

Why would anybody need CL?

C. Mungall¹, F. Neuhaus^{2,3}

¹Lawrence Berkeley Laboratory

²National Institute of Standards and Technology

³University of Maryland Baltimore County

ICBO, July 2011

- represent knowledge in a computer *interpretable* way
- enable separation of knowledge from software
- are sets of axioms in a given ontology language

expressivity vs. computational properties

Restrict expressivity of FOL to guarantee good computational properties

- OWL DL: a decidable fragment of first-order logic
- OWL EL: consistency, subsumption decidable in polynomial time
- OWL QL: very fast instance retrieval (logspace)

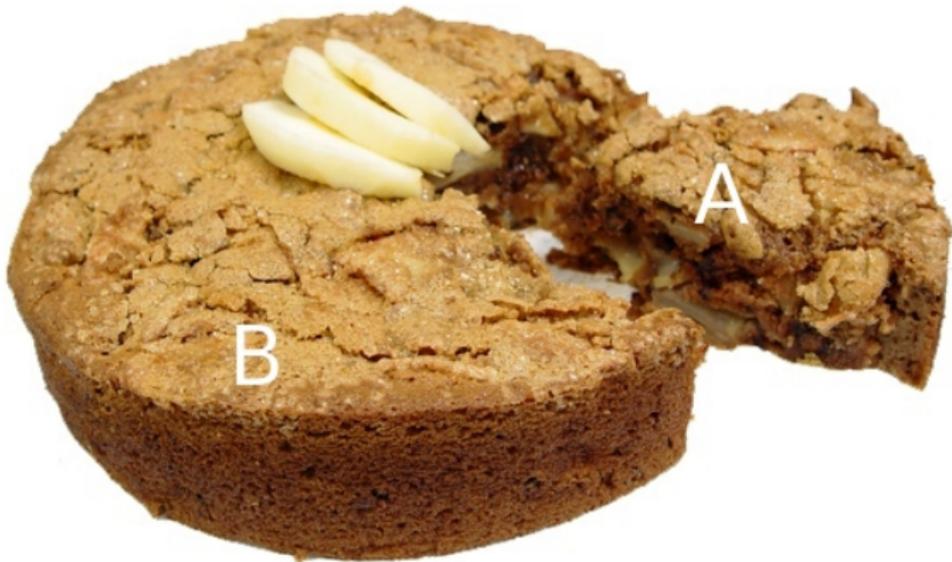
Example – ternary relations cannot be expressed directly

Nebraska is between South Dakota and Kansas.



Some things are impossible to express

If A is part of a whole then, there is remainder B that does not overlap with A.



OWL is probably the right choice if ...

- you have a classification problem
- you deal with a static situation
- you have only binary relations
- you need no 'modal' operators

In many other situations OWL might be the best solution



... but sometimes OWL does not fit

- need for variables
- recursive definitions

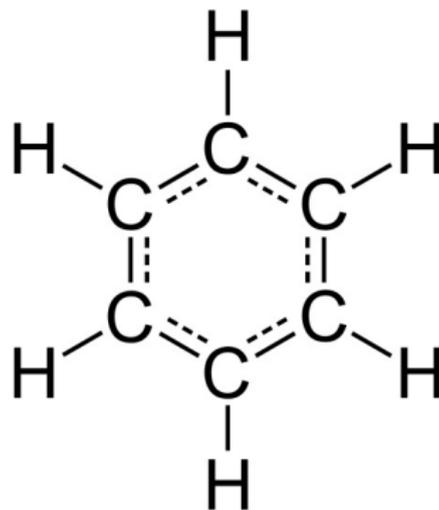


Mixing transitivity and cardinality

- Every protein dimer has `_part` exactly 2 protein
- Every mad-max complex has `_part` exactly 1 mad protein and has `_part` exactly 1 max protein
- Problem: OWL-DL does not allow mixing of transitivity and cardinality
- Solution: Use a different (non-transitive) relation
- Hurts interoperability

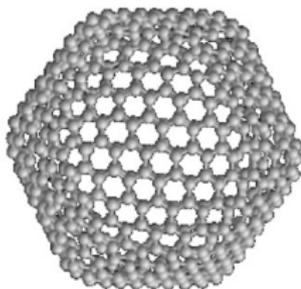
- Every bone of X is part of the skeleton of X
- We would like to have rules and variables such that we can say:
 - (bone and part_of some ?X) SubClassOf (part_of some skeleton-of-?X)

Cyclic structures



Need for variables to 'close the ring'

Fullerenes (Janna Hastings)



Can be defined with
Monadic Second Order
Logic (MSOL)

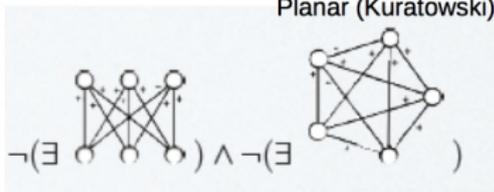
Cubic

$$\text{Cubic} \Leftrightarrow \forall x. \exists! 3y. \text{edge}(x, y)$$

3-connected

$$\forall x, y. \text{Connected_subgraph}(V \setminus \{x, y\})$$

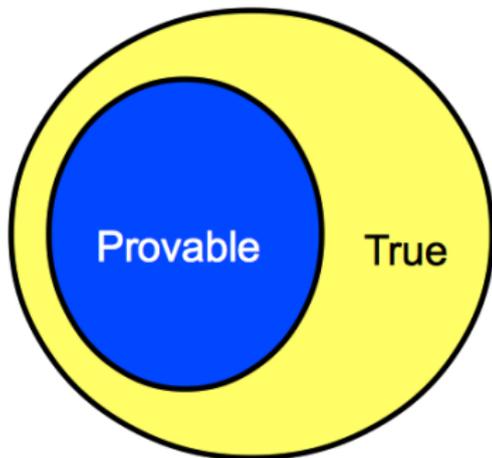
Planar (Kuratowski)



Probabilities (Melissa Haendel)

- 5% of adipocyte develop from neural crests
- 95% of adipocyte develop from pre-adipocyte

- Very expressive language
- Not decidable
- Not even a complete proof theory



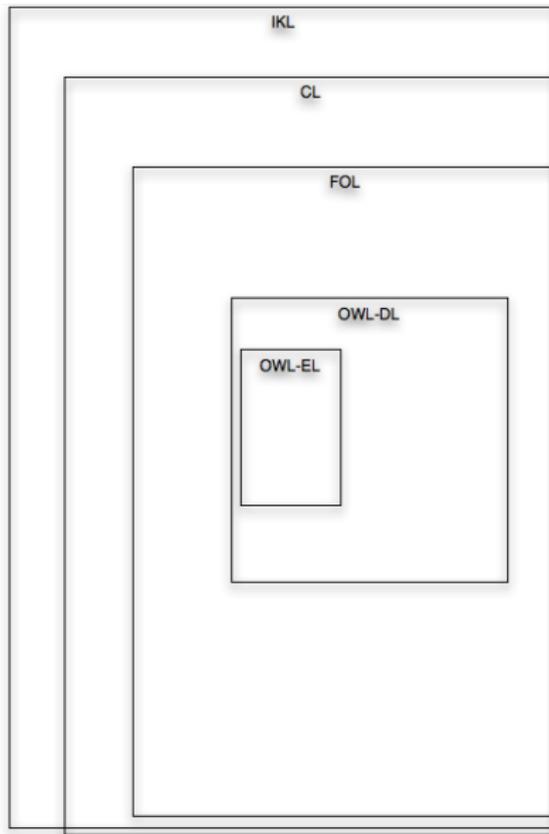
- Too limited language might lead to
 - loss of important information
 - bad ontology design
- Expressive language
 - usually harder to learn
 - bad (worst case) computational behavior

Available FOL reasoners are typically

- academic projects
- designed to compete on mathematical problems
- not designed to work on large axiom sets

Small change in axioms might cause them to get lost and never return

There are many choices (1)



There are many choices (2)

