"Why & when to use a classifier?"

Alan L Rector

OWL - or rather OWL-Lite and OWL-DL - have a semantics specifically designed to the use of "reasoners" that are computationally tractable.   Although OWL reasoning can be used in different ways, its original and still primary use is to classify the ontology - i.e. to check the consistency of class definitions and to organise them into a poly-hierarchy. Why use a reasoner to do this rather than organising the hierarchy manually? If the class hierarchy is organised when the ontology is authored, is a reasoner still needed at run time? What is the difference between "classification" as performed by OWL (and other description logic) reasoners and "querying"? We argue that:

a)   Using an automatic classifier demonstrably improves the quality of the resulting hierarchies both when originally authored, and more importantly, during routine maintenance and evolution.. The more complex and densely interconnected the ontology, the more the improvement. Whether complex multihierarchies can be maintained manually.

b)   The discipline of normalisation - or untangling - made possible by the use of automatic reasoning makes it much easier to modularise and maintain ontologies.

c)   Using automatic classification can reduce the overall effort required for both authoring and, particularly maintenance, although the initial investment may be greater.

d)   Most applications will use "pre-coordinated" classifications at run time. The reasoner is used only while authoring and maintaining the ontology.

e)   A modest number of applications require ontologies that are so combinatorially large that they cannot be pre-enumerated and must be compiled on the fly from re-usable fragments.

f)   Existing OWL reasoners are concerned almost entirely with analysing the class hierarchy and are inappropriate for reasoning about large numbers of instances for both theoretical reasoning.

In most cases the questions required to reason about classes are best formulated as "open world" classification; questions about instances as "closed world queries" A series of examples will be provided.