CARO 2.0

David Osumi-Sutherland

Department of Genetics, University of Cambridge, UK

Keywords: CARO, anatomy, ontology, OWL, PATO

The Common Anatomy Reference Ontology (CARO) [1] has been used extensively as an upper ontology for many anatomy ontologies (Teleostei, Hymenoptera, zebrafish, Drosophila, plants, even Dictyostelium). However, its application has not been very consistent.

There are a number of factors contributing to this inconsistency. Firstly, many of CARO's definitions are quite opaque. In particular, many use specialist terms (e.g. multi-tissue aggregate) that are not defined elsewhere in CARO or in references associated with the definitions. Secondly, it contains no equivalent class definitions or declarations of disjointness, although the original paper [1] makes clear that CARO was intended to be pairwise disjoint. Consequently, CARO cannot be used with a reasoner to aid ontology building via auto-classification and consistency checking. Thirdly, almost all users of CARO have re-implemented its terms in their ontology, rather than use an import system, leading to deviations from CARO in local implementations. Recently there has been increased interest in the development of multi-species anatomy ontologies, including efforts to build or refine anatomy ontologies for arthropods and vertebrates and plants. This work has highlighted many of the problems with CARO and its inconsistent application and made fixing them an urgent priority for the groups involved.

I will present details of a draft revised CARO [2], developed in OWL, that aims to correct these issues. In the new draft:

(a) Definitions have been simplified;

(b) Equivalent class definitions (intersections in OBO) have been used wherever possible, for example 'material anatomical entity' is defined as *EquivalentTo*: ('anatomical entity' *that* **has_quality** *some* PATO:mass);

(c) Disjointness has been widely declared, for example: 'material anatomical entity' *DisjointWith* 'immaterial anatomical entity';

(d) Useful terms missing from CARO, such as 'multicellular anatomical structure', have been added and defined;

(e) Some use has been made of the increased expressiveness of OWL over OBO, for example: 'multicellular anatomical structure' is defined as *EquivalentTo*: ('anatomical structure' *that* (**has_component** *min* 2 cell)), where **has_component** is a non-transitive *SubPropertyOf* **has_part.**

References

1. Haendel, M.A., Neuhaus, F., Osumi-Sutherland, D.J., Mabee, P.M., Mejino Jr., J.L.R., Mungall, C.J. and Smith, B. (2007) CARO – The Common Anatomy Reference Ontology: Principles and Practice. In Burger, A., Davidson, D. and Baldock, R.A. (eds), *Anatomy Ontologies for Bioinformatics*. Springer-Verlag.

2. caro_2.owl is available from:
   http://tinyurl.com/6a595gj
   (Loading requires a file of PATO terms: http://tinyurl.com/6dq6f22). Instructions for Protégé setup for viewing can be found here: http://tinyurl.com/69sxt3r.