Role-based representation and inference of biochemical processes

Christian Bölling¹, Michel Dumontier², Michael Weidlich³ and Hermann-Georg Holzhütter¹

¹Charité Universitätsmedizin Berlin, Germany
²Carleton University, Ottawa, Canada
³Humboldt-Universität Berlin, Germany
motivation

- computational approaches to study biochemistry require machine accessible representations of biochemical knowledge

- existing representations adopt formal semantics to varying degrees

- limited expressivity and ability for reasoning
Dumontier (2008) Proc. 5th OWLED Workshop on OWL: Experiences and Directions
biochemical processes viewed as *microscopic molecular events*

- individuals of classes describing chemical species represent *single molecules*

- individuals of classes describing biochemical processes represent single molecular events, i.e. *directed transitions of a chemical system from an initial to a terminal state* involving individual molecules
main axis of classification: stoichiometry

\[ aA + bB \rightarrow xX + yY \]

EquivalentTo:
\[
(\text{realizes exactly } a \text{ (reactant\_role and (has\_bearer some A))}) \text{ and }
(\text{realizes exactly } b \text{ (reactant\_role and (has\_bearer some B))}) \text{ and }
(\text{realizes exactly } x \text{ (product\_role and (has\_bearer some X))}) \text{ and }
(\text{realizes exactly } y \text{ (product\_role and (has\_bearer some Y))}) \text{ and }
(\text{realizes exactly } (a+b) \text{ reactant\_role and })
(\text{realizes exactly } (x+y) \text{ product\_role})
\]

\[ \text{glc + atp } \rightarrow \text{ g6p + adp} \]

EquivalentTo:
\[
(\text{realizes exactly } 1 \text{ (reactant\_role and (has\_bearer some glc))}) \text{ and }
(\text{realizes exactly } 1 \text{ (reactant\_role and (has\_bearer some atp))}) \text{ and }
(\text{realizes exactly } 1 \text{ (product\_role and (has\_bearer some g6p))}) \text{ and }
(\text{realizes exactly } 1 \text{ (product\_role and (has\_bearer some adp))}) \text{ and }
(\text{realizes exactly } 2 \text{ reactant\_role and })
(\text{realizes exactly } 2 \text{ product\_role})
\]
classification by overall stoichiometry

Equivalence:
(realizes exactly 1 (reactant_role and (has_bearer some gluc)))
and
(realizes exactly 1 (reactant_role and (has_bearer some ATP)))
and
(realizes exactly 1 (product_role and (has_bearer some G6P)))
and
(realizes exactly 1 (product_role and (has_bearer some ADP)))
and
realizes exactly 2 reactant_role and
realizes exactly 2 product_role
... and catalysis

EquivalentTo:
(realizes exactly 1 (reactant_role and (has_bearer some glc))) and
(realizes exactly 1 (reactant_role and (has_bearer some ATP))) and
(realizes exactly 1 (product_role and (has_bearer some G6P))) and
(realizes exactly 1 (product_role and (has_bearer some ADP))) and
realizes exactly 2 reactant_role and
realizes exactly 2 product_role and
(realizes exactly 1 (enzyme_role and (has_bearer some HK)))
... and mechanism

EquivalentTo:
(\textit{realizes exactly 1 (reactant\_role and (has\_bearer some glc))}) and
(\textit{realizes exactly 1 (reactant\_role and (has\_bearer some atp))}) and
(\textit{realizes exactly 1 (product\_role and (has\_bearer some g6p))}) and
(\textit{realizes exactly 1 (product\_role and (has\_bearer some adp))}) and
\textit{realizes exactly 2 reactant\_role and
realizes exactly 2 product\_role and
(\textit{realizes exactly 1 (enzyme\_role and (has\_bearer some hk))}) and
(\textit{has\_proper\_part \ldots})}
partonomy of processes
process types

A + B → (AB)

catalyzed by multi-enzymes

“pathways”
shared reactants & products: sequence of processes
location of processes and transport

glc + atp $\rightarrow$ g6p + adp in the cytosol of human liver cells

EquivalentTo:
(realizes exactly 1 (reactant_role and (has_bearer some glc))) and
(realizes exactly 1 (reactant_role and (has_bearer some atp))) and
(realizes exactly 1 (product_role and (has_bearer some g6p))) and
(realizes exactly 1 (product_role and (has_bearer some adp))) and
(realizes exactly 2 reactant_role) and
(realizes exactly 2 product_role) and
(occurs_in some (cytosol and (part_of some (cell and (part_of some (liver and (part_of some Homo_sapiens)))))))
location of processes and transport

EquivalentTo:
(realizes exactly 1 (reactant_role and
(has_bearer some (2kg and (located_in
some mitochondrion))))) and
(realizes exactly 1 (reactant_role and
(has_bearer some (mal and (located_in
some cytosol))))) and
(realizes exactly 1 (product_role and
(has_bearer some (2kg and (located_in
some cytosol))))) and
(realizes exactly 1 (product_role and
(has_bearer some (mal and (located_in
some mitochondrion))))) and
realizes exactly 2 reactant_role and
realizes exactly 2 product_role
inference of process and entity characteristics
inference of process and entity characteristics

- location of reaction participants:
  biochemical_process(?p), occurs_in(?p,?l),
  has_participant(?p,?o) -> located_in(?o,?l)

- sequence of processes within same location:
  product_role(?r1), reactant_role(?r2), has_bearer(?r1,?o),
  has_bearer(?r2,?o), realizes(?p1,?r1),
  realizes(?p2,?r2), occurs_in(?p1,?l), occurs_in(?p2,?l)
  -> directly_precedes(?p1,?p2)
outlook

- integrate with ontological descriptions of chemical entities, cell-types, tissues
- implement design patterns / rules for classification according to taxonomies of chemicals, locations, reaction mechanisms,…
- add support for quantitative description of reactions / participants
- expand on other molecular interactions
- use for biochemical network reconstruction, integration of pathway datasets, inference of molecular associations…
Acknowledgments

• Michel Dumontier
• Michael Weidlich
• Matthias König
• Andreas Hoppe
• Hermann-Georg Holzhütter