Petri Net Representations in Metabolic Pathways

Venkatramana N. Reddy*, Michael L. Mavrovouniotis*, and Michael N. Liebman†
*Chemical Engineering Department and Institute for Systems Research
A.V. Williams Bldg. (115)
University of Maryland
College Park, MD 20742, U.S.A.
and
†Amoco Technology Company
Amoco Research Center, MC F-2
Naperville, IL 60566, U.S.A.

vnreddy@src.umd.edu, mlmavro@src.umd.edu, mliebman@nap.amoco.com

Abstract

The present methods for representing metabolic pathways are limited in their ability to handle complex systems, incorporate new information, and to provide for drawing qualitative conclusions from the structure of pathways. The theory of Petri nets is introduced as a tool for computer-implementable representation of pathways. Petri nets have the potential to overcome the present limitations, and through a multitude of properties, enable the preliminary qualitative analysis of pathways.

Introduction

Living cells are composed of a wide array of compounds, and chemical reactions that occur simultaneously. A complete understanding of the behavior of these reactions is possible only through a complete analysis, and "To understand the molecular logic of cells we must be able to consider biomolecules and their interactions in both qualitative and quantitative terms" (pp. 14, Lehninger 1982). A qualitative analysis of the behavior of these reactions constitutes the qualitative study of *metabolic pathways*.

A *Metabolic pathway* within an organism, is a series of enzymatic reactions consuming certain metabolites and producing others. Often these metabolites participate in more than one metabolic pathway, forming a complex network of reactions. The analysis of this complex network involves an enormous amount of data from biochemistry and cell physiology, such as metabolic intermediates and their properties, biochemical reactions, properties of enzymes from different sources, regulation of gene expression, etc., as well as principles from physical sciences. The complexity of the problem and the limitations in existing data can be gauged by the fact that, "... the 1972 'Enzyme Nomenclature' (Elsevier, Amsterdam 1973) lists 1770 enzymes which are 'well characterized', and estimations of the enzymes in single cells run up to 5,000-10,000." (G. Michal, Biochemical Pathways Index). It is also difficult to interpret such information from the traditional representation of metabolic pathways as a two-dimensional pictorial drawing, without ambiguity. The inherent complexity involved in representation and analysis requires the use of computers. There are substantial difficulties in the computer-representation, integration, and application of the very diverse forms of information. We need to codify all the necessary information in a coherent and structured manner, to enable a computer-representation and reasoning using this metabolic knowledge.

The existing information on some of the metabolic pathways is not complete, for example there may be unknown catalyzing components, uncertainty about the role of known components, unreliable experimental data, etc. A complete quantitative analysis in this scenario is impractical and it is more justifiable to rely on whatever useful information we can derive from a qualitative analysis. The important issues in a qualitative analysis are the selection of appropriate descriptions for whole sets of pathways, the selection of operations that can be used to combine these sets and identify qualitative properties and recurring pathway structures from them. The descriptions of sets of pathways should limit the computational complexity and make results much easier to comprehend. In this context, if we initially ignore details of the dynamics (exact kinetic information) of the individual reactions and concentrate on the inter-connectivity between the metabolites through their direct mechanisms, we could model a metabolic network as a *"discrete-event" system.* The proposed methodology is based on *Petri nets*, a modeling tool for discrete event dynamic systems from a varied, and growing, number of fields which include Computer Science (Yau & Caglayan 1983), Control Engineering (Ichikawa & Hiraishi 1987, Yamalidou & Kantor 1991), Manufacturing Systems (Martinez et al. 1986) and Information Science (Berthelot & Terrat 1982, Diaz 1982).

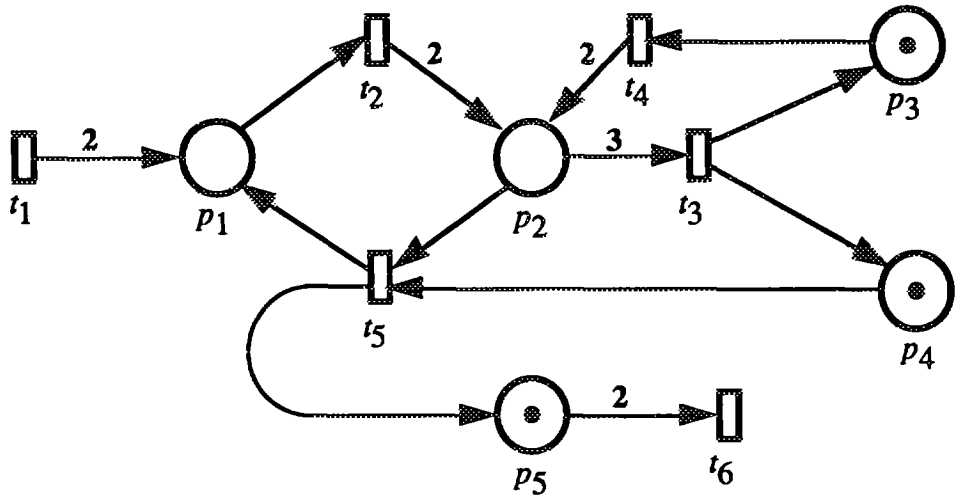

Fig. 1. A Petri net graph with *places*, *transitions* and *arcs*.

The purpose of this paper is to introduce a method of representation of metabolic pathways as Petri nets, and illustrate some useful properties that can be determined from such a representation.

Petri nets

The essential concepts in Petri net theory relevant to our analysis are summarized below. This is merely an introduction, and a complete review of the theory and applications are available in literature (Murata 1989; Peterson 1981).

A Petri net (PN) is a graph (Fig. 1.) formed by two kinds of nodes, called *places* (p_j) and *transitions* (t_j). Directed edges, called *arcs*, connect places to transitions, and transitions to places. A non-negative integer number of *tokens* is assigned to each place, and it can vary based on the state of the Petri net. Each arc has a weight, a positive non-zero integer, assigned to it. Pictorially, places are represented by circles, transitions by bars or boxes, arcs by lines ending in an arrow, and tokens as black dots placed in the circles. Generally if there is no arc-weight specified on the graph we assume it to be equal to one.

Each transition (event) is associated with a finite number of input places (pre-conditions) and output places (post-conditions). 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 Fig. 1., the transitions t_1 and t_4 are enabled, while the rest are not. An enabled transition can fire, depositing tokens in its output places, again their number determined by the arc-weights. A transition with no output places, called a *sink transition*, can fire when enabled consuming the tokens from its input places. Fig. 2. shows the same PN after the enabled transitions are fired, in turn enabling other transitions. The state of a PN, which is the number of tokens present in the individual places, is denoted by M and called the *marking* of the net. The initial marking of a PN is denoted by M_0. Thus, the firing of a sequence of transitions may change the marking of the net. A marking M is *reachable* from M_0 if there exists some firing sequence that accomplishes this change.

Mathematically a Petri net is represented as PN = (P, T, E, W, M_0) ; where

$P = \{ p_1, p_2, p_3, \ldots p_m \}$ is a finite set of places

$T = \{ t_1, t_2, t_3, \ldots t_n \}$ is a finite set of transitions

$E \subseteq (P \times T) \cup (T \times P)$ is a set of arcs

$W: E \rightarrow \{ 1, 2, 3, \ldots \}$ is a weight function

$M_0: P \rightarrow \{ 0, 1, 2, \ldots \}$ 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 which can be used to specify and manipulate the state of the system. We can examine the dynamic behavior of a Petri net, with m places and n transitions, if we formulate a state-space equation of the type

$$M_k = M_{k-1} + A^T u_k , \quad k = 1, 2, 3, \ldots \qquad (1)$$

where, k represents a point in a firing sequence; M_k an m × 1 vector, the marking after the kth firing; u_k an n × 1 vector, the *control vector* indicating the

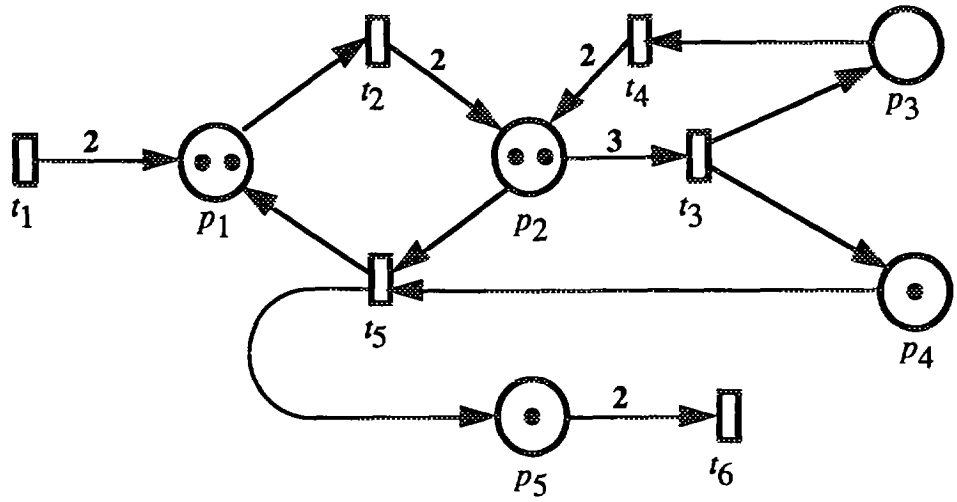

Fig. 2a. Marking after firing enabled transitions t_1 and t_4 from Fig. 1.

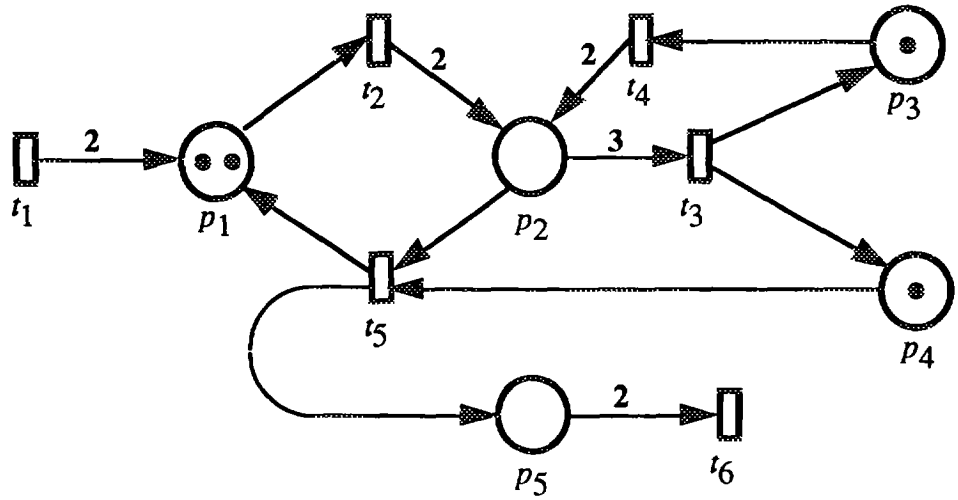

Fig. 2b. Marking after firing in sequence t_2, t_3, t_5, and t_6 from Fig. 2a.

transition fired at the kth firing; and A an n × 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.

If a particular marking M_n is reached from the initial marking M_0, through a *firing sequence* $\sigma = \{ u_1, u_2, u_3, \ldots u_n \}$, and the state-equations are summed for all the firings in this σ, we obtain

$$M_n = M_0 + A^T \sum_{k=1}^{n} u_k \qquad (2)$$

if we define $x = \sum_{k=1}^{n} u_k$, we obtain

$$M_n - M_0 = A^T x \qquad (3)$$

or

$$A^T x = \Delta M \qquad (4)$$

where x, an n × 1 vector, is called the *firing count vector*. The element i in x indicates the number of times transition i must fire in order to transform M_0 to M_n.

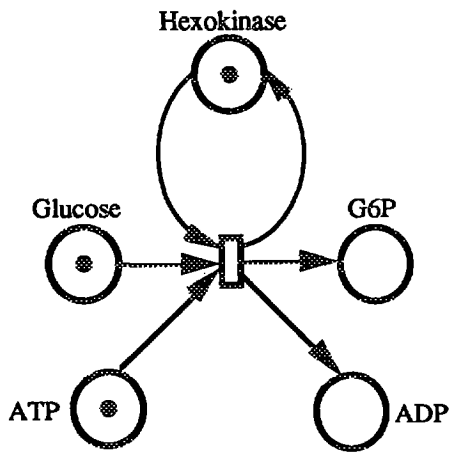

Fig. 3a. Each place represents a biological compound. G6P: Glucose-6-Phosphate.

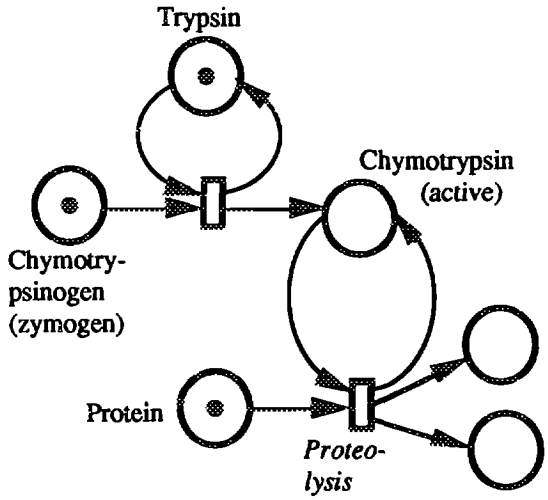

Fig. 3b. In this case two places represent forms of the same compound, one is the zymogen and the other the enzyme.

It is now evident that a necessary condition for the reachability of M_n from M_0 is the existence of a solution to Eqn. (4), the elements of x being non-negative rational or integer numbers. Although the rational solutions for x have no physical interpretation, since fractional firings of transitions are not realizable, they can be transformed to integer numbers by multiplying Eqn. (4) with an appropriate scalar constant. This does not alter in any way our condition for the reachability of M_n from M_0.

Metabolic Pathways

Representation

The representation of the essential components in a metabolic pathway, using Petri net terminology, is the first step in modeling the metabolic network as a discrete-event system. Petri nets, among other methods, have been reported (Kohn & Letzkus 1983) as graph theoretical tools for the representation of biochemical reaction systems. Although the representation of the system as a Petri net is not unique, proper choice of mapping is essential in determining the utility and applicability of the method.

For metabolic pathways, places would represent compounds (such as metabolites, enzymes, cofactors etc.) participating in a bioreaction. We could have one place represent one compound or we could have two or more places represent the same compound, based on their physical attributes or distinct functions. For example, in Fig. 3a, each place represents one biological compound, whereas in Fig. 3b, two places represent the same compound, due to their difference 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 (or organelle), we could have different places represent the same compound. For example, the ATP pools inside and outside the mitochondrion in a cell are different and their relative concentrations 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.

In the same way we assign transitions to represent individual reactions or a series of forward reactions when the intermediary compounds are not of primary interest. Tokens indicate the presence of a compound. Arc-weights represent the stoichiometry of reactions, and the direction of an arc is based on the thermodynamic feasibility of the reaction. The example in Fig. 4b, illustrates these mappings. To keep the figures simpler, we allow the same place to appear more than once in the drawing of a PN, such as ATP in Fig. 4b. From the viewpoint of Petri net simulation and properties, we consider all of the appearances of ATP in Fig. 4b. to represent only one place.

A representation of a pathway should allow 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. Thus, it is possible to examine the difference in behavior of the pathways in

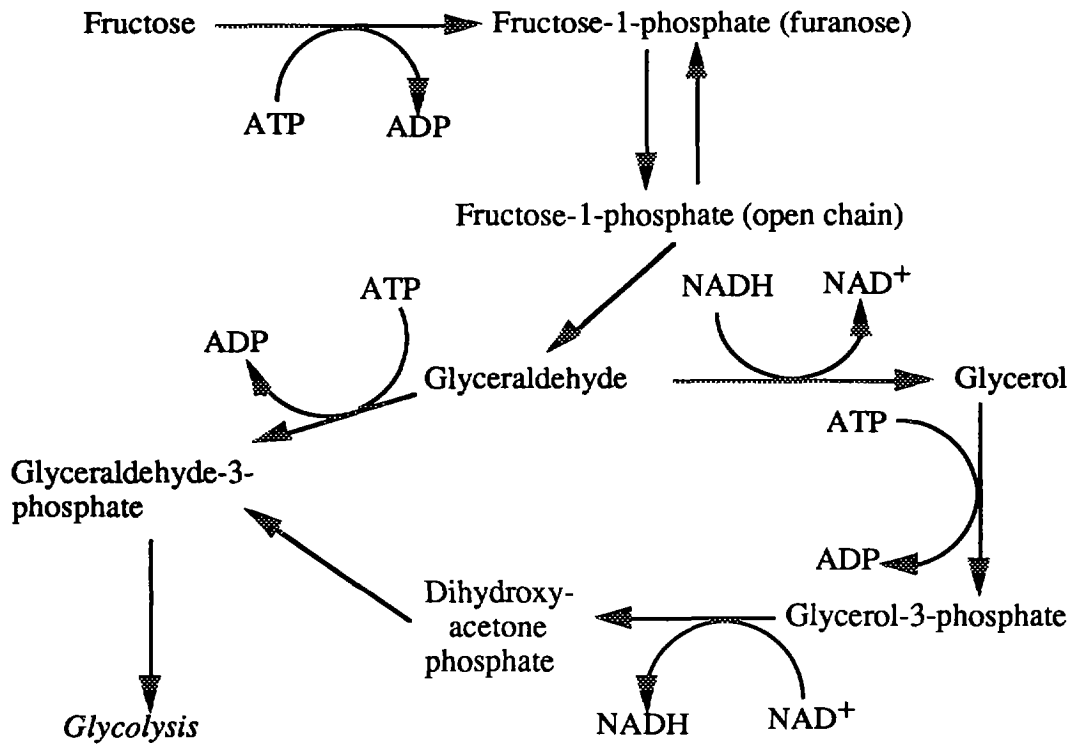

Fig. 4a. Fructose metabolism in liver (pp. 454, Voet and Voet 1990). All the reactions are catalyzed by enzymes.

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.

The present knowledge of metabolic reactions is incomplete and we would like a representation which can be extended upon the present state of knowledge, without significant deviation from the existing structure. Petri nets have this extendibility embedded in the methodology. A transition represents an event which requires certain pre-conditions and results in some post-conditions if the event actually occurs. If we now consider this event as a combination of other events, we can visualize a transition that emulates a Petri sub-net. Any modification to this sub-net is reflected in the behavior of the original transition. This can be well understood in the context of an Object Oriented Programming framework, where one object inherits attributes of other objects that it contains. For instance, in Fig. 4b, the transitions have places corresponding to the metabolites involved in the reaction. Even though there may be other compounds participating in the reaction, such as enzymes, coenzymes, metal ions, etc., they are implicit in the action of the transition. However, in the event that new information is to be incorporated, it is a simple task to modify the sub-net characterized by the transition, to reflect the change in our state of knowledge.

Properties of Pathways

The representation of metabolic pathways as Petri nets is of value because the properties of Petri nets can be applied to the pathway, leading to relevant observations.

Liveness. A Petri net is said to be *live*, if from any marking reachable from M_0 it is possible to fire any transition in the net through some further firing sequence. For a metabolic pathway this is a condition that, in any state of the biotransformation, all the individual reactions are potentially active. If the PN of Fig. 4b. is live, then one mole of each of ATP, NAD^+, NADH, and Fructose is sufficient to enable each enzyme to act at least once.

Reachability. A necessary condition for a marking M_n to be reachable from the initial marking M_0 is the existence of a non-negative solution to Eqn. (4). This condition becomes sufficient if we add the following requirement. All markings in the firing sequence from M_0 to M_n must be *coverable*, *i.e.* there must exist a

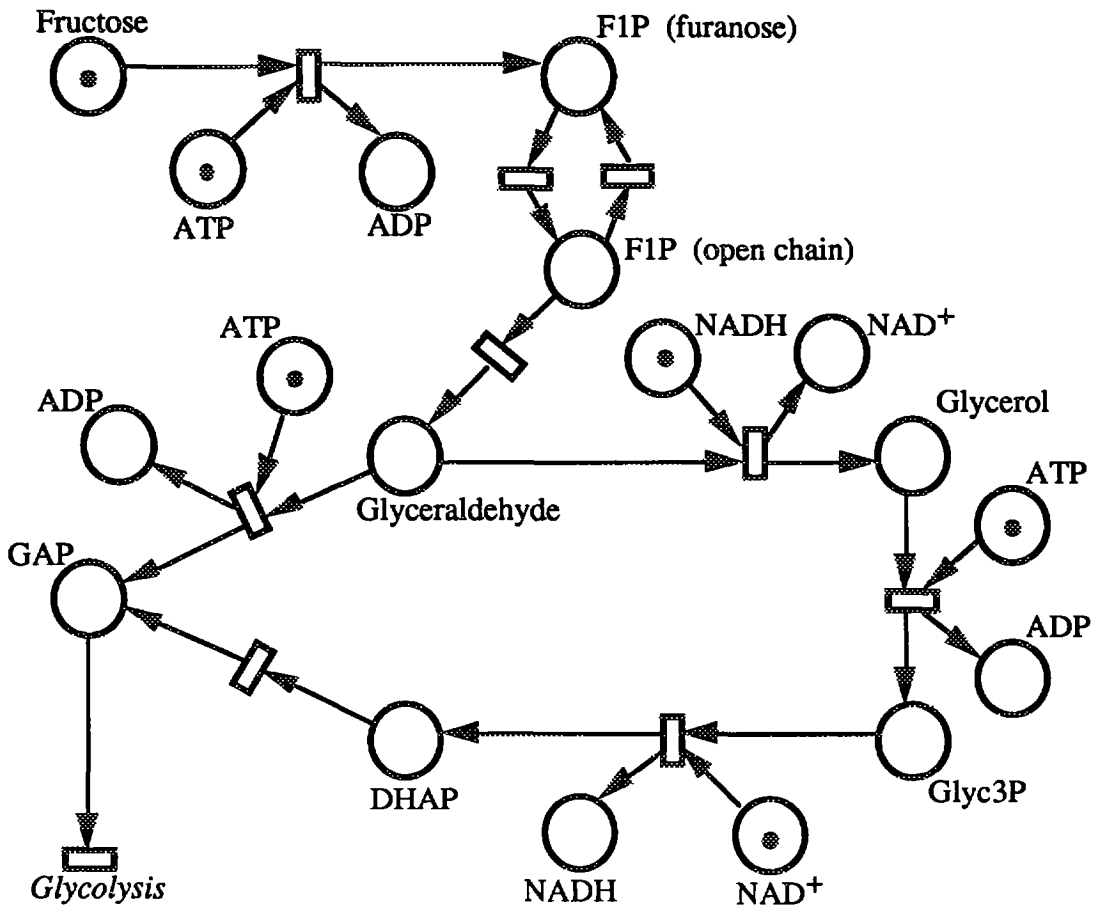

Fig. 4b. PN representation of pathway in Fig. 4a. There is an one-to-one mapping between places and metabolites.

minimum marking such that the transitions in the firing sequence from M_0 to M_n are enabled.

The reachability of a marking from some other marking, in the PN model of a metabolic pathway, determines the possibility of formation of a specified set of product metabolites from another set of reactant metabolites, by some sequence of reactions which is dictated by the PN's firing sequence(s). The sufficiency condition can be envisioned as the necessity for the presence of certain essential compounds, like the enzymes to particular reactions or their required cofactors. In the PN model of Fructose metabolism, the path to Glycolysis is possible since there is at least one possible firing sequence from Fructose to GAP (*viz.* Fructose, F1P, Glyceraldehyde, GAP, and finally Glycolysis)

Reversibility. A Petri net is said to be *reversible* if the initial marking M_0 is reachable from all other possible markings in the set of markings reachable from M_0. A marking M' is a home state if it is reachable from all other markings in the set of possible markings of a Petri net.

Most metabolic pathways are not literally reversible, due to the thermodynamic irreversibility of many reactions. However an alternate set of reactions may reproduce the precursor metabolites. This constitutes a regulatory function which can shift the metabolism in either direction, with the aid of pertinent enzymes, depending on the need for the compounds. Reversibility of a Petri net is a property which could test the possibility of such an alternate set of reactions. One such reversible pair of pathways which can be modeled and analyzed with this property is the Glycolysis-Gluconeogenesis system.

Fairness. There are different definitions to fairness, but the one that is of relevance is *Unconditional* (or *Global*) *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 M_0 is globally fair.

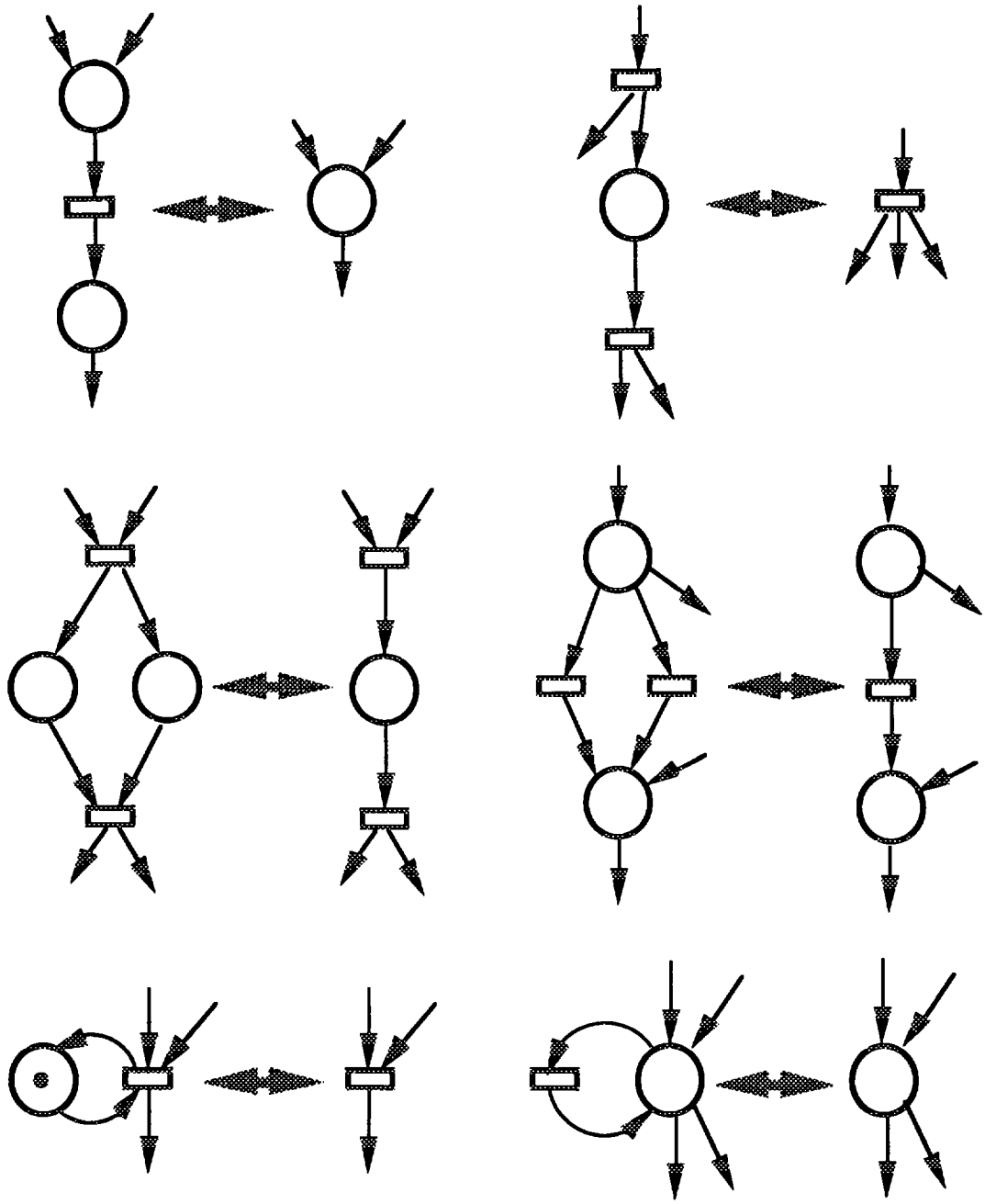

Fig. 5. Some examples of structural reductions that are possible in PN graphs

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 infinite amounts. Although not completely elucidated, an important feature of the blood-coagulation pathway is the ability of the pathway to be activated by a relatively small tissue injury, which results in a nearly continuous cascade of reactions, leading to a Fibrin cross-linkage at the injury site. Intuitively, we can expect that the PN of the blood-coagulation pathway is globally fair.

Structural reduction. Large Petri nets can be reduced by the substitution of certain combinations of places and transitions with smaller units (Fig. 5.) without sacrificing the original properties of the net (pp. 553, Murata 1989). This is particularly useful in our context, since metabolic pathways are complex networks and

any concession in the structure leads to a simpler analysis. We could view this as a method for model reduction.

Invariants. *S-invariants* are the solutions to the equation

$$Ay = 0\,, \qquad y \geq 0. \tag{5}$$

where, A is the incidence matrix and y is an $m \times 1$ vector. The non-zero entries in y, constitute the *support* of a S-invariant. The support is the set of places whose token count does not change with any firing sequence from M_0. The support of an S-invariant is analogous to the set of compounds in a pathway that do not undergo any net change in the course of a biotransformation.

T-invariants are of solutions to the equation

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

where, x is an $n \times 1$ vector. The *support* of a T-invariant denotes the firing count vector of transitions which have to fire, from some M_0, in order to return the Petri net to the same M_0. Supports to T-invariants give an insight into the direct mechanisms which are necessary to form a cyclic pathway.

A support is *minimal* if it does not contain another, non-trivial support. The invariant corresponding to a minimal support is called the *minimal-support invariant.* All invariants of a Petri net are a linear combination of all possible minimal-support invariants. Minimal-support invariants are the smallest sets of reactions which may function as futile cycles in a given bioreaction system.

Discussion

We have described a method of representation for modeling metabolic pathways as discrete-event systems. The method delineates the dynamics of the individual reactions, and emphasizes a discrete system-oriented approach. The modeling of biochemical systems by Petri nets is appealing because it is simple in its application, is visually comprehensible, and allows computer manipulation. Another advantage in Petri net representation is that it can be further extended or modified to accommodate specific attributes required for modeling of different systems. For instance, the inclusion of *inhibitor arcs* in our Petri net representation of a pathway, which disable a transition if the input place to that transition is marked, can model the behavior of enzyme inhibition and feed-back regulation. *High-Level Nets* (Genrich & Lautenbach 1981) and *Timed Petri Nets* (pp. 570, Murata 1989) are results of such extensions.

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 would 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. 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.

The properties of the Petri net of a pathway also provide valuable information on appropriate means of intervention. For example, an analysis of a Petri net of a metabolic pathway which is in continuous operation, might yield a set of transitions (reactions) and/or markings (essential biological components) which culminate in such behavior. 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.

In conclusion, the representation of metabolic pathways by Petri nets is a promising modeling approach and should enable many types of qualitative analysis of pathways.

References

Berthelot, G., and Terrat, R. 1982. Petri Nets Theory for the Correctness of Protocols. *IEEE Trans. Commun.* COM-30(12): 2497-2505.

Diaz, M. 1982. Modeling and Analysis of Communication and Cooperation Protocols using Petri Net based Models. *Comput. Networks* 6: 419-441.

Genrich, H. J., and Lautenbach, K., 1979. The Analysis of Distributed Systems by means of Predicate/Transition-nets. *Lecture Notes in Computer Science, vol. 70, Semantics of Concurrent Computation.* New York: Springer-Verlag.

Ichikawa, A., and Hiraishi, K., 1987. Analysis and Control of Discrete Event Systems represented by Petri Nets. *Lecture Notes in Control and Information*

Sciences, vol. 103, Discrete Event Systems: Models and Application. New York: Springer-Verlag.

Kohn, M. C., and Letzkus, W.J., 1983. A Graph Theoretical Analysis of Metabolic Regulation. *J. Theor. Biol.* 100: 293-304.

Lehninger, A. L., 1982. *Principles of Biochemistry.* Worth Publishers, Inc.

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

Murata, T., 1989. Petri Nets: Properties, Analysis and Applications. *Proceedings of the IEEE* 77(4).

Peterson, J. L., 1981. *Petri Net Theory and the Modeling of Systems.* New Jersey: Prentice-Hall, Inc.

Voet, D., and Voet, J. G., 1990. *Biochemistry.* John Wiley & Sons, Inc.

Yamalidou, E. C., and Kantor, J. C., 1991. Modeling and Optimal Control of Discrete-event Chemical Processes using Petri Nets. *Computers Chem. Eng.* 15(7): 503-519.

Yau, S. S., and Caglayan, M. U., 1983. Distributed Software System Design representation using Modified Petri Nets. *IEEE Trans. Software Eng.* SE-9(6): 733-745.