Petri Net representations in systems biology

J.W. Pinney[1], D.R. Westhead and G.A. McConkey
Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, U.K.

Abstract

The mathematical structures known as Petri Nets have recently become the focus of much research effort in both the structural and quantitative analysis of all kinds of biological networks. This review provides a very brief summary of these interesting new research directions.

Representation in systems biology

In order to perform any kind of computational analysis on a biological system, we must first encode the relevant information in a way which can be represented as simply as possible in the computer's memory. This problem is well known in bioinformatics and has been studied extensively using existing knowledge representation techniques from computer science. Bioinformatic representations can be loosely classified into four main groups: (i) Ontologies of biological entities, processes and functions, which are starting to be widely used in genome annotation and related projects [1]. A biological ontology can be thought of as a network of categories, connected using *is a kind of* and other relationships, which allows its users to classify a gene, protein, compound or pathway within a *knowledge base*. (ii) Database models, which define the way biological information is stored in a database [2]. A database model may be derived from an ontology, in which case it shares the same logical structure but must be realized in a particular database management paradigm (e.g. relational or object-oriented). Database models support simple types of query, often via a query language such as SQL (structured query language). However, they are usually inappropriate for performing complex network analyses such as pathway discovery. (iii) Structural network models for analysis of the structural properties of biological networks [3]. These models are usually based on a graph representation, with nodes standing for a certain type of entity (e.g. compounds in a metabolic network) and edges standing for relations between them (e.g. co-occurrence of two compounds in a metabolic reaction). A structural network model of a system will generally be held in the computer's memory rather than used as a database itself, and will be used for the execution of some algorithms on the network in order to discover a structural property of the system [4]. (iv) Quantitative analysis models for simulation and analysis of the time-dependent properties of a network [5]. These models are traditionally based on systems of ordinary differential equations. More recent developments have focused on modelling concurrency in biological networks using computer science techniques such as the pi-calculus process algebra [6].

Key words: biochemical network, modelling, Petri Net, simulation, systems biology.
[1]To whom correspondence should be addressed (e-mail john@bioinformatics.leeds.ac.uk).

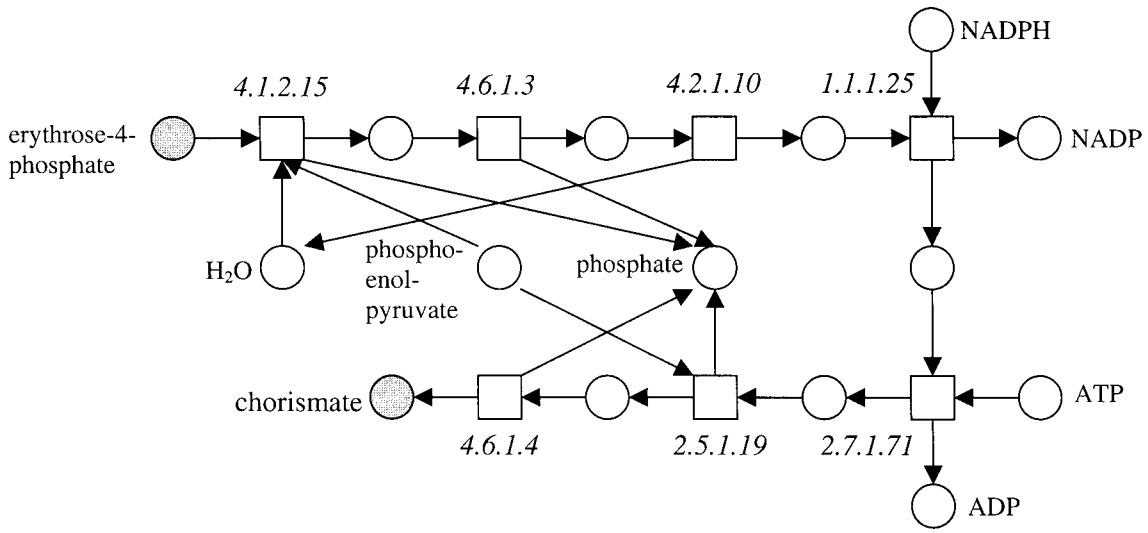
Petri Nets

One approach to the representation of concurrent systems is based on a mathematical concept called a *Petri Net*, which was developed in the early 1960s by Carl Adam Petri [7] and has subsequently been adapted and extended in many directions [8]. The simplest kind of Petri Net is a bipartite digraph, i.e. a graph with two types of node and directed arcs which connect nodes of different types, as shown in Figure 1. The two types of node are called *place nodes*, represented as circles, and *transition nodes*, represented as boxes; hence this type of net is also known as a *place-transition net*. The arcs may be labelled with an integer weight, but if unlabelled are assumed to have a weight equal to 1.

Places may be *marked* by an integer number of *tokens*. The overall state of a system of n places is represented by a vector of size n consisting of the markings on each place. The arcs connected to a transition node define sets of *input places* and *output places* for that transition. In a simple Petri Net, a transition is *enabled* if all of its input places have a marking equal to or greater than the weight of the arc connecting that place to the transition. When a transition is enabled, it may be *fired* to remove a number of tokens from each input place equal to the weight of the connecting input arc, and create a number of new tokens at each output place equal to the weight of the connecting output arc.

Many extensions to the simple Petri Net model have been developed for various modelling and simulation purposes [9]. These *high-level nets* include the following. (i) Hierarchical Petri Nets, which allow composition relations where a previously defined net is represented by a single place or transition in a new net. (ii) Hybrid Petri Nets, which incorporate places which may take continuous values instead of integer numbers of tokens. (iii) Timed Petri Nets, in which places and/or transitions may be assigned deterministic time delays. (iv) Stochastic Petri Nets, in which places and/or transitions may be assigned delays which are given by a probability distribution. (v) Coloured Petri Nets, which allow tokens to have internal structure and transitions to

Figure 1 | A Petri Net representation of the shikimate pathway, a metabolic pathway from erythrose 4-phosphate to chorismate
Place nodes (circles) represent metabolites whereas transition nodes (boxes) represent reactions. Reactions are labelled by their EC numbers.



have more complex firing rules. These nets have sufficient descriptive power to be used to execute sophisticated logic programs.

Examples of applications

Petri Nets of various different kinds have been used in several studies in systems biology, both as structural network models for qualitative analyses and as quantitative models using high-level nets.

Structural analysis

In previous studies of metabolic networks using Petri Nets, place nodes are usually used to represent biochemical entities (enzymes, compounds, ions etc.) and transition nodes to represent reactions, as shown in Figure 1 [10]. This formulation is very attractive since it gives a simple and intuitive representation of a metabolic network, with the stoichiometry of each reaction encoded by the arc weights of its transition node. Minor extensions to the simple Petri Net may be used to model catalysis and inhibition using special kinds of arc.

Computations on this type of net have recently been used in algorithms to discover novel metabolic pathways between two compounds [11]. In addition, the *net invariants* which may be computed using Petri Net theory have real-world meaning as the flux modes and conservation relations in the metabolic network, so are important for the rigorous determination of the metabolic capabilities of a system [12]. Outside the realm of metabolism, Petri Nets have also been used as part of a structural model for complex biological processes [13].

These qualitative methods have the advantage that they only require information about the stoichiometry and reversibility of the constituent reactions, rather than relying on the availability of measurements of kinetic parameters. However, overcoming the combinatorial explosion associated with larger networks is recognized as a difficult challenge for these new approaches to metabolic analysis.

Quantitative analysis

The high-level Petri Nets outlined above have recently been applied to quantitative analyses in many areas of systems biology.

An extension of the Hybrid Petri Net has been used in the representation and simulation of gene regulation networks and signal transduction pathways [14]. This formulation has now been realized as a software package, Genomic Object Net (http://www.genomicobject.net), which allows the user to study the time evolution of concentrations of species within the system in much the same way as is done in other simulation packages based on the ordinary differential equation formulation. Coloured Petri Nets have also been used to simulate the time evolution of metabolite concentrations in a metabolic pathway [15].

These applications are good illustrations of the suitability of high-level Petri Nets for the quantitative analysis of complex biological systems. It seems likely that they will prove useful in the future as bioinformaticians develop ever more integrated representations for biological networks.

References

1 Bard, J. (2003) Bioessays **25**, 501–506
2 Wittig, U. and de Beuckelaer, A. (2001) Brief Bioinform. **2**, 126–142
3 Deville, Y., Gilbert, D., van Helden, J. and Wodak, S.J. (2003) Brief Bioinform. **4**, 246–259
4 van Helden, J., Wernisch, L., Gilbert, D. and Wodak, S.J. (2002) in Bioinformatics and Genome Analysis (Mewes, H.-W., Seidel, H. and Weiss, B., eds), pp. 245–274, Springer-Verlag, Berlin

5 de Jong, H. (2002) J. Comput. Biol. **9**, 67–103
6 Regev, A., Silverman, W. and Shapiro, E. (2001) Pac. Symp. Biocomput. 2001, 459–470
7 Petri, C.A. (1962) PhD Thesis, Institut für Instrumentelle Mathematik, Bonn
8 Murata, T. (1989) Proc. IEEE **77**, 541–580
9 Bernardinello, L. and de Cindio, F. (1992) A Survey of Basic Net Models and Modular Net Classes, Springer Verlag, Berlin
10 Reddy, V.N., Liebman, M.N. and Mavrovouniotis, M.L. (1996) Comput. Biol. Med. **26**, 9–24
11 Küffner, R., Zimmer, R. and Lengauer, T. (2000) Bioinformatics **16**, 825–836
12 Zevedei-Oancea, I. and Schuster, S. (2003) In Silico Biol. **3**, in the press
13 Peleg, M., Yeh, I. and Altman, R.B. (2002) Bioinformatics **18**, 825–837
14 Matsuno, H., Tanaka, Y., Aoshima, H., Doi, A., Matsui, M. and Miyano, S. (2003) In Silico Biol. **3**, in the press
15 Genrich, H., Küffner, R. and Voss, K. (2001) Int. J. STTT **3**, 394–404

Received 12 July 2003