Japet: an integrated Tool for recreating KEGG data into hierarchical and Petri net analysis

Albert Gevorgyan, Dept. of Bioinformatics, University of Applied Sciences, Berlin, Germany
Prof. Dr. -Ing. Monika Heiner, Dept. of Computer Science, Brandenburg University of Technology, Cottbus, Germany
Prof. Dr. Ina Koch, Dept. of Bioinformatics, University of Applied Sciences, Berlin, Germany

Introduction

Petri nets have been successfully used in systems biology for modelling metabolic and regulatory pathways since the last eighties [4]. The Kyoto Encyclopedia of Genes and Genomes (KEGG) [2] provides an enormous amount of biological data, which could be used for creation of Petri nets. The Integrated Net Analyzer (INA)[3] is acknowledged as a powerful tool for Petri net analysis. The Hierarchical Decomposition Algorithm (HDA) [1] is a useful method for dependency analysis of the pathways.

Objectives

The objective of the work is the development of a dedicated Petri net tool, adapted for the solution of biological problems, by providing:

- Transformation of KEGG data [2] into hierarchical Petri nets
- Implementation of the hierarchical analysis of pathways
- An editor with built-in layout methods and an equation parser
- An interface to the INA (Integrated Net Analyzer), which allows the Petri net analysis of the pathways
- A relational Petri net database with the possibility of cross-net queries

Methods

- **Petri net representation [4]**: metabolites correspond to places, reactions to transitions, stoichiometric numbers to edge weights, enzymes are objects linked to transitions.
- **Hierarchical decomposition algorithm (HDA) [1]**, transforms the net into a tree of clusters. Each cluster is a transition-bordered subnet, interpreted as one sub-process, usually the metabolism of a single or several substances. The clusters may be coarsed to transitions or isolated to nets. (Fig. 1, 2).
- **Hierarchy based layout**: a novel algorithm, based on the hierarchical decomposition, was developed. The cluster tree is unfolded recursively in horizontal and vertical directions.

Implementation

- The program is written in pure Java.
- SAX and HTML parsers are used to parse the data from the KEGG database (subdivisions PATHWAY and LIGAND). The pathways can be downloaded and parsed immediately from the internet.
- An object-oriented hierarchy of the Petri net components was developed.
- HDA: input: incidence matrix -> output: hierarchical cluster tree, O(n^3)
- Layout methods: UnfolfdX: input: hierarchy tree -> output: two-dimensional array of knots (scala), O(n). UnfoldY: input: scala -> output: layout image, O(n).
- Postgresql was used for the design of the relational Petri net database.

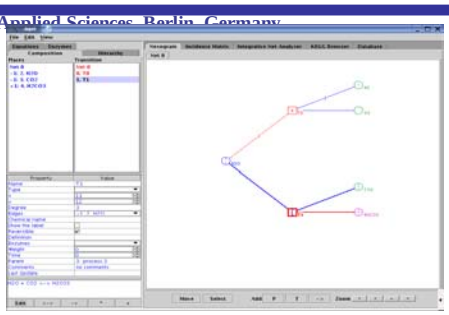

Fig. 3: A simple net consisting of two reactions entered as an equation to the window in the left-down corner. The places are divided to internal (blue), sources (green) and sinks (magenta). The reversible reactions are marked by two small squares. The direction of the edges is also distinguished by the color, and the edge weights are shown. The options „Show labesl“ and „Show tokens“ are chosen.

Conclusions

- The program allows downloading KEGG pathways and their automatic transformation into Petri nets with biological features (reversible reactions, sources and sinks).
- The nets get a layout and may be edited, additional biological information is shown online.
- The equation parser allows to enter chemical equations, transforming them to Petri net components
- The implementation of Hierarchical decomposition algorithm enables dependency analysis of the pathways, and the hierarchy based layout presents a convenient tool for reading complicated nets (by coarsing or isolating the clusters).
- The interface to INA offers the possibility of qualitative analysis of the created nets.
- The relational Petri net database provides a new qualitative level of topological and Petri-net analysis of the biosystems.

References

[1] Gagneur, J. and Jackson, D. B. and Casari. G. 2003 *Hierarchical Analysis of Dependency in Metabolic Networks.* Biofinformatics

[2] Kanehisa, M. and Goto. S. 2000. *KEGG: Kyoto Encyclopedia of Genes and Genomes*. Nucleic Acids Res.

[3] Starke, P. H. and Roch, S. *INA - The Integrated Net Analyzer.* http://www.informatik.hu-berlin.de

[4] Heiner, M. and Koch, I. 2004. *Petri Net Based Model Validation in Systems Biology*. Proc. 25th ICATPN, LNCS 3099

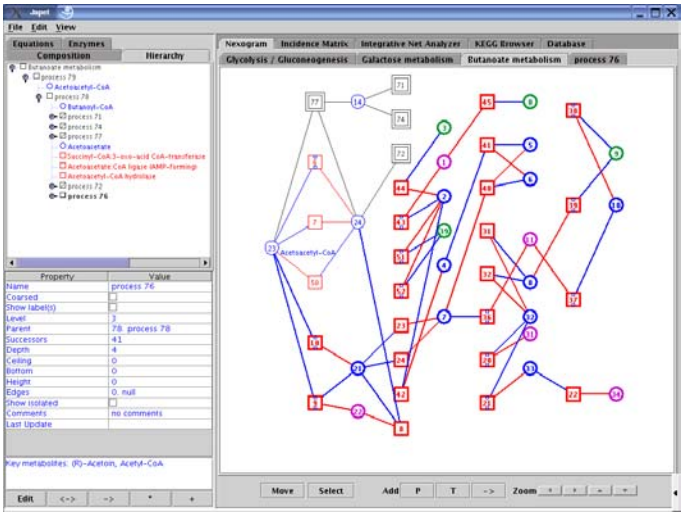

Figure 1: Butanoate metabolism.The hierarchical tree and the component properties are shown in the left part of the frame. The cluster, representing the metabolism of (R)-Acetoin and Acetyl-CoA, is selected. Four smaller clusters are coarsed (black).

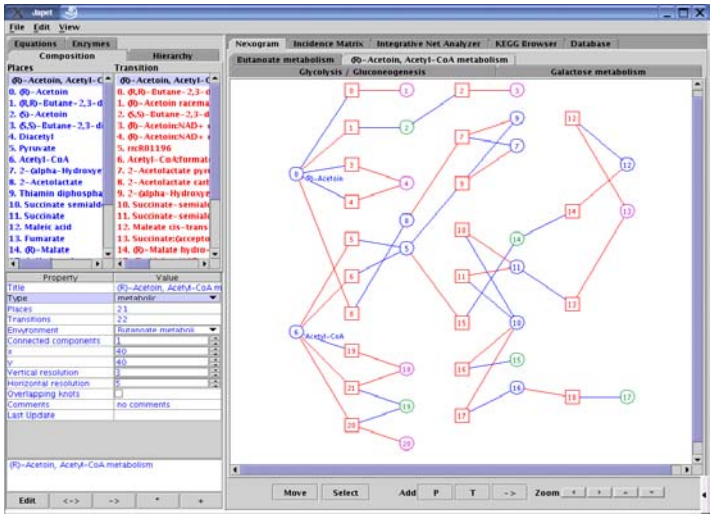

Figure 2: the same cluster is isolated to a new Petri Net. The lists of places and transitions are shown. In both layouts the key metabolites are placed in the leftmost column. The multiple document view allows working at several nets at once.

SPONSORED BY THE
Federal Ministry of Education and Research
Projektträger Jülich
This work was done within the context of the
Nationales Genomforschungsnetz