, as it is common in boolean networks.
It is possible to model the behavior of entities by an explicit formulation of the underlying biological processes, for example an oscillation of a protein level by modeling a negative feedback loop delayed by transport via the core membrane. At the other hand one could force entities to behave in a particular way by *defining* their behavior with appropriate rules and without explicitly modeling real biological processes. This is for example very useful when a certain behavior of entities can be observered experimentally but not yet explained adequately by a model, while at the same time the modeling of the observed behavior is crucial as it affects other parts of the system. Additionally, replacing the extensive elaboration of biological processes by simpler systems mimicking their behavior also allows a hierarchical modeling (application 2, figures 6 and 7).

The described approach (PNFL) is currently improved and extended, including a GUI suited for model building, defining fuzzy sets, formulation of FLS rule sets and visualizing simulation runs and results. The implementation will also support concurrent simulations of biological systems in several cells. A prototype system was successfully applied during different developmental stages to several small test systems, like an in-silico network ([ZDGS01, ZGSD03]), typical network motifs (e.g. feed-forward loops, switches) and several oscillator models (Higgins-Sel'kov, minimal mitotic, coupled oscillators; [KHK+05]). As a larger application a model of the EGF signal transduction pathway as defined in [LZLP06] was evaluated by replacing mass action kinetics by fuzzy logic systems.

References

[CCNG+00] K. C. Chen, A. Csikasz-Nagy, B. Gyorffy, J. Val, B. Novak, and J. J. Tyson. Kinetic Analysis of a Molecular Model of the Budding Yeast Cell Cycle. *Mol. Biol. Cell*, 11(1):369–391, January 2000.

[Cha07] C. Chaouiya. Petri net modelling of biological networks. *Brief Bioinform*, 8(4):210–219, July 2007.

[GFG+06] D. Gilbert, H. Fuß, X. Gu, R. Orton, S. Robinson, V. Vyshemirsky, M. J. Kurth, C. S. Downes, and W. Dubitzky. Computational methodologies for modelling, analysis and simulation of signalling networks. *Brief Bioinform*, 7(4):339–353, November 2006.

[GKV01] 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 (STTT)*, 3(4):394–404, 2001.

[KBG07] A. Kremling, K. Bettenbrock, and E. D. Gilles. Analysis of global control of Escherichia coli carbohydrate uptake. *BMC systems biology*, 1(42), 2007.

[KHK+05] E. Klipp, R. Herwig, A. Kowald, C. Wierling, and H. Lehrach. *Systems Biology in Practice. Concepts, Implementation and Application.* WILEY-VCH, 2005.

[KK04] B. Kofahl and E. Klipp. Modelling the dynamics of the yeast pheromone pathway. *Yeast*, 21(10):831–850, July 2004.

[Lee90a] C. C. Lee. Fuzzy logic in control systems: fuzzy logic controller. I. *Systems, Man and Cybernetics, IEEE Transactions on*, 20(2):404–418, 1990.

[Lee90b] C. C. Lee. Fuzzy logic in control systems: fuzzy logic controller. II. *Systems, Man and Cybernetics, IEEE Transactions on*, 20(2):419–435, 1990.

[LGN+07] C. Li, Q. W. Ge, 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, January 2007.

[LZLP06] D. Y. Lee, R. Zimmer, S. Y. Lee, and S. Park. Colored Petri net modeling and simulation of signal transduction pathways. *Metabolic Engineering*, 8(2):112–122, March 2006.

[Men95] J. M. Mendel. Fuzzy logic systems for engineering: a tutorial. *Proceedings of the IEEE*, 83(3):345–377, 1995.

[MPLD04] J. Mandel, N. M. Palfreyman, J. A. Lopez, and W. Dubitzky. Representing bioinformatics causality. *Brief Bioinform*, 5(3):270–283, January 2004.

[MTA+03] H. Matsuno, Y. Tanaka, H. Aoshima, A. Doi, M. Matsui, and S. Miyano. Biopathways representation and simulation on hybrid functional Petri net. *In silico biology*, 3(3):389–404, 2003.

[Mur89] T. Murata. Petri nets: Properties, analysis and applications. *Proceedings of the IEEE*, 77(4):541–580, 1989.

[OSV+05] R. J. Orton, O. E. Sturm, V. Vyshemirsky, M. Calder, D. R. Gilbert, and W. Kolch. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *Biochem J*, 392(Pt 2):249–261, December 2005.

[RLM96] V. N. Reddy, M. N. Liebman, and M. L. Mavrovouniotis. Qualitative analysis of biochemical reaction systems. *Computers in biology and medicine*, 26(1):9–24, January 1996.

[SEJGM02] B. Schoeberl, C. Eichler-Jonsson, E. D. Gilles, and G. Müller. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. *Nat Biotechnol*, 20(4):370–375, April 2002.

[Zad65] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[Zad96] L. A. Zadeh. Fuzzy logic = Computing with words. *IEEE Transactions on Fuzzy Systems*, 4(2):103–111, 1996.

[ZDGS01] D. E. Zak, F. J. Doyle, G. E. Gonye, and J. S. Schwaber. Simulation studies for the identification of genetic networks from cDNA array and regulatory activity data. *Proc. 2nd Intl. Conf. Systems Biology*, pages 231–238, 2001.

[ZGSD03] D. E. Zak, G. E. Gonye, J. S. Schwaber, and F. J. Doyle. Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: insights from an identifiability analysis of an in silico network. *Genome Res*, 13(11):2396–2405, November 2003.