

The Mereological Structure of Chemical Reactions: Implications for Biochemical Simulation

William Andersen
University of Maryland
Department of Computer Science
College Park, MD 20742
USA
and Ontology Works
1132 Annapolis Road
Odenton, MD 21113
USA

andersen@ontologyworks.com

Ursula Kummer
Bioinformatics and Computational Biochemistry
Group
European Media Laboratory
Schloss-Wolfsbrunnenweg 33
69118 Heidelberg
Germany
ursula.kummer@eml.villa-bosch.de

ABSTRACT

Improvement of experimental techniques is resulting in a dramatic acceleration of accumulation of data in biology and biochemistry. In order to make sense of these data computational methods have to be developed and improved. One of the computational fields which will help to interpret e.g. gene and protein expression data as well as kinetic data from experiments is the simulation of biochemical processes. These simulations will heavily depend on the quality of biochemical databases.

Even though there will probably always be a need for well trained experts who actually carry out the simulations, the sheer mass of data and possible applications calls for automation to a certain extent. However, automation can only be improved if the data stored in the respective databases represent reality as closely as possible and if the need for human interpretation of the data is minimized. Normally, any ambiguities are compensated for in the research community by researchers working in the shared context of a research group or a research community. However, computer systems that are designed to store, exchange, and process information on chemical reactions do not share the context of their human users and designers.

Ambiguities in the representation of biochemical data are often encountered. One example is the notation used to represent (bio)chemical reactions which is crucial for the successful use of simulations in the biochemical context. There is a clear lack of distinction between mass reactions and elementary reactions. However, this distinction is important, since there are properties of mass reactions which elementary reactions do not have, e.g. an elementary reaction cannot be catalyzed.

Inadequate data representations based on this blurred distinction may lead to the inappropriate application of certain simulation methods, thus hampering the automation of simulation in the (bio)chemical context. In this poster we propose a partial ontology of (bio)chemical reactions, emphasizing the ontological structures needed to support (and to some extent justify) uses of various simulation formalisms for

systems of (bio)chemical reactions. This approach has the additional advantage of making implicit assumptions about the nature of (bio)chemical reactions explicit which yields significant advantages over common representations of these reactions developed to date.