

LinKBase® and SNOMED: some distinct features and impact on NLP

Maria van Gorp, PhD, Marnix Holvoet, Mariana Casella dos Santos, MD
Language and Computing NV, Sint-Denijs-Westrem, Belgium
Tel: +32-(0)9-2808400; <http://www.landc.be>; {marjan, marnix, mariana}@landc.be

ABSTRACT

In this paper a description is presented in which the architectural, lexical and mapping differences are foregrounded between two compositional systems, both operating in the health care domain: LinkBase® and SNOMED. Based on these distinctive features, repercussions on NLP applications are exemplified and briefly discussed.

1. INTRODUCTION

The use of an ontology as a resource to access and aggregate several different types of medical data for a range of purposes inside healthcare information systems has demonstrated significant advantages. Nevertheless, this very variability of the healthcare information to be reconciled within and across different healthcare organizations, as well as the diversity of information systems accessing this information, imposes a challenge on the identification of the ontology of choice. This document intends to describe a direct comparison between two medical domain ontologies, The Systemized Nomenclature of Medicine (SNOMED)^{1,2,3} and LinKBase®⁴, from the perspective of their applicability to healthcare information systems and their embedded requirements. This comparison focuses on structural and lexical aspects, as well as differences in mapping methodology.

2. STRUCTURAL DIFFERENCES

2.1 Relationships and the principles behind them

SNOMED and LinKBase® are compositional systems: ontologies in which concepts can be specialized through combinations with other concepts. Both are based on Description Logics and contain binary relationships that interconnect the concepts. To enable semantic reasoning, a consistent meaning of the relationships is indispensable.

2.1.1 Relationships in LinKBase®

Concepts in LinKBase® are interrelated by a set of 383 relationship types that are structured in a multi-parented hierarchy, in which both the formal realistic ontological implications and the linguistic aspects of the relationships are taken into account. Most relationships are based on theories⁵, that deal with topics such as mereology and topology^{6,7}, time and causality⁸ and models for semantics driven natural language understanding^{9,10}. The large set of relationship types allows LinKBase® to define the sometimes subtle semantic differences between concepts (figure 1).

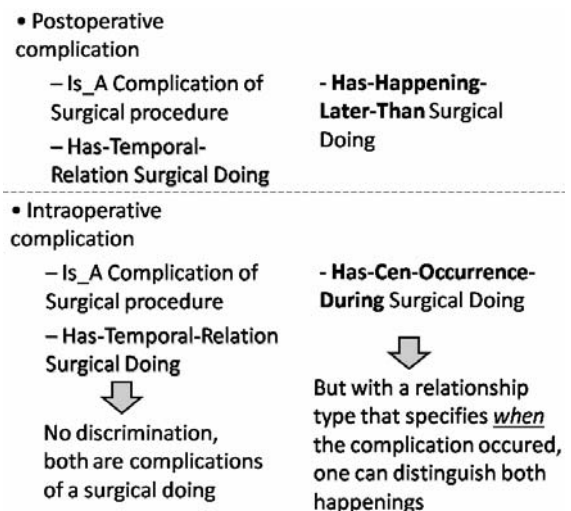


Figure 1- Detailed relationships types are needed to define subtle semantic differences

A large set of relationship types allows definition of subtle semantic differences as for example the relationship to time, which is needed to discriminate an intraoperative with a postoperative complication

In LinKBase®, consistency is maintained throughout the entire system for all types of relations by enforcement of an ontological principle named the “Principle of 100 % true relationships”⁴. According to this principle concepts can only receive a specific relationship if that relationship is true for *all* subclasses and instances of that concept. For example, in LinKBase®, “meningitis” will never be a subclass of “infectious disease” since it is not always the result of an infection. The concept “infectious

meningitis”, on the other hand, is always caused by an infection and is a subclass of both “infectious disease” and “meningitis”.

2.1.2 Relationships in SNOMED

The set of relationship types in SNOMED is much smaller as compared to LinKBase®. SNOMED contains 50 relationship types and although a more restrictive, smaller set of relationship types might yield an easier to manage and utilize ontology for a user, a more granular set of relationship types allows the introduction of unique formal definitions to a much larger set of concepts in the ontology. For example in SNOMED the concepts ‘intraoperative care’ and ‘postoperative care’ cannot be defined, because there are no relationships which are specific enough to relate two procedures in different time aspects. In LinKBase® on the other hand these concepts can be defined by adding a temporal relation to the concept ‘surgical procedure’, the former relation being ‘occurs-during’ and the latter ‘occurs-after’.

The SNOMED relationship types are divided into three types: defining, refining and additional relationship types. Only the former type ‘defining’ are used to insert truly ontological information used, as its name states, to logically define concepts. ‘Refining’ relationships on the other hand can be seen more as application supporting artifacts which allow concepts to be connected to ‘qualifiers’ in specific instances¹¹. For example, the concept ‘pneumonia’ contains the defining link *is-a* to ‘lung disorder’ and the refining link *clinical course* to ‘courses’. This latter relation allows for the relating of specific instances of pneumonia to one of the qualifier values under ‘courses’ as for example ‘acute’ or ‘chronic’. As in the example just described, for most concepts, the refining link inside the ontology/terminology itself, is an empty one. To the concept ‘pneumonia’, *clinical course* ‘courses’ does not add any useful information except that ‘pneumonia’ can have a ‘course’.

In order to cope with this distinction on the essence and use of relationship types SNOMED divides them in the three categories mentioned above. Also, when creating logical formal definitions for its concepts, only the ontologically based relations (‘defining’) are allowed to be used (see section 2.3). Although this restriction allows for the co-existence of both ontological and non-ontological information in the same syntax (i.e. binary relations), it does introduce problems when a refining characteristic becomes definitional for a given concept. For example the concept “acute inflammation” should ontologically be defined by being an ‘inflammation’ which has an ‘acute’ course. This becomes impossible given the ‘refining’ characteristic of the relationship type

‘*clinical course*’. The compositional model of SNOMED as such becomes limited to the specific boundaries of its relationship types and their characteristics. This impacts reasoning capabilities (see section 2.3) and applications which would rely on this resource such as decision support, run-time semantic interoperability (e.g. in messaging) and Natural Language Processing (NLP) and Natural Language Understanding (NLU) or more sophisticated Information Extraction.

SNOMED, just as LinKBase, aims at creating ‘defining’ relationships that are 100 % true. However, due to its structure and automated methods of hierarchy creation, inconsistencies are created (i.e. relationships that do not fulfil the “principle of 100 % true relationships”). A clear example is shown in figure 2 where ‘amputation of foot’ is incorrectly subsumed by ‘limb amputation’. These errors are the result of the SNOMED strategy to use Description Logic to create hierarchies based on other hierarchies. Although this is a valid method, the lack of specificity in relationship types creates a problem: by basing the procedure-hierarchy on the body part-hierarchy the ‘amputation of foot’ error was created.

The example in figure 2 is directly related to the fact that SNOMED has only one relationship type

```

SNOMED-CT : 362958002 : PROCEDURE BY SITE (PROCEDURE)
  SNOMED-CT : 118689006 : PROCEDURE ON BODY PART (PROCEDURE)
    SNOMED-CT : 118949002 : PROCEDURE ON EXTREMITY (PROCEDURE)
      SNOMED-CT : 363187007 : LIMB OPERATION (PROCEDURE)
        SNOMED-CT : 239573003 : LIMB AMPUTATION (PROCEDURE)
          SNOMED-CT : 367499009 : AMPUTATION OF FOOT (PROCEDURE)

```

Figure 2 - Incorrect subsumption in SNOMED
‘*amputation of foot*’ is incorrectly subsumed by ‘*limb amputation*’ since a foot is part of the limb, not the limb in total

to relate a procedure with the procedure site: the relationship type ‘*procedure-site*’. According to formal ontological theories¹² there are several ways through which a given procedure may be related to a body part. The body part might be removed or placed during the procedure (e.g. excision or transplant), it might be altered structurally (e.g. incision), it might receive another structure (e.g. implant) etc. For each of these cases the participation of the body part is different, and only for some of these cases ontological reasoning properties, like transitivity, (see section 2.2) apply. As such, applying transitivity over all instances of the ‘*procedure-site*’ relationship creates hierarchical inconsistencies like the one described above. This occurs due to the fact that not all instances of this relationship type within SNOMED are in fact transitive. In order to make relationship generation via reasoning, either at production time to expand the ontology, or at run-time to connect information to the ontology, it is necessary to have

relationship types logically defined and founded solely on sound ontological theory.

2.2 Top structure: Presence of Upper- and Mid-layer

Another structural aspect which determines an ontology's reasoning power and usability within specific information systems is the nature of its 'system of classification'. Here we refer to 'system of classification' when discoursing about the perspective into reality through which the concepts are organized or classified within the ontology. The 'system of classification' refers to the ontology's top structure and top classes which subsume the more granular ontological content. Ontological systems of classification may reside at three different layers: 1. The upper-layer, comprising of a small set of the most general classes formalized to be sufficient to described all that exists (examples of these top classes are 'process' (subsuming for example a 'hand-shake' or a 'surgical procedure') or 'property' (subsuming for example 'temperature')) , 2. The mid-layer, comprising of more specific classes, which are shared by different domains (for example 'temperature' (subsumed for example by 'fever' (medical domain) or 'atmosphere air temperature' (aviation domain)), 3. The domain specific layer, comprising those classes specific to a given domain and organized according the perspective this domain takes at the world.

A domain ontology in information sciences is commonly a data model, holding a set of concepts and relationships between those concepts for a particular domain of interest, and represents or reflects 'reality' through that model of domain. It is used to reason about the objects within that domain. Both LinKBase® and SNOMED are medical domain ontologies, dealing with those aspects which make up the world of medicine. Formal ontology implies that the model is governed by strict logical (formal) axioms; in the case of LinKBase®, the mereological, axiomatic scheme is applied, which results in a structure characterized by: reflexivity (a concept A is part of itself), anti-symmetry (two distinct concepts cannot be part of each other) and, transitivity: if concept A has a transitive relation to concept B, and B has the same transitive relation to concept C, then A has also this same transitive relation to concept C.

The difference between the two resources in respect to this structural aspect pertains to the type of system of classification or top layer structure that each of them applies. Structurally SNOMED is a shared, health care classification system generated and applied by the medical domain and its actors. Its main branches (18 totally) and embedded top nodes or concepts are derived from a

strictly medical classification perspective and reflect those entities through which the domain of medicine views and divides its realm (examples are Clinical Finding (main branch), Procedures (main branch), Body Structure (supporting branch), Substance (supporting branch), Organism (supporting branch), Context-Dependent Category (bridging branch), and Qualifier Value (bridging branch). LinKBase® applies an upper-layer and a mid-layer ontology at the top of its medical classification. The upper-layer is comprised of classes derived from the Basic Formal Ontology (BFO)⁵, while the mid-layer is comprised of those generic domain unspecific concepts which connect the domain specific layer to the upper-layer concepts⁴ (Examples of upper-layer concepts are (1) 'perdurants' (processes) or concepts with a time component, and (2) 'endurants' or concepts without a time component. Examples of mid-layer concepts are 'temperature' or 'movement'). From a usability perspective, the presence of an upper-and mid-layer influences on the degree of intelligence and specificity of logical reasoning and/or other process automation algorithms which can be applied over this ontology. Data integration and warehousing, semantic operability and NLP applications have the frequent need of mapping terminology (either in databases and terminological systems, data entry systems or free text documents) to the concepts in the ontology. Given the large degree of ambiguity in medical terminology, it is difficult to decide upon mappings between different concepts when performing terminological matching. The presence of an upper-layer ontology can strongly assist in this process. Figure 3 shows the example of an ambiguity which can be solved by combining language syntactical information with upper level ontological information: the term 'transposition' can either mean transposition as a surgical procedure (transposing a given part of the body to another part) (figure 3, panel A), or a transposition as a pathological structure (a part of the body which is constitutively displaced) (figure 3 panel B). The fact that the term in the sentence in figure 3A functions as the object of the verb 'repair' which in its turn instantiates the concept 'surgical repair'; and that this concept is subsumed by the upper level concept 'process' which can only have 'substances'¹³ as its participants, allows the deduction that the concept in question is the 'transposition' as a pathological structure due to the fact that this concept is subsumed by the upper-layer concept 'substance'. Opposed to the scenario above, in figure 3B the term transposition is related to the term 'technique' as if this latter was one of its ways through or one of its parts; this combined with the fact that a 'technique' is instantiated by processes and that processes can only have other processes as its parts, allows us to deduce that 'transposition' in this case must be a process and

therefore disambiguate to the surgical procedure¹⁴.

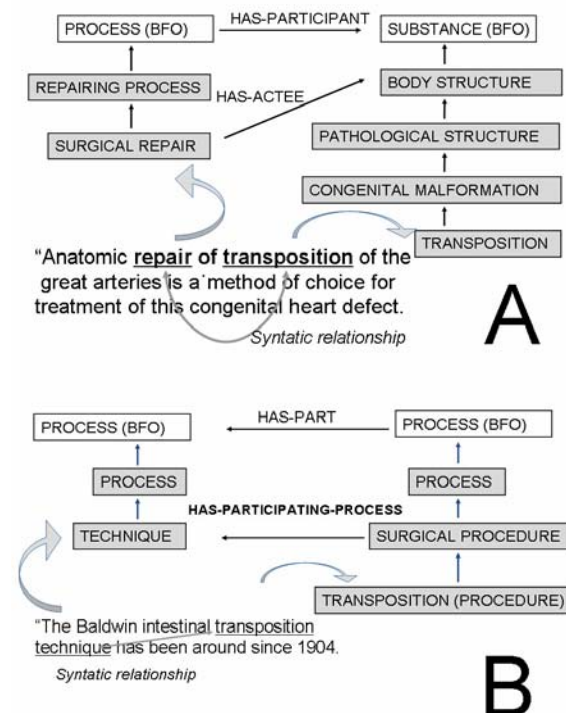


Figure 3 - Example of Ontology-based Terminological Disambiguation: 'transposition' as a congenital abnormality See text for details

Another consequence of the lack of an upper-and mid-layer structure is the need to create so called 'ad hoc hierarchies' to place those elements which do not pertain to any of the main domain classes, but which are nevertheless still important to represent the domain. An example of these is the main branch 'Qualifier Value' in the SNOMED system. The concepts under this hierarchy are only represented as values of an attribute or context used to represent other concepts or nodes inside SNOMED. As such nothing else can be understood or deduced about the essence and meaning of these concepts besides the fact that they can be used to 'qualify' other concepts. For example the concepts 'entrance' or 'exit' in SNOMED are classified only as 'any hazardous entity' without the ontological information of them actually being 'openings' (i.e. structures with a hole). Given the heterogeneity of the elements classified in such ad hoc top nodes, little to no formal reasoning can be applied safely over this content since no ontological property can be generalized and implied for it.

2.3 Methods and principles behind full definitions

Ontologies written in description logics¹⁵, such as the case of both ontologies being compared in this document, rely on an artifact called a 'formal definition' in order to apply reasoning and as such

explore the ontology's intelligence inside information systems. Formal definitions in description logics are elements composed of a (sub)set of a concept's relationships towards other concepts (both hierarchical and horizontal), which are supposed to uniquely define this given concept. For this particular ontological element, the distinction between the two ontologies here in question is given by the principles which govern the assignment of a formal definition to a concept, as well as by the methods used to insert and/or generate these formal definition assignments.

According to SNOMED a formal definition of a concept comprises "the set of relationships which together define that concept plus an indication of whether this definition fully-defines the concept (i.e. whether the concept is primitive or fully-defined)"¹¹, where **all** present defining relationships are used in the computation of this definition. For a defining relationship SNOMED understands those relationships which "are always known to be true"¹⁶ (see section 2.1 in this document). When a concept is marked as 'fully defined' in SNOMED it implies that all 'necessary' relationships required to uniquely define the given concept in SNOMED's terms have been asserted to that concept.

In counterpart to the computational method of asserting formal definitions of SNOMED described above, the formal definitions in LinKBase® are asserted manually by human domain experts in 100% of the cases. This allow for the introduction of an extra notion in the principle which govern this formal definition assignment. Besides the notion of the 'necessary' relationships, the notion of the 'sufficient' relationships is also taken into account. For LinKBase® the smallest set possible of relationships which is 'sufficient' to define a concept comprises the concept's full definition, instead of the complete set of defining relationships such as it is with SNOMED. Figure 4 shows an example of the distinction between a formal definition of a concept as assigned by SNOMED (top panel) compared to what the definition would be if the notions of necessary and sufficient would be applied (lower panel). By applying the notion of 'sufficient', the set of relationships which comprise a formal definition becomes smaller, resulting in a reduction of computing time when reasoning over the ontology, since fewer conditions must be cross-checked. In addition, data retrieval is enhanced since the set of conditions to fulfill in order to result into correct subsumption, is smaller.

Another distinction in the nature of formal definitions between these two ontologies is given by the degree of specificity of their relationship types. LinKBase® is very granular with respect to the creation of relationship types, allowing the creation of a new relationship type for each unique

NECESSARY

IS_A SNCT: 128332003: CONGENITAL ANOMALY OF DIGESTIVE ORGAN (DISORDER)
 AND
 IS_A SNCT: 118926004: DISORDER OF BILE DUCT (DISORDER)
 AND
 IS_A SNCT: 363026004: CONGENITAL ANOMALY OF BODY CAVITY (DISORDER)
 AND
 SNCT FINDING SITE: SNCT: 28273000: BILE DUCT STRUCTURE (BODY STRUCTURE)
 AND
 SNCT ASSOCIATED MORPHOLOGY
 SNCT: 107656002: CONGENITAL ANOMALY (MORPHOLOGIC ABNORMALITY)
 AND
 IS_A SNCT: 363024001: CONGENITAL ANOMALY OF ABDOMEN (DISORDER)
 AND
 SNCT OCCURRENCE SNCT: 255399007: CONGENITAL (QUALIFIER VALUE)

NECESSARY AND SUFFICIENT

IS_A SNCT: 49755003: MORPHOLOGICALLY ABNORMAL STRUCTURE (MORPHOLOGIC ABNORMALITY)
 AND
 SNCT OCCURRENCE SNCT: 255399007: CONGENITAL (QUALIFIER VALUE)
 AND
 SNCT FINDING SITE SNCT: 28273000: BILE DUCT STRUCTURE (BODY STRUCTURE)

Figure 4 - Necessary versus Necessary AND Sufficient conditions for assigning formal definitions to concepts See text for details

notion of how two concepts can be related in the real world¹⁷. SNOMED in counterpart is more restrictive, allowing only for a specific set of attributes (i.e. relationship types in SNOMED's terminology) to be used for particular hierarchies¹⁶ (see section 2.1.2.) In summary, an ontology with a vast set of formal definitions has an exceptional capacity to reason both inside itself and over any external data that is connected, mapped or pointed to the ontology.

3. LEXICAL DIFFERENCES

The connection between an ontology and its lexicon is given by the assignment of terms (natural language descriptions) in the lexicon to the concepts and relationships within the ontology, where the ontological content become the symbolized elements while their lexical counterpart are the symbols. Although the content, in terms of the nature or types of descriptions, of both LinKBase's® and SNOMED's lexicons is quite similar, the distinction between these two resources, regarding this particular aspect, is given by the assignments of language descriptions to the ontological content. In addition, the lexicon of SNOMED and LinKBase® differ in their grammatical content.

3.1 Principles behind and nature of synonym assignments

LinKBase® follows a defined set of principles¹⁷ for term assignment which intends to assert that every term assigned to an ontological element is strictly a natural language representation (synonym) for this specific element. SNOMED in counterpart allows for the association of not only terms that represent strictly the given ontological element in question, but also of other terms which are somehow related to this given element either from a taxonomic standpoint (i.e. more generic terms assigned to more specific concepts) or from a clinical perspective (e.g. terms representative of a symptom assigned to concepts representative of a disorder which can manifest with this given symptom). An example of this distinction showed in figure 5, which demonstrates the collection of terms for the concept which represents a "common cold" in both ontologies. While in LinKBase® all terms assigned

Concept COMMON COLD:

SNOMED concept 82272006	LinKBase® concept 3422018
<input type="checkbox"/> head cold	<input type="checkbox"/> common cold
<input type="checkbox"/> common cold (disorder)	<input type="checkbox"/> common head cold
<input type="checkbox"/> infective rhinitis	<input type="checkbox"/> ccold
<input type="checkbox"/> acute infective rhinitis	<input type="checkbox"/> cold
<input type="checkbox"/> acute nasopharyngitis	<input type="checkbox"/> virus cold
<input type="checkbox"/> Cold	<input type="checkbox"/> common cold syndrome
<input type="checkbox"/> common cold	
<input type="checkbox"/> infective nasopharyngitis, nos	
<input type="checkbox"/> acute coryza	
<input type="checkbox"/> acute nasopharyngitis, nos	
<input type="checkbox"/> acute rhinitis	
<input type="checkbox"/> acute nasal catarrh	
<input type="checkbox"/> infective nasopharyngitis	

Figure 5 - Term assignments in LinKBase® versus SNOMED This figure shows the collection of terms assigned to the concept 'Common Cold' both in LinKBase® (right) and in SNOMED (left) and exemplifies the difference in principles behind these assignments.

to this concept are strict synonyms in natural language for this notion, in SNOMED other related terms like "infective rhinitis" or "acute coryza" are also assigned. The former term "infective rhinitis" is actually related to the concept which represents a "common cold" from a taxonomic perspective (i.e. common cold is one of the types of infective rhinitis), while the latter is related from a clinical perspective (i.e. acute coryza is one of the symptoms which may be present in the presentation of a common cold).

But how does this distinction affect the usability of each of these ontologies? There are mainly three use cases where we can see a direct influence of one choice of term assignment vs. the other described above. From a search and retrieval perspective one can see a benefit from the broader methods of term assignment of SNOMED, since the user has a higher probability of finding a related concept for the information he/she intends to

encode given the larger spectrum of related terms to search for. On the other hand this larger spectrum of term assignment introduces more ambiguity as well as an 'indirect hierarchical inconsistency' (when the ontology is viewed from the lexical standpoint), which complicates processes for automated data integration, (semi)automated creation of mappings between other external resources to be reconciled via the reference ontology of choice, as well as for semantic interoperability between information systems relying on this ontology for their connectivity.

The usability of the ontology as terminological and intelligence resource inside NLP and NLU applications is also highly influenced by the principles applied behind terminological assignment. While broader (non strict) synonym assignment allows for simple uses of the ontology in NLP, such as indexing, it will not yield satisfactory results in more complex NLP/NLU information extraction applications. This is due to the fact that information extraction applications need to make use of the ontology to understand the very concept to which the text refers to, rather than other related content, and then place this identified concept into the context which surrounds it in the given text. Clinical associations in this case for example are highly disruptive, as it might lead the application to conclusions not necessarily true for the situation described by the given text being processed. For example if a mention of "acute coryza" is found in text when using SNOMED as terminological resource, it would be directly associated with the concept of "common cold". Nevertheless, in the particular text or instance in question, the "acute coryza" could as well be due to an allergy or to another disorder, which makes the association to "common cold" incorrect yielding erroneous results.

3.2 Specialized versus Non-Specialized lexicon

A computerized lexicon connected to a given ontology can contain different degrees of information about the terms or vocabulary it contains. SNOMED's lexicon is what we would call 'non-specialized' lexicon, constituting mainly from a collection of terms but without any extra grammatical information about these terms or the way they connect together. In counterpart LinKBase® is connected to what we would call a 'specialized' lexicon, containing for each of its terms (or lexemes in this case) information such as their part of speech or their inflections.

This distinction is vital when considering the ontology for use within NLP/NLU applications. A non-specialized lexicon will allow for the access of

the ontology only from a basic indexing or keyword tagging capability. More sophisticated NLP/NLU applications frequently make use of a syntactic parser, which makes the availability of grammatical information about the terms or in other words of a 'specialized' lexicon mandatory.

4. MAPPINGS AND MAPPING METHODOLOGY

Terminologies and classifications are used for different purposes and have different structures and content. To allow exchange of information between different data sources, they need to be connected. For this reason, LinKBase® and SNOMED are linked to 3rd party terminologies such as the International Classification of Diseases (ICD)¹⁸ or the Logical Observation Identifiers Names and Codes (LOINC®)¹⁹. Although both are linked to the original style and structure of these terminologies, they have a different strategy for building a bridge between them. Their solutions to deal with differences in granularity, an important but complex step²⁰, especially differ.

The mapping relationship between a LinKBase® concept and a concept or term in an external system is always of complete 'identity'. If needed, additional concepts are created to solve differences in granularity. In contrast, SNOMED solves differences in granularity with the creation of 'narrow to broad' or 'broad to narrow' relationships or no relationships at all, e.g. the concept SNCT2 : 276792008 : PULMONARY HYPERTENSION WITH EXTREME OBESITY (DISORDER), has a 'narrow to broad' relationship to the concepts: ICD-9-CM : 416.8 : OTHER CHRONIC PULMONARY HEART DISEASES and ICD-9-CM : 278.00/ : OBESITY, UNSPECIFIED. Needless to say, that the SNOMED approach results in an incorrect representation of the other terminologies. For example, the LOINC® codes for the two most common tests to diagnose pertussis in humans are LOINC® 548-8 and 549-6²¹. Although these LOINC® codes represent an assay involving a culture to test for the presence of Bordetella pertussis, both are linked to the SNOMED codes for the *organism* Bordetella pertussis and the *condition* pertussis since no such 'Bordetella test culture' exists in SNOMED. In LinKBase® however, this difference in granularity is solved by the creation of concepts that represent identical tests and/or cultures as in LOINC® and to link these to the condition and organisms involved. This method not only allows the correct representation of LOINC® and other terminologies in LinKBase®, but also allows for the reusability of existing mappings, the ability to cross map several data sources simultaneously and the ability to

transpose divergent levels of granularity between external information sources²⁰. The LinKBase® system allows physicians to enter a diagnostic term, to find its relationships based on its exact meaning and to select the terminology system they wish to use during a patient encounter. There is no need to search multiple terminologies, LinKBase® suffices, since the 3rd party terminologies are fully mapped, based on a consistent meaning without ambiguities.

CONCLUSION

In this paper, we have outlined some distinct features between SNOMED and LinkBase®. Currently, based on their differences in architectural and lexical approach, the principles behind these, together with their different mapping methodology, LinKBase® seems to be more suitable for use in NLP/NLU engineering. However, the comparison also provides a possibility to accommodate SNOMED to a level that it can be integrated in NLP technology and be used for the analysis of free text, as is the case for LinKBase®.

References

1. Côté, R.A., Robboy, S. (1980). Progress in medical information management: Systematized Nomenclature of Medicine (SNOMED), JAMA, 243; 756-762.
2. Spackman, K.A., Campbell, K.E. & Cote, R.A. (1997), SNOMED RT: a reference terminology for healthcare, proc AMIA Annu Fall Symp, 640-644.
3. SNOMED; <http://www.snomed.org/>
4. M. van Gorp, M. Decoene, M. Holvoet, M. Casella dos Santos. Proceedings of KRMed2006, LinKBase, a Philosophically-inspired Ontology for NLP/NLU Applications <http://ftp.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-222/krmed2006-p08.pdf>
5. Smith, Basic formal Ontology (BFO) <http://ontology.buffalo.edu/bfo/>
6. B. Smith, A. C. Varzi., Fiat and Bona Fide Boundaries, Proc COSIT-97 Springer-Verlag, 1997: 103-119.
7. Smith, Data and Knowledge Engineering 20, <http://ontology.buffalo.edu/smith/articles/meretopology.htm>, 1996.
8. F. Buekens, W. Ceusters, G. De Moor, The Explanatory Role of Events in Causal and Temporal reasoning in Medicine, Met Inform Med 32, 1993: 274-278.
9. W. Ceusters, F. Buekens, T. Deray, A. Waagmeester, The distinction between linguistic and conceptual semantics in medical terminology and its implications for NLP-based knowledge acquisition, Met Inform Med 37, 1998: 327-333.
10. Bateman, Ontology construction and natural language. Proceedings of International Workshop on Formal Ontology, Padua (Italy), 1993: 83-93.
11. SNOMED CT Abstract Logical Model and Representational Forms – November 2006 Revision to External Draft Guide http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/Technical_Docs/abstract_models_and_representational_forms.pdf
12. Degen, W., Herre, H., What is an Upper Level Ontology?, Workshop on Ontologies 2001, Vienna
13. B. Smith. On Substance, Accidents and Universals: In Defense of Constituent Ontology. Philosophical Papers 26, 105-127, 1997.
14. P. Grenon, B. Smith. SNAP and SPAN: Towards Dynamic Spatial Ontology. Spatial Cognition and Computation 4(1), 69-103, 2004.
15. F. Baader, D. Calvanese, D. McGuinness, D. Nardi, P. Patel-Schneider, The description logic handbook: theory, implementation, and applications. Cambridge University Press, New York, NY, 2003
16. SNOMED CT User Guide – January 2007 release http://www.ihtsdo.org/fileadmin/user_upload/Docs_01/Technical_Docs/snomed_ct_user_guide.pdf
17. Flett, M. Casella dos Santos, W. Ceusters Some Ontology Engineering Processes and their Supporting Technologies, in: Gomez-Perez A, Benjamins VR (eds.) Ontologies and the Semantic Web, EKAW2002, Springer 2002, 154-165.
18. International Classification of Diseases, Ninth Revision (ICD-9) <http://www.cdc.gov/nchs/about/major/dvs/icd9des.htm>
19. Logical Observation Identifiers Names and Codes (LOINC®) <http://www.regenstrief.org/loinc/>
20. M. Casella dos Santos, C. Dhaen, D. Decraene, M. van Gorp, T. Deray, The methodology behind the military health system conceptual framework and core ontology, 2005. http://www.landcglobal.com/images/TSB_Methodology.pdf
21. R. Wurtz, ELR, LOINC, SNOMED, and Limitations in public Health, WHP 0042-A, 2005