MPath2PN - Translating metabolic pathways into Petri nets

Marta Simeoni

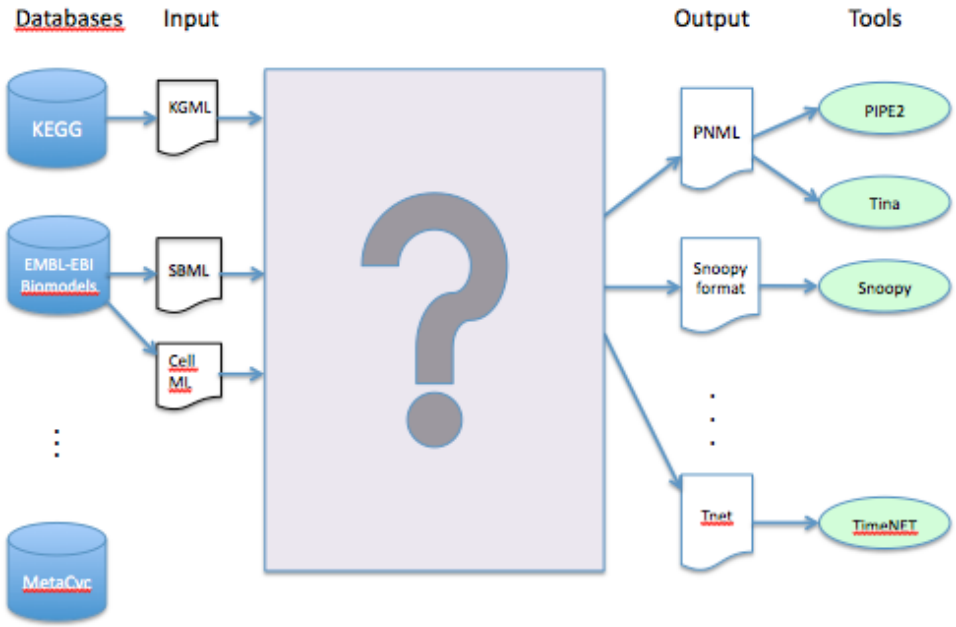
Università Ca' Foscari Venezia

Workshop BioPPN 2011 - Newcastle upon Tyne, June 20, 2011

Joint work with
P. Baldan (Univ. Padova), N. Cocco, F. De Nes (Univ. Venezia),
M. Llabrés Segura (Univ. de les Illes Balears), A. Marin (Univ. Venezia)

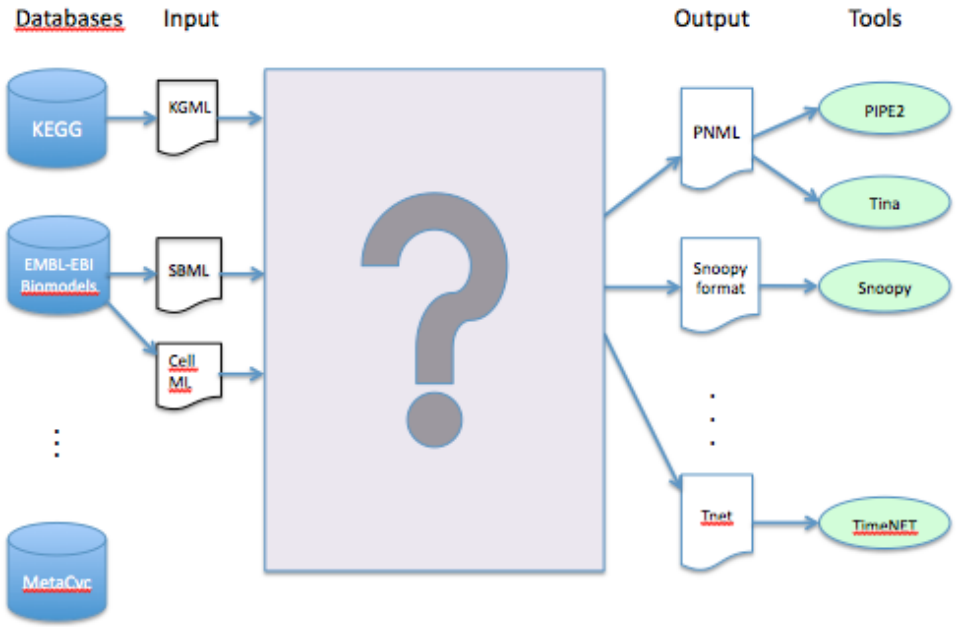
Framework

Modelling metabolic pathways with Petri nets (PNs)



Framework

Modelling metabolic pathways with Petri nets (PNs)



IDEA: offering automatic translations to close the gap between pathways and PN tools

Metabolic pathways

Metabolism: the chemical system which generates the essential components for life

- Metabolic pathways:
  - subsystems dealing with some specific function
  - represented as a network of chemical *reactions* catalised by one or more *enzymes* where some molecules (*reactants or substrates*) are transformed into others (*products*)
  - the *stoichiometric matrix* identifies the pathways components and their relations
  - kinetics represented by the rate equation associated with each reaction

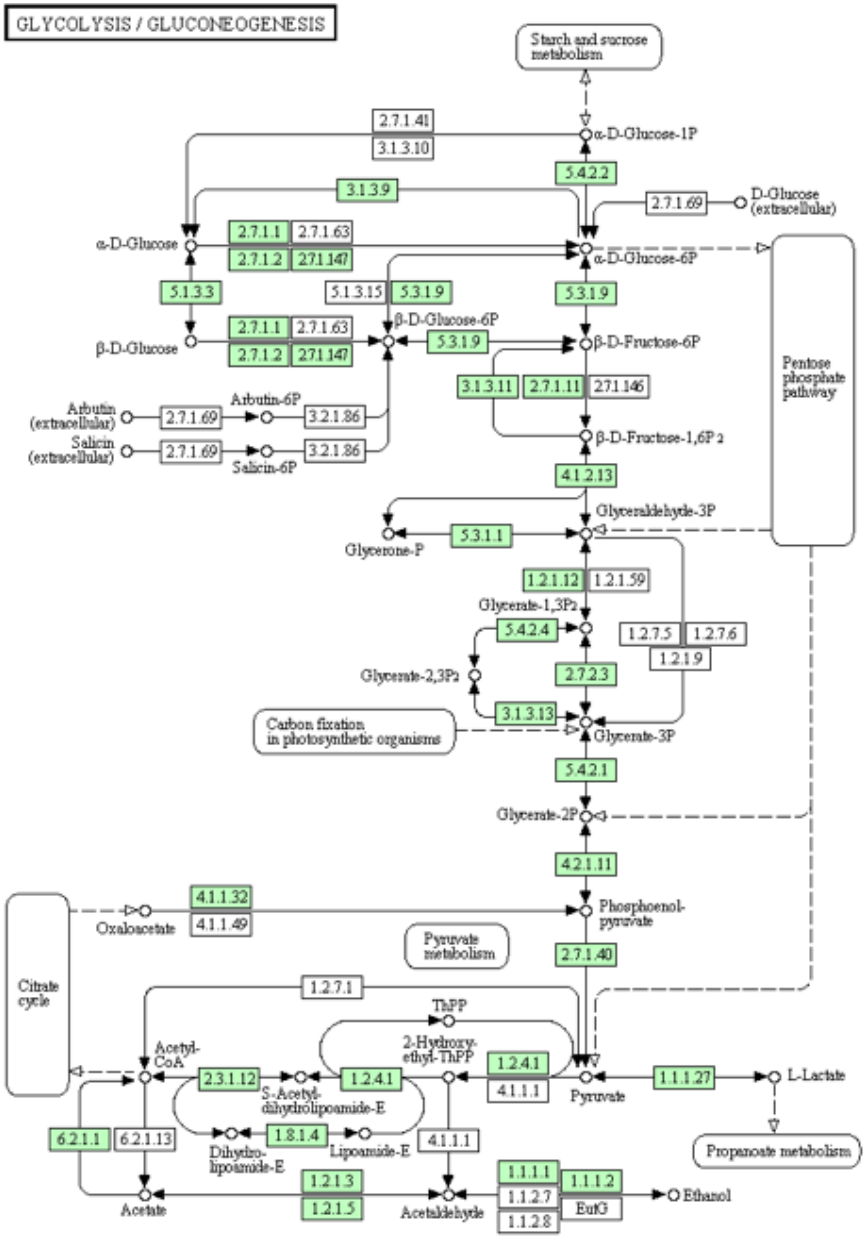
Sources of metabolic data

Metabolic pathways information are collected in many different databases:

- KEGG PATHWAYS
- Biomodels
- MetaCyc
- Reactome
- TRANSPATH (part of BIOBASE)
- BioCarta

KEGG pathways

- At present it contains around 93000 pathways
- Pathways are represented by maps with additional information
- Models are coded in KGML (KEGG Markup Language)
- A web service for querying the KEGG system from users programs is available



Petri net representation of metabolic pathways

Metabolic pathways can be naturally modelled with Petri nets:

- Places are associated to molecular species (metabolites, enzymes)
- Transitions correspond to chemical reactions
  - Input places are substrates
  - Output places are products
- The incidence matrix of the PN is identical to the stoichiometric matrix of the system of chemical reactions
- The number of tokens in each place of the PN indicates the amount of substance associated with that place.

Petri net representation of metabolic pathways

Some modelling choices have to be taken:

- **Enzymes**

  They are taken and then released by the reactions.
- **Ubiquitous substances**

  Once assumed to be constant, they can be omitted in the PN model.
- **External metabolites**

  They can be represented in different ways, with different impacts on the resulting net.
- **Reversible reactions**

  They are decomposed into two distinct reactions, a forward one and a backward one, leading to two corresponding transitions in the PN model.

The tool MPath2PN

It provides a way of automatically transforming a metabolic pathway into a corresponding PN

- Input pathways expressed in one of the various existing formalisms (e.g. KGML, SBML)
- Output PN models expressed in one of the specific formalisms for PN tools (e.g. PNML, PIPE2, SNOOPY, INA, TimeNET)

The tool MPath2PN

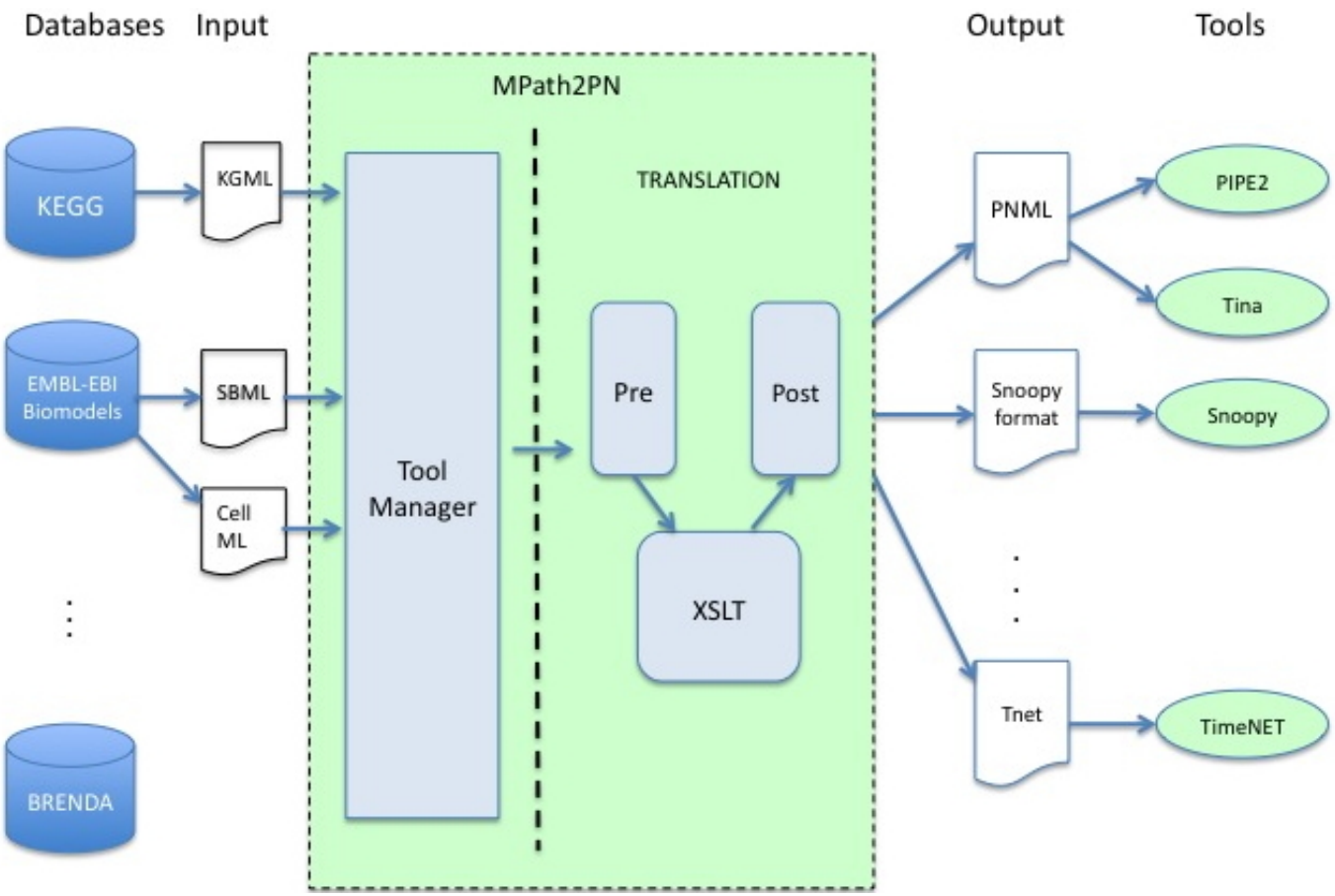
It provides a way of automatically transforming a metabolic pathway into a corresponding PN

- Input pathways expressed in one of the various existing formalisms (e.g. KGML, SBML)
- Output PN models expressed in one of the specific formalisms for PN tools (e.g. PNML, PIPE2, SNOOPY, INA, TimeNET)

The current prototype implements two translations:

- from KEGG to PIPE2 with or without ubiquitous substances

Structure of MPath2PN



Structure of MPath2PN

- Most of the descriptions of metabolic pathways and of PNs are based on XML formats
  $\Rightarrow$ MPath2PN produces the translations by using XSLT
- Each translation requires the definition of an appropriate style sheet XSL
- However, the plain XSLT translation is often not sufficient
- For this reason the tool performs each translation in three main steps:
  - pre-treatment - to add information to the input file
  - XSLT translation
  - post-treatment - to add information to the output file

From KEGG to PIPE2

Modelling choices taken by the implemented translations:

- Enzymes: they are not represented
- Ubiquitous substances: both alternatives are taken into consideration
- External metabolites: they are places where connected transitions either all consume or all produce tokens
- Reversible reactions: they can be distinguished by their identifiers

From KEGG to PIPE2

A KGML file representing a metabolic pathway contains qualitative information on:

- Compounds and reactions
  - substrates and products
  - type of reaction: reversible/irreversible

From KEGG to PIPE2

A KGML file representing a metabolic pathway contains qualitative information on:

- Compounds and reactions
  - substrates and products
  - type of reaction: reversible/irreversible $\Rightarrow$ only present in the KGML file

From KEGG to PIPE2

A KGML file representing a metabolic pathway contains qualitative information on:

- Compounds and reactions
  - substrates and products
  - type of reaction: reversible/irreversible $\Rightarrow$ only present in the KGML file

... but

- Enzymes and ubiquitous substances are not present
- Reaction equations are not complete:
  - ubiquitous substances are not mentioned
  - stoichiometric values are not specified.

From KEGG to PIPE2

A KGML file representing a metabolic pathway contains qualitative information on:

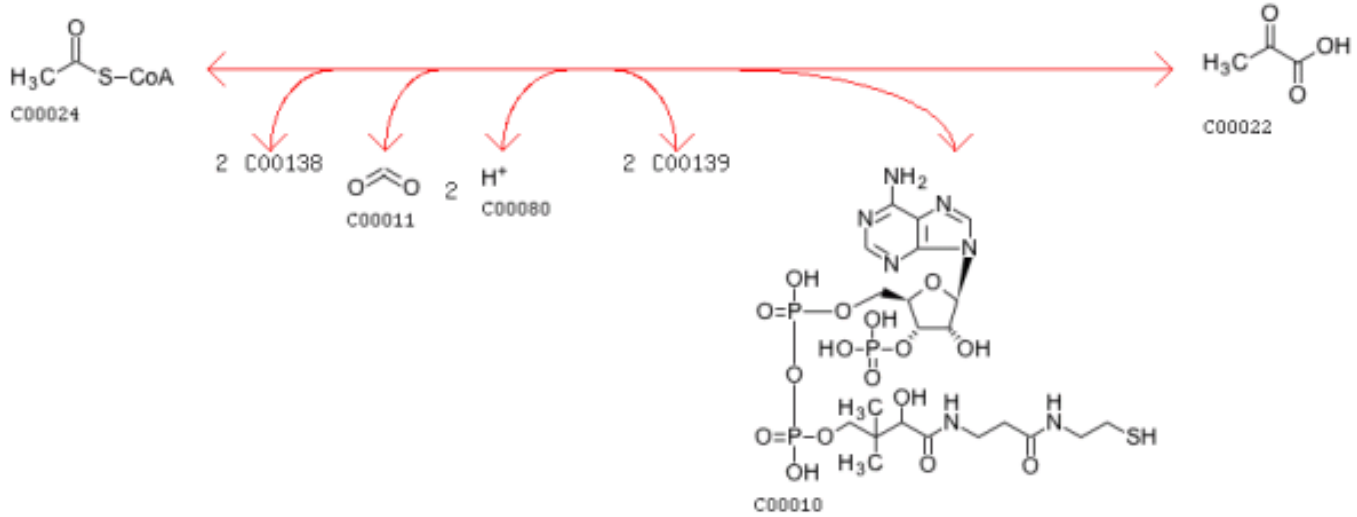
- Compounds and reactions
  - substrates and products
  - type of reaction: reversible/irreversible $\Rightarrow$ only present in the KGML file

... but

- Enzymes and ubiquitous substances are not present
- Reaction equations are not complete:
  - ubiquitous substances are not mentioned
  - stoichiometric values are not specified. $\Rightarrow$ must be retrieved separately

From KEGG to PIPE2

Entry	R01196 Reaction
Name	pyruvate:ferredoxin 2-oxidoreductase (CoA-acetylating)
Definition	2 Reduced ferredoxin + Acetyl-CoA + CO2 + 2 H+ <=> 2 Oxidized ferredoxin + Pyruvate + CoA
Equation	2 C00138 + C00024 + C00011 + 2 C00080 <=> 2 C00139 + C00022 + C00010



From KEGG to PIPE2

From KEGG to PIPE2 without ubiquitous substances:

- Pre-treatment: not needed
- XSLT
- Post-treatment: adding the stoichiometric values

From KEGG to PIPE2 with ubiquitous substances:

- Pre-treatment: adding the ubiquitous substances to the compounds and the reactions equations
- XSLT
- Post-treatment: adding the stoichiometric values

Conclusions and future work

- MPath2PN: tool for translating metabolic pathways into corresponding PNs, coping with different input and output formats
- Aim: systematic reuse of the PN tools also for analysis and simulation of metabolic pathways
- The current prototype implements two translations from KEGG to PIPE2
- It is available at www.dsi.unive.it/∼simeoni/MPath2PNtool.tgz
- Further translations to be included are:
  - KGML ⇒ INA
  - SBML ⇒ PNML
  - SBML ⇒ Snoopy
  - SBML ⇒ TimeNet