An Overview of Data Models for the Analysis of Biochemical Pathways

Yves Deville[1], David Gilbert[2], Jacques van Helden[3], Shoshana Wodak[3]

[1] Computing Science and Engineering Department
Université catholique de Louvain
Place Saint-Barbe 2, B-1348 Louvain-la-Neuve, Belgium
deville@info.ucl.ac.be

[2] Bioinformatics Research Centre
Department of Computing Science
University of Glasgow
17 Lilybank Gardens, Glasgow G12 8QQ, Scotland, UK
drg@brc.dcs.gla.ac.uk

[3] Unité de Conformation des Macromolécules Biologiques.
Université Libre de Bruxelles
50 av. F.D. Roosevelt, B-1050 Bruxelles, Belgium
{jvanheld,shosh}@ucmb.ulb.ac.be

Abstract

Various forms of data models can be used for the analysis of biochemical pathways such as metabolic, regulatory, or signal transduction pathways. This paper overviews and classifies the different forms of data models found in the literature, and describes how these models have been used in the analysis of biochemical pathways.

The quantity of available information on biochemical pathways for different organisms is increasing very rapidly, and it has now become possible to perform detailed analyses of metabolic pathway structures for entire organisms. However, such analyses face difficulties due to the nature of the databases which are often heterogeneous, incomplete, or inconsistent. This makes pathway analysis a challenging problem in system biology and in bioinformatics.

In this overview, we concentrate on models of network structure, focusing on the analysis of existing information, collected from experiments and stored in databases. We overview and classify the different forms of data models found in the literature using a unified framework. We describe how these models have been used in the analysis of biochemical pathways. This enables us to underline the strengths and weaknesses of the different approaches, and at the same time highlights some relevant future research directions.