Alignment between domain ontologies and SNOMED: three case studies

L. Mazuel, J. Charlet

AP-HP, STIM, Paris, France; INSERM, UMR S 872, eq. 20, Paris, France.

Abstract and Objective

Experiments on ontologies show increasingly clearly that the latter are only capable of representing small domains correctly and consensually. Hence, domain ontologies have been developed for particular applications, whereas reference ontologies tend to be used to draw together the results of specific applications. Here, we present the analysis and discussion of an alignment between three domain ontologies created by the I NS E R M UMR_S 872, Éq. 20 (OntoPneumo, OntoHTA and OntoReaChir) and the French translation of SNOMED v3.5. We propose a categorization of non-adherent decisions.

Keywords:

Ontology, Terminology alignment, Reference terminology, Inter face terminology, SNOMED.

Material and method

SNOME D v3.5 is a multidimensional structured classification that has been translated into French and currently contains 116,000 concepts and (including synonyms) a total of around 150,000 terms. It is organized into 11 axes (Terminology, Diagnosis, etc.), each of which is structured hierarchically.
The construction of the OntoPneumo ontology made intensive use of the terminological resources in the modelled domain. This ontology currently contains 1,113 concepts but does not use a top-ontology (i.e. using vocabulary from graph theory, OntoPneumo is currently a forest with 25 disjoint connected subgraphs).

OntoHTA arose from a research project on the determinants of medical reasoning, which resulted in the construction of an initial ontology and a suggested update of the clinical data entry forms in the field of arterial hypertension. At present, this ontology is a strictly taxonomic monohierarchy which complies with the principles of differential semantics and structures a total of 503 concepts.

Post-surgery intensive care is a specialized medical domain in the management of post-surgical complications and in traumatology]. Furthermore, OntoReaChir is an ontology of 2,135 taxonomically ranked concepts and a hierarchy of 200 relations.

There are various alignment methods available, the use of which depends on the ontological situations and formalisms encountered. We first use automatic method to produce a draft alignment, using morphosyntaxic methods.

Results

The final numbers of alignments after manual adjustment are given in Table 1. The additional, manual alignment helps markedly increase the number of matches (and even doubles it for OntoHTA).

Table 1-Alignments

Ontology	Cpts num.	Final mapping num.
OntoPneumo	1113	756
OntoHTA	503	210
OntoReaChir	3135	1152

Conclusion

We have presented here the results of a case study on the alignment of three domain ontologies with the SNOMED v3.5 terminology. The distribution of these alignments demonstrates the utility of these domain ontologies, relative to direct use of a generic model. In fact, the specialty ontologies are more appropriate in terms of the internal granularity and the level of detail for the most specific concepts. In addition, the ontological formalism is more complete and better designed for the definition of specific concepts generated in a post-coordination step - meaning that whole chunks of specialty ontologies are not at all represented in SNOMED v3.5.

A natural extension of this work would involve studying the alignments with snomed-ct; although the latter has links with SNOMED v3.5, our alignments could be directly reused as a working database. Furthermore, the OntoHTA ontology was based on SNOMED-CT (in contrast to the other two ontologies, which used ICD-10). Hence, this step would probably generate some very interesting results.

Secondly, concepts generated by post-coordination are not included in the domain ontologies in this article. However, it would be interesting to see how these particular concepts might align with SNOMED v3.5 and what the consequences would be.

Lastly, the need to develop domain ontologies and the difficulty of alignment prompt us to suggest that alignment with SNOMED during the ontology-building phase itself would be useful.