

Ontologies and classification of chemicals: can they help each other?

Andrew Buxton
UDC Editorial Team, UK

Classes of chemicals

- **By chemical composition**
e.g. compounds of nitrogen and oxygen
- **By uses**
e.g. dyes, narcotics
- **By physical properties**
e.g. solids, electrical conductors
- **By chemical properties**
e.g. stable in air

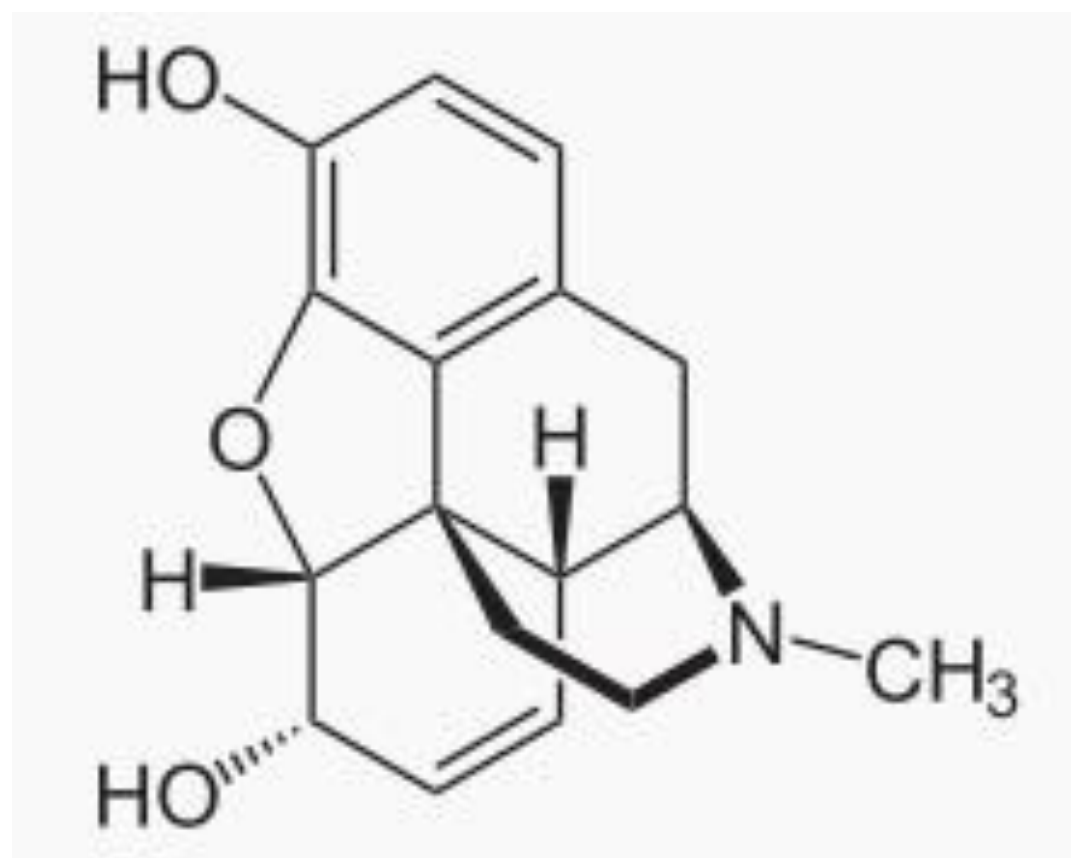
Advantages of a synthetic classification

- Retrieval by part of the structure
e.g. chlorine compounds, alcohols
- Don't need to list every chemical in CAS Registry
- Accommodate new substances

Synthetic classification

| NAME | FORMULA | UDC | A synthetic classification |
|----------------------------|--|-------------|----------------------------|
| Ferric chloride | FeCl_2 | 546.722'131 | eFe o+2 eCl |
| Ferrous chloride | FeCl_3 | 546.723'131 | eFe o+3 eCl |
| Toluene (methylbenzene) | $\text{C}_6\text{H}_5\cdot\text{CH}_3$ | 547.533 | a6 m1 c1 m1 |
| Dimethylbenzene | $\text{C}_6\text{H}_4(\text{CH}_3)_2$ | 547.534.2 | a6 m1 c1 m2 |
| Dinitrobenzene | $\text{C}_6\text{H}_4(\text{NO}_2)_2$ | | a6 m1 (eN eO m2) m2 |

Morphine



Ontologies

“An ontology is a controlled vocabulary that describes objects and the relations between them in a formal way, and has a grammar for using the vocabulary terms to express something meaningful within a specified domain of interest” (Jermeý & Browne, 2004: 94).

EXAMPLE

ChEBI - Chemical Entities of Biological Interest

Example relations in ChEBI

acetone *is a* methyl ketone

acetone *is a* propanone

sodium chloride *is a* inorganic sodium salt

caffeine monohydrate *has part* caffeine

warfarin *has role* rodenticide

warfarin *has role* anticoagulant

Graphical display of hierarchy in ChEBI

Ontology Lookup Service... x Ontology Lookup Service... x

www.ebi.ac.uk/ontology-lookup/browse.do?ontName=ChEBI

Suggested Sites Web Slice Gallery Download Video Andrew's homepage

Other bookmarks

It is now possible to exclude obsoleted terms from the auto-completion list when searching for terms. Simply uncheck the *Include obsolete terms* box and any terms that have been marked as obsolete will no longer be returned in the search results.

August 2006: SOAPAction added to WSDL

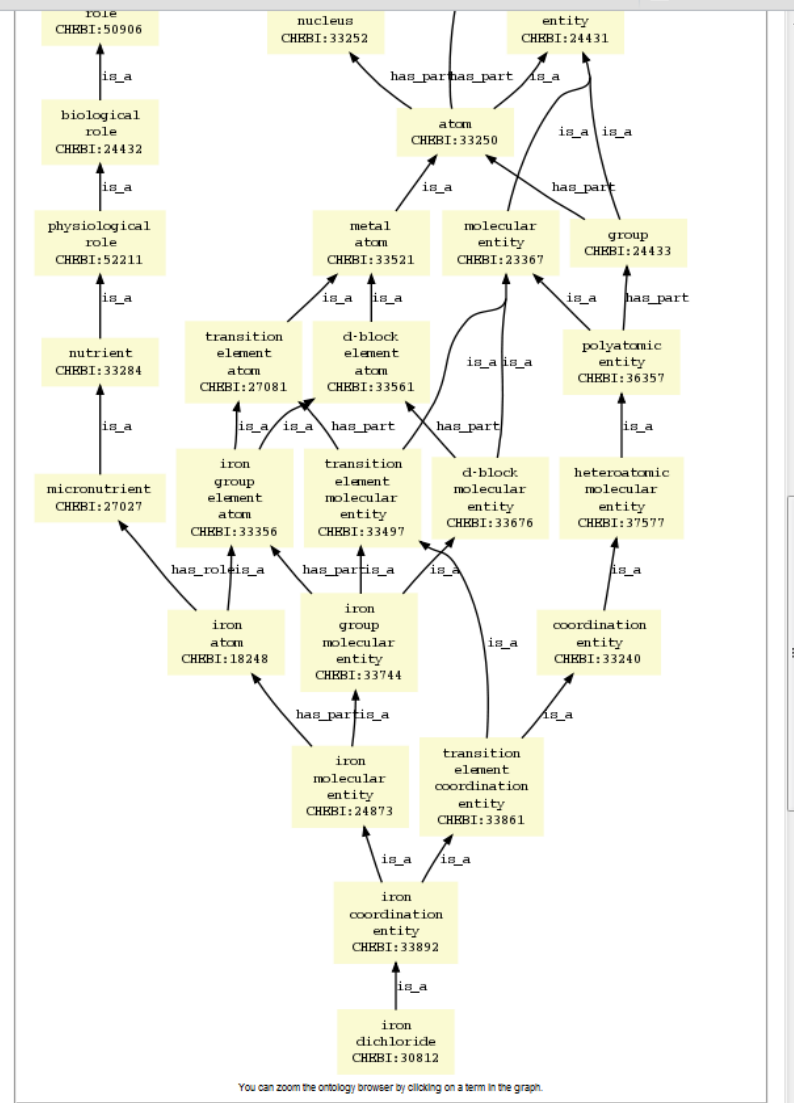
The SOAPAction element has been added to the WSDL service declaration to improve interoperability with PHP and other languages that require it.

July 2006: New webservice deployed

As announced in the previous news item, the OLS webservice has been redeployed to make it compliant with the latest WSDL specifications and also improve its interoperability. If you are experiencing problems with the new webservice code, please do not hesitate to contact us for support.

June 2006: Notification of WSDL change

The OLS webservice will be undergoing change that will not be backwards-compatible and might break existing code at the end of June. The current WSDL defines multiple operations with the same name but with different parameters. While such method overloading was allowed under previous WSDL specifications, it will no longer be supported. Furthermore, to maximize interoperability, the WSDL will change from a RPC-style service to a document/wrapped style service. As such, users of the OLS webservice are encouraged to update their code to use the new WSDL, which will be published on June 30th.



How can classifications help ontologies?

1. Providing a ready-made hierarchy (problems with a synthetic classification?)
2. Providing a notation
3. Providing terms from other disciplines

How can ontologies help classifications?

1. Providing hierarchies

| | |
|---|---|
| <p>chemical entity molecular entity main group molecular entity p-block molecular entity carbon group molecular entity organic molecular entity natural product carbohydrate sugar monosaccharide aldose aldohexose glucose</p> | <p>54 Chemistry 547 Organic chemistry 547.4 Multivalent acyclic compounds. Acyclic compounds with mixed functions 547.45 Aldehyde alcohols. Ketone alcohols 547.454 Carbohydrates 547.455 Simple sugars or monosaccharides 547.455.6 Hexoses 547.455.62 Aldohexoses with six carbon atoms 547.455.623 Glucose</p> |
| <p>Hierarchy in ChEBI</p> | <p>Hierarchy in UDC</p> |

Differences found between ChEBI and UDC

1. UDC is lacking some recently produced chemicals and types of chemicals
2. UDC chemistry is lacking some biologically important substances
3. ChEBI divides compounds between p-block and transition elements.
4. UDC divides organic compounds into those with 1 functional group and those with >1
5. UDC divides inorganic compounds of an element firstly by valency

How can ontologies help classifications?

2. Providing a subject index

e.g. synonyms, specific compounds not in the schedules

3. Providing access by roles, e.g. narcotics, herbicides

4. Access via alternative hierarchies

5. Graphical tools

Conclusions

- Both classifications and ontologies have their uses
- Making links between classifications and ontologies, such as through CAS registry numbers or perhaps including UDC numbers in collaboratively-produced ontologies, offers the possibilities of mash-ups that could combine the best features of both.