

# Why do it the hard way? The Case for an Expressive Description Logic for SNOMED

Alan Rector, Sebastian Brandt

School of Computer Science, University of Manchester, Manchester M13 9PL,

(rector | brandt @cs.manchester.ac.uk

Since SNOMED-RT/CT was originally formulated in the early to mid 1990s, there have been major developments in logic-based formalisms, ontology design and associated tools. Combined with the increase in computing power in the past two decades, these developments mean that many of the restrictions that limited SNOMED's original formulation and schemas no longer need apply. We contend that future development of SNOMED would be made easier if a more expressive formalism and more modern tools were adopted.

The difficulties in the existing structure of SNOMED have been well documented. For example, Bodenreider (1) examined the specialization hierarchy of SNOMED classes. Schulz discussed 'relationship groups' (2) and a broad range of other ontological problems along with potential remedies (3). Schulz suggested a modest extension of SNOMED's formalism to one with more clearly defined semantics (EL+) but which still lacks true negation and disjunction. We argue here that judicious use of a more expressive language, OWL 1.1<sup>1</sup>, is now practical and would bring great benefits including:

- A uniform, clear and understandable schema for all concepts used in clinical records, including context and negation.
- Elimination of the need for special mechanisms to deal with context, partonomy, and role groups.
- More effective leveraging of the underlying logical representation to organise and quality assure the SNOMED hierarchies.
- Improved ability to recognise semantic equivalence between post-coordinated and pre-coordinated expressions and between "observables" with "values" and the corresponding "findings."
- Improved ability to modularise and segment SNOMED for specific purposes
- Access to the tools and techniques being developed by the wider Semantic Web and OWL communities.

In outline, the proposals are:

- To represent all concepts used in clinical records (findings, observables, and procedures)

uniformly as fully defined "situations" that include any context required and that deal with negation explicitly and formally.

- To represent all sites explicitly as to whether they refer to the site in its entirety or to the disjunction of the site and its parts.
- To define observables and related findings in such a way that the classifier can be used to recognise the equivalence between a situation involving an observable with a given value and the corresponding finding of the observable with that value – e.g., between an observable of "blood pressure" qualified by "elevated" and a finding of "elevated blood pressure".
- To organise the stated form as a set of modules that can be separated for specific applications.

Details of the proposed mechanisms are described in the extended version of this paper and in (4, 5).

Although the effort to migrate any large software object should not be underestimated, most of the proposed changes would cause few changes to the schemas except for "Situations with specific context," which are known to be problematic. (However, the proposed analysis would identify many errors to be corrected.) The effort would be more than repaid by providing a more regular and consistent system that would improve usability and simplify software development and query formulation. We argue that a feasibility study using a modest subset of around 25K concepts should be an urgent priority for the SNOMED community.

## References

1. Bodenreider O, Smith B, Kumar A, Burgun A. Investigating subsumption in SNOMED CT: An exploration into large description logic-based biomedical terminologies. *AI in Medicine*. 2007;39:183-195.
2. Schulz S, Hanser S, Hahn U, Rogers J. The semantics of procedures and diseases in SNOMED CT. *Meth Inf Med*. 2006;45:354-358.
3. Schulz S, Suntisrivaraporn B, Baader F; SNOMED CT's Problem List: Ontologists' and logicians' therapy suggestions.; *Medinfo 2007*: IOS Press; 802-806.
4. Rector A, Qamar R, Marley T; Binding ontologies & coding systems to electronic health records and messages. 2006; *Formal Biomedical Knowledge Representation (KR-MED 2006)*
5. Rector AL; What's in a code: Towards a formal account of the relation of ontologies and coding systems. 2007; *Medinfo 2007*: Brisbane, Australia: IOS Press; 730-734.

<sup>1</sup> <http://www.webont.org/owl/1.1/>