Chemical Knowledge Representation with Description Graphs and Logic programming

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF OXFORD

Despoina Magka, Boris Motik, Ian Horrocks

Motivation

✔ ChEBI ontology: Chemical Entities of Biological Interest

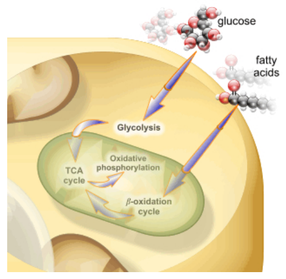
ChEBI EMBL-EBI

✔ High quality annotation and taxonomy of chemical compounds

Applications

Drug discovery

Elucidation of metabolic pathways



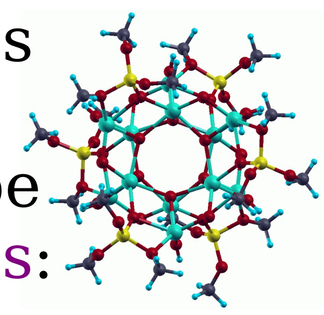
Objective

✔ ChEBI is manually curated

✔ Currently contains ~27,000 entries

✔ Growth rate 1,500 per curator per year

✔ Biologically interesting entities possibly > 1,000,000

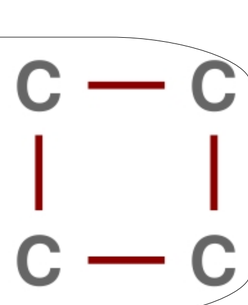


✔ Each new molecule needs to be classified under chemical classes:

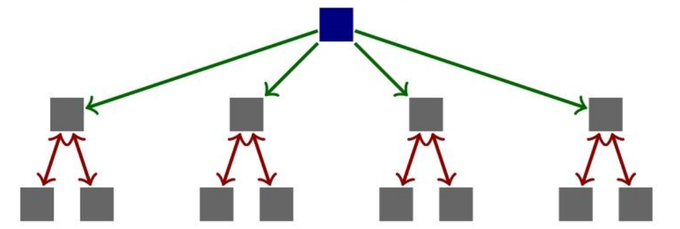
- ➢ Is dinitrogen inorganic?
- ➢ Is acetylene a hydrocarbon?
- ➢ Is benzaldehyde a benzenoid?
- ➢ Does cyclobutane have a four-membered ring?

✔ Automate chemical classification!

How to classify cyclobutane?



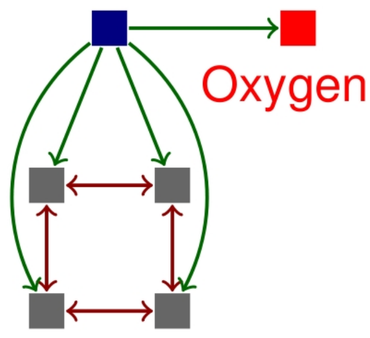
OWL cannot precisely describe cycles due to the tree-model property



Canonical model

Does cyclobutane have a four-membered ring? ✗

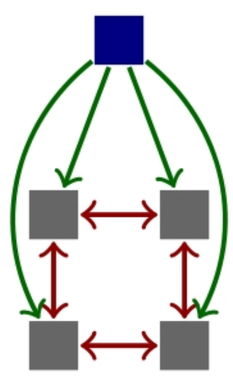
FOL semantics of OWL with Description Graphs allows for models with additional structure



Extra model

Is cyclobutane a hydrocarbon? ✗

Interpret Description Graphs under Logic Programming



Only model

Does cyclobutane have a four-membered ring? ✓

Is cyclobutane a hydrocarbon? ✓

Results

✔ Expressive and decidable formalism for modelling structured objects

✔ Ensure decidability via a novel acyclicity condition

✔ Encouraging results of a prototypical implementation