

Reviewers' Comments (blue font) and responses (black font) to
A Unified Framework for Biomedical Terminologies and Ontologies

Werner Ceusters, Barry Smith

Reviewer 1:

The authors address one of the most important topics in the area of knowledge engineering: the reusability and interoperability of ontological resources. These objectives require a coordinated approach to developing ontologies. First, all domain specific ontologies should be based on a common upper level ontology. Secondly, ontologies should not overlap but should rather be joined under the same upper level ontology. And last but not least, all ontologies should be developed using the same universal constructs built on a clear understanding what these really "mean". In their paper the authors give a comprehensive overview of the different ontological dimensions which need to be respected when defining an ontology. All of them are explained by examples and integrated into a sound theoretical framework.

With regards to the application in real world environments I am not convinced if this will work because even for experts in this area it not always easy to understand the subtle distinctions which are made in this paper.

→ That is a matter of training and of becoming an expert in the method proposed here.

However, it is important to bring the ideas of the authors to the attention of the scientific audience.

Reviewer 2:

This paper proposes a framework for the analysis of "the different sorts of redundancy and ambiguity" in existing terminological systems (e.g., SNOMED CT). The three components of the framework are described (different purposes of definitions; types of entities; levels of reality), and their use is illustrated on examples from SNOMED CT.

This is a very important topic for the conference; the paper is clearly written and structured; the framework is well structured and builds on a strong line of work by the authors (with a negative side-effect on the coverage of related literature though, since references are highly centered on the authors' previous work: 12/19 are self-references).

→ This is a paper about some fundamental aspects of ontological realism, an area in which not that many scholars are active contributors (in contrast to the large group of researchers that are applying the principles).

However, this reviewer considers that the authors' view on the role of synonyms is fundamentally incorrect.

→ No, our view is not wrong, as we will demonstrate.

In their analysis of SNOMED CT's concept 'Fractured nasal bones (disorder)' and the 'synonyms' provided for it, ...

→ Here is already a first argument: this reviewer puts sneer quotes around the word 'synonyms' and is absolutely right in doing so: with him, we argue that all terms listed are indeed not synonyms of each other.

... the authors mix up (i) natural language words and expressions (these 'synonyms') and (ii) concepts (classes, entities) and their composition.

→ We do not mix up these notions, as witnessed by the provided self-references. To make it more obvious that we don't, we added the following sentence: " To this end, we must provide a framework for ontological analysis of terms in legacy terminologies that will support adequate mappings especially for those terms that, because they are declared as 'synonyms', are associated with single 'concepts' under the terminological view. Such terms must be mapped separately wherever they refer – on face value – to entities of different types."

The former obey natural language mechanisms, while the latter may conform to a formal framework such as that proposed in this paper. All the expressions listed as synonyms for 'Fractured nasal bones (disorder)' can be used to express the fact that there is a disorder, more specifically a fracture. In such an expression, 'Fractured nasal bones' does not denote bones, but a fracture. The same holds of sentence "Fractured nasal bones.", for instance in the Discharge Diagnosis section of a discharge summary.

→ That these terms can be used in expressions to assert some fact, does not make them synonyms. Furthermore, the reviewer mixes up 'disorder' and 'presence of a disorder'

In other words, such expressions are not to be construed as a way to define or specify the concept/class/entity 'Fractured nasal bones (disorder)', ...

→ That is correct. We did not argue otherwise in this paper.

... but as different ways a language user might express the presence of this concept. Therefore, the consequences derived by the authors ("it now being clear that they cannot continue to be designated as synonyms") do not hold.

→ In the above, the reviewer mixes up the presence of a concept in a terminology with the first-order entity in reality that can be denoted when such a concept is used.

CLASSES are introduced to handle the ambiguity in SNOMED's 'Finger structure' example, but seem to be much more general than what is needed for this: the authors may want to motivate their choice not to design a more specific device to address naming ambiguity. A consequence too is that 'Finger structure' is not the best example of an Ad Hoc Class (AHC).

→ We did not understand this comment. In the standard approach, there is nothing else than 'concept' and 'class' while here we provide a taxonomy of different sorts of classes, a rather very specific device we would argue.

In conclusion, this reviewer considers that the application of the authors' ontological framework to the analysis of SNOMED CT's redundancy and ambiguity is inadequate in several places.

→ The reviewer did (in the above) not give any argument for his claim: he only discussed our alleged wrong view of what synonyms are supposed to be, not inconsistencies or flaws in our delineation of the various sorts of entities that we describe to remove ambiguities.

Other details:

UNVs: is 'generic entities' a synonym of 'universals'?

→ No, a universal is one specific sort of generic entity. We made that more clear in the paper and used now 'portions of reality' instead to highlight that some portions of reality are not entities.

PARs: "Such entities are represented in typical ontology resources by nouns or noun phrases.": (1) note that 'general terms' in the above examples are also nouns or noun phrases;

→ That is correct. We removed the sentence.

(2) more specifically, all the examples here are 'named entities'; could the authors cite non-named-entity examples of particulars in the medical domain?

→ Sure, we added the following sentence: " Some PARs are what linguists would describe as 'named entities', but the majority – a liver cell in Mary, the fracture in her leg, and so forth – are not."

"Finger structure": by 'all fingers', do the authors mean "all fingers in the world" or "all fingers in a specific patient"?

→ Well, ask SNOMED. These sorts of ambiguities are we exactly talking about and we describe a framework for resolution

"[19]. Few, if any, existing biomedical terminologies exhibit these characteristics." Wouldn't the authors of [19] refer to their experience in annotating the GALEN ontology with natural language terms in different languages?

→ We don't understand what the reviewer tries to convey.

Reviewer 3:

This paper addresses three important but only loosely related themes, which are, to different degrees important in the context of biomedical ontology and knowledge representation: (i) ontology mapping / alignment, a highly relevant topic; (ii) a philosophical theory of representational units, which goes much beyond what is really needed for biomedical ontology engineering, and (iii) the problem of semantically close terms and concepts.

→ Interesting comment. We believe however that the problems we have identified over so many years in so many terminologies are for a large part due to the failure to see the three themes that we address as being closely related.

(i) The orthogonality discussion reaches short when it comes to certain annotation requirements. For instance, semantic annotations in the context of biomedical language processing would rather require a mid-granular, mid-level ontology encompassing basic biomedical terms like cell, organ, protein, i.e. the domain specific types of entities below BFO.

→ Sure, we did not claim otherwise.

There is no ontology in the OBO foundry that really provides these terms in a consistent and comprehensible way.

→ They are getting there, however. Nowhere we said that the OBO ontologies are perfect. We discussed principles here, not OBO ontologies.

Furthermore, annotators will have to solve the problem how to annotate those linguistic expressions that do not refer to biological reality, e.g. objects or processes of speculative existence. Therefore the focus on biological reality may be too restrictive.

→ That is accounted for by the 'levels of reality': speculations and hypotheses are L2.

(ii) The theory of representational units, as developed by the authors, is mostly plausible (with certain restrictions, e.g. is the cyclic definition of human (P2) really a good example?).

→ An example that fails its purpose, is indeed not good. We replaced it by another one.

However, good reasons are missing for many of the introduced types of representational units in the context of biomedical informatics. Recurring to speculative and “inexisting” objects, I had expected to see them included into a special subcategory of “defined class”. For instance, if there is SDC “Human born in Belgium”, what about the SDC “Human born on the moon?” Isn't there a need for this kind of representation? Even knowing that currently this class is empty, there may be scientific discourse about it, which may matter for annotation tasks.

→ This topic has been addressed in two earlier papers about 'negative findings'. Also in even earlier papers about the shortcomings of the concept-based approach.

(iii) Finally, the discussion on broken noses and fractures of noses only peripherally addresses the representational unit typology introduced before. What is at stake here is rather the phenomenon of ontologically distinct (and logically disjoint) but nevertheless closely related classes. Such “duals” are characterized by the fact that they mutually imply each other.

→ The goal is to make such distinctions, however close, explicit.

In order to be suited for publication in MEDINFO the paper should be revised in order to achieve a more uniform and consistent picture, possibly at the expense of some degree of detail in the theory of representational units.

→ We did so.