Biochemical Reaction Pathway Analysis Using Petri Net

N. Balasubramanian and Chuei-Tin Chang
Department of Chemical Engineering, National Cheng Kung University
Tainan, Taiwan- 70101, ROC, Fax: 886-6-234-4496, E mail: ctchang@mail.ncku.edu.tw

Abstract
A systematic procedure has been developed to construct Petri net (PN) models for the biochemical reaction networks. The component models are first constructed for characterizing the elementary reactions and gene regulation mechanisms. These components are then assembled to model the entire reaction network. All possible reaction pathways leading to the desired product have been identified by the reachability analysis. Simulation studies have also been carried out to analyze the dynamic behavior of Glycolysis reaction system.

1. Introduction
Any biochemical process involves a complex network of elementary reactions. The throughput of a biochemical reaction is dependent upon its carbohydrate intake and also the enzyme catalyst activity. In the biochemical reaction network, a pathway is defined as a set of reactions that converts an input carbohydrate into an output metabolite. The reaction pathway can be manipulated with a proper selection of enzymes. The performance of a biochemical reaction system can thus be improved by genetic modification. For example, the gene regulation mechanisms can be altered to extend an existing pathway to obtain a new product, to create new arrays of enzymatic activities that synthesize a novel structure, to shift the metabolite flow to a desired product and accelerate the rate-determining step [Baily, 1991; Nielson, 2001]. Consequently, the development of an accurate model should be considered as a critical step in designing any biochemical reaction system.

There has been considerable attention in the literature on the issue of biochemical reaction network modeling. Conventionally, these reactions are modeled with a system of differential equations, e.g. the Mechaleis Menton equation [Chen *et al.*, 1999]. The main drawbacks of this approach are (1) the computation demand is overwhelming and (2) it is difficult to represent the reaction and enzyme generation mechanisms explicitly. In recent years, the Petri net (PN) has emerged as a promising alternative to model the biochemical reaction networks. This is because of the fact that it possesses the mathematical formalism to model, analyze and simulate discrete systems with inherent concurrency. Reddy *et al.* [1996] initiated the first attempt in this area. A discrete model has been proposed for representing the reaction mechanisms that yield various metabolites. Hofestadt and Thelen [1998] have modified this model by incorporating catalytic and kinetic information in the Petri net. In the mean time, Goss and Peccound [1998] modeled gene regulatory network using a stochastic Petri net. Koch *et al.* [1999] studied the pentose phosphate reaction cycle using time dependent Petri net. Matsuno *et al.* [2000] proposed a λ page switch mechanism with hybrid Petri net. However, since the scopes of these studies are restricted to either pathway identification or kinetics analysis, the available PN models are really unsuitable for design applications. It is thus the intention of the present work to develop a systematic procedure to integrate *both* metabolic reaction and gene regulation mechanisms in a more comprehensive model.

2. Development of the component models
The ordinary PN structure is a 5 tuple: PN =(P, T, F, W, M_0), where P = $\{p_1,p_2, \ldots.p_m\}$ is a set of places; T = $\{t_1,t_2, \ldots.t_n\}$ is a set of transitions; F $\subseteq$ (PXT) $\cap$ (T XP) is a set of arcs; W:F$\rightarrow${1,2,…}is arc weight; M_o:p $\rightarrow${0,1,2,3,…} is the initial marking. A formal definition of the ordinary PN and its properties can be found in Murata [1989]. In a discrete-event system, each condition can be modeled by a place and each event can be modeled by a transition. In a later

study, Drath *et al.* [1998] proposed the hybrid Petri net to model the dynamic behavior in various man-made and natural systems. Other than the basic elements used in ordinary PN, the token quantities can be real numbers and the weights can be continuous functions of the token quantities in the input places.

Generally speaking, two types of mechanisms are involved in a biochemical reaction network: the production of required enzymes through protein biosynthesis and the production of biochemical products. A biochemical reaction is usually catalyzed with a specific enzyme. Each enzyme is produced in a distinct protein synthesis process. Following are brief descriptions of the corresponding component models:

2.1. Biochemical reaction mechanisms

Let us consider a simple catalytic biochemical reaction as an example. The reaction can be modeled with the Petri net in Figure 1. The reactant is associated with a continuous input place. Its token numbers represents the reactant concentration. On the other hand, the availability of enzyme is represented with a discrete input place. Each reaction step is denoted by a time-delayed transition and the rate of reaction is reflected in the arc weights. A transition is enabled only when the reactant reaches sufficient concentration and the enzyme is available. The concentrations of metabolites are very important as they directly influence the reaction activity.

2.2. Gene regulation mechanisms

The genetic information storehouse, DNA, controls the protein synthesis process. The enzyme-producing process starts when DNA receives the demand for protein requirement and passes information to RNA polymerase. Once informed, RNA polymerase produces three types of functional RNA molecules: ribosome RNA (rRNA), messenger RNA (mRNA) and transport RNA (tRNA). The functional RNA molecules then undergo further processing to produce the desired protein.

The component model for the gene regulation mechanism is shown in Figure 2. A token is placed in the protein demand place when an enzyme is required. This activates RNA polymerase to produce functional RNA molecules i.e. mRNA, rRNA and tRNA. A token in the rRNA place will produce ribosomes. Part of the ribosome molecules convert the messenger RNA to produces a token in the aminocycle demand place. On receiving the information transport RNA generate aminacycle synthestate (atRNA) which reacts with ribosome to give the desired protein.

3. The Petri net for modeling biochemical reaction systems

The Petri net for modeling biochemical reaction systems can be constructed by combining the component models developed in the previous section. For the illustration purpose, the Glycolytic cycle has been considered in the present work, which is an important step in the cell metabolism. Table 1 lists the individual elementary reactions while the symbols for metabolites/enzymes are defined in Table 2. A component model has been developed for each individual reaction. These components were assembled according to the reaction scheme presented in Table 1 and then combined with the Petri net representing the enzyme-producing mechanisms. Figure 3 gives the complete PN system model for Glycolyis cycle in which the glucose is converted into lactate through series of reactions.

4. Systematic generation of reaction pathways

One of the potential applications of a reaction system model is to construct all possible reaction pathways from a given reactant to the desired product. Reachability is a fundamental property of the discrete PN for studying the firing sequence to reach the final state from a given initial condition. Let p and t denotes respectively the number of places and transitions in a Petri net. A marking $\mathbf{M}_k$ of PN is a (p x 1) column vector of nonnegative integers [Murata, 1989]. The j^{th} entry

in $\mathbf{M}_k$ (j = 1, 2, 3,..., p) is the number of tokens in the j^{th} place. A marking $\mathbf{M}_n$ is said to be reachable from an initial marking $\mathbf{M}_0$ if there exist a firing sequence which transforms $\mathbf{M}_0$ to $\mathbf{M}_n$. The firing sequence is denoted by $s = \mathbf{M}_0\ t_1\ \mathbf{M}_1\ t_2\ \mathbf{M}_2\ t_3 \ldots t_n\ \mathbf{M}_n$ or simply $s = t_1\ t_2\ t_3 \ldots t_n$. The set of all possible markings reachable from $\mathbf{M}_0$ in a PN is denoted by $R(\mathbf{M}_0)$ and is referred to as the *reachability set*. The reachability set of a PN can be represented by a tree called *reachability tree* whose nodes are the markings of the PN under consideration and whose arcs represent the possible changes in state resulting from the firing of transitions. Towards constructing the reachability tree, some modifications must be introduced in the system model in Figure 3. In particular, the continuous places should all be converted to discrete places and appropriate arc weight of two has been assigned wherever branching reactions occurs. For example, an arc weight of two has been assigned to the arc connects G6P and reaction transition to produce Ru5P. In addition, the component model for gene regulatory mechanisms should be excluded from reachability analysis as it does not have any role in reaction pathway identification. To produce the tree, a token should be first placed in the place representing the raw material and let the transitions to fire sequentially in the modified system PN. The complete reachability tree can be constructed according to the history of token movements in all possible firing sequences. Consequently, the reaction pathways, i.e. the firing sequences leading to the final product, can also be identified from this analysis (see Figure 4).

5. Simulation studies of the Glycolysis reaction system

In this work, an attempt has been made to check the validity of the developed model with simulation studies. All the reaction rates in this work are described with the Michaleis Menton kinetic expression and the delay time associated with every timed transition is 5 sec. The required enzyme is generated whenever there is a demand. Say for example, the enzyme hexakinase is produced once the token number in the place representing Glucose reaches a threshold value. Then the transition representing reaction is enabled to produce G6P. Figure 5 shows the model prediction of glucose concentration. Notice that the experimental data reported in Theobald *et al.* [1997] is also presented for comparison purpose. This result shows that the present model can be used to describe the dynamic behaviors of biochemical reactions in the Glycolytic cycle satisfactorily.

6. Conclusion

As mentioned before, it is the intention of the present study to integrate both the biochemical reactions and gene regulation mechanisms in a comprehensive model. A systematic procedure has been developed to construct the Petri nets for modeling elementary biochemical reactions. The PNs for the individual reactions are then assembled together to produce a complete reaction network model. All the possible reaction pathways have been identified on the basis of reachability analysis. The gene regulatory mechanisms for the generation of required enzymes have been modeled separately and then attached to the PN representing reaction network. Simulation studies have also been carried out for the glycolysis cycle. By comparing with the experimental data, it can be observed that the biochemical reaction systems can indeed be properly modeled with the Petri nets.

7. References

Bailey, J., 1991, Toward a science of Metabolic Engineering, *Sci.*, 252, 1668.

Chen, T., H.L. He and G.M. Church, 1999, Modeling gene expression with differential equations., *Pacific symp. Bio-comput.*, 29.

Goss, P.J.E. and J. Peccouds, 1998, Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets' *Proc. Natl. Acad. Sci* 6750

Hofestädt, R. and S. Thelen, 1998, Quantitative modeling of biochemical networks, *http:/www. bioifo.de/isb/1998/01/006main/html/*

Koch, I., S. Schuster and M. Heiner, Simulation and analysis of metabolic networks by time

dependent Petri nets, 1999, German Conference of Bioinformatics Poster Abstract, *http://www.bioinfo.de/isb/ gcb99/poster/koch/*

Drath, R., 1998, Visual Object Net++, *http://www.daimi.au.dk/PetriNets/tools/complete_de.html*.

Matsuno, H. and A. Doi, 2000, Hybrid Petri net representation of gene regulation network, *Proc. Pac. Symp. Bio-comput*, 338.

Murata, T., 1989, Petri nets: Properties, analysis and applications, *Proc. IEEE*, *5*41.

Nielson, J., 2001, Metabolic Engineering, *Appl. Biotech., 55*, 263.

Reddy, V.N., M. N. Libeman and M.L. Mavrovouniotis, 1996, Qualitative analysis of biochemical reaction systems. *Comput. Biol. Med.*, 9, 26.

Theobald, U., W. Mailinger, M. Baltes, M. Rizzi, and M. Reuss, 1997, In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: I. Experimental observation. *Biotech .and Bioeng.*, 55, 305.

Table 1: Stoichiometry of the reactions involved in the Glycolysis process

1. $Glu + ATP \xrightarrow{HK} G6P + ADP$
2. $G6P \xrightarrow{PGI} F6P$
3. $F6P + ATP \xrightarrow{PFK} FDP + ADP$
4. $FDP \xrightarrow{ALD} DHAP + GAP$
5. $DHAP \xrightarrow{TPI} GAP$
6. $GAP + NAD \xrightarrow{GAPDH} 13PG + NADH$
7. $13PG + ADP \xrightarrow{PGK} 3PG + ATP$
8. $3PG \xrightarrow{PGM} 2PG$
9. $2PG \xrightarrow{ENL} PEP$
10. $PEP + ADP \xrightarrow{PK} PYR + ATP$
11. $PYR + NADH \xrightarrow{LDH} LAC + NAD$
12. $G6P + NADP \xrightarrow{PGDH} NADPH + Ru5P$
13. $Ru5P \xrightarrow{XPI} Xu5P$
14. $Ru5P \xrightarrow{RPI} R5P$
15. $R5P + Xu5P \xrightarrow{TKI} S7P + GAP$
16. $S7P + GAP \xrightarrow{TKI} E4P + F6P$
17. $E4P + Xu5P \xrightarrow{TKI} GAP$
18. $GSSG + NADHP \xrightarrow{GLTR} NADP + GSH$
19. $GSH \xrightarrow{GLTO} GSSG$

Table 2: The metabolites and enzymes involved in the Glycolysis process.

Symbol	Metabolite/Enzyme	Symbol	Metabolite/Enzyme
Glu	Glucose	HK	Hexokinase
G6P	Glucose 6 Phosphate	PGI	Phosphoglucose isomerase
F6P	Fructose 6 Phosphate	PFK	Phosphofructokinase
FDP	Fructose 1,6 bi phosphate	ALD	Adolase
GAP	Gleceraldehyde 3 Phosphate	GAPDH	Gleceraldehyde phosphate dehydeogenase
13PG	1,3 Diphoshoglycerate	PGK	Phosphoglycerate kinase
3PG	3 Diphosphoglycerate	ENL	Enolase
2PG	2 Diphosphoglycerate	PK	Pyruvate kinase
PEP	Phosphoenolpyruvate	LDH	Lactate dehydrogenase
PYR	Pyruvate	PGDH	Glucose 6 phosphate dehydrogense
LAC	Lactate	RPI	Ribose 5Phosphate isomerase
Ru5P	Ribulose 5Phosphate	XPI	Xylulose 5Phosphate epimerase
Xu5P	Xylulose 5Phosphate	TKI	Transketolase
S7P	Sedoheptulose 7Phosphate	GLTR	Glutathione reductase
E4P	Erythose 4Phosphate	TPI	Triosephophate isomerase
ADP	Adenosine Diphosphate	PGM	Phophoglycerate mutase
ATP	Adenosine Triphosphate	TPI	Triphosphate isomerase
NAD	Nicotinamide adenine dinucleotide	GLTO	Glutathione oxidation reaction
NADH	Nicotinamide adenine dinucleotide Phosphate	NADP	Nicotinamide adnine dinuclotide oxidized form
DAHP	Dihydroxyacetone phosphate	NADHP	Nicotinamide adnine dinuclotide reduce form
GSH	Glutathione	GSSH	Glutathione disulfide
R5p	Ribose 5Phosphate		

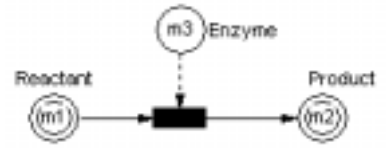

Figure 1: PN model for catalytic biochemical reaction

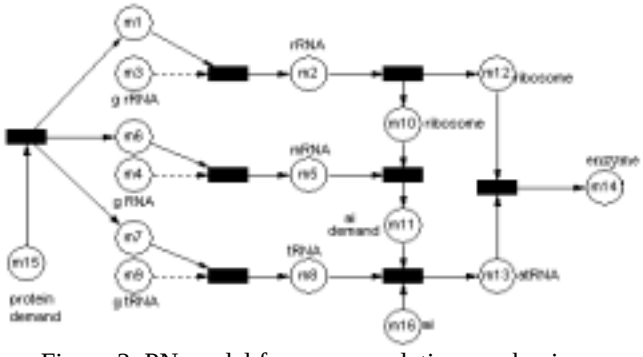

Figure 2: PN model for gene regulation mechanism

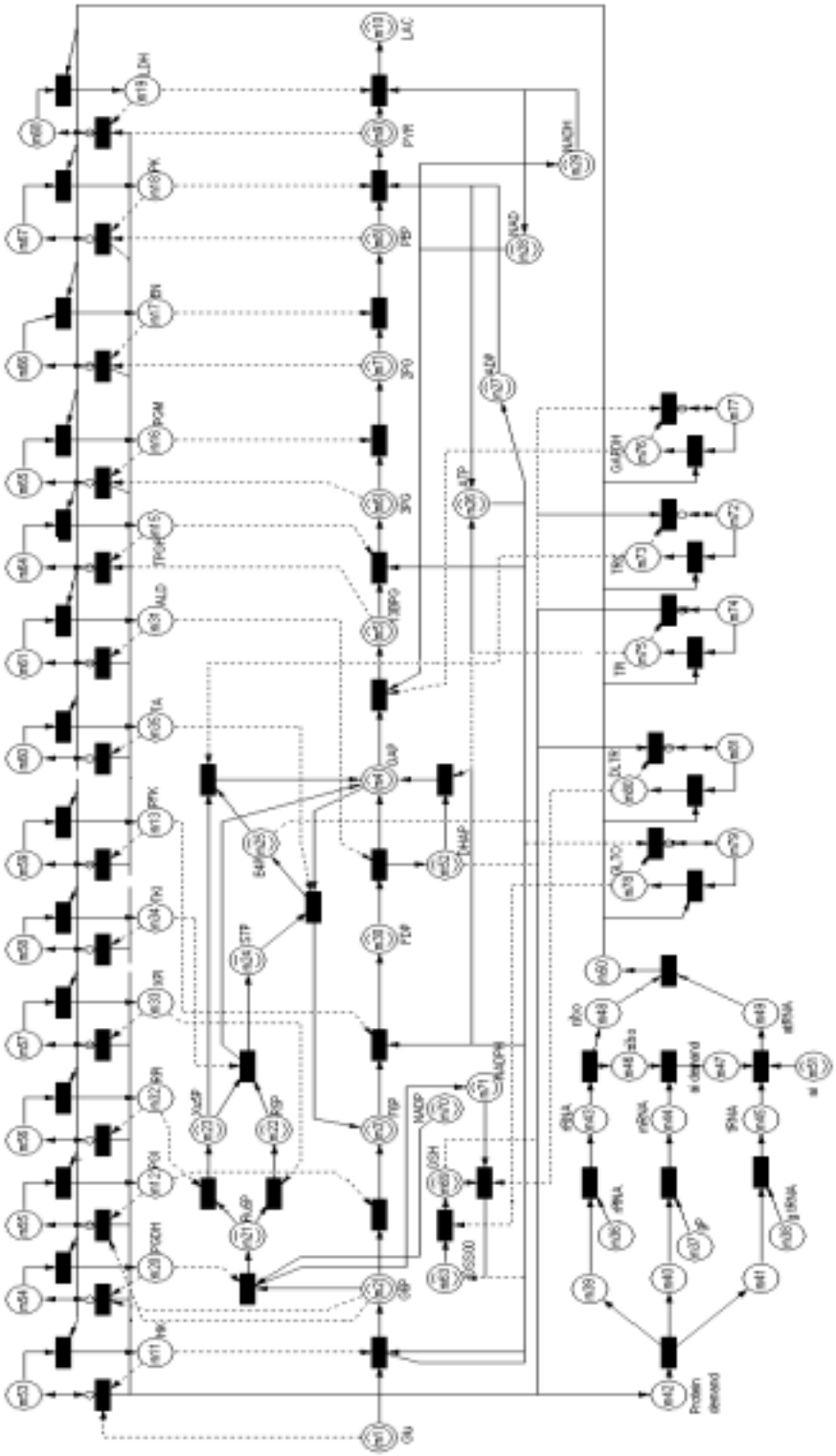

Figure 3: PN model for the glycolysis reaction system

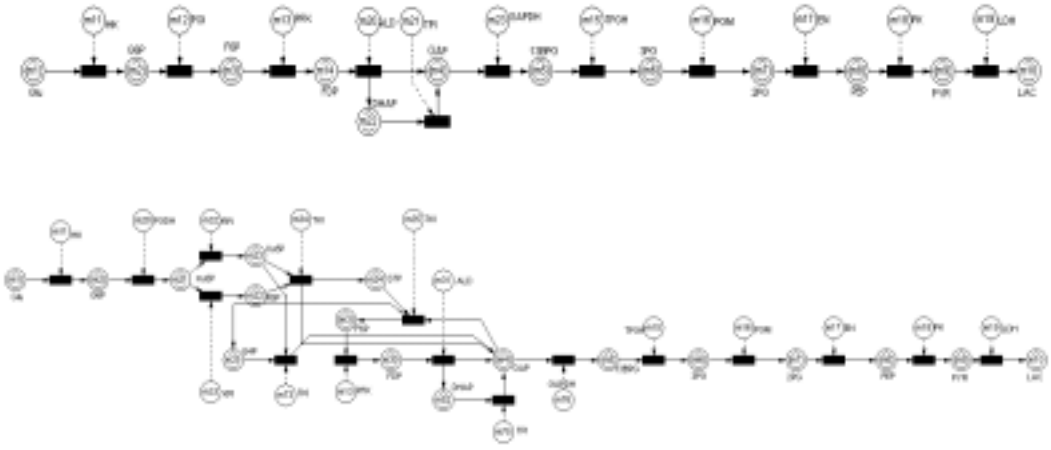

Figure 4: Reaction pathways involved in the glycolysis reaction cycle

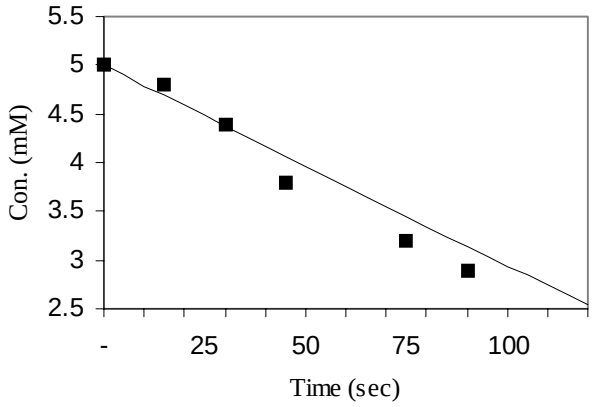

Figure 5 Model prediction of glucose concentration in glycolysis cycle, ——: model prediction
❑: experimental data [Theobald *et al.*, 1997]