THESIS REPORT

Master's Degree

Modeling Biological Pathways:
A Discrete Event Systems Approach

by V. N. Reddy
Advisor: M.L. Mavrovouniotis

M.S. 94-4

ISR
INSTITUTE FOR SYSTEMS RESEARCH

Sponsored by
the National Science Foundation
Engineering Research Center Program,
the University of Maryland,
Harvard University,
and Industry

ABSTRACT

Title of Thesis: Modeling Biological Pathways: A discrete event systems approach

Name of degree candidate: Venkatramana N. Reddy

Degree and Year: Master of Science in Chemical Engineering, 1994

Thesis directed by: Dr. Michael L. Mavrovouniotis
Associate Professor
Department of Chemical Engineering

A discrete-event systems approach is proposed for the modeling of biochemical reaction systems. The approach is based on Petri nets, which are particularly suited to modeling stoichiometric transformations, i.e., the inter conversion of metabolites in fixed proportions. Properties of Petri nets and methods for their analysis are presented, along with their interpretation for biological systems. An example of the human erythrocyte metabolism is presented to illustrate the concepts of the methodology

MODELING BIOLOGICAL PATHWAYS: A DISCRETE EVENT SYSTEMS APPROACH

by

Venkatramana N. Reddy

Thesis submitted to the Faculty of the Graduate School
of The University of Maryland in partial fulfillment
of the requirements for the degree of
Master of Science
1994

Advisory Committee:

Associate Professor Michael L. Mavrovouniotis, Chairman/Advisor
Professor Evanghelos Zafiriou
Assistant Professor William E. Bentley

DEDICATION

To my parents

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to my advisor Dr. Michael Mavrovouniotis for his guidance and encouragement in my graduate study. I would like to thank Dr. Evanghelos Zafiriou and Dr. William Bentley for serving on my thesis committee. Their comments on the thesis are much appreciated. I am also grateful to Dr. Michael Liebman of Amoco Research Center for the very informative discussions I had with him. I would like to acknowledge the support and encouragement that I received from Dr. David Elliott in the initial stages of my research into the subject.

I would like to thank the Institute for Systems Research, University of Maryland, for providing me with the financial support and the remarkable facilities.

I would also like to thank my family for their support and my friends who have made my stay at Maryland an enjoyable and rewarding experience.

This work was supported in part by the National Library of Medicine (grant R29 LM05278).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The modeling, simulation, and analysis of biological pathways require the integration of a large volume of diverse data from the biological, chemical and physical sciences, if one attempts a complete and quantitative analysis. The amount of information necessary for a complete analysis of a biochemical system can be summarized by Figure 1.1. Usually, the problem of modeling a complex biochemical system involves data that are incomplete and unreliable.

With a traditional quantitative model, such as a system of differential equations of the form $dx/dt = f(x)$, which describes the change of a vector (x) of metabolite concentrations based on enzyme kinetics, one cannot tolerate even a single unknown parameter in the function $f(x)$: The model cannot

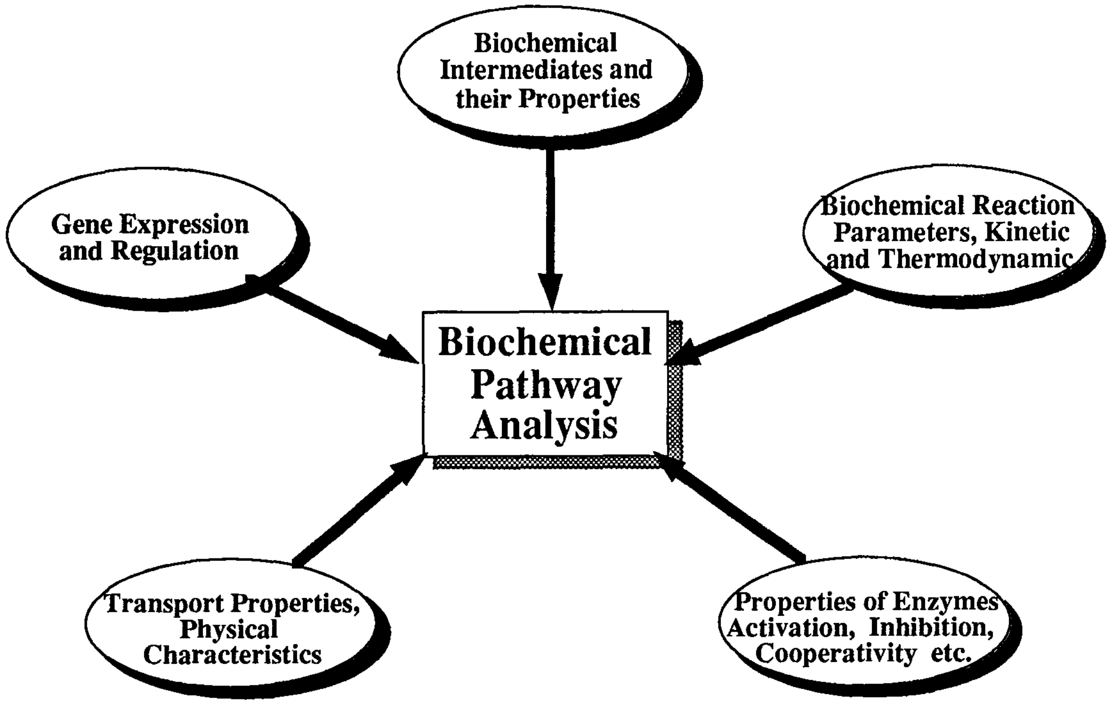


Figure 1.1: Quantitative information for the analysis of biochemical pathways

be simulated to obtain any results, unless all the parameters appearing in the kinetic expressions (in the right hand differential equations, $\mathbf{f(x)}$) are known.

In other cases, many parameters are not constant; thus, in order to draw a general conclusion about the behavior of a pathway, one would have to simulate many combinations of values for these parameters. Suppose, for example, that the concentrations of the enzymes participating in a pathway are known to vary within certain ranges, and we are interested in the general characteristics of the pathway within these ranges. In simulating the quantitative model, a large number of combinations of values for the enzyme concentrations might be tested to establish the general pattern.

While various techniques (such as sensitivity analysis) exist to facilitate work with quantitative models in the face of uncertain parameters, one way to address these problems is to begin with a practical qualitative model of the system and avoid the quantitative analysis - in fact, avoid the construction of a complete quantitative model altogether.

A qualitative analysis allows us to draw preliminary conclusions about the biological pathway such as the influence of particular reactions, metabolites or pathway segments on the overall system behavior. In drawing any *a priori* conclusions about the behavior of a pathway, a qualitative method that excludes conducting detailed simulations and analysis of the system is essentially independent of the detailed parametric information (i.e., rate constants, cooperativity indices, etc.) of the pathway.

A new approach to the qualitative modeling of a pathway is presented. The method incorporates the use of a discrete event methodology for the representation and analysis of bioreaction networks. The reactions and other biological processes are modeled as discrete events and analyzed by applying Petri net modeling and analysis techniques to this representation of the pathway.

In the following chapter we discuss the subject of biochemical reaction systems, the methods for analysis and some qualitative properties that are of interest to researchers. Chapter 3 contains an almost comprehensive introduction to the theory and properties of Petri Nets and can be skipped by a reader with a prior knowledge of the methodology.

Chapter 4 demonstrates the application of the discrete event modeling strategy to metabolic pathways; the oxidative and energy metabolism of the human erythrocyte cell is presented as an example. We conclude with the discussion of certain aspects of Petri nets which are of importance and future directions in this field. Appendix A contains the complete working to the example in Chapter 4 and Appendix B is a collection of information sources for current trends in Petri net theory, applications and software tools available.

Chapter 2

Biochemical Reaction Systems

The field of Chemical Engineering has within its domain different areas of research; one such area is in the biological sciences. Biochemical Engineering is recognized as a major topic of research as well as industrial innovation. The analysis of reaction systems pertaining to the biological environment and the specific pathways of the organism is of importance. Therefore the development of analytical methods for modeling these reaction systems is essential. Traditional approaches to the analysis of chemical reaction systems have been supplemented with newer methods that deal with properties specific to biochemistry. The applications of biochemical engineering and some of the analytical methods are outlined here.

2.1 Biological and Biochemical Applications

2.1.1 Chemical Industry

The food, pharmaceutical and biochemical industries are the main consumers and suppliers of products derived from biological organisms [1]. Fermentation, production of dairy products (cheese, yogurt, etc.), sugar substitutes are all biochemical processes. The production of proteins, hormones and steroids are essential in the pharmaceutical industry. All these processes require an in-depth study of the underlying biochemical reaction system for the design of efficient bioreactors and process units.

2.1.2 Biotechnology

The enzymes and pathways from plant and animal tissue cultures are used for the biotransformation of compounds that have potential uses in the medical field. The cultures also serve as a source to produce metabolites in the pharmaceutical industry, natural dyes, flavoring and fragrances. To improve the efficiency of these processes we would need to investigate the pathways that result in the biotransformation. The identification of essential metabolites, steady-states and pathway dynamics fall within the domain of pathway analysis.

Another method of increasing the productivity of a bioprocess is by genetically engineering the organism to achieve a higher yield. A knowledge of the regulation and control mechanism of the biological pathway is necessary to identify potential sites for modification. Genetically engineered microbial cultures are used in the efficient manufacture of antibiotics such as penicillin and streptomycin.

2.1.3 Medical Applications

The study of metabolic pathways plays an important role in drug research because it is necessary to first understand the reaction pathways that are affected by the drug, directly and indirectly. It is important to know the effect of modification of reaction steps and depletion of metabolites on the overall reaction network.

Phenylketonuria (PKU), a disease that leads to mental retardation, occurs due to the deficiency of the enzyme that converts phenylalanine to tyrosine. The identification of sites in a metabolic pathway that result in such diseases, caused by inborn errors in metabolism, would be useful so that the relevant enzyme or metabolite can be substituted or suitably modified. This is particularly useful in the pharmaceutical industry for designing site-directed drugs to target mutant enzymes.

2.2 Methods of Analysis

2.2.1 Qualitative versus Quantitative

The traditional method of analysis of complex systems was to model the system with a set of differential equations that described (or mirrored) the behavior of the system within a range of variables and parameters. The model thus developed requires the use of a vast amount of data and often the data is insufficient, inconsistent or unreliable for an accurate analysis. A quantitative model is usually very complex incorporating many parameters and variables that calls for an enormous computational effort to completely solve the model. The model remains very specific to the system under study, thus the analysis of another problem with similar attributes warrants a new model; a quantitative model cannot support generality, only precision. Dynamic models described by a set of differential equations cannot accurately capture the non-linearity prevalent in most biochemical systems. Non-linear kinetics is modeled by suitable approximations, for instance, linearization at a steady state that results in the model being valid only near the steady state.

Qualitative models, on the other hand, do not emphasize on the prediction of the system *per se*, but on understanding the behavior of the system. Qualitative models allow us to draw preliminary conclusions about the behavior of the system without the need for detailed simulation. At times the qualitative model is sufficient for prediction of the system or provides an overview for a more systematic application of quantitative methods. In a qualitative model the variables are limited and not as complex, therefore do not require extensive computational analysis. Qualitative models are usually simple in construction, easy to understand and can be modified according to the analytical needs of the modeler.

2.2.2 Dynamic simulation

Most dynamic simulation techniques used for biochemical reaction systems are modeled, like most engineering systems, with a set of ordinary differential equations that describe the system performance. These quantitative models aim to predict precise values of the variables of interest based upon the parameters and extraneous constants introduced by the modeler. The model attempts to simulate the system closely to coincide with laboratory values within reasonable experimental error.

Biochemical reactions can be modeled by the single substrate single enzyme Michaelis-Menten kinetics and work sufficiently well for elementary reactions. Many biochemical reactions involve more than one substrate, sometimes multiple enzymes and in addition are under the influence of activators (or inhibitors). This makes the modeling of the system very complex and computationally impractical. In an attempt to reduce the complexity of the systems, compartmentalization by way of structured cell models (Cornell model, Palsson model) has been widely used.

2.2.3 Biochemical Systems Theory

Biochemical Systems Theory [27, 34] developed by Michael Savageau in the early '70s has found wide application in the modeling, regulation and control of biochemical reactions. The method developed is capable of handling nonlinear systems and regulatory behavior of enzymes. A power law formulation is used to express the system in terms of the concentrations and their kinetic orders. The equations are linear in logarithmic space and can be solved to yield the steady states. Since the kinetic orders are descriptive of the apparent rate information, the parameters are subject to interpretation of experimental data introducing uncertainty to the model. The method utilizes the rate information and therefore can be used to analyze the dynamic response of the system.

2.2.4 Graph Theoretic Methods

The graphical models explored by Kohn and Letzkus [18] for modeling metabolic networks consisted of Kauffman binary networks [15], Signal flow graphs [26] and Bond graphs [31]. Derived from all of the above graph models and later modified by Kohn and Lemieux [17] a new model called MetaNets was introduced. The method is based on maintaining the biochemical reaction structure that is visually intuitive in a graphical model consisting of functional nodes and interconnecting arcs. The nodes consist of "chemnodes" (metabolite, enzyme, nucleic acid) and "relnodes" (substrate, activator, inhibitor, spontaneous) to represent the relation between the chemnodes. The arcs serve to transfer the information flow. The model does not possess the capability to capture certain properties of biochemical systems such as transport processes, thereby losing generality in its representation.

The method serves to identify the potential feedback sites and controlling enzymes in the network without the necessity for quantitative information. The actual sites of regulation can be decided with the aid of quantitative information analyzed in terms of signal strengths passing though the feedback path of the model. This again requires kinetic information such as concentrations, equilibrium constants, cooperativity indices, etc.

2.2.5 Metabolic Control Theory

Metabolic control theory is a method of relating properties, such as flux or concentration of metabolites, to certain parameters of the network, such as concentration of enzymes, to study the effects of perturbations in these parameters on the overall system behavior [7]. The theory developed independently by Kacser & Burns [14] and Heinrich & Rapaport [9] has since been modified and applied by others [6, 24, 29].

The methodology involves the calculation of control coefficients that reflect the sensitivity of the reaction system to perturbations in the parameters. The theory can be used to analyze the regulatory properties of the biochemical network like Biochemical Systems Theory, but cannot be used in the analysis of the stability or dynamics unlike Biochemical Systems Theory [3,27]. Although structural analysis of the systems by this method is possible [24], the calculation of the control coefficients and elasticity coefficients requires quantitative enzymatic data.

2.3 Qualitative and Structural Questions

In an analysis of a biochemical system it would be useful to determine certain properties of the network without having to build a complete quantitative model of the system. A dynamic response of the pathway to changes in relevant parameters would be extremely useful in designing bioreactors and in predicting responses of the systems under controlled conditions. However, building such models is a painstaking effort and often too complex without incorporating simplified assumptions. The simulation itself would be computationally intensive notwithstanding the completeness and reliability of the experimental data used for estimating parameters in the model.

If on the other hand a qualitative analysis that is simpler, generalized and capable of analysis without quantitative data is adopted, it would aid in a preliminary analysis of the system avoiding the complexity of a complete quantitative model.

A complex quantitative model is not practical if instead a qualitative model is capable of providing descriptions of the system at a lesser cost. A qualitative model can identify the key compounds (metabolites, enzymes, activators, etc.) necessary for the biotransformation; existence of unbounded metabolites capable of accumulation and those that exist in invariant concentrations; behavior that results in cyclic transformations (futile cycles); effect on overall behavior of the pathway due to a modification on a part of the pathway and; identification of classes of behaviors based on the network structure of the system; to name a few. These properties would serve as effective markers for an in-depth investigation into the dynamics of the system and in some cases would answer most questions that may not require a quantitative analysis at the outset.

Chapter 3

Petri Nets

Petri nets are a mathematical and computational tool for the modeling and analysis of discrete event systems. Petri nets offer a formal way to represent the structure of a discrete-event system, simulate its behavior, and draw certain types of general conclusions on the properties of the system. The methodology has wide applications in a number of varied fields such as Computer Science [36], Control Engineering [11, 35], Manufacturing Systems [19] and Information Science [2, 4].

The essential concepts in Petri net theory are outlined in this section. Further details on the theory and applications of Petri nets are available elsewhere in literature [20, 21, 25].

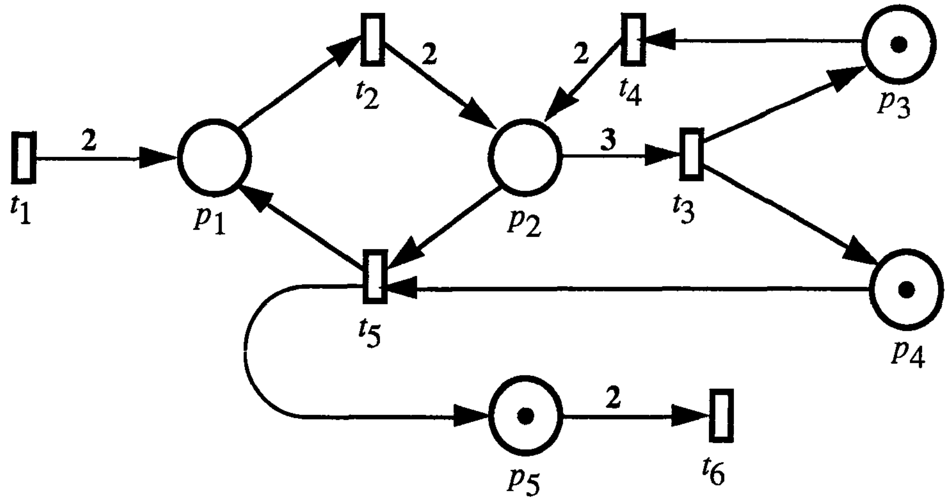


Figure 3.1: A Petri net graph with places, transitions and arcs

3.1 Definitions

A Petri net is a graph (Figure 3.1) formed by two kinds of nodes, called places and transitions. Directed edges, called arcs, connect places to transitions, and transitions to places. For the sake of convenience, the presence of multiple arcs between a single place and a single transition is represented by a single arc with an arc-weight: We associate a non-zero positive integer equal to the number of implied connecting arcs with this one weighted arc.

A positive integer number of tokens may be assigned to each place; these numbers of tokens form the state of the Petri net (which will be defined as a marking, below). The state of the net changes to another (after each event) as the discrete event system operates.

Pictorially, places are represented by circles, transitions by boxes, arcs by lines ending in an arrow, and tokens as black dots placed in the circles. Generally if there is no arc-weight explicitly specified on the graph it is

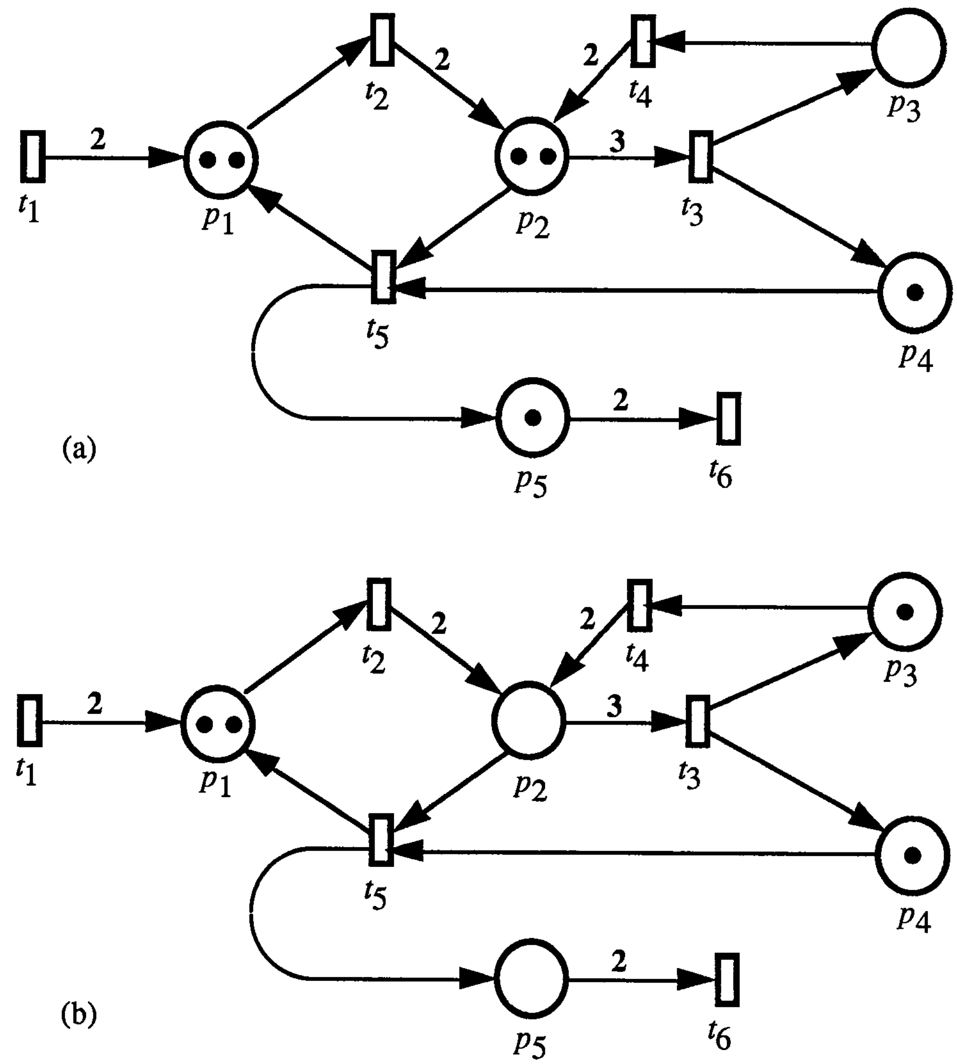


Figure 3.2: (a) Marking after firing enabled transitions t_1 and t_4 from Figure 3.1 (b) Marking after firing in sequence t_2, t_3, t_5, and t_6 from Figure 3.2(a)

assumed to be equal to one.

3.2 Execution

Each transition is associated with a finite number of input places and output places. From the perspective of modeling a discrete event system, it is necessary to satisfy a set of pre-conditions (defined by the input places) before an event (transition) may occur; the event results in a set of post conditions (output places).

In a Petri net a transition is enabled when the number of tokens in its input places is greater than or equal to the weights on the arcs connecting the places to the transition. A transition with no input places, called a source transition, is always enabled. In Figure 3.1, the transitions t_1 and t_4 are enabled, while the rest are not.

An enabled transition can fire, consuming tokens from its input places and depositing tokens in its output places; the numbers of tokens consumed and produced are determined by the arc-weights. The firing of transitions can be understood as the movement of tokens, from one place to another, through the transitions. A transition with no output places, called a sink transition, can fire when enabled consuming the tokens from its input places.

Figure 3.2 shows the same Petri net graph from Figure 3.1 after the firing of several transitions. The firing of one enabled transition may deposit tokens in the input places of another transition – thus enabling that transition to fire in turn. In Figure 3.1 t_1 and t_4 are enabled, hence one possible firing sequence could begin with transition t_1 firing and depositing two tokens in place p_1; then t_4 firing, consuming one token from p_3 and depositing two tokens in p_2 (Figure 3.2(a)). Similarly, the firing sequence t_2, t_3, t_5 and t_6, starting from a marking of Figure 3.2(a) will result in the marking of Figure 3.2(b).

3.3 Marking

The state of a Petri net is determined by the number of tokens present in each place of the net. The marking **M** of a Petri net is a vector whose elements correspond to the number of tokens present at each place of the Petri net. Thus, the size of the vector **M** is equal to the number of places in that net. The execution of the Petri net changes the marking by decreasing tokens in certain places and increasing them in other places. The initial state of a Petri net before execution is called the <u>initial marking</u> $\mathbf{M}_0$.

3.4 Mathematical Definition

A Petri net is represented as PN = $(\mathbf{P}, \mathbf{T}, \mathbf{E}, \mathbf{W}, \mathbf{M}_0)$, where:

$\mathbf{P} = \{ p_1, p_2, p_3, \mathrm{K}\ \ p_m \}$ is a finite set of places
$\mathbf{T} = \{ t_1, t_2, t_3, \mathrm{K}\ \ t_n \}$ is a finite set of transitions
$\mathbf{E} \subseteq (\mathbf{P} \times \mathbf{T}) \cup (\mathbf{T} \times \mathbf{P})$ is a set of arcs
$\mathbf{W} : \mathbf{E} \rightarrow \{ 1, 2, 3, \mathrm{K}\ \}$ is a weight function
$\mathbf{M}_0 : \mathbf{P} \rightarrow \{ 0, 1, 2, \mathrm{K}\ \}$ is the initial marking
P and **T** being disjoint sets

Since the Petri net is presented as a model for a discrete event system, it is helpful to have a system of equations that can be used to specify and manipulate the state of the system.

For a Petri net with m places and n transitions, we can formulate [23] a state equation of the type

$$\mathbf{M}_k = \mathbf{M}_{k-1} + \mathbf{A}^T \mathbf{u}_k \ , \quad k = 1, 2, 3, \mathrm{K} \qquad (3.1)$$

The index k represents a point in a firing sequence. For each k, $\mathbf{M}_k$ represents an m $\times$ 1 vector, the marking after the kth firing; $\mathbf{u}_k$ an n $\times$ 1 vector, the control vector indicating the transition fired at the kth firing; and $\mathbf{A}$ an n $\times$ m matrix, the incidence matrix whose elements a_{ij} denote the change in the number of tokens in place j due to the firing of transition i. The control vector $\mathbf{u}_k$ is simply the unit vector, containing the entry of 1 in the position corresponding to the transition that fired, and 0 everywhere else. The matrix $\mathbf{A}$ describes the weights on the arcs, with an entry of $a_{ij} = 0$ describing the absence of an arc altogether between transition i and place j.

If a particular marking $\mathbf{M}_n$ is reached from the initial marking $\mathbf{M}_0$, through a firing sequence $\sigma = \{ \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \dots \mathbf{u}_n \}$, and the state-equations are summed for all the firings in this σ, we obtain

$$\mathbf{M}_n = \mathbf{M}_0 + \mathbf{A}^T \sum_{k=1}^{n} \mathbf{u}_k \tag{3.2}$$

We define $\mathbf{x} = \sum_{k=1}^{n} \mathbf{u}_k$, as an n $\times$ 1 vector, called the firing count vector. The element i in $\mathbf{x}$ indicates the number of times transition i must fire to transform $\mathbf{M}_0$ to $\mathbf{M}_n$. Substituting the definition of $\mathbf{x}$ in Equation 3.2, we obtain

$$\mathbf{M}_n - \mathbf{M}_0 = \mathbf{A}^T \mathbf{x} \tag{3.3}$$

or

$$\mathbf{A}^T \mathbf{x} = \Delta \mathbf{M} \tag{3.4}$$

For example, from the Petri net graph of Figure 3.1 we obtain the incidence matrix **A** and the initial marking $\mathbf{M}_0$ as follows

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 0 & 0 \\ 0 & -3 & 1 & 1 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{bmatrix} \qquad \mathbf{M}_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

At $k = 1$, t_1 fires K

$\therefore\ \mathbf{u}_1 = [1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0]^T$ and $\mathbf{M}_1 = [2\ \ 0\ \ 1\ \ 1\ \ 1]^T$

At $k = 2$, t_4 fires K

$\therefore\ \mathbf{u}_2 = [0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 0]^T$ and $\mathbf{M}_2 = [2\ \ 2\ \ 0\ \ 1\ \ 1]^T$

and so on.

3.5 Behavioral Properties of Petri nets

Behavioral properties of Petri nets are those properties which depend on the initial marking $\mathbf{M}_0$ of the Petri net. Below, we follow the definition of each behavioral property with a brief comment on the interpretation in the context of a Petri net model.

3.5.1 Liveness

A Petri net is said to be live if all transitions are potentially firable for all reachable markings of the net i.e., if from any marking reachable from $\mathbf{M}_0$ it is possible to fire any transition in the net through some further firing sequence.

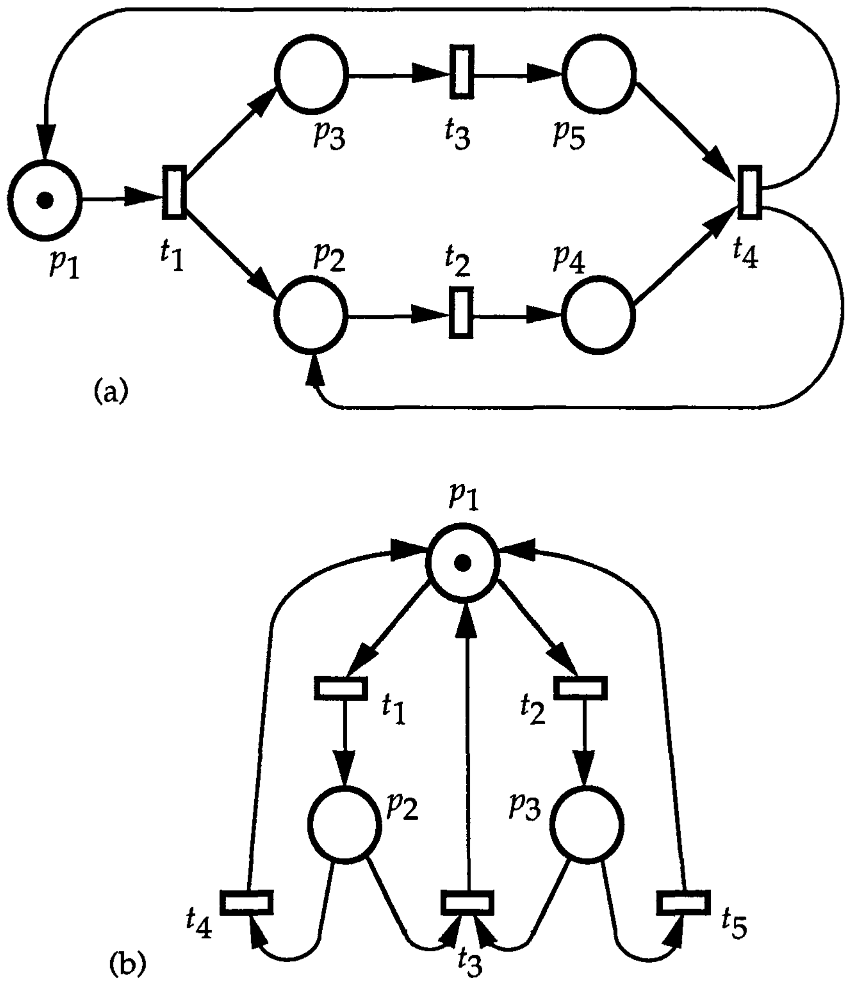


Figure 3.3: (a) Petri net that is live, non reversible, unbounded and globally fair (b) Petri net that is not live, reversible, bounded and not globally fair

The net in Figure 3.3(a) is live whereas the net in Figure 3.3(b) is not, since the transition t_3 is dead. Testing for liveness is a stringent criterion for most real life systems, rather a more practical solution is to test for the absence of deadlocks (*i.e.* transitions that are not enabled) in the net. One method of solution for this property is discussed in Section 3.8.3.

3.5.2 Reachability

A marking $\mathbf{M}_n$ is reachable from the initial marking $\mathbf{M}_0$ if there exists a firing sequence that can accomplish the change. A necessary (although not sufficient) condition for a marking $\mathbf{M}_n$ to be reachable from the initial marking $\mathbf{M}_0$ is the existence of a non-negative solution to Equation 3.4.

The sufficiency condition is yet to be derived for a general class of Petri nets. The only conclusive method of determining the reachability of a marking in a Petri net is by enumerating all of the markings. A general method adopted to tackle problems of such complexity is by reducing the problem to other problems that are more familiar, for instance the reachability problem is reducible to the liveness problem [8].

3.5.3 Boundedness

A Petri net is bounded if the number of tokens in each place is finite for any marking reachable from the initial marking. Boundedness of a Petri net is a property that can be easily determined by examining the reachable markings of the Petri net by way of the Reachability Tree (detailed later in this chapter). The bound on a place, in a bounded Petri net, is determined by the maximum token count for that place in the set of reachable markings. For example, in Figure 3.4 the set of all reachable markings from $\mathbf{M}_0$ is $\{(1, 0, 0, 0), (0, 1, 0, 0), (0, 0, 1, 0), (0, 1, 1, 0), (0, 0, 2, 0), (0, 0, 0, 1)\}$ hence, the bounds on the set of places $\{p_1, p_2, p_3, p_4\}$ are $\{1, 1, 2, 1\}$ respectively.

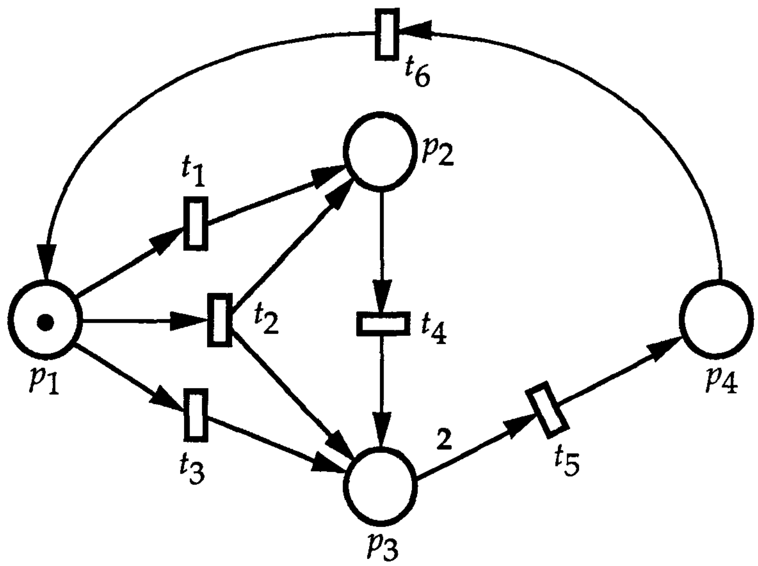


Figure 3.4: A bounded Petri net with the bounds {1, 1, 2, 1} on the places

3.5.4 Reversibility

A Petri net is said to be reversible if the initial marking $\mathbf{M}_0$ is reachable from all other possible markings in the set of markings reachable from $\mathbf{M}_0$.

The net in Figure 3.3(b) is reversible since the marking $\mathbf{M}_0 = \{1, 0, 0\}$ is reachable from all other markings {0, 1, 0} and {0, 0, 1} of the net; conversely the net in Figure 3.3(a) is not reversible.

3.5.5 Fairness

If a firing sequence σ is finite or if every transition in σ occurs infinitely often, then σ is globally fair. A Petri net is globally fair if every σ in all possible markings from $\mathbf{M}_0$ is globally fair. This suggests the capability of continuous firing of the transitions in the net.

The Petri net from Figure 3.3(a) is an example of a globally fair net since any infinite firing sequence starting from a marking reachable from $\mathbf{M}_0$, contains an infinite occurrences of every transition in the net.

3.6 Structural Properties of Petri nets

The properties that do not depend on the marking $\mathbf{M}_0$ of the net, but only on the structure or connectivity of the net come under this category.

3.6.1 Structural Boundedness

A Petri net is structurally bound if it is bounded for any finite initial marking $\mathbf{M}_0$. It is shown [20] that a Petri net is structurally bound *iff* there exists an m-vector $\mathbf{y}$ of positive integers such that $\mathbf{Ay} \leq 0$.

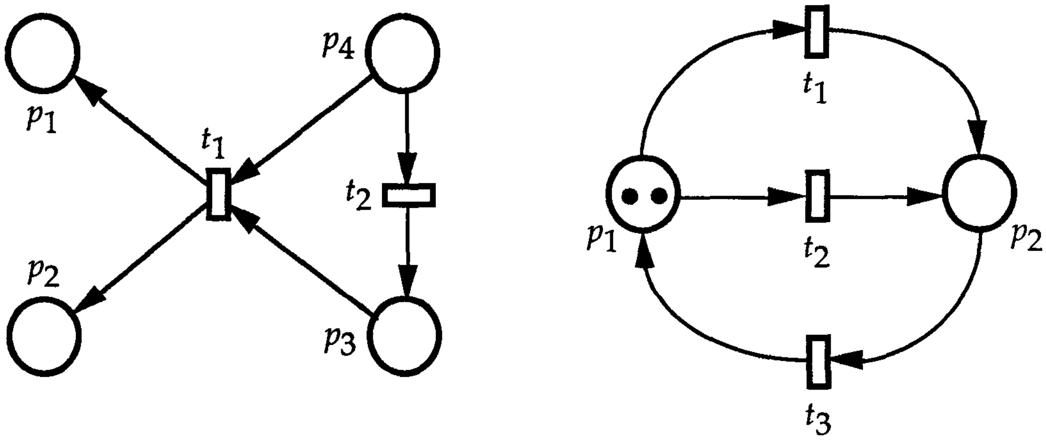
The bounds on the places can be determined by the the expression $\mathbf{M}(p) \leq (\mathbf{M}_0^T\mathbf{y})/\mathbf{y}(p)$ where $\mathbf{M}(p)$ is any reachable marking for place p and $\mathbf{y}(p)$ is the pth element in the solution vector $\mathbf{y}$.

3.6.2 S-invariants

These are defined by the solutions to the equation

$$\mathbf{Ay} = 0 \tag{3.5}$$

Here, $\mathbf{y}$ is an m $\times$ 1 vector. The non-zero entries in $\mathbf{y}$ ($\mathbf{y} \geq 0$), constitute the <u>support</u> of an S-invariant. The support is the set of places whose token count does not change with any firing sequence from $\mathbf{M}_0$. In other words the equation $\mathbf{M}^T\mathbf{y} = \mathbf{M}_0^T\mathbf{y}$ holds for all $\mathbf{M}$ reachable from $\mathbf{M}_0$.



(a) (b)

Figure 3.5: (a) Petri net with S-invariants such that the token count in a set of places is constant (b) Petri net with T-invariants that identify firing sequences that return the net to the initial marking $\mathbf{M}_0$

As an example let us consider the net in Figure 3.5 (a), the S-invariants derived from the net are $\mathbf{y}_1 = [1 \quad 1 \quad 1 \quad 1]^T$, $\mathbf{y}_2 = [2 \quad 0 \quad 1 \quad 1]^T$ and $\mathbf{y}_3 = [0 \quad 2 \quad 1 \quad 1]^T$. Therefore, the token count in the place combinations $\{p_1+p_2+p_3+p_4,\ 2p_1+p_3+p_4,\ 2p_2+p_3+p_4\}$ remains unaltered for all transition firings from $\mathbf{M}_0$.

3.6.3 T-invariants

These invariants are the solutions to the equation

$$\mathbf{A}^T\mathbf{x} = 0 \ , \quad \mathbf{x} \geq 0 \tag{3.6}$$

The support of a T-invariant, defined by $\mathbf{x}$, is the set of transitions that have to fire, from some $\mathbf{M}_0$, to return the Petri net to the same $\mathbf{M}_0$.

Consider the net in Figure 3.5(b), the minimal set of T-invariants are

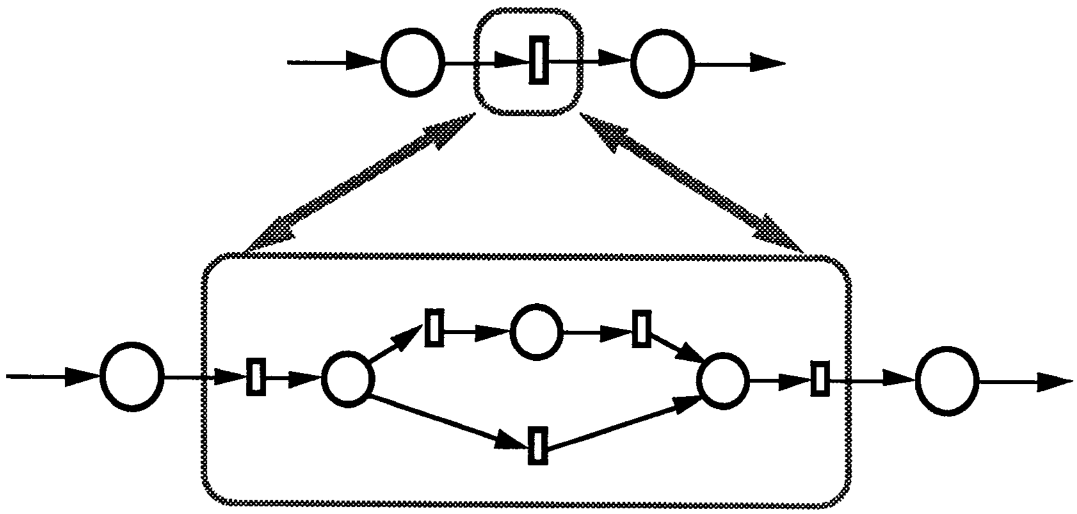


Figure 3.6: A transition as a representation of a sub-net

$\mathbf{x}_1 = [1 \quad 0 \quad 1]^T$ and $\mathbf{x}_2 = [0 \quad 1 \quad 1]^T$; all other T-invariants can be obtained by linear combinations of this minimal set. Therefore, the firing sequences $\sigma_1 = \{t_1, t_3\}$ and $\sigma_2 = \{t_2, t_3\}$ and all other combinations of these would result in the same marking as the initial marking $\mathbf{M}_0$.

3.7 Features of Petri net models

We discuss below a few additional features of Petri net models that are relevant in their analysis or in their application to bioreaction systems.

3.7.1 Extendibility

If a high level of abstraction is initially adopted to construct a simpler Petri net, the net can be subsequently extended upon the initial structure with the modification of relevant sections of the net [32]. This modification need not involve changing the structure of the complete net. For instance, a transition can be visualized as the representation of a Petri sub-net

(Figure 3.6) and any modification to this sub-net is reflected in the behavior of the original transition. This feature is particularly useful in cases where the present knowledge is incomplete, and we would like a representation that can be extended upon the present state of knowledge without significant deviation from the existing structure.

3.7.2 Abstraction

Petri nets allow a certain level of abstraction in the representation of biological reaction systems. This corresponds, in many cases, to a process which is the reverse from that mentioned in the previous paragraph. For example, if a part of a Petri net model is not of primary interest to us or is not of direct consequence to our analysis, it is possible to collapse this information to a smaller representation without the loss of behavior from the original net.

3.7.3 Classification

Ordinary Petri nets can be classified according to their structural properties to facilitate the analysis of the model. Structural classification includes the subclasses State Machines (SM), Marked Graphs (MG), Free-Choice (FC) nets, Extended Free-Choice (EFC) nets, and Asymmetric-Choice (AC) nets [20]. In each of the restricted forms of Petri nets, a few particular properties can be established conclusively [16].

3.7.4 Modification

Ordinary Petri nets can be modified to suit the specific needs of the system to be modeled. Some of the modifications applied to Petri nets that are in literature are Timed Petri nets, Stochastic Petri nets and High level Petri nets.

Timed Petri nets have transition firings that are associated with a deterministic time delay which is useful in terms of modeling and simulation of any system that consists of processes with finite execution times, e.g. manufacturing systems. Stochastic Petri nets are a different form of timed nets where the time-delay is a randomly distributed variable, potential applications include the simulation of those systems where the arrival, waiting and execution times are important, e.g. queuing networks. Timed Petri nets are discussed in detail in [22].

Colored Petri Nets are an important class of High level Petri nets. The Colored Petri net incorporates colored tokens and functional weights on the arcs. The enabling and firing of the transitions is determined by the color and the weight function on the input and output arcs. Detailed theory and application can be found in [13].

3.8 Methods of Analysis

3.8.1 Reachability Tree

The Reachability tree enumerates the possible states of a Petri net starting from a specified initial marking $\mathbf{M}_0$. Each node of the tree denotes the reachable marking $\mathbf{M}$ starting from $\mathbf{M}_0$. The branches from a node denote the possible transition firings from that node. The branches in the tree are terminated if a marking results in a deadlock (i.e., no enabled transitions exist) or if a marking is repeated.

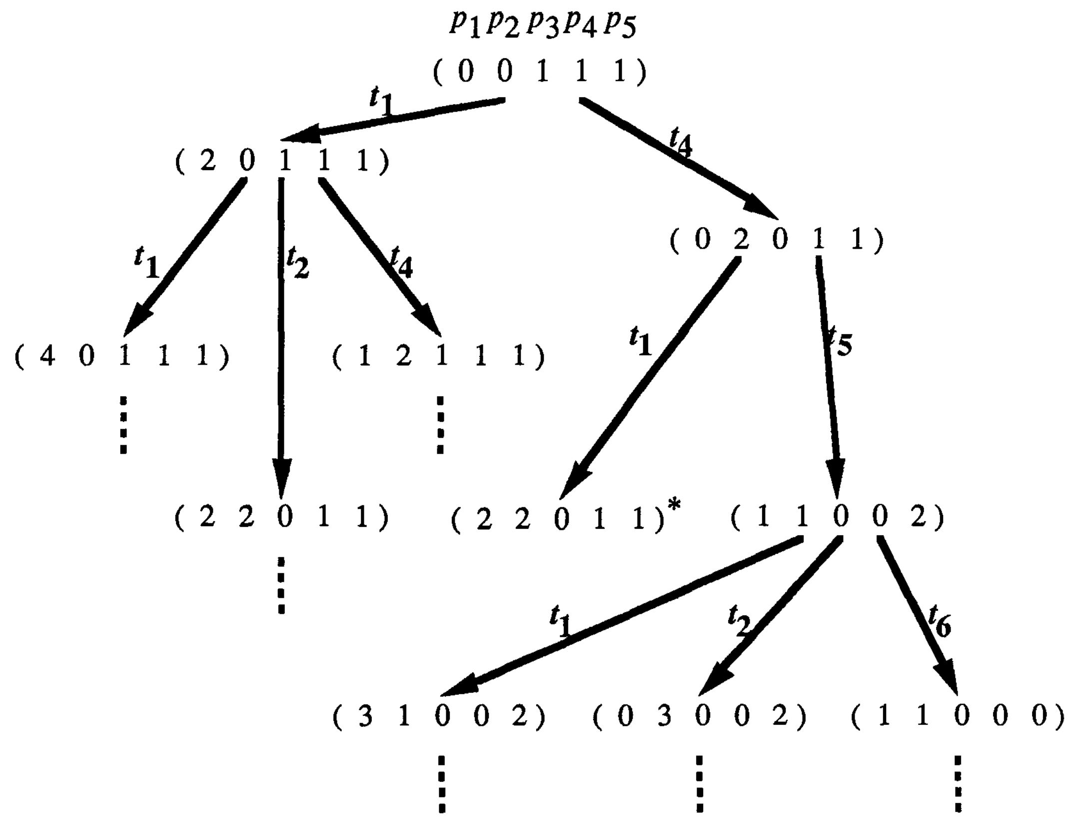


Figure 3.7: Reachability Tree for the Petri net from Figure 3.1

The Reachability tree for the Petri net graph in Figure 3.1 is shown in Figure 3.7. The root of the tree is the node (0 0 1 1 1) representing the initial marking, t_1 and t_4 are the possible firings leading to the respective markings (2 0 1 1 1) and (0 2 0 1 1), and so forth. The node (2 2 0 1 1)* is a repeated marking and branching from that node is terminated.

We can also choose to terminate branching when a marking supersedes one of its ancestors, i.e., it has elements greater than or equal to its ancestor marking (on an element by element comparison). This would be indicated through special notation that flags some of the entries in the marking as unbounded.

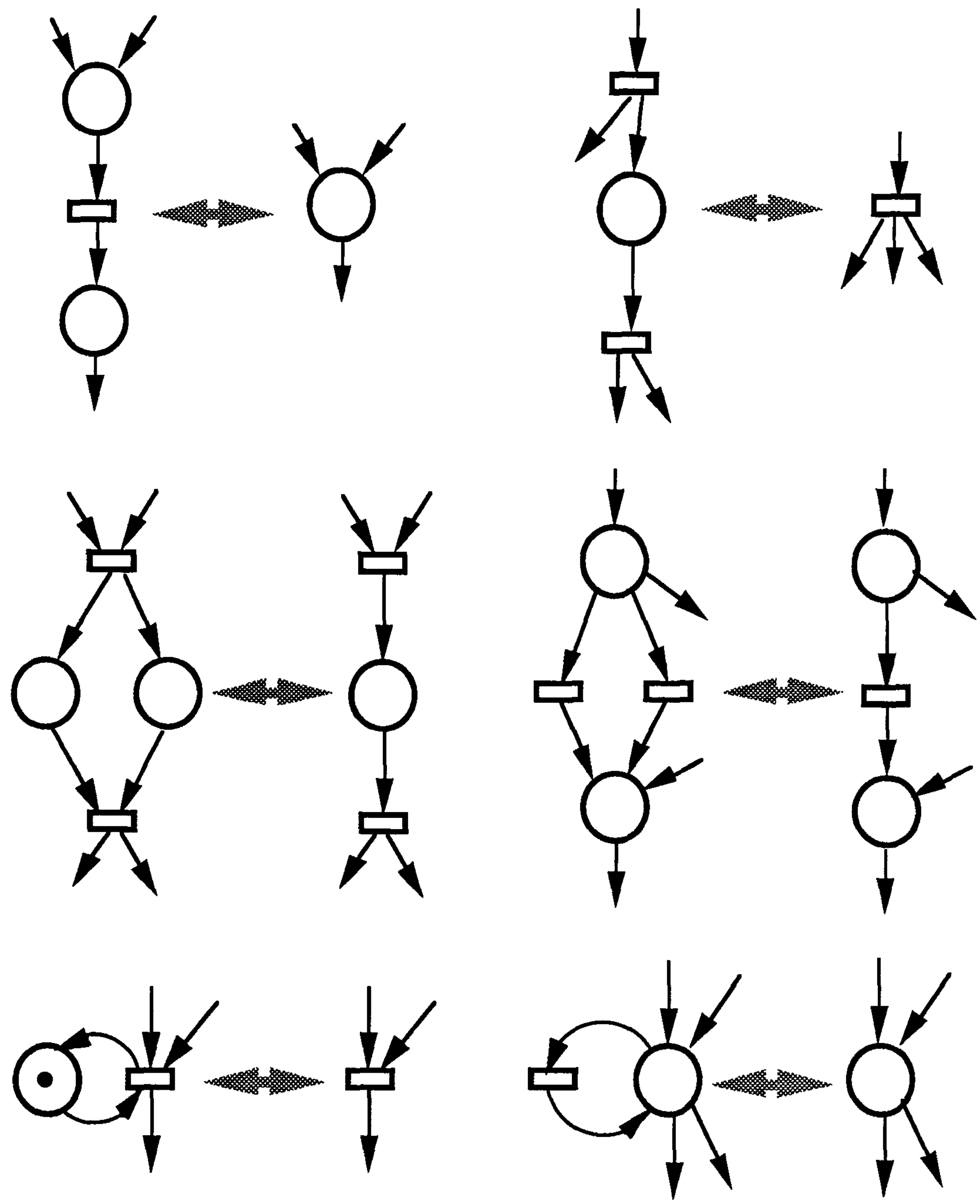


Figure 3.8: Some examples of structural reductions that are possible in Petri net graphs

It is possible to answer many questions regarding the properties of Petri nets, such as reachability, liveness, boundedness, etc., by an extended search of the reachability tree.

Often, complex systems result in a Reachability tree that is very large requiring exponential time and memory to solve the problem. However, the analysis of smaller nets is simple and convenient with this method.

3.8.2 Structural Reduction

Large Petri nets can be reduced to smaller nets by the substitution of certain combinations of places and transitions (Figure 3.8) without sacrificing the original properties of the net [20]. Though only simple reduction rules have been discussed here other steps of transformation of subnets to single places or transitions is discussed elsewhere [30]. This can be viewed as a method for model reduction that lessens the complexity of the original network. Reducing the size of the net is of importance not only for reducing the complexity of the system but also in achieving a reduction in the memory usage on a computer.

A different approach to the structural reduction of a Petri net is by partitioning the net into smaller nets. The sub-nets preserve the characteristics of the original net [5], hence the analysis of the sub-nets would lead to similar results but with lesser effort.

3.8.3 Matrix Methods

The matrix representation of a Petri net model can be used in the algebraic analysis of properties such as reachability. A necessary condition for the reachability of a marking $\mathbf{M}$ from the initial marking $\mathbf{M}_0$ is the existence of a solution to Equation 3.4, the elements of $\mathbf{x}$ being non-negative integer numbers. The sufficiency condition is a more difficult task and the reachability for a general Petri net can conclusively be deduced from the Reachability Tree.

The liveness of a Petri net is a difficult problem [21]. The non-existence of deadlock markings – a marking where there are no more enabled transitions – is pertinent to the analysis of biological pathways since it addresses the issue of the non-existence of active reactions. An algebraic method for determining the non-existence of deadlock markings has been illustrated in [12]. These methods of solution are easier to implement on a computer than the Reachability Tree method and therefore are more appealing.

3.8.4 Structural Mapping

Structural mapping is the systematic comparison of structures of the Petri net models. It is possible to identify structural similarities of biological pathways with this method. Intuitively, we expect Petri nets with the same structure to exhibit similar dynamic behavior, hence common structural links between Petri net models of pathways would predict similar behavior among the pathways. Pathway models can be compared to elucidate the behavior of one pathway from the known behavior in the other pathway, if they have common structural links.

3.9 Summary

This chapter serves only as an introduction to the theory and applications available on Petri nets. The interested reader should refer to the literature listed in Appendix B to gain more insights into the present state of research in this area.

Chapter 4

Modeling of Metabolic Pathways

4.1 Representation

The proposed approach to the qualitative modeling of a pathway incorporates the use of a discrete event methodology for the representation and simulation of bioreaction networks. The reactions and other biological processes are modeled as discrete events and analyzed by applying Petri net modeling and analysis techniques to this representation of the pathway. The properties of Petri nets are useful in drawing qualitative conclusions about the behavior and structure of biological pathways.

The representation of the essential components in a biological pathway, using Petri net terminology, is the first step in modeling the metabolic network as a discrete event system [23].

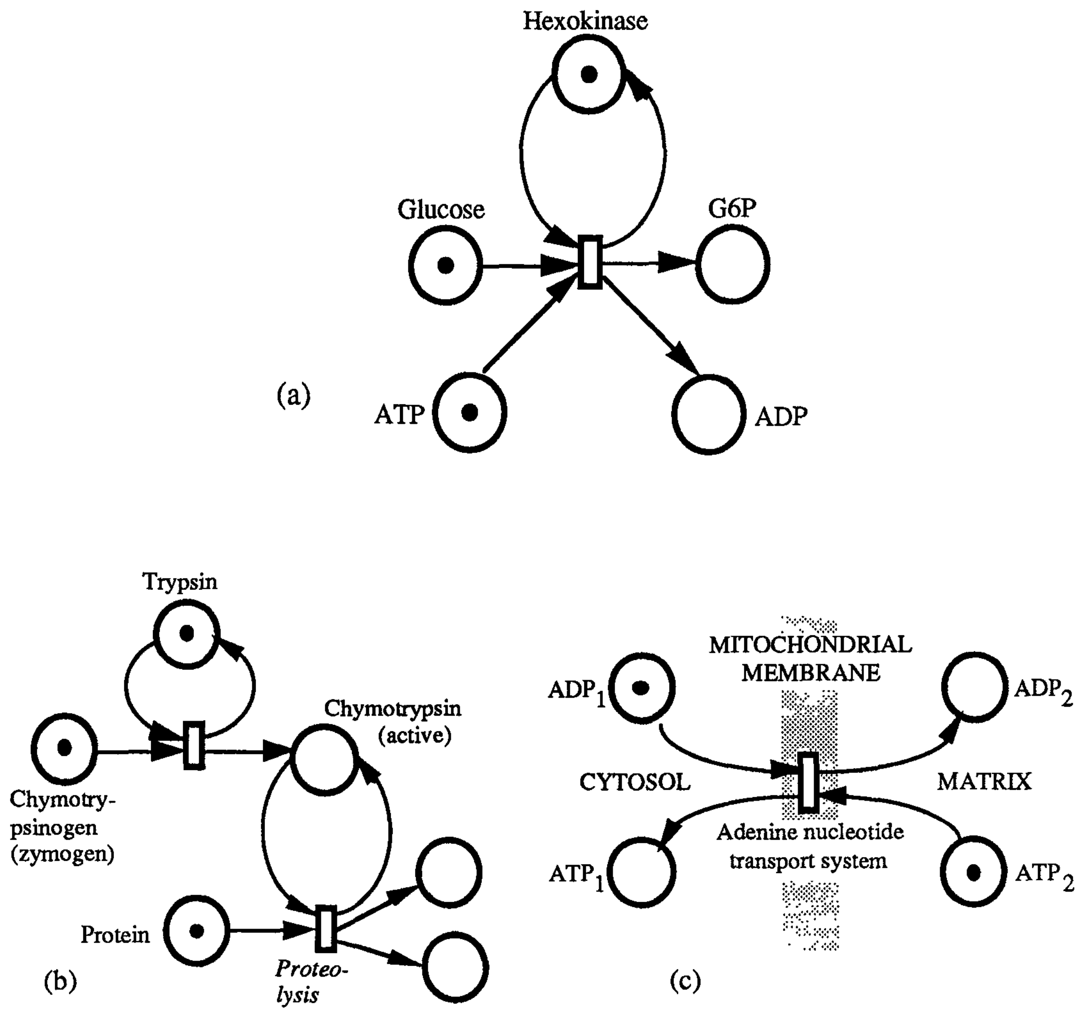


Figure 4.1: (a) Each place represents a biological compound (b) & (c) Two different places represent the same compound

For biological pathways, places would represent compounds (such as metabolites, enzymes, cofactors, etc.) participating in the biochemical system. Tokens indicate the presence of a compound.

Instead of having just one place represent each compound, it may be necessary to prescribe two or more places to represent the compound, when there are alternative physical attributes, changes in the activity of the

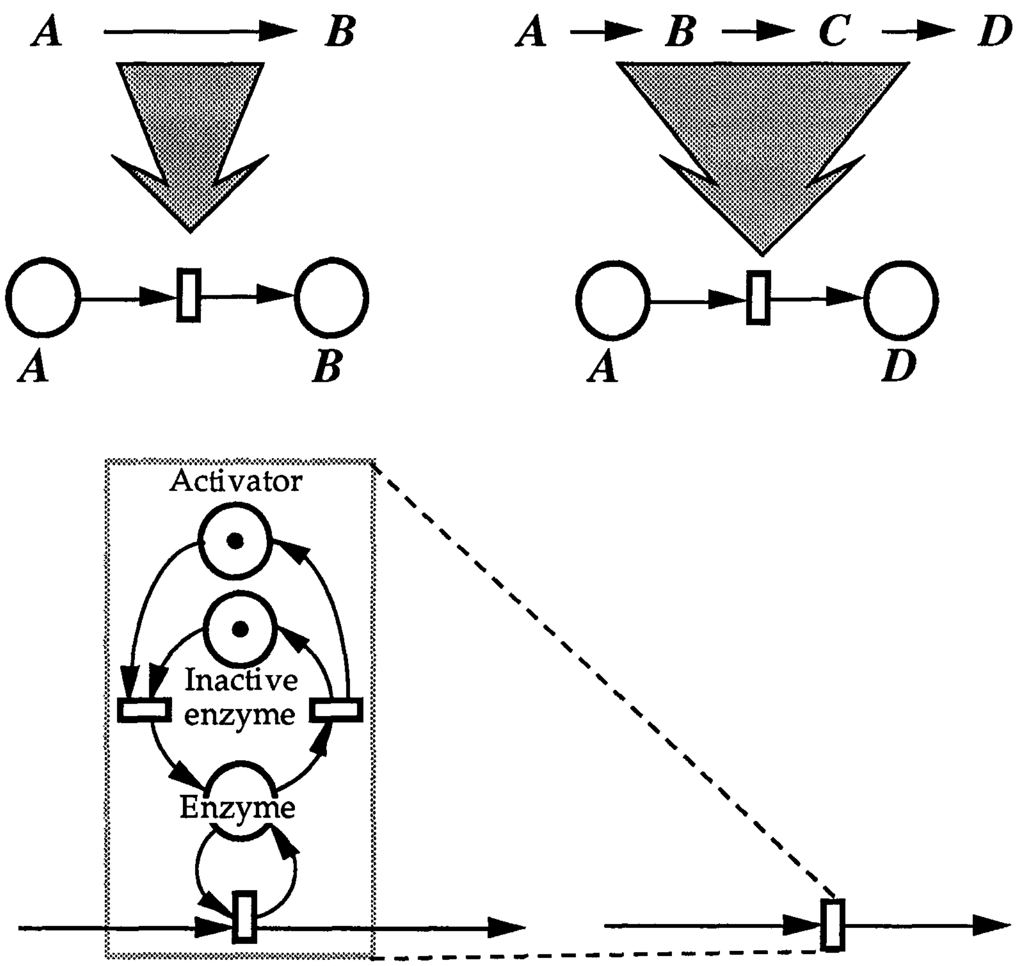


Figure 4.2: A transition can represent a single reaction, a chain of forward reactions or an abstraction of a subnet

compound, or distinct biological functions.

For example, in Figure 4.1(a), each place represents one biological compound, whereas in Figure 4.1(b), two places represent the same compound because a difference arises in activities: One place represents the inactive zymogen and the other the active enzyme Chymotrypsin. As another example, if we would like to distinguish between compounds based on their location in the cell, we could have different places represent the same compound. For instance, in Figure 4.1(c), the ATP pools inside and

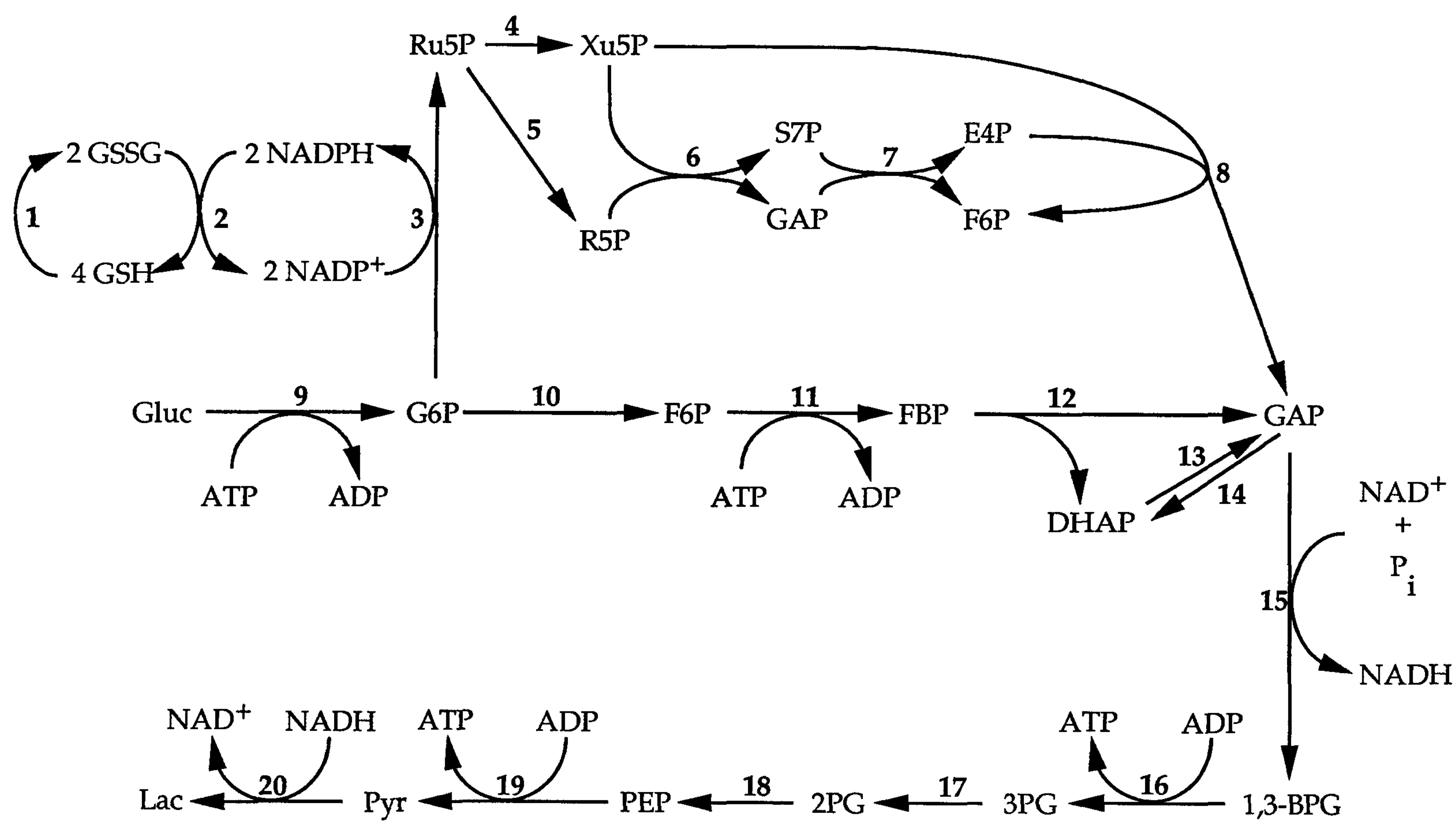


Figure 4.3: The combined metabolism of the glycolytic and the pentose phosphate pathways of an erythrocyte cell. Tables 4.1 & 4.2 provide the list of participating enzymes and metabolites.

Transition mapping for the Petri net model			
Index	**Enzyme / Reaction**	**Index**	**Enzyme / Reaction**
1	Glutathione oxidation reaction	2	*glutathione reductase*
3	G6P oxidation reactions	4	*ribulose-5-phosphate epimerase*
5	*ribulose-5-phosphate isomerase*	6	*transketolase*
7	*transaldolase*	8	*transketolase*
9	*hexokinase*	10	*phosphoglucose isomerase*
11	*phosphofructokinase*	12	*aldolase*
13	*triosephosphate isomerase* (forward reaction)	14	*triosephosphate isomerase* (backward reaction)
15	*glyceraldehyde-3-phosphate dehydrogenase*	16	*phosphoglycerate kinase*
17	*phosphoglycerate mutase*	18	*enolase*
19	*pyruvate kinase*	20	*lactate dehydrogenase*

Table 4.1: Mapping between reactions in the pathway and transitions in the Petri net model

outside the mitochondrion in a cell are different and their relative concentrations are determined through a selective transport process. Hence, we could have two places, one representing the compound inside and the other representing the compound outside the mitochondrion.

As would be natural, we assign transitions to represent individual reactions (Figure 4.2). A series (chain) of forward reactions or an abstraction of a subnet could be represented as a single transition, if desired, provided that the intermediary compounds are not of primary interest. Arc-weights represent the stoichiometry of reactions, and the direction of an arc is based on the thermodynamic feasibility of the reaction.

Place representation for the Petri net model			
Index	**Metabolite / Compound**	**Index**	**Metabolite / Compound**
1	Orthophoshpate (**P_i**), ionic form	2	Adenosine triphosphate (**ATP**)
3	Adenosine diphosphate (**ADP**)	4	Nicotinamide adenine dinucleotide, oxidized form (**NAD^+**)
5	Nicotinamide adenine dinucleotide, reduced form (**NADH**)	6	Nicotinamide adenine dinucleotide phosphate, oxidized form (**$NADP^+$**)
7	Nicotinamide adenine dinucleotide phosphate, reduced form (**NADPH**)	8	Glutathione disulfide (**GSSG**)
9	Glutathione (**GSH**)	10	Ribulose-5-phosphate (**Ru5P**)
11	Xylulose-5-phosphate (**Xu5P**)	12	Ribose-5-phosphate (**R5P**)
13	Sedoheptulose-5-phosphate (**S7P**)	14	Glyceraldehyde-3-phosphate (**GAP**)
15	Erythrose-4-phosphate (**E4P**)	16	Fructose-6-phosphate (**F6P**)
17	Glucose (**Gluc**)	18	Glucose-6-phosphate (**G6P**)
19	Fructose bisphosphate (**FBP**)	20	Dihydroxyacetone phosphate (**DHAP**)
21	1,3-Bisphosphoglycerate (**1,3-BPG**)	22	3-Phosphoglycerate (**3PG**)
23	2-Phosphoglycerate (**2PG**)	24	Phosphoenolpyruvate (**PEP**)
25	Pyruvate (**Pyr**)	26	Lactate (**Lac**)

Table 4.2: Mapping between metabolites in the pathway and places in the Petri net model

4.2 Metabolic Pathways in Erythrocytes

4.2.1 Pathway Overview

The pathways in the erythrocyte (red blood cell) are numerous as in any mammalian cell, but the common pathways under study are the oxidative pentose phosphate pathway and the main glycolytic pathway [10, 28]. The relative importance placed on this particular system stems from

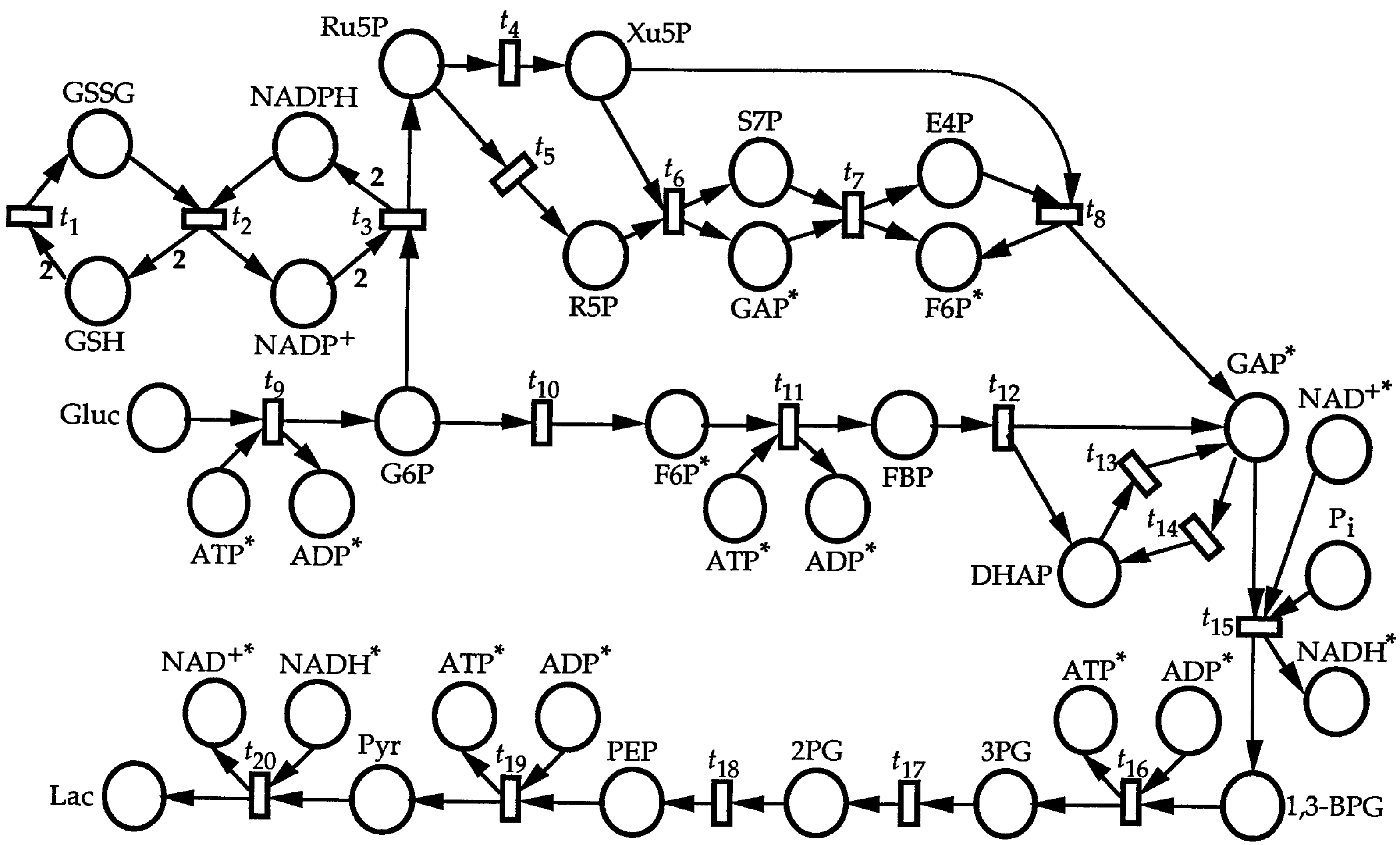


Figure 4.4: Petri net graph of the erythrocyte metabolism shown in Figure 4.3

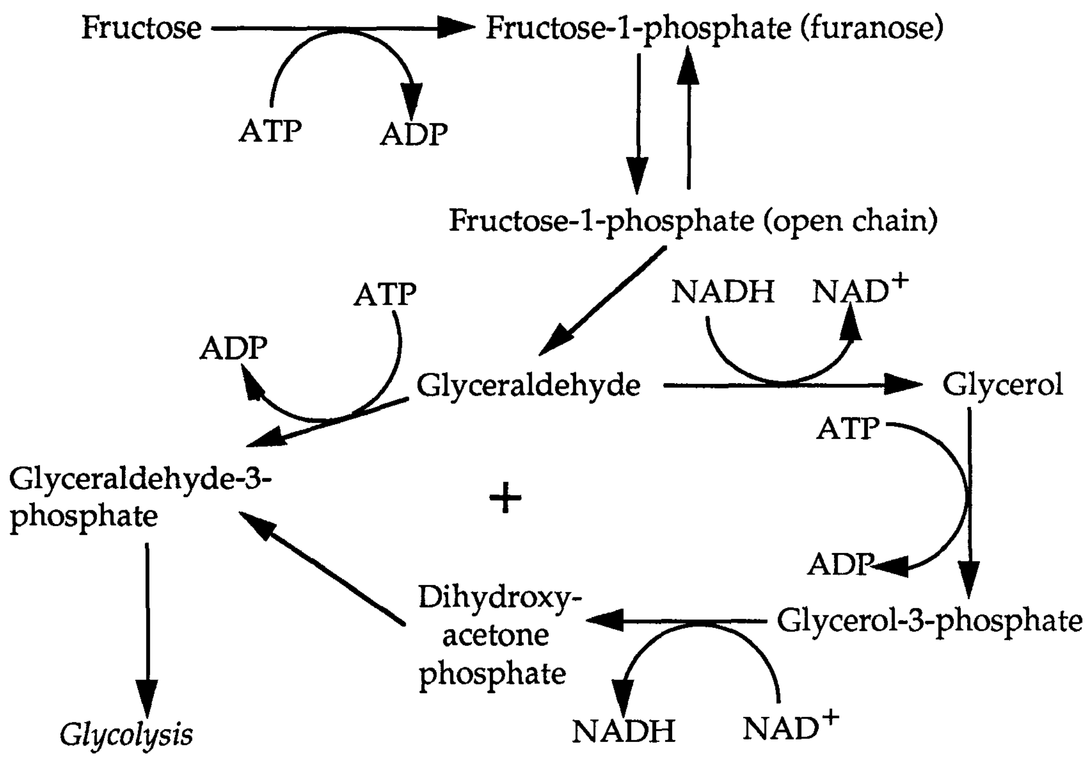


Figure 4.5: The metabolism of fructose in muscle cells

the fact that one of the many widespread diseases in humans is due to defects in this metabolic pathway. The familiarity with this well-studied mechanism among most readers with an elementary knowledge in biochemistry is an added advantage to using this set of pathways as an illustrative example of the methods presented hereafter.

The overall pathway as shown in Figure 4.3, in conjunction with Table 4.1 & Table 4.2, define the various reactions occurring in the cell that use glucose as the substrate and produce lactate as the product under heavy energy loads, such as in brisk muscle activity. Also, a steady supply of NADPH is required to regenerate GSH that is essential in sustaining cell integrity by reducing harmful peroxides produced in the cell [33].

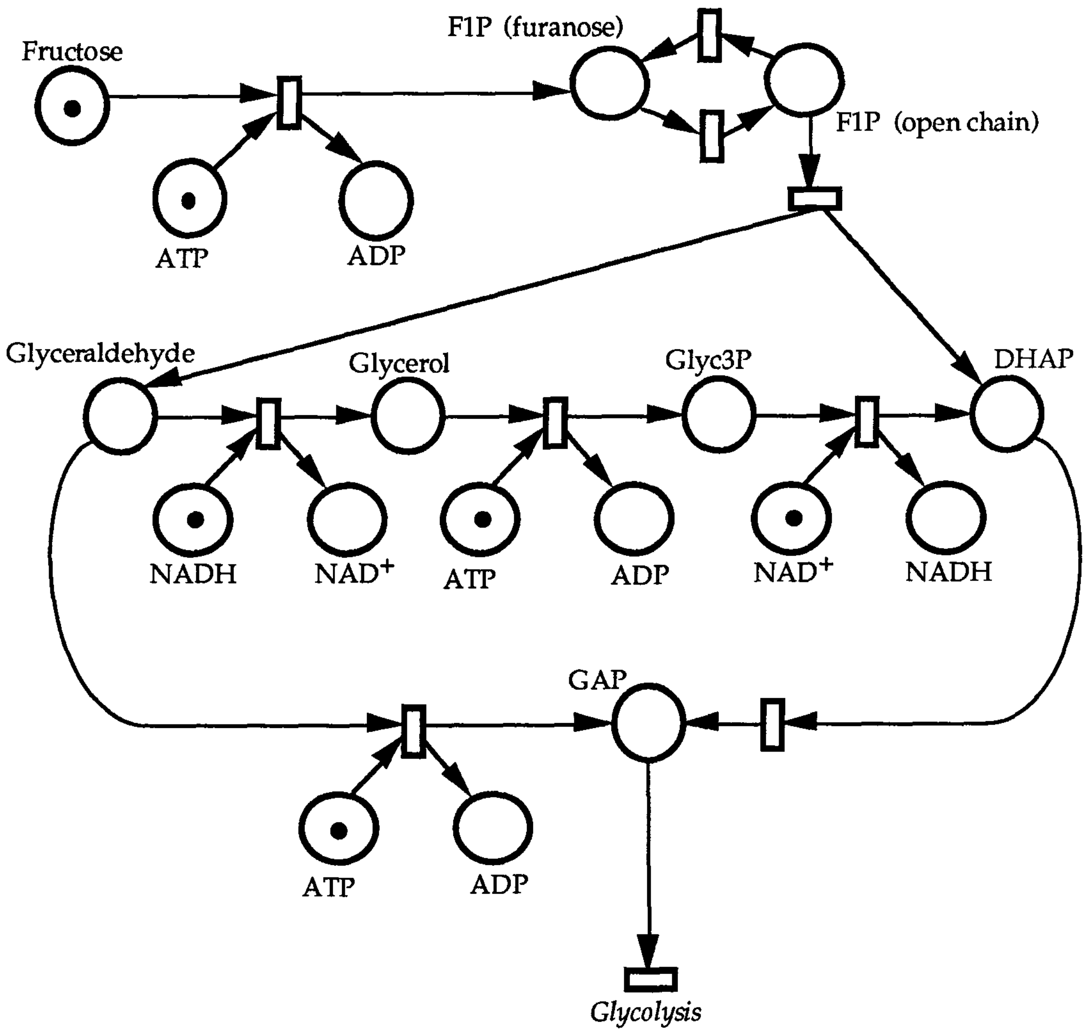


Figure 4.6: Petri net model of the fructose metabolism

4.2.2 Petri Net Representation

The Petri net graph of the pathway illustrated in Figure 4.3 is shown in Figure 4.4. The place and transition mapping between the pathway and the model is defined according to the method of representation explained earlier and is listed in Table 4.1 & Table 4.2. It is important to emphasize at this point that although the places labeled with an asterisk (*) occur more than once in the Petri net graph they in fact represent only one place with

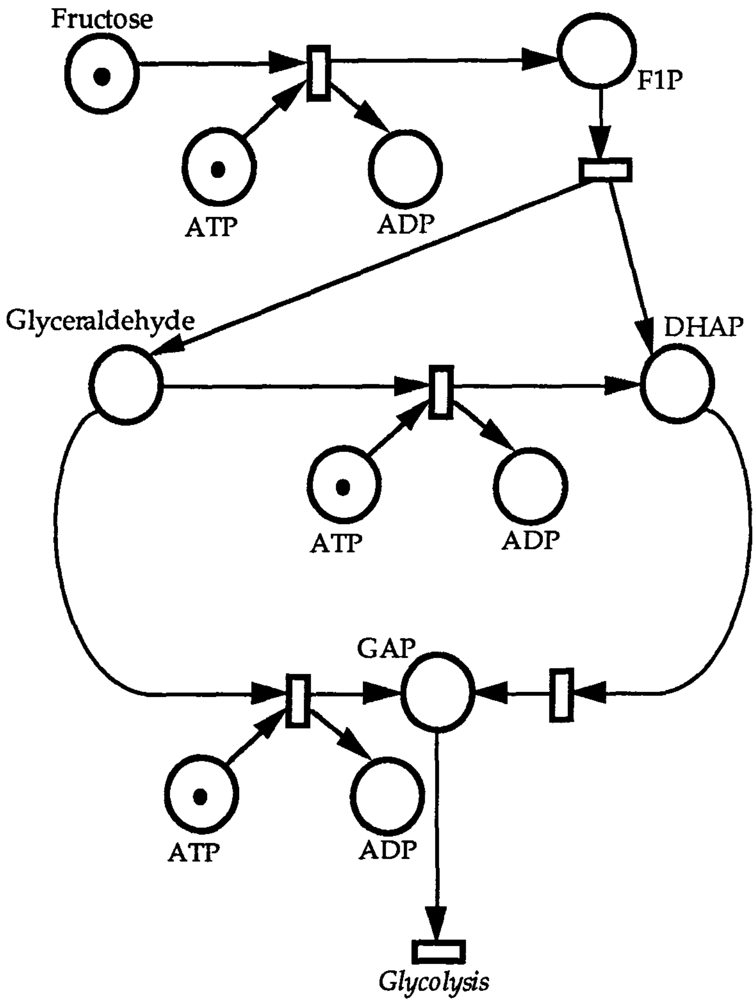


Figure 4.7: Abstracted model of the fructose metabolism

the appropriate label, for example ATP, ADP and F6P. The equilibrium reaction (reaction indices 13 & 14) is represented by two separate transitions to facilitate the modeling of the directionality of the reaction. The presence of the enzymes associated with each reaction is implicit in the representation of the transition unless otherwise noted.

4.3 Analysis

The qualitative analysis of metabolic pathways largely depends on the information that is required from the system. For instance, if we choose to concentrate our analysis on properties that are derived from the structural connectivity of the reactions we would analyze the Petri net model of the pathway based on structural properties (Sec. 3.7) of Petri nets defined earlier. Supplement to such an analysis we would also analyze the behavioral properties (Sec. 3.6) of the Petri net model to obtain functional qualities of the system. We illustrate a few of such properties with an application in the context of a biological pathway.

4.3.1 Model Reduction

4.3.1.1 Abstraction

Petri nets allow us the flexibility to abstract certain information in the net which is not essential for the analysis of the overall pathway. For example, consider the pathway shown in Figure 4.5, the corresponding abstract Petri net model in Figure 4.6 has information that is implicit in the structure of the net such as the presence of enzymes, activators, and co-factors. Abstraction of detailed information allows the analysis of large and complex systems without the loss of overall properties of the system.

The information of enzymatic activity can be abstracted by way of the subnet illustrated in Figure 4.2. The representation of the enzymes becomes necessary when we are interested in studying the influence of such enzymes and their associated activators on the behavior of the pathway as a whole.

In Figure 4.4 the oxidative reaction of glutathione with the participating peroxide is abstracted to a single transition. The abstraction follows from the assumption that if the availability of peroxides and the corresponding enzyme is unlimited then the only required condition for the activity of the reaction is the presence of glutathione.

4.3.1.2 Structural Reduction

This property leads to a reduction in the number of intermediate metabolites in the pathway. The method allows us to eliminate certain combinations of places and transitions recognized as structural reductions, defined earlier, such that the change does not affect the overall behavior of the model in terms of liveness and boundedness of the Petri net model of the metabolic pathway. This serves as a method of model reduction which could be a first step in reducing the complexity in the Petri net model.

The Petri net in Figure 4.6 is shown to illustrate the application of the reduction rules to the whole net to yield the simpler net in the Figure 4.7. The rules applied here are from the simple reduction rules derived in Chapter 3.

4.3.1.3 Partitioning

Partitioning of a large Petri net into smaller and simpler Petri nets is equivalent to the analysis of individual pathways in a biotransformation. The properties derived for those subnets (pathways) that are of interest can be obtained independent of the structure of the overall Petri net model. This allows us to analyze each subnet individually. The behavior of the overall Petri net model is influenced by the properties of the constituent subnets. Although there are no formal methods to deduce the behavior of the whole net from the behavior of the subnets, it is of interest to derive such methods to aid in the analysis of complex metabolic pathways. It is also of interest to derive theorems and methods to formally define rules under which a combination of properties of the subnets would yield the properties of the net as a whole. The ongoing research in this area as yet has no definite partitioning rules for general Petri nets. However, partitioning is possible for Petri nets that satisfy certain structural rules.

One method that allows such partitioning by splitting transitions is presented in [5]. The method can be used to test for liveness, safeness, boundedness, and reachability of simpler subnets that contain only places that are of interest. In the erythrocyte example, say, if we were only interested in the products of the pentose phosphate pathway in the overall metabolism, then we could split the transitions at *hexokinase* and GAPDH to yield the subnet of interest.

4.3.2 Qualitative Inference

4.3.2.1 Accumulation of Metabolites

In a large network of reactions a method that can indicate those reactants that can potentially accumulate under certain conditions is helpful as an initial pointer to target the metabolites. The property of boundedness of Petri nets can provide such information by the method defined in Section 3.6.1. Testing for structural boundedness can indicate whether the Petri net is bounded for a given $\mathbf{M}_0$. Therefore a Petri net that is not bounded implies that certain metabolites in the pathway can accumulate indefinitely.

As an example consider the model from Figure 4.4. The equations are set up as follows (Appendix A)

$$\mathbf{Ay} \leq 0 \quad , \mathbf{y} > 0 \tag{4.1}$$

The Petri net is bounded *iff* there exists a solution to the above. One such solution to the equation is $\mathbf{y}$ = [1, 3, 3, 4, 4, 5, 5, 10, 5, 16, 16, 16, 25, 7, 14, 16, 18, 17, 15, 7, 6, 5, 4, 3, 2, 1 $]^T$. The bound on each place can now be calculated given the initial marking as follows.

$$\mathbf{M}(p) \leq (\mathbf{M}_0^T\mathbf{y})/\mathbf{y}(p) \tag{4.2}$$

For our case, if we consider the initial marking $\mathbf{M}_0$ = [1, 3, 0, 3, 0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0$]^T$, the bounds on the set of places P = { $p_2, p_4, p_6, p_8, p_{22}, p_{26}$} are { 32, 24, 19.2, 9.6, 19.2, 96} respectively. These bounds are for any general class of Petri nets and are therefore conservative; however this does provide a hard bound on the concentrations (moles) the corresponding metabolites can ever accumulate under all possible conditions of the pathway, given the set of initial concentrations. A method to improve on this bound is by considering the minimal support S-invariants [20] demonstrated later in this section.

In the Petri net model of a pathway the support of an S-invariant determines the set of compounds whose total net concentrations remain unchanged in the course of a biotransformation. This may happen with compounds that act in a catalytic capacity. For example, the unbound form of enzymes, along with all its bound (or inactivated) forms, may collectively represent an S-invariant. This will occur if there is no production of new enzyme, but merely association/dissociation and activation/inactivation events. S-invariants may also occur for currency metabolites (such as ATP, ADP, and AMP) if, in the system being modeled, consumption of one member of the family is always accompanied by production of another (in an equal number of moles).

The S-invariants for our example can be computed (Appendix A) and are as follows,

$\mathbf{y}_1 = [1, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1]^T$

$\mathbf{y}_2 = [-1, -1, 0, 1, 0, 0, 0, 0, 0, -1, -1, -1, -1, -1, -1, -1, 0, -1, -2, -1, -1, 0, 0, 0, 1, 0]^T$

$\mathbf{y}_3 = [1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 2, 1, 2, 1, 1, 1, 0, 0]^T$

$\mathbf{y}_4 = [0, 0, 0, 0, 0, 0, 0, 2, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

$\mathbf{y}_5 = [0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

$\mathbf{y}_6 = [0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

$\mathbf{y}_7 = [0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

The equality $\mathbf{M}_0{}^T.\mathbf{y} = \mathbf{M}^T.\mathbf{y}$ is satisfied for a given initial marking $\mathbf{M}_0$ and all reachable markings $\mathbf{M}$. This can be thought of as a weighted sum of the tokens present in the support of the invariants. The property of conservation is evident in the above statement and is demonstrated by the fact that the set of invariants $\{\mathbf{y}_7, \mathbf{y}_6, \mathbf{y}_5\}$ includes the "currency" metabolites ATP, ADP; NAD^+, NADH; and $NADP^+$, NADPH which is intuitive since the combined moles of the pairs have to remain constant due to their inter convertibility. The other set of invariants is not as obvious and requires careful consideration. The invariant $\mathbf{y}_4$ corresponds to the GSSG, GSH pair of compounds. The reduction of GSSG produces two moles of GSH and in the oxidative reaction two moles of GSH combine to form GSSG. Thus to conserve the token count among the pair the number of tokens has to be weighted to reflect the reaction stoichiometry. The invariant $\mathbf{y}_3$ corresponds to the set of compounds {P_i, ATP, Ru5P, Xu5P, R5P, S7P, GAP, E4P, F6P, G6P, FBP, DHAP, 1,3-BPG, 3PG, 2PG, PEP} which indicates the conservation of phosphate groups among these set of compounds in the pathway. The invariant formed by the sum of basis invariants $\mathbf{y}_1 + \mathbf{y}_2 + \mathbf{y}_3$ corresponds to a straight chain {P_i, 1,3-BPG, 3PG, 2PG, PEP, Pyr, Lac} that does not have any external input arcs to the places. The number of tokens in such a chain of places remains constant no matter which transition fires in the Petri net.

The minimal support invariants can be used to improve on the bounds of places (Appendix A) and the values for the set of places $P = \{p_2, p_4, p_6, p_8, p_{22}\}$ are {3, 3, 2, 1, 4}. As can be seen, the bounds are tighter and give an accurate description of the upper bounds on the token count for selected places given the initial marking.

4.3.2.2 Deadlock Free Pathways

A general method for determining the property of liveness for a Petri net has not been developed due to the complexity of the problem, however there are methods to determine the nonexistence of deadlock markings [12]. This is important from the point view of modeling biochemical pathways to determine the possibility for the existence of any metabolic blocks that could hinder the progress of the reactions.

4.3.2.3 Continuous Operation

A pathway that has the property of global fairness suggests the existence of a state of continuous operation, starting from an initial state, without outside intervention. This could result in the formation of certain compounds in unrestricted amounts. Fairness is a property that is difficult to test in a general Petri net, but we can relax the conditions to test for properties that yield similar behavior in a Petri net. For example, repetitiveness in a Petri net indicates the presence of a marking such that there exists some firing sequence in which some (or all) transitions occur infinitely often. This can be interpreted as a condition where a set of reactions (or the entire pathway) can be in a state of continuous operation. This condition can be further relaxed to yield the test for T-invariants that indicates the presence of cyclic firing sequences. For the model in the example, there exists only one T-invariant that is, $\mathbf{x} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0]^T$. Indeed, this corresponds to the equilibrium reaction (indices 13 &14) between DHAP and GAP that can exhibit such cyclic behavior under certain conditions.

4.4 Summary

The analysis of metabolic pathways in a discrete event methodology using Petri nets as a tool has been explained. The use of such methods would be of immense importance to researchers who would like to gain preliminary insights into the behavior of metabolic pathways, even in the absence of quantitative data, at a relatively low cost in terms of effort and computation. The analysis of even large complex networks can be handled with the same set of simple structural and behavioral properties defined. The applicability of Petri nets among different fields has prompted concurrent research that could lead to the development of newer algorithms and theories which in turn could be used across the different areas of research. In fact many of the methods that have been described have their sources from different areas of research spanning computer science to information theory to flexible manufacturing systems. Evidently, a concerted effort among researchers from different areas is required to unlock the potential of this emerging field.

Chapter 5

Conclusion

The methodology of Petri nets demonstrates the use of qualitative methods as useful preliminary analysis tools for biological pathways. The method is easy to implement and visually comprehensible. Although most properties can be determined for certain classes of Petri nets, the solution to the liveness, reachability, and boundedness problems are still very complex for the general Petri net structure [21]. The issue of complexity and decidability of problems in the general class of Petri nets is a matter of concern; for practical purposes, one may strive to model the system (or parts of the system) through one of the restricted categories of Petri nets, reducing the complexity issues.

Petri nets can also be extended by suitable modifications (Section 3.7.4)

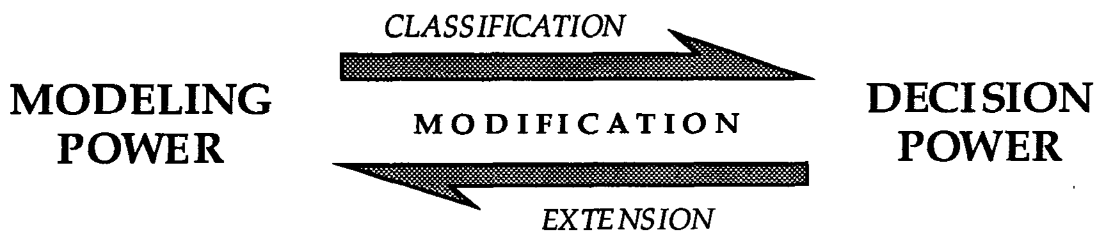


Figure 5.1: Effect of modification of Petri nets on the modeling and decision (analytical) powers

to the definition to enhance the modeling power of the net. In fact, the extension of Petri nets by the inclusion of inhibitor arcs (i.e., a transition is only enabled when the input place connected by this arc does not contain any tokens) can increase the modeling power of Petri nets to that of a Turing machine [21]. Naturally, the extension of the modeling power exacerbates undecidability and complexity obstacles. This increase in modeling power can also lead to a decrease in the analytical power of Petri nets (Figure 5.1) which is the main identifying feature of the methodology. Thus, an important task is the identification of appropriate narrow classes of nets that combine reasonable expressive power for the domain of biochemical pathways without undue complexity burdens.

A representation in a Petri net model allows a certain level of abstraction to enable us to look at the pathway from a different 'perspective'. In a group of pathways which are made up of similar compounds but behave in a dissimilar manner (dependent on their functionality), it may be possible to identify certain common patterns of interactions (templates) among the similar compounds, by some means of abstraction. The hypothesis is that these pathways might have diverged from their primitive templates to their present state, the divergence being task-dependent

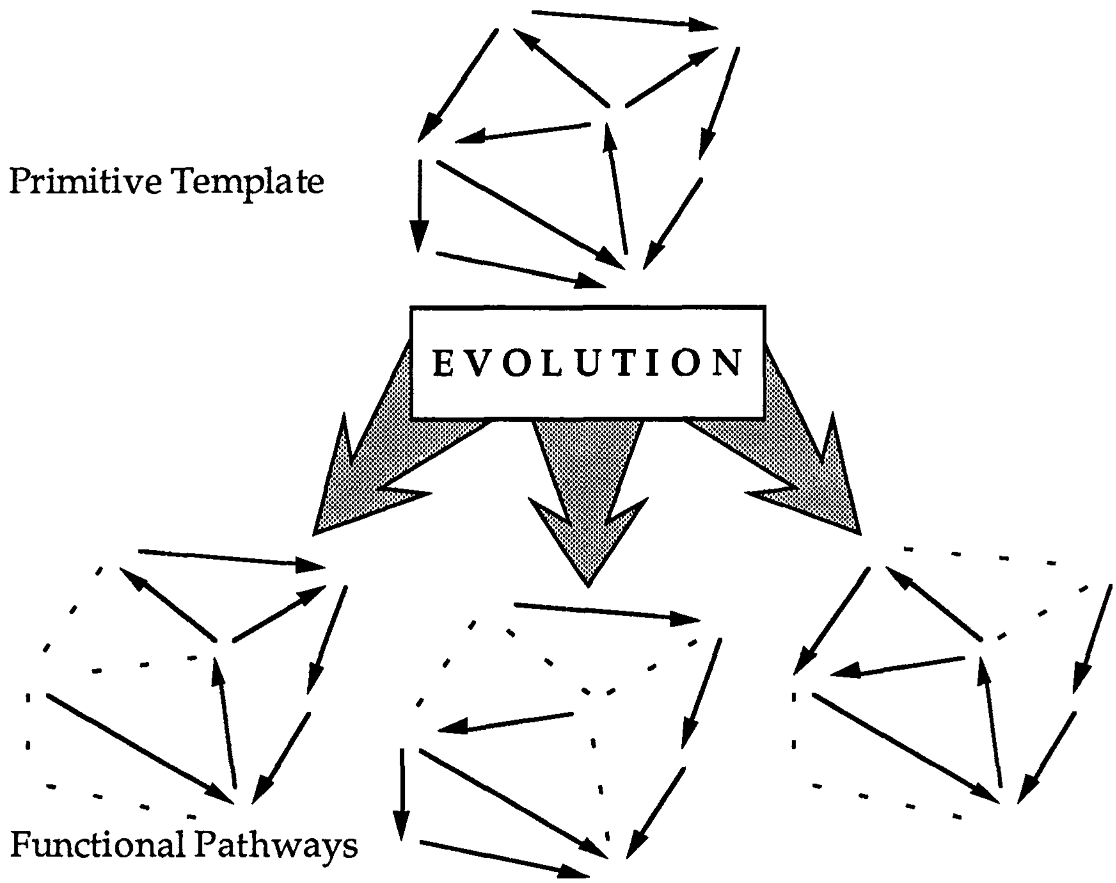


Figure 5.2: Functionally dissimilar pathways that evolved from the same structurally primitive template

(Figure 5.2). Thus, it is possible to examine the difference in behavior of the pathways in relation to structural modification of the templates. An example of a set of pathways which may have this homology is the cascade mechanism of activation of serine proteases in the blood-coagulation pathway, the complement system, and the differentiation in the fruit fly.

In addition to the properties defined earlier, a qualitative analysis of a Petri net representation of a pathway would entail comparison of Petri net models. Formally this could be carried out through optimal mapping of one Petri net to another, such that the maximum number of place-transition relations are preserved. For example, we can require that each place (biological component) of one pathway should correspond to precisely one component in the other pathway, or we can allow one-to-many mappings; a variation could be a mapping applied to transitions. Other comparisons of pathways would include mappings which preserve particular Petri net properties or sub-nets, for instance the amplification property of the cascade mechanism displayed in the blood coagulation pathway could be a common template that is also observable in the cascade mechanism of the complement system. Also, we can examine the differences in the properties of two Petri nets by a direct comparison of their structures, and relate these to differences in structures (*i.e.* biological components and transformations) of the metabolic pathways represented by the Petri nets. In effect we can determine the role of each difference in the biological structure through the effect it has on the properties of the Petri nets. An appropriate modification then, of only these reactions or components, is required to either enhance or terminate this behavior of the system. The possibilities would also encompass the identification and regulation of alternate routes of biotransformation, *e.g.* to overcome metabolic blocks due to defective enzymes, or to obtain greater yields in bioprocesses.

Appendix A

Structural Bounds and Invariants

The example presented in Chapter 4 models the metabolism in an erythrocyte. The corresponding petri net model is given in Figure 4.4. The structural boundedness, S-invariants and T-invariants of this model are calculated from the corresponding incidenxe matrix of the petri net.

The incidence matrix **A** as defined in section 3.4 is shown on the following page. The rows in **A** indicate the transitions and the columns the places in the petri net.

The net is structurally bound *iff* the following condition is statisfied:

$\exists\, y > 0, \quad \mathbf{A}y \leq 0$

$$
\mathbf{A} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -2 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\
0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1
\end{bmatrix}
$$

The incidence matrix A for the Petri net model of the erythrocyte example

Therefore one such integer value solution to this set of linear inequalities is the m-vector $\mathbf{y} = [1, 3, 3, 4, 4, 5, 5, 10, 5, 16, 16, 16, 25, 7, 14, 16, 18, 17, 15, 7, 6, 5, 4, 3, 2, 1]^T$. In our case, if we consider the initial marking $\mathbf{M}_0 = [1, 3, 0, 3, 0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$ the structural bounds on the places can be calculated by the expression.

$$\mathbf{M}(p) \leq (\mathbf{M}_0^T\mathbf{y})/\mathbf{y}(p)$$

where $\mathbf{M}(p)$ is any reachable marking in place p and $\mathbf{y}(p)$ is the pth element in the vector $\mathbf{y}$. Therefore, $\mathbf{M}_0{}^T\mathbf{y} = 96$, and the bounds on each of the places are tabulated below:

Place	1	2	3	4	5	6	7	8	9
Bound	96	32	32	24	24	19.2	19.2	9.6	19.2
Place	10	11	12	13	14	15	16	17	18
Bound	6	6	6	3.84	13.71	6.86	6	5.33	5.65
Place	19	20	21	22	23	24	25	26	
Bound	6.4	13.71	16	19.2	24	32	48	96	

The S-invariants are the integer value solutions to the system of linear equations, $\mathbf{Ay} = 0$.

Solving this set of equations we obtain the set of vectors,

$\mathbf{y}_1 = [1, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1]^T$

$\mathbf{y}_2 = [-1, -1, 0, 1, 0, 0, 0, 0, 0, -1, -1, -1, -1, -1, -1, -1, 0, -1, -2, -1, -1, 0, 0, 0, 1, 0]^T$

$\mathbf{y}_3 = [1, 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 2, 1, 2, 1, 1, 1, 0, 0]^T$

$\mathbf{y}_4 = [0, 0, 0, 0, 0, 0, 0, 2, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

$\mathbf{y}_5 = [0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

$\mathbf{y}_6 = [0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

$\mathbf{y}_7 = [0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]^T$

It is to be noted that these are the basis solutions to the system of equations and any linear combination of these bases form an invariant.

Analogously, the T-invariants are the integer value solutions to the system of linear equations, $\mathbf{A}^T\mathbf{x} = 0 \;\; \mathbf{x} \geq 0$.

The only solution that is possible for the given system is,

$\mathbf{x} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0]^T$

Structural bounds can be improved upon by the method described in [20]. The bounds are calculated for the minimal support invariants by the expression $\mathbf{M}(p) \leq \mathrm{Min}\,[(\mathbf{M}_0^T\mathbf{y}_i)/\mathbf{y}_i(p)]$, where the minimum is taken over all non-negative minimal support s-invariants $\mathbf{y}_i$ such that $\mathbf{y}_i\ (p) \neq 0$. The bounds calculated for the non-negative $\mathbf{y}_i$ are tabulated below:

Place	2	4	6	8	22	26
Bound	3	3	2	1	4	Indeter-minate

Comparing with the previously obtained bounds on these places we observe that the bounds are considerably improved.

Appendix B

Information Sources

Petri Net mailing list

The steering committee of the International Conference on Application and Theory of Petri Nets offers a mailing group to researchers interested in Petri Nets. The mailing group is maintained at the Computer Science Department, Aarhus University, Denmark. The group includes mail related to all aspects of the Petri net community; current research, conference announcements, discussions and in general sharing of information related to Petri nets.

To subscribe, send an e-mail to the following address: "PetriNets-request@daimi.aau.dk". You will receive further instructions on subscription.

Literature

In addition to the standard reference books on the subject there are other sources for current research on the subject such as,

"Application and Theory of Petri Nets," Springer–Verlag, Berlin, Germany.

"Advances in Petri Nets," Springer–Verlag, Berlin, Germany.

The individual volumes contain selected papers from the annual International Conference on Application and Theory of Petri Nets. These are published under the series "Lecture Notes in Computer Science" edited by G. Goos and J. Hartmanis.

An extensive bibliography is presented in, "Drees, S. *et al*, Bibliography of Petri Nets, in: *Advances in Petri Nets 1987*, Ed. G. Rozenberg, LNCS 266, Springer–Verlag, New York, NY, 1987, pp. 309–451."

Software Tools

Efficient modeling of complex systems with the methodology of Petri nets requires the use of appropriate computaitonal tools. Some of the currently available software is listed below; a complete list of tools is available via anonymous ftp from the file: petri.crim.ca/petri/PETRI-TOOLS. The description of software tools are presented with complete functionality, hardware environment and availability.

CPN/AMI:

Supports Place/Transition nets and Colored nets, graphical interface available on the Macintosh platform while analysis is done on a Sun Sparc workstation.

PROD:

This is a reachability analysis tool by generating a reachability graph. Available for UNIX and MS-DOS operating systems.

XSimNet:

A simulation program for extended (timed, inhibitor arcs, colored etc.) Petri nets. Contains a XWindows graphical interface and runs on UNIX operating system.

References

1. Bailey, J. E. and Ollis, D. F., Biochemical Engineering Fundamentals, McGraw–Hill Book Co., New York, NY, 1986.

2. Berthelot, G. and Terrat, R., Petri Nets Theory for the Correctness of Protocols, *IEEE Transactions in Communication*, vol. COM-30, no. 12, 1982, pp. 2497–2505.

3. Cornish-Bowden, A., Metabolic Control Theory and Biochemical Systems Theory: Different Objectives, Different Assumptions, Different Results, *Journal of Theoretical Biology*, vol. 136, 1989, pp. 365–377.

4. Diaz, M., Modeling and Analysis of Communication and Cooperation Protocols using Petri Net based Models, *Comput. Networks*, vol. 6, 1982, pp. 419–441.

5. Fahmy, H. M. A., Analysis of Petri Nets by Partitioning: Splitting

Transitions, *Theoretical Computer Science*, vol. 77, 1990, pp. 321–330.

6. Fell, D. A. and Sauro, H. M., Metabolic Control and its Analysis: Additional Relationships between Elasticities and Control Coefficients, *European Journal of Biochemistry*, vol. 148, 1985, pp. 555–561.

7. Fell, D. A., Metabolic Control Analysis: a survey of its theoretical and experimental development, *Biochemical Journal*, vol. 286, 1992, pp. 313–330.

8. Hack, M., Decidability Questions for Petri Nets, Ph.D. dissertation, Dept. of Elec. Eng., MIT, Cambridge, MA, Dec. 1975.

9. Heinrich, R. and Rapoport, T. A., A Linear Steady-State Treatment of Enzymatic Chains, *European Journal of Biochemistry*, vol. 42, 1974, pp. 89–95.

10. Holzhütter, H-G., Schuster, R., Buckwitz, D. and Jacobasch, G., Mathematical Modelling of Metabolic Pathways Affected by an Enzyme Deficiency, *Biomedica Biochimica Acta*, vol. 49, no. 8-9, 1990, pp. 791–800.

11. Ichikawa, A. and Hiraishi, K., Analysis and Control of Discrete Event Systems Represented by Petri Nets, in: *Lecture Notes in Control and Information Sciences vol. 103, Discrete Event Systems: Models and Application*, Springer–Verlag, New York, NY, 1987.

12. Ivanov, N. N., Algebraic Method of Determining Nonexistence of Deadlock Markings in Petri Nets, *Automation and Remote Control*, vol. 52, no. 7, 1991, pp. 986–989.

13. Jensen, K., in: *Colored Petri Nets Vol. 1*, W. Brauer, G. Rozenberg, A. Salomaa, (Eds.), Monographs on Theoretical Computer Science, Springer–Verlag, New York, NY, 1992.

14. Kacser, H. and Burns, J. A., The Control of Flux, in: *Control of Biological Processes*, D. Davies (Ed.), Cambridge University Press, Cambridge, MA, 1973, pp. 65–104.

15. Kauffman, S. A., Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets, *Journal of Theoretical Biology*, vol. 22, 1969, pp. 437–467.

16. Kemper, P. and Bause, F., An Efficient Polynomial-Time Algorithm to Decide Liveness and Boundedness of Free-Choice Nets, in: *Application and Theory of Petri Nets 1992*, Ed. K. Jensen, LNCS 616, Springer–Verlag, New York, NY, 1992.

17. Kohn, M. C. and Lemieux, D. R., Identification of Regulatory Properties of Metabolic Networks by Graph Theoretical Modeling, *Journal Theoretical Biology*, vol. 150, 1991, pp. 3–25.

18. Kohn, M. C. and Letzkus, W. J., A Graph-theoretical Analysis of Metabolic Regulation, *Journal of Theoretical Biology*, vol. 100, 1983, pp. 293–304.

19. Martinez, J., Alla, H. and Silva, M., Petri Nets for the Specification of flexible Manufacturing Systems, in: *Modeling and Design of Flexible Manufacturing Systems,* Elsevier Science Publ., New York, NY, 1986.

20. Murata, T., Petri Nets: Properties, Analysis and Applications, *Proceedings of the IEEE*, vol. 77, no. 4, 1989, pp. 541–580.

21. Peterson, J. L., Petri Net Theory and the Modeling of Systems, Prentice–Hall, Inc., New Jersey, NJ, 1981.

22. *Proceedings of the International Workshop on Timed Petri Nets*, Torino, Italy, July 1-3 1985.

23. Reddy, V. N., Mavrovouniotis, M. L. and Liebman, M. N., in: *Proc. of the First Intl. Conf. on Intell. Sys. for Mol. Biol.;* Hunter, L.; Searls, D.;

Shavlik, J., Eds.; AAAI/MIT Press: Menlo Park, CA, 1993.

24. Reder, C., Metabolic Control Theory: A Structural Approach, *Journal of Theoretical Biology*, vol. 135, 1988, pp. 175–201.

25. Reisig, W., Petri Nets: An Introduction, Springer–Verlag, New York, NY, 1985.

26. Robichaud, L. P. A., Boisvert, M. and Robert, J., Signal Flow Graphs: Applications, Prentice–Hall Inc., Englewood Cliffs, NJ, 1962.

27. Savageau, M. A., Biochemical Systems Analysis, Addison–Wesley Publ. Co., Inc., Reading, MA, 1976.

28. Schuster, R., Jacobasch, G. and Holzhütter, H-G., Mathematical Modelling of Metabolic Pathways Affected by an Enzyme Deficiency, *European Journal of Biochemistry*, vol. 182, 1989, pp. 605–612.

29. Sen, A. K., Topological Analysis of Metabolic Control, *Mathematical Biosciences*, vol. 102, 1990, pp. 191–223.

30. Suzuki, I. and Murata, T., A Method for Stepwise Refinements and Abstractions of Petri Nets, *Journal of Computer and System Sciences*, vol. 27, no. 1, 1983, pp. 51–76.

31. Thoma, J. U., Introduction to Bond Graphs and their Applications, New York, NY, Pergamon Press, 1975.

32. Vallete, R., Analysis of Petri Nets by Stepwise Refinements, *Journal of Computer and System Sciences*, vol. 18, 1979, pp. 35–46.

33. Voet, D. and Voet, J. G., Biochemistry, John Wiley & Sons, New York, NY, 1990.

34. Voit, E. O., Optimization of Integrated Biochemical Systems, *Biotechnology and Bioengineering*, vol. 40, 1992, pp. 572–582.

35. Yamalidou, E. C. and Kantor, J. C., Modeling and Optimal Control of Discrete-event Chemical Processes using Petri Nets, *Computers and Chemical Engineering,* vol. 15, no. 7, 1991, pp. 503–519.

36. Yau, S. S. and Caglayan, M. U., Distributed Software System Design representation using Modified Petri Nets, *IEEE Transactions on Software Engineering,* vol. SE-9, no. 6, 1983, pp. 733–745.