Pergamon

Comput. Biol. Med. Vol. 26, No. 1, pp. 9–24, 1996

0010-4825/96 $15.00 + 0.00

0010-4825(95)00042-9

QUALITATIVE ANALYSIS OF BIOCHEMICAL REACTION SYSTEMS

VENKATRAMANA N. REDDY,* MICHAEL N. LIEBMAN† and MICHAEL L. MAVROVOUNIOTIS*‡

*Dept. of Chemical Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208–3120, U.S.A.; and †Vysis, Inc., 3100 Woodcreek Drive, Downers Grove, IL 60515, U.S.A.

(*Received* 6 *March* 1995; *and in revised form* 31 *July* 1995)

Abstract—The qualitative analysis of biochemical reaction systems is presented. A discrete event systems approach is used to represent and analyze bioreaction pathways. 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. The properties and methods for the analysis of Petri nets, along with their interpretation for biochemical systems, are presented. As an example, the combined glycolytic and pentose phosphate pathway of the erythrocyte cell is presented to illustrate the concepts of the methodology.

Petri nets Biochemical pathways Modeling Discrete event Qualitative analysis

INTRODUCTION

In order to simulate and analyze a biochemical pathway one must research and incorporate in a model many characteristics and parameters that determine the behavior of the pathway. This requires the integration of a large volume of diverse data from the biological, chemical and physical sciences, if one attempts a complete quantitative analysis. Several methods for the analysis of biochemical reaction systems seek to address particular aspects of the modeling problem. Biochemical Systems Theory (BST) [1, 2] developed by Michael Savageau can be used to study the regulation, optimization and control of biochemical reactions. A power law approximation is used to express system behavior in terms of the concentrations and other parameters, such as kinetic orders and rate constants. The graphical models explored by Kohn and Letzkus [3] for modeling metabolic networks consisted of Kauffman binary networks [4], signal flow graphs [5] and bond graphs [6]. Derived from these graph models, a new model called MetaNets was introduced by Kohn and Lemieux [7]. The MetaNets method is based on maintaining the biochemical reaction structure in a graphical model consisting of functional nodes and interconnecting arcs. The method serves to identify the potential feedback sites and controlling enzymes in the network, without quantitative information. However, the model does not capture certain biochemical processes, such as transport mechanisms, thereby losing generality in its representation. Metabolic Control Theory (MCT) is a method for relating properties, such as flux or concentration of metabolites, to parameters of the network, such as concentrations of enzymes, to study the effect of perturbations in these parameters on the overall system behavior [8]. The theory developed independently by Kacser and Burns [9] and Heinrich and Rapaport [10] has since been modified and applied by others [11–13].

The methods described above include parameters and constants estimated quantitatively from experimental data. Sometimes, modeling a complex biochemical system involves data that are incomplete, uncertain or unreliable. For instance, a model of a

‡ Author to whom correspondence should be addressed.

complex biochemical pathway is likely to involve many parameters describing kinetics and regulation; the absence of even one of these parameters may prevent evaluation of the model. Also, the primary data that are used to estimate model parameters sometimes contain experimental errors and inconsistencies. In addition, some parameters used in the model are not constant, e.g. the concentration of enzymes participating in a pathway may vary within a certain range. It is difficult to arrive at an accurate quantitative model for a biochemical pathway under these circumstances. Techniques suited for individual reaction mechanisms do not suffice since we require methods that are scalable to pathways of arbitrary size: Comprehensive experimental and calculated values are usually available for single reactions (or a small set of reactions) that are deemed important, but a complex pathway requires a much larger volume of data which is unlikely to be available. The problem of analyzing complex biochemical pathways in the absence of detailed accurate data remains unresolved and there is a need for a method based primarily on qualitative data rather than on quantitative parameters.

Given the sparseness and uncertainty of data, it is practical to explore methods that do not require detailed quantitative data to reach general conclusions on the behavior of biochemical pathways. A qualitative method that excludes detailed simulations of the system is essentially independent of parametric information of the pathway, such as rate constants and cooperativity indices. A qualitative analysis allows us to draw preliminary conclusions about the biochemical pathway such as the influence of particular reactions, metabolites or pathway segments on the overall system. We would like to adopt a method that can identify, for example, key compounds (metabolites, enzymes, activators, etc.) necessary for a biotransformation; metabolites capable of accumulation in unbounded amounts and those that exist in invariant amounts; behavior that results in cyclic transformations (futile cycles); potential effects on the overall behavior of a pathway due to modification of a portion of the pathway; and classes of behavior based on the reaction network structure of the system.

A new method of qualitative analysis of biochemical pathways is presented in this article. The technique incorporates the use of a discrete event methodology for the representation and analysis of biochemical reaction networks. The reactions and other biological processes are modeled as discrete events and analyzed by applying Petri net theory and properties [15,16]. A brief overview of the Petri net theory relevant to the method is discussed in the next section.

PETRI NET THEORY

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 applications in a number of fields such as control engineering, manufacturing systems and computer science.

The essential concepts in Petri net theory are outlined in this section. Detailed theory and applications of Petri nets are available elsewhere in literature [17–19].

Definitions

A Petri net is a directed graph (Fig. 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 non-negative 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).

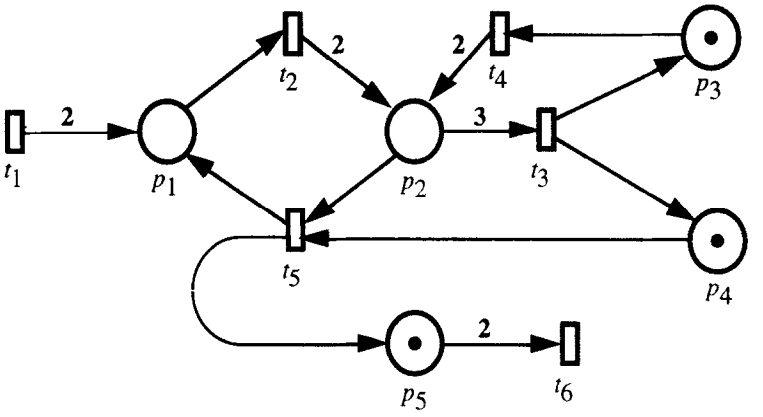

Fig. 1. A Petri net graph with places, transitions and arcs.

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 assumed to be equal to one.

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 of size $m \times 1$, where m is the number of places, whose elements correspond to the number of tokens present at each place of the Petri 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 *initial marking* $\mathbf{M}_0$.

Execution

Each transition is associated (through arcs) with a finite number of input places and output places. In many types of models of discrete event systems, it is necessary to satisfy a set of pre-conditions (defined by the input places for Petri nets) before an event (transition for Petri nets) may occur; the event results in a set of post conditions (output places for Petri nets).

In a Petri net a transition is *enabled* when the number of tokens in each input place is greater than or equal to the weight on the arc connecting that place to the transition. A transition with no input places, called a *source transition*, is always enabled. In Fig. 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 2 shows the same Petri net graph from Fig. 1 after firing 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 Fig. 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 (Fig. 2(a)). Similarly, the firing sequence t_2, t_3, t_5 and t_6, starting from a marking of Fig. 2(a), will result in the marking of Fig. 2(b).

State space representation

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.

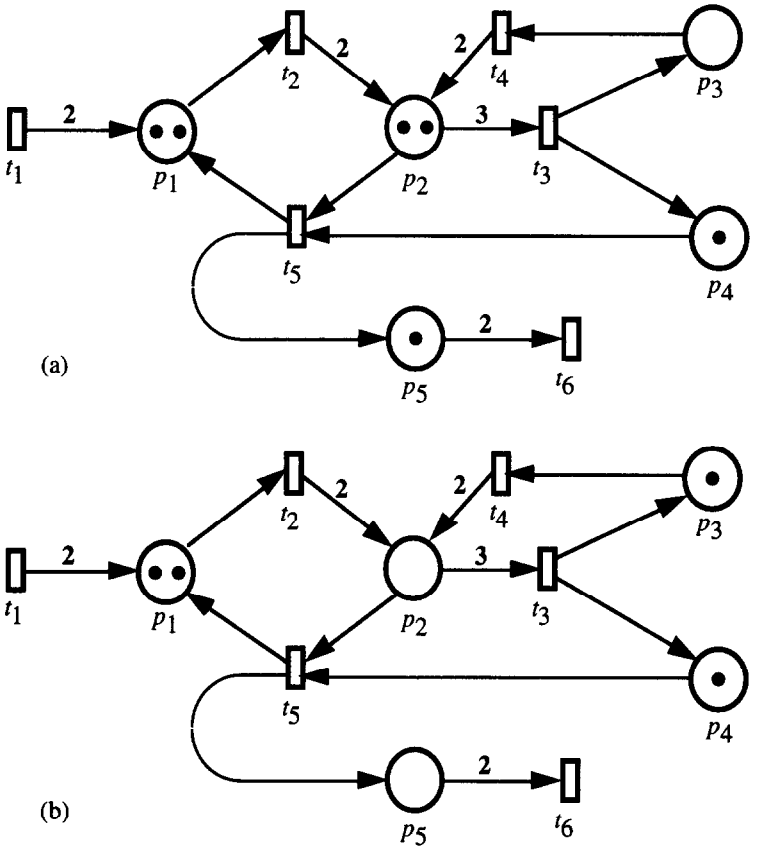

Fig. 2. (a) Marking after firing enabled transitions t_1 and t_4 from Fig. 1 (b) Marking after firing in sequence t_2, t_3, t_5 and t_6 from Fig. 2(a).

For a Petri net with m places and n transitions, we can formulate [20] a state equation of the type

$$\mathbf{M}_k = \mathbf{M}_{k-1} + \mathbf{A}^T \mathbf{u}_k, \qquad k = 1, 2, 3, \ldots \tag{1}$$

The index k represents a state 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 $\mathbf{M}_0$, though a firing sequence $\sigma = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \ldots \mathbf{u}_n\}$, and the state-equations are summed for all the firings in this σ, we obtain

$$\mathbf{M}_n = \mathbf{M}_0 + \mathbf{A}^T \sum_{k=1}^{n} \mathbf{u}_k \tag{2}$$

We define $\mathbf{x} = \Sigma_{k=1}^{n} \mathbf{u}_k$, 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 (2), we obtain

$$\mathbf{M}_n - \mathbf{M}_0 = \mathbf{A}^T \mathbf{x} \tag{3}$$

or

$$\mathbf{A}^T \mathbf{x} = \Delta \mathbf{M}. \tag{4}$$

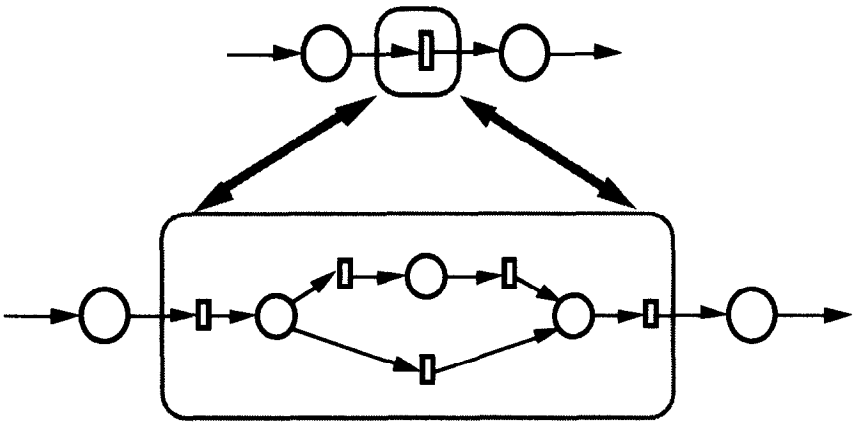

Fig. 3. A transition as a representation of a sub-net.

For example, from the Petri net graph of Fig. 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 . . .

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

At $k=2$, t_4 fires . . .

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

and so on.

Petri net properties relevant to the biochemical system

Extendibility. If a high level of abstraction is initially adopted to construct a simpler Petri net, the net can be subsequently extended from the initial structure by modifying relevant sections of the net [21]. This modification does not involve changing the structure of the complete net. For instance, a transition can be visualized as the representation of a Petri subnet (Fig. 3) and any modification to this subnet 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 without significant deviation from the existing structure.

Abstraction. Petri nets allow abstraction in the representation of biochemical 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 in our analysis, it is possible to collapse this information to a smaller representation.

Structural reduction. In addition to the above, large Petri nets can be reduced to smaller nets by substituting certain combinations of places and transitions following specific rules (Fig. 4) without sacrificing the original properties of the net [17]. Reducing the size of the net is of importance not only for reducing the complexity of the system but also in achieving more efficient computational analysis.

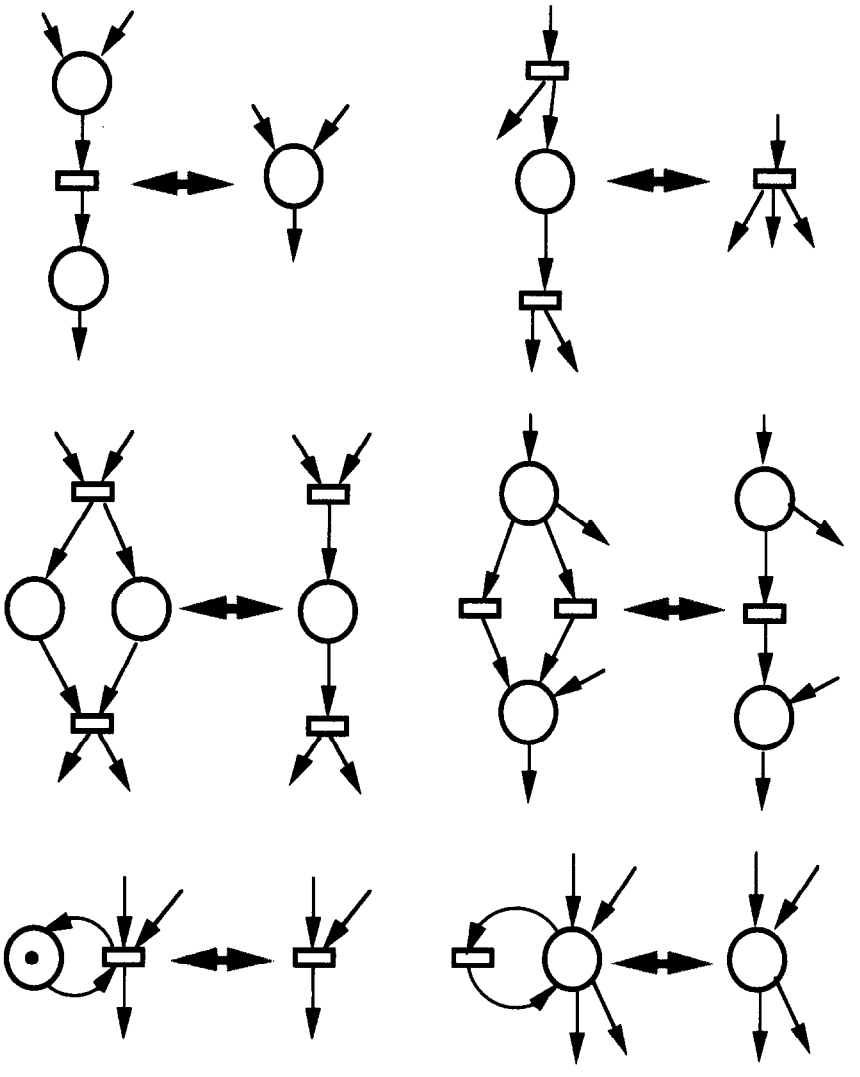

Fig. 4. Some examples of structural reductions that are possible in Petri net graphs.

Boundedness, S-invariants and T-invariants. These properties do not depend on the initial marking $\mathbf{M}_0$ of the net, but only on the structure or connectivity of the net.

A Petri net is said to be bounded if all reachable markings of the net are such that the number of tokens in each place is bounded by a finite value. A Petri net is structurally bounded if 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 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 $\mathbf{y}$.

S-invariants are defined by the solutions to the equation

$$\mathbf{Ay} = 0. \tag{5}$$

The non-zero entries in $\mathbf{y}$ constitute the set of places whose total token count does not change with any firing sequence from $\mathbf{M}_0$ and is called the *support* of the invariant. In other words the equation $\mathbf{M}^T\mathbf{y} = \mathbf{M}_0^T\mathbf{y}$ holds for all $\mathbf{M}$ reachable from $\mathbf{M}_0$.

T-invariants are the solutions to the equation

$$\mathbf{A}^T\mathbf{x} = 0, \qquad \mathbf{x} \geq 0. \tag{6}$$

The solution vector $\mathbf{x}$ is the set of transitions that have to fire, from some $\mathbf{M}_0$, to return the Petri net to the same $\mathbf{M}_0$.

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 a subsequent firing sequence. Liveness is too stringent a criterion for most real life systems; a more practical solution is to test for the absence of deadlocks (i.e. transitions that are not enabled) in the net.

The significance of these properties in the analysis of biochemical pathways will be discussed in detail in the latter section where we present an example.

BIOCHEMICAL PATHWAYS

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 properties of Petri nets are useful in drawing qualitative conclusions about the behavior and structure of biochemical pathways.

The representation of the essential components in a biochemical pathway, using Petri net terminology, is the first step in modeling the metabolic network as a discrete event system [20].

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

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

For example, in Fig. 5(a), each place represents one biological compound, whereas in Fig. 5(b), two places are used to represent a difference in activities: One place represents

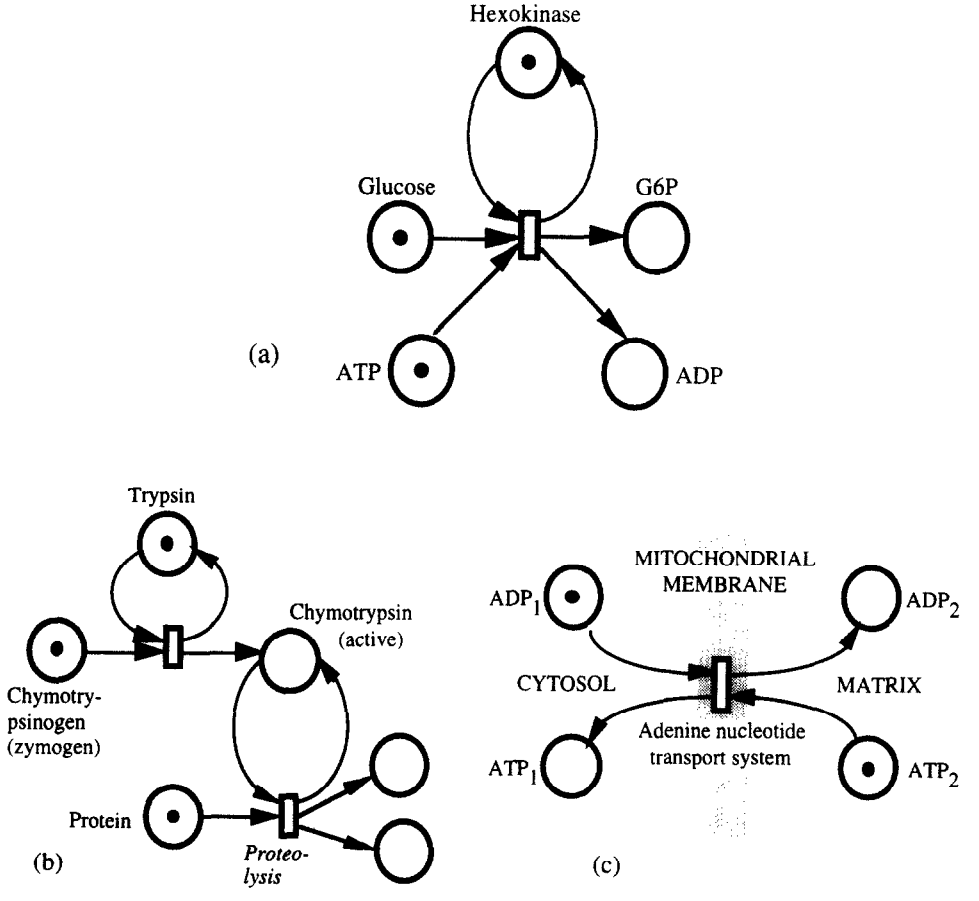

Fig. 5. Correspondence of places to biological components.

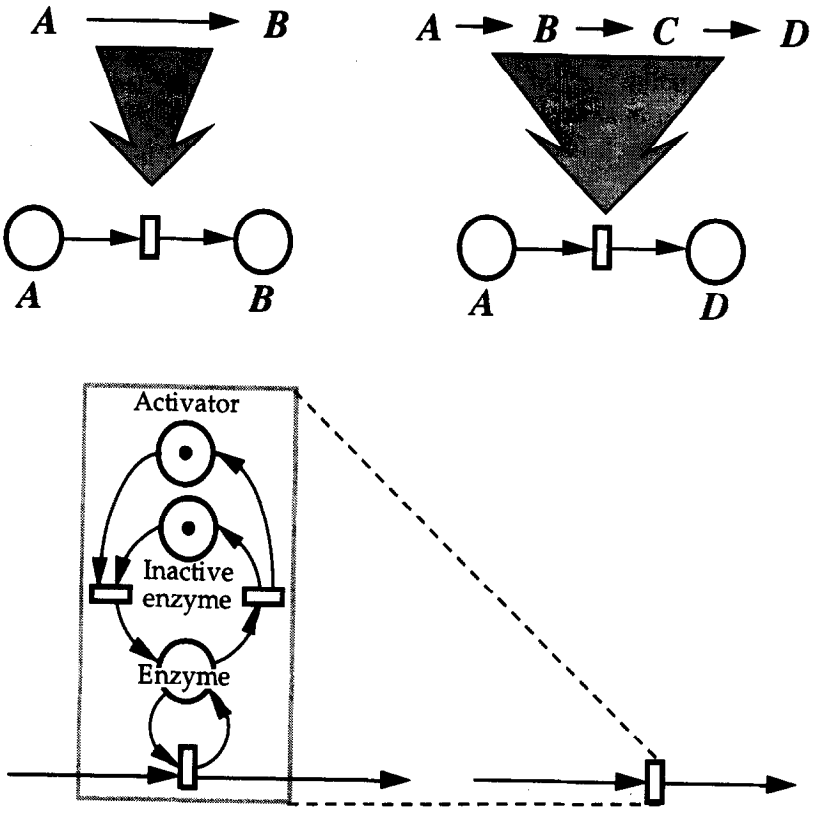

Fig. 6. A transition can represent a single reaction, a chain of forward reactions or an abstraction of a subnet.

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 Fig. 5(c), the ATP pools inside and 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 (Fig. 6). 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 either on the thermodynamic feasibility or the physiological tendency of the reaction.

EXAMPLE: METABOLIC PATHWAYS IN ERYTHROCYTES

Pathway overview

The pathways in an erythrocyte (red blood cell) are numerous as in any mammalian cell; the small set of pathways used as an illustrative example in this study includes the oxidative pentose phosphate pathway and the main glycolytic pathway [22, 23].

The overall pathway as shown in Fig. 7, in conjunction with Tables 1 and 2, defines 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 [24].

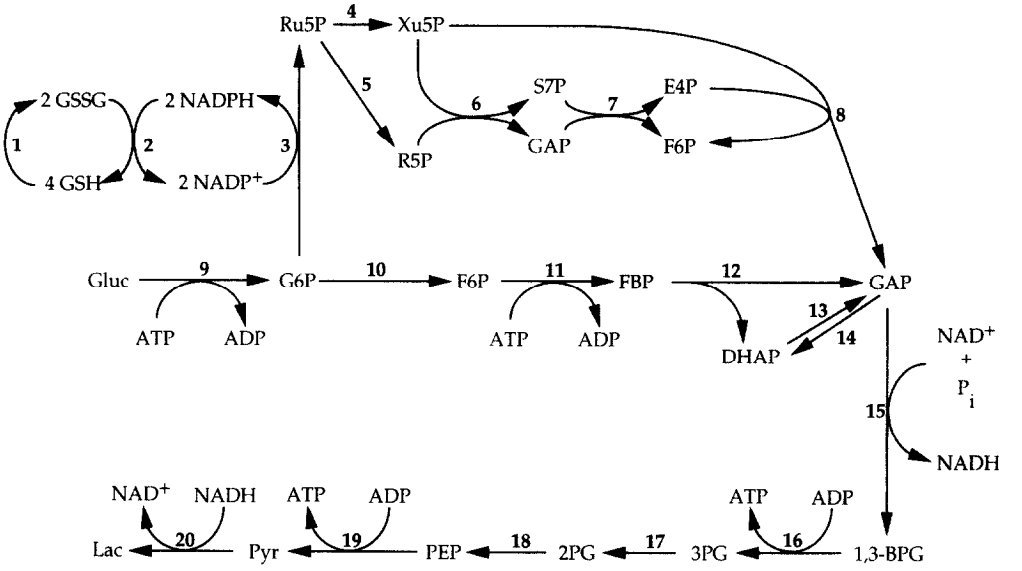

Fig. 7. The combined metabolism of the glycolytic and the pentose phosphate pathways of an erythrocyte cell.

Petri net representation

The Petri net graph of the pathway of Fig. 7 is shown in Fig. 8. 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 Tables 1 and 2. It is important to emphasize at this point that although the places labeled with an asterisk (*) occur more than once in the depiction of the Petri net, they in fact represent only one place with the appropriate label, for example ATP, ADP and F6P. The equilibrium reaction (reaction indices 13 & 14) is represented by two separate transitions since individual Petri nets must have a predefined directionality. The presence of the enzymes associated with each reaction is implicit in the representation of the transition unless otherwise noted.

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 of Petri nets defined earlier. To supplement such an analysis, we would also determine behavioral properties of the Petri net model to obtain functional qualities of the system. We illustrate a few of the properties applied in the context of a biochemical pathway.

Table 1. Mapping between reactions in the pathway and transitions in the Petri net model

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 2. Mapping between metabolites in the pathway and places in the Petri net model

Place mapping for the Petri net model			
Index	Metabolite/compound	Index	Metabolite/compound
1	Orthophosphate (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	Erythose-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**)

Model reduction. 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 Fig. 7, the corresponding abstract Petri net model in Fig. 8 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 losing the overall properties of the system.

The information of enzymatic activity can be abstracted by way of the subnet illustrated in Fig. 6. The representation of the enzymes becomes necessary when we are interested in studying the details of the influence of enzymes (such as their associated activators) on the behavior of the pathway as a whole. In Fig. 8 the oxidative reaction of glutathione with the participating peroxide is abstracted to a single transition (t_1). 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.

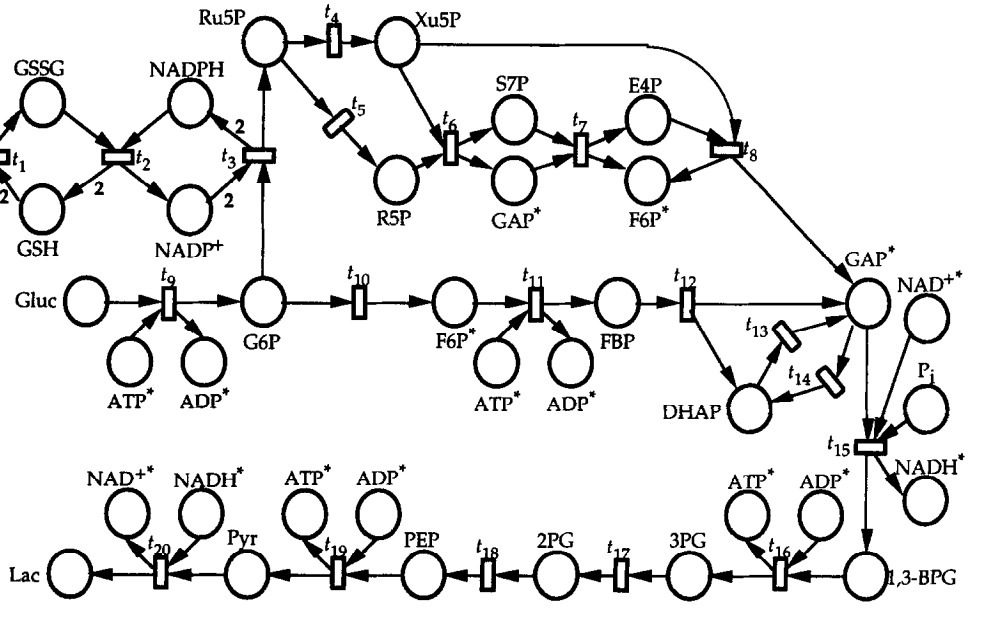

Fig. 8. Petri net graph of the erythrocyte metabolism shown in Fig. 7.

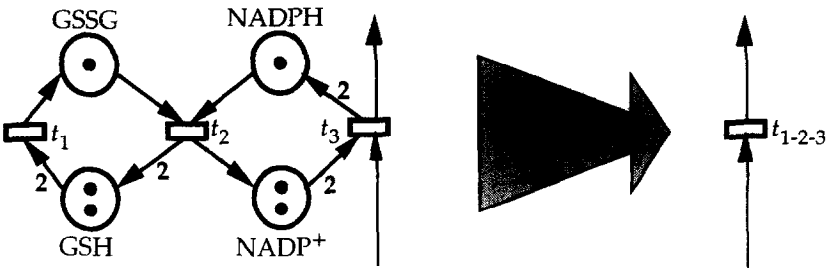

Fig. 9. Reduction of the subnet structure to a single transition.

Structural reduction: this process reduces the number of intermediate metabolites represented in the pathway. The method allows us to eliminate certain combinations of places and transitions, defined earlier, such that the change does not affect the overall behavior of the model (in terms of specific Petri net properties, such as liveness and boundedness).

In Fig. 8, if we assume that the compounds GSH, GSSG, NADPH and $NADP^+$ are always present in proportions that permit the oxidative reactions to proceed, then the structure can be reduced to a single transition as shown in Fig. 9. The rules applied here are from the simple reduction rules presented in Fig. 4.

Qualitative inference. Accumulation of metabolites (Boundedness): industrial bioprocesses usually aim to produce a particular metabolite in the maximum amount achievable; in other biological systems, the accumulation of some intermediates may be toxic, and therefore undesirable. Thus, in a large network of reactions, the identification of those metabolites that can potentially accumulate under certain conditions is useful. The boundedness of a Petri net, for a given $\mathbf{M}_0$, provides a criterion to identify such metabolite accumulation.

As an example, consider the model from Fig. 8. The equations are set up as:

$$\mathbf{A}\mathbf{y} \leqslant 0, \qquad \mathbf{y} > 0 \tag{7}$$

The Petri net is bounded if there exists a solution to the above set. 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:

$$\mathbf{M}(p) \leqslant (\mathbf{M}_0^T \mathbf{y})/\mathbf{y}(p), \tag{8}$$

where $\mathbf{M}(p)$ is any reachable marking in place p and $\mathbf{y}(p)$ is the pth element in the vector $\mathbf{y}$. For our case, if we consider the initial marking $\mathbf{M}_0 = [1, 3, 0, 3, 0, 2, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 3, 0, 0, 0, 0, 0, 0, 0, 0]^T$, the bounds on the set of places $P = \{p_2, p_4, p_6, p_8, p_{22}\}$ are {32, 24, 19.2, 9.6, 19.2}, respectively. The bounds for any general class of Petri nets are conservative; therefore the concentrations (moles) of the corresponding metabolites may not reach these bounds under any possible conditions of the pathway, given the set of initial concentrations. A method to improve this bound is by considering the minimal support S-invariants [17] demonstrated later in this section.

Invariant proportions (S-invariants): in the Petri net model of a pathway the non-zero entries of an S-invariant determine the set of compounds whose total net concentrations remain unchanged in the course of a biotransformation. This property occurs 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 from Equation (5) based on the incidence matrix for the Petri net of Fig. 8 and are as follows:

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

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

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

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

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

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

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

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 invariant. 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, as would be expected, since the combined moles of the pairs must remain constant. 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 2 moles of GSH and in the oxidative reaction 2 moles of GSH combine to form GSSG. Thus to conserve the token count within 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 this 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 by the method described in [17]. The bounds are calculated for the minimal support invariants by the expression $\mathbf{M}(p) \leq \text{Min}\,[(\mathbf{M}_0^T\mathbf{y}_i)/\mathbf{y}_i(p)]$, 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 values calculated for the set of places $P = \{p_2, p_4, p_6, p_8, p_{22}\}$ are {3, 3, 2, 1, 4}. This gives a better description of the upper bounds on the token count for selected places given the initial marking.

Continuous operation (T-invariants): T-invariants of a pathway indicate the presence of cyclic firing sequences. This can be interpreted as a condition where a set of reactions (or the entire pathway) can be in a state of continuous operation. For the model in the example, there exists only one T-invariant, $\mathbf{x} = [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0]^T$. Indeed, this corresponds to the reversible reaction (indices 13 & 14) between DHAP and GAP. No other cycles are present in this Petri net. In other cases, however, more complicated T-invariants may arise.

Deadlock free pathways (Liveness): a general method for determining the property of liveness for a Petri net has not been developed; however there are methods to determine

the nonexistence of deadlock markings [25], which are markings where no transition is enabled. This is important, from the point of view of modeling biochemical pathways, because it determines the possibility of metabolic blocks that could hinder the progress of the reactions.

A method described by Ivanov [25] determines the existence of deadlock markings in a Petri net. The method generates all (reachable and unreachable) deadlock markings using a logical predicate function, and then tests for their reachability using the matrix representation of the Petri net.

An outline of the method follows. A logical function F_i is defined based on the token count of the input places to transition i, such that $F_i = 1$ would indicate that the transition is not enabled. The equality to 1 of the conjunction K, formed by all $F_i (i = 1 \ldots n)$, would indicate the non-existence of enabled transitions (i.e. a deadlock) in the Petri net for some given marking. The expression K is reduced by identities of Boolean algebra and simplification rules [25]. The complete set of marking vectors, $\mathbf{M}(K)$ that would result in $K = 1$ is constructed. This is the set of all deadlock markings of the Petri net.

The S-invariants of a Petri net are used to describe the invariant properties of the Petri net and can be used to eliminate those markings that are not reachable. All reachable markings from some $\mathbf{M}_0$ must satisfy $\mathbf{M}^T \cdot \tilde{\mathbf{Y}} = \mathbf{M}_0^T \cdot \tilde{\mathbf{Y}}$, with $\tilde{\mathbf{Y}}$ an $m \times k$ matrix, (where $k = n\text{-}r$ and r is the rank of $\mathbf{A}$) formed by the S-invariants of the Petri net. Further analysis is performed to find those marking vectors that are actually reachable, using the matrix representation of the Petri net. If the equation $\mathbf{M} - \mathbf{M}_0 = \mathbf{A}^T\mathbf{x}$ has no solution on the set $\mathbf{M}(K)$, then the Petri net has no deadlock markings reachable from that $\mathbf{M}_0$. For Petri nets that have S-invariants, only those markings of the set $\mathbf{M}(K)$ that satisfy the invariant properties need be checked for the reachability condition.

The analysis was applied to the erythrocyte glycolysis example with different initial markings. The analysis shows that it is possible to identify those initial markings that result in a deadlock in the Petri net of the pathway. This translates, in the case of biochemical pathways, to the identification of proportions of starting substrate concentrations that could potentially lead to an insurmountable block in the course of the pathway. The analysis assumes that the proportions between the metabolites are finite and are not altered except by the normal progress of the reactions of the pathway. The procedure applied to the initial condition $\mathbf{M}_0$ (Table 3) shows the existence of deadlocks in the pathway. One such reachable deadlock marking is $\mathbf{M}_1$ that shows the formation of Xu5P in the same proportion as the substrate Gluc and the unavailability of Ru5P for subsequent reaction. This is one outcome among many reachable deadlock markings, such as $\mathbf{M}_2$ and $\mathbf{M}_3$. The initial condition $\mathbf{M}_0'$ does not lead to any deadlock markings and the justification would be the cyclic nature of the reversible reaction (13 and 14) between DHAP and GAP that could continue indefinitely. The important aspects of this analysis lie in predicting the existence of deadlocks in a pathway structure based on the presence of metabolites in certain proportions. One can systematically identify those proportions between metabolites that could potentially lead to deadlocks.

SUMMARY

This article provided an introduction to the analysis of metabolic pathways as discrete event systems, using the methodology of Petri nets. The use of discrete event and other qualitative methods is very useful for gaining preliminary insights into the behavior of biochemical pathways, even in the absence of quantitative data. The analysis of large complex networks can be handled with the same set of simple structural and behavioral properties.

The Petri net methodology is often useful in the analysis of complex biochemical pathways when it is necessary to determine the role of particular reactions. The properties of a Petri net model of the pathway may provide a preliminary test of possible experimental outcomes. In a pathway with several species and reaction steps, it is difficult to screen for instances where elimination of a reaction might block the pathway or result in the accumulation of a species. It is not necessary for a biochemist to

Table 3. Deadlock markings M_1, M_2, M_3 are reachable from the initial marking M_0. M_0' does not lead to any deadlock marking

Place	Metabolite	M_0	M_1	M_2	M_3	M_0'
1	P_i	**6**	**6**	**6**	**6**	**6**
2	**ATP**	**3**	0	0	0	**3**
3	**ADP**	0	**3**	**3**	**3**	0
4	**NAD^+**	**6**	**6**	**6**	**6**	**6**
5	**NADH**	0	0	0	0	0
6	**$NADP^+$**	**2**	**2**	**2**	**2**	**2**
7	**NADPH**	0	0	0	0	0
8	**GSSG**	**1**	**1**	**1**	**1**	**1**
9	**GSH**	0	0	0	0	0
10	**Ru5P**	0	0	0	0	0
11	**Xu5P**	0	**3**	0	0	**3**
12	**R5P**	0	0	**3**	0	**2**
13	**S7P**	0	0	0	0	0
14	**GAP**	0	0	0	0	0
15	**E4P**	0	0	0	0	0
16	**F6P**	0	0	0	**3**	**3**
17	**Gluc**	**3**	0	0	0	0
18	**G6P**	0	0	0	0	0
19	**FBP**	0	0	0	0	0
20	**DHAP**	0	0	0	0	0
21	**1,3-BPG**	0	0	0	0	0
22	**3PG**	0	0	0	0	0
23	**2PG**	0	0	0	0	0
24	**PEP**	0	0	0	0	0
25	**Pyr**	0	0	0	0	0
26	**Lac**	0	0	0	0	0

experimentally investigate the behavior of the complete reaction system when it is possible to deduce properties of interest with this method. For example, in a biochemical pathway one expects a change in reaction rates as a result of a nonfunctional enzyme. In the Petri net model of the pathway, one would examine the S-invariants containing the input or output places of the non-firing transition (nonfunctional enzyme). The invariants would contain those places that are likely to be affected by the change. Quantitative evaluation can be confined to the set of metabolites contained in the S-invariants to investigate experimentally the change in reaction rates.

The methodology of Petri nets demonstrates the use of qualitative methods as useful preliminary analysis tools for biochemical pathways. The method is easy to implement and visually comprehensible. However, the solution to some Petri net problems, notably the determination of liveness, reachability, and boundedness, has not been achieved for the general Petri net structure [18]. 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 structural classes [17] of Petri nets, which allow easier determination of properties.

The modeling power of Petri nets can also be extended by suitable modifications to the basic definition. 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 Turing machines [18]. Naturally, the extension of the modeling power exacerbates undecidability and complexity obstacles. This increase in modeling power leads to a decrease in the

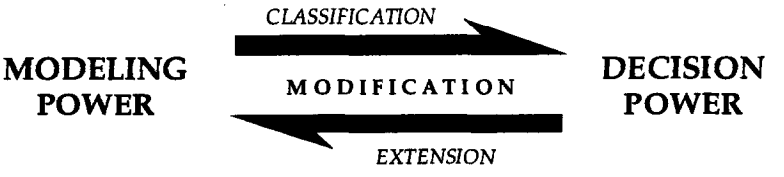

Fig. 10. Effect of modification of Petri nets on the modeling and decision (analytical) powers.

analytical power of methods for determining properties of Petri nets (Fig. 10) which is an essential 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.

Acknowledgements—This work was supported in part by the National Library of Medicine, Bethesda, MD (LM05278) and the Institute for Systems Research, University of Maryland, College Park, MD (NSFD CDR/8803012).

REFERENCES

1. M. A. Savageau, *Biochemical Systems Analysis*. Addison-Wesley, Reading, MA (1976).
2. E. O. Voit, Optimization of integrated biochemical systems, *Biotechnol. and Bioengng* **40**, 572–582 (1992).
3. M. C. Kohn and W. J. Letzkus, A graph-theoretical analysis of metabolic regulation, *J. theor. Biol.* **100**, 293–304 (1983).
4. S. A. Kaufmann, Metabolic stability and epigenesis in randomly constructed genetic nets, *J. theor. Biol* **22**, 437–467 (1969).
5. L. P. A. Robichaud *et al.*, *Signal Flow Graphs: Applications*. Prentice-Hall, Englewood Cliffs, NJ (1962).
6. J. U. Thoma, *Introduction to Bond Graphs and their Applications*. Pergamon Press, New York (1975).
7. M. C. Kohn and D. R. Lemieux, Identification of regulatory properties of metabolic networks by graph theoretical modeling, *J. theor. Biol.* **150**, 3–25 (1991).
8. D. A. Fell, Metabolic control analysis: a survey of its theoretical and experimental development, *Biochem. J.* **286**, 313–330 (1992).
9. H. Kacser and J. A. Burns, The control of flux, in *Control of Biological Processes*, D. Davies (Ed.), pp. 65–124 Cambridge University Press, Cambridge, MA (1973).
10. R. Heinrich and T. A. Rapaport, A linear steady-state treatment of enzymatic chains, *Eur. J. Biochem.* **42**, 89–95 (1974).
11. D. A. Fell and H. N. Sauro, Metabolic control and its analysis: additional relationships between elasticities and control coefficients, *Eur. J. Biochem.* **148**, 555–561 (1985).
12. C. Reder, Metabolic control theory: a structural approach, *J. theor. Biol.* **135**, 175–201 (1988).
13. A. K. Sen, Topological analysis of metabolic control, *Math. Biosci.* **102**, 191–223 (1990).
14. A. Cornish-Bowden, Metabolic control theory and biochemical systems theory: different objectives, different assumptions, different results, *J. theor. Biol.* **136**, 365–377 (1989).
15. V. N. Reddy, *Modeling biological pathways: a discrete event systems approach*, M. S. Thesis, University of Maryland, College Park, MD, May (1994).
16. V. N. Reddy, M. L. Mavrovouniotis and M. N. Liebman, Modeling biological pathways: a discrete-event systems approach, in *Molecular Modeling*, T. F. Kumosinski and M. N. Liebman (Eds), pp. 221–234 ACS Symposium Series Vol. 576. ACS Books, Washington, D.C., (1994).
17. T. Murata, Petri nets: properties, analysis and applications, *Proceedings of the IEEE* **77**, 541–580 (1989).
18. J. L. Peterson, *Petri Net Theory and the Modeling of Systems*. Prentice-Hall, Englwood Cliffs, NJ (1981).
19. W. Reisig, *Petri Nets: An Introduction*. Springer-Verlag, New York (1985).
20. V. N. Reddy, M. L. Mavrovouniotis and M. N. Liebman, Petri net representations in metabolic pathways, in *Proc of the First Intl. Conf. on Intell. Sys. for Mol. Biol.* L. Hunter, D. Searls and J. Shavlik (Eds). AAAI/MIT Press, Menlo Park, CA (1993).
21. R. Vallete, Analysis of Petri nets by stepwise refinements, *J. Comput. Syst. Sci.* **18**, 35–46 (1979).
22. H-G. Holzhütter, R. Schuster, D. Buckwitz and G. Jacobasch, Mathematical modelling of metabolic pathways affected by an enzyme deficiency, *Biomedica Biochimica Acta* **49**, 791–800 (1990).
23. R. Schuster, G. Jacobasch and H-G. Holzhütter, Mathematical modelling of metabolic pathways affected by an enzyme deficiency, *Eur J. Biochem.* **182**, 605–612 (1989).
24. D. Voet and J. G. Voet, *Biochemistry*. Wiley, New York (1990).
25. N. N. Ivanov, Algebraic method of determining nonexistence of deadlock markings in Petri nets, *Automation Remote Control* **52**, 986–989 (1991).

About the Author—VENKATRAMANA N. REDDY received his Bachelor of Technology in Chemical Engineering from the Indian Institute of Technology, Madras in 1991. He received his M.S. in Chemical Engineering from the University of Maryland, College Park in 1994. He worked as a Research Assistant at the Institute for Systems Research, College Park from 1992 to 1994. He is currently pursuing his Ph.D. at Northwestern University. His research interests are in the area of control and identification of biochemical processes; and computational modeling of metabolic reaction systems. Mr Reddy is a member of the American Institute of Chemical Engineers.

About the Author—MICHAEL MAVROVOUNIOTIS is an Associate Professor of Chemical Engineering at Northwestern University in Evanston, Illinois. He received his Diploma of Engineering from the National Technical University of Athens (Greece) in 1984, and his Ph.D. from the Massachusetts Institute of Technology in 1989. He served on the faculty of the Institute for Systems Research and the Chemical Engineering department of the University of Maryland, College Park, advancing to the rank of Associate Professor before joining Northwestern in 1993. His research interests lie in the general area of computer-aided engineering of chemical and biochemical processes. His work includes new computer-based methods for the estimation of

thermodynamic and physical properties of chemical and biochemical compounds from their molecular structure; qualitative analysis of the metabolism and the role of enzymes and metabolites; methods for understanding the structure, properties, and behavior of complex chemical reaction systems; techniques for the reduction of complex dynamic models with multiple time-scales; and the design and control of chemical processes. His work on an object-oriented environment for process engineering received the 1987 Best Paper award from the journal *Computers and Chemical Engineering*, his algorithm for the synthesis of biochemical pathways received the 1992 Best Paper award from the same journal. He is also the recipient of the 1991 Ted Peterson award of the Computing and Systems Technology division of the American Institute of Chemical Engineers.

About the Author—MICHAEL N. LIEBMAN is currently Director of Bioinformatics, Vysis, Incorporated, of Downers Grove, Illinois. Michael also serves as an Adjunct Associate Professor of Molecular and Cellular Biochemistry and member of the Graduate Faculty of Loyola University Stritch School of Medicine. Michael's research focuses on modeling clinical disease and disease processes, biochemical pathways and processes and protein structure and function using novel methods of data abstractions, Petri nets and neural nets and fourier transform infrared spectroscopy. His Ph.D. was received from Michigan State University in 1977 (Physical Chemistry) in the area of protein X-ray crystallography and followed with postdoctoral work in tRNA crystallography at the University of Wisconsin-Madison. He established the Laboratory for Molecular Modeling at the Fox Chase Cancer Center (Philadelphia) (1979–81) and was a Revson Scholar and Associate Professor in the Departments of Pharmacology, Biophysics and Physiology at Mount Sinai School of Medicine in New York (1981–88). Michael was involved in the founding of ImClone Systems, Inc. before joining Amoco Technology Company's Biotechnology Division as Senior Scientist in 1988. He transferred to ATC's Strategic Planning and Development Department as Program Manager for Bioinformatics in 1993. Vysis, Inc. was created as a wholly-owned, standalone business of Amoco Technology on 1 January 1995. Michael participated in the Illinois Working Group on Biotechnology, and is a member of Governor Edgar's Biotechnology Advisory Council, also serving on the Board of Directors of the Illinois Alliance for Biotechnology, American Cancer Society (DuPage), Central DuPage Hospital Strategic Planning Committee and several oversight committees for Amoco Corporation.