Ontology Engineering Assignment 3

Adriaan Louw (53031377)

July 1, 2018

1 Question 1

Snomed CT (Systematized Nomenclature of Medicine Clinical Terms) is a ".. the most comprehensive clinical terminology system" (Héja, Surján, & Varga, 2008, p1). It uses the formal top-level ontology DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) as its base (Héja et al., 2008, p2).

1.1 History

SNOMED CT can be traced back to 1965 when the College of American Pathologists published the Systematized Nomenclature of Pathology (SNOP). SNOP covered a large number of known medical conditions, since it was designed for the classification of surgical and autopsy diagnoses (Côté & Robboy, 1980).

This was further expanded into SNOMED in 1977. Unlike SNOP, SNOMED was a collaboration between various societies and not just the College of American Pathologists. It included various medical societies from the United States, England and Canada (Côté & Robboy, 1980).

SNOMED RT was a new logic based version published in 2000, in collaboration with Kaiser Permanente (NIH, 2018; SNOMED International, 2018a). SNOMED RT was merged with Clinical Terms Version 3 (CTV3) and released as SNOMED CT in January 2002. (SNOMED International, 2018a)

In 2007 the International Health Terminology Standards Development Organization gained the intellectual rights to all SNOMED versions (NIH, 2018).

1.2 Purpose & Use

According to IHTSO (2015) there are 3 types of implementations for SNOMED CT:

- *Clinical Records*: These implementations manage records of patients. Such systems can vary insize from specialized departments to large organizations. These systems provide means of entering new SNOMED CT expressions, storing and retrieving them. (IHTSO, 2015)

- *Knowledge Representation*: SNOMED CT is used in knowledge resources like decision support protocols, electronic reference books and clinical guidelines. In these implementations, SNOMED CT expressions are used to define concepts and relationships. In more sophisticated uses queries can be generated on stored information, like in a clinical decision support system.(IHTSO, 2015)

- *Aggregation and analysis*: The process of gathering data for later analysis.(IHTSO, 2015)

Using an ontology like SNOMED CT comes with a lot of benefits. It creates a " ...consistent way of indexing, storing, retrieving and aggregating clinical data across specialties and sites of care" (Cornet & De Keizer, 2008, p2).

Data can be easily reused for other purposes. For example if the organisation also use ICD codes, SNOMED CT can generate these codes based on the diagnosis entered in the SNOMED CT enabled system(Lee, Cornet, Lau, & De Keizer, 2013, p92).

A SNOMED CT Query language is being developed by SNOMED International. It does not build On W3C Web Ontology Language(OWL). But it does allow for SNOMED CT to be transformed into OWL using a perl script (Héja et al., 2008, p218).

1.3 Ongoing Maintenance

There have been a number of issues identified in implementations of SNOMED CT. Certain terms have been found to be ambiguous (Lee et al., 2013, p92). Like the term for "cold", does it refer to the "common cold" ("82272006|Common cold (disorder)|") or a "cold injury" ("11925005|Effects of reduced temperature (disorder)|")? Issues with hierarchical relationships have also been found. No subsumption relationship exists between the concepts "69973000|Vascular anomaly of eyelid (disorder)|" and "193966008|Eyelid vascular anomalies (disorder)|". The latter is only a congenital occurrence. even though both are conditions of the eye lid (Lee et al., 2013, p92). Additionally issues were found where the use of hyphens, full stops and commas we used inconsistently (Lee et al., 2013, p92).

There have been calls to IHTSDO for more guidance with respect to how to create subsets of SNOMED CT where the domain is large. For example reason for admittance (Lee et al., 2013, p92) A subset is a meaningful fragment of the ontology that encompasses everything related to a specific domain. This makes it easier for implementers an users to use. This fragment needs to be as small as possible but still contain all concepts that are needed for the specific domain (Patrick et al., 2008, p25).

Large ontologies like SNOMED CT, that is in constant use, requires constant maintenance hence a new version of the ontology is published every 6 months (SNOMED International, 2018b). This leads to an an additional challenge for clinicians. After every update the hierarchy has changed therefore the queries could also need to change over time (Lee et al., 2013, p92).

1.4 Conclusion

SNOMED CT is officially used in over 50 countries. But there are not a lot of implementations. It still has some issues to work out. But it has been of benefit to some (Lee et al., 2013, p93).

References

Cornet, R., & De Keizer, N. (2008). Forty years of SNOMED: A literature review. In *Bmc med. inform. decis. mak.* (Vol. 8, pp. 1–6). Retrieved from https://bmcmedinformdecismak.biomedcentral.com/track/pdf/10.1186/1472-6947-8-S1-S2 doi: 10.1186/1472-6947-8-S1-S2

Côté, R. A., & Robboy, S. (1980). Progress in Medical Information Management. *Jama*, *243*(8), 756–762. Retrieved from http://jama.jamanetwork.com/article.aspx?articleid=368815 doi: 10.1001/jama.1980.03300340032015

Héja, G., Surján, G., & Varga, P. (2008). Ontological analysis of SNOMED CT. In *Bmc med. inform. decis. mak.* (Vol. 8, pp. 1–5). Copenhagen, Demark. Retrieved from https://0-link-springer-com.oasis.unisa.ac.za/content/pdf/10.1186%2F1472-6947-8-S1-S8.pdf doi: 10.1186/1472-6947-8-S1-S8

IHTSO. (2015). *SNOMED CT ® Technical Implementation Guide 2015* (Tech. Rep.). Author. Retrieved from https://doc.ihtsdo.org/download/doc_TechnicalImplementationGuide_Current-en-US_INT_20150131.pdf

Lee, D., Cornet, R., Lau, F., & De Keizer, N. (2013). A survey of SNOMED CT implementations. *J. Biomed. Inform.*, *46*, 87–96. Retrieved from https://ac.els-cdn.com/S1532046412001530/1-s2.0-S1532046412001530-main.pdf?_tid=1c161356-6d6f-4e0a-94ae-4e46254f39ed&acdnat=1527329697_01f0958e81e1e12380178df1ec05f63b doi: 10.1016/j.jbi.2012.09.006

NIH. (2018). *Overview of SNOMED CT.* U.S. National Library of Medicine. Retrieved 2018-05-26, from https://www.nlm.nih.gov/healthit/snomedct/snomed_overview.html

Patrick, J., Wang, Y., Budd, P., Rector, A., Brandt, S., Rogers, J., ... Vazirnezhad, B. (2008). Developing SNOMED CT subsets from clinical notes for intensive care service. *Heal. Care Informatics Rev. Online*, *12*(3), 25–30. Retrieved from https://cdn.ymaws.com/www.hinz.org.nz/resource/collection/0F09C2E4-7A05-49FB-8324-709F1AB2AA2F/F38_Patrick.pdf

SNOMED International. (2018a). *History Of SNOMED CT.* Retrieved 2018-05-26, from https://www.snomed.org/snomed-ct/what-is-snomed-ct/history-of-snomed-ct

SNOMED International. (2018b). *How is SNOMED CT created?* Retrieved 2018-05-26, from https://www.snomed.org/snomed-ct/what-is-snomed-ct/how-is-snomed-ct-created

1 Question 2

1.1 Nature and Purpose of Foundational Ontologies

Foundational ontologies, also sometimes called upper or large scale ontologies, focuses on modelling very general and primitive objects and relationships (Borgo & Masolo, 2009). These ontologies are domain agnostic and try to model "the semantics of the real world" (Conesa, Storey, & Sugumaran, 2010).

According to Borgo and Masolo (2009), foundational ontologies have the following characteristics: "(i) have a large scope, (ii) can be highly reusable in different modeling scenarios, (iii) are philosophically and conceptually well founded, and (iv) are semantically transparent and (therefore) richly axiomatized."

Numerous foundational or upper level ontologies exist. For example SUMO (Pease, Niles, & Li, 2002), DOLCE (Borgo & Masolo, 2009), BFO (Arp & Smith, 2008), BORO (De Cesare, Henderson-Sellers, Partridge, & Lycett, 2015), and UFO (De Cesare et al., 2015).

Some believe that there will eventually be only a couple of foundational ontologies. Where people will prefer to share ontologies rather than have to translate between them constantly (Pease et al., 2002).

1.1.1 Advantages of using an ontology

Firstly, using a foundational ontology lets the ontology designer use concepts and models that have been thoroughly tested and designed by experts in the field (Pease et al., 2002). This speeds up development of the ontology and reduces errors, thus reducing cost. This can be put another way. According to De Cesare et al. (2015) foundational ontologies provide "general architectural infrastructure for roles".

Secondly, using a foundational ontology can reduce the amount of rework that needs to be done on an ontology due to changing requirements. Such ontologies anticipate changes because they have been designed to handle challenging modelling situations (Pease et al., 2002).

Using a foundational ontology allows interoperability, to a certain extent, between different domain ontologies that are interoperable (compliant) with the foundational ontology (Niles, 2001).

1.1.2 Differing philosophical positions

The designers of a foundational ontology need to decide which philosophical standpoints the new ontology will reflect. Here we will only briefly mention some of the more significant ones.

An ontology can be seen as either descriptive or revisionary. A descriptive ontology builds its models based on the assumed ontological structure of natural language and human thought. These ontologies do not limit their categories to "philosophical or scientific paradigms" (Masolo, Borgo, Gangemi, Guarino, &

Oltramari, 2003). Whereas revisonary ontologies limit all assumption to those that can be regarded as scientific.

There are 2 differing views when it comes to the nature of time. The endurantist sees an entity (as an individual or thing) as always completely or "wholly" present at every moment in time, where the perdurantist believes the same entity has different constituent parts at different times(Borgo & Masolo, 2009). In this way entities "endure" for the endurantists irrespective of time. As an example, say we want to model an individual named *Jane* in the 2 using the 2 different views on time. In the endurant representation the individual named *Jane* would always be present in the ontology irrespective of time. In the perdurantist representation *Jane* can only be partially present at any time. For instance from her birth to her death (De Cesare et al., 2015).

An ontology can also be regarded as either multiplicative or reductionist. Reductionist ontologies aim to have as few as possible primitives. They regard the reduction of complexity as a high priority. Multiplicative ontologies regards the need for expressiveness higher than that of the reduction of complexity and therefor tend to have many more primitives than reductionist ontologies (Masolo et al., 2003).

1.1.3 Areas that need improving

Conesa et al. (2010) list the following areas in which current foundational ontologies needs to be improved.

- Documentation in foundational ontologies needs to be improved. Including which domains the ontology covers.
- There is a lack of standardised graphical tools to interact with these ontologies.
- The ability to "search and summarize" concepts is also inhibited by the lack of tools.

1.2 SUMO

Numerous foundational ontologies were combined to create the Suggested Upper Merged Ontology (SUMO)(Pease et al., 2002). According Pease et al. (2002), amongst those are ontologies from ITBM-CNR, Stanford KSL and content based on the works of Sowa (Sowa, 2000), Guarino and colleagues (Borgo, Guarino, & Masolo, 1996), Allen (Allen, 1984) and Smith (Smith, 1996). Niles (2001) describes the main issues that were encountered during the merging process.

There are 11 sections with documented interdependencies. These include sections on structure of relation, entity abstraction, graph theory, and units of measure (Pease et al., 2002).

In SUMO, entities are grouped into 2 main categories: Physical and Abstract. Physical entities are entities that have positions in space and time. By making the 2 concepts under physical(Object and Process) disjoint, SUMO allows an endurantist approach. Abstract then includes all other entities (those with no time and space coordinates). This includes relationships.

Niles (2001) believes that since SUMO is open source and supported by the IEEE, SUMO has the advantage over commercial ontologies like Cyc. For instance, since it is not proprietary, it is safe to assume more people will use it.

1.3 DOLCE

DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) is another foundational ontology. The originators of DOLCE believe there should be a library of foundational ontologies. This library is called the WonderWeb Foundational Ontologies Library (WFOL). Developers who need a foundational ontology should be able to choose an ontology that best suit their needs and assumptions. DOCLE was intended to be a "reference" ontology. In other words it is intended to be used to in comparison with other ontologies, comparing how relationships are implemented and which fundamental assumptions were made (Masolo et al., 2003).

DOLCE can classify an entity as either an Endurant or a Perdurant. From the name it is a descriptive ontology and focussed on the ontological meaning in natural language. Additionally DOLCE regards itself as multiplicative (Masolo et al., 2003).

1.4 BFO

The aim of BFO is to assist in the integration of scientific data (Arp & Smith, 2008) and was developed at the IFOMIS institute in Leipzig (Masolo et al., 2003). BFO contains both endurant and perdurant entities while following a multiplicative approach (Arp, Smith, & Spear, 2015). Endurant entities are named continuants and perdurant entities are called occurants.

References

Allen, J. (1984). Towards a general theory of action and time. *Artificial Intelligence*, *23*, 123–154. Retrieved from `http://www.cs.ucf.edu/ lboloni/Teaching/EEL6938_2007/papers/Allen-GeneralTheoryActionTime.pdf` doi: 10.1016/0004-3702(84)90008-0

Arp, R., & Smith, B. (2008). Function , Role , and Disposition in Basic Formal Ontology. *Nature*, *2*(713), 1–4. Retrieved from `http://precedings.nature.com/documents/1941/version/1/files/npre20081941-1.pdf` doi: 10.1038/npre.2008.1941.1

Arp, R., Smith, B., & Spear, A. D. (2015). *Building ontologies with basic formal ontology.* Mit Press.

Borgo, S., Guarino, N., & Masolo, C. (1996). A pointless theory of space based on strong connection and congruence. *KR*, *96*, 220–229.

Borgo, S., & Masolo, C. (2009). Foundational choices in DOLCE. *Handb. Ontol. (Second Ed.*, 361–382. Retrieved from https://www.researchgate.net/profile/Stefano_Borgo/publication/225213719_Foundational_choices_in_DOLCE/links/00b495373a8bd031c3000000/Foundational-choices-in-DOLCE.pdf doi: 10.1007/978-3-540-92673-3

Conesa, J., Storey, V. C., & Sugumaran, V. (2010). Usability of upper level ontologies: The case of ResearchCyc. *Data and Knowledge Engineering*, *69*(4), 343–356. Retrieved from http://dx.doi.org/10.1016/j.datak.2009.08.002 doi: 10.1016/j.datak.2009.08.002

De Cesare, S., Henderson-Sellers, B., Partridge, C., & Lycett, M. (2015). Improving Model Quality through Foundational Ontologies: Two Contrasting Approaches to the Representation of Roles. In *International conference on conceptual modeling* (pp. 304–314). Retrieved from https://bura.brunel.ac.uk/bitstream/2438/11978/1/Fulltext.pdf

Masolo, C., Borgo, S., Gangemi, A., Guarino, N., & Oltramari, A. (2003). *WonderWeb Deliverable D18* (No. June). Laboratory For Applied Ontology - ISTC-CNR. Retrieved from https://www.researchgate.net/publication/245584936_WonderWeb_Deliverable_D18_Ontology_Library

Niles, I. (2001). Origins of The IEEE Standard Upper Ontology. *International Joint Conference on Artificial Intelligence*, 4–10. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.76.8966&rep=rep1&type=pdf http://www.researchgate.net/publication/2374394_Origins_of_The_IEEE_Standard_Upper_Ontology/file/9fcfd511f0fd592888.pdf%5Cnhttp://www.ontologyportal.org/pubs/IJCAI2001.pdf

Pease, A., Niles, I., & Li, J. (2002). The Suggested Upper Merged Ontology: A Large Ontology for the Semantic Web and its Applications. *Work. notes AAAI-2002 Work. Ontol. Semant. web*, *28*, 7–10. Retrieved from http://www.aaai.org/Papers/Workshops/2002/WS-02-11/WS02-11-011.pdf

Smith, B. (1996). Mereotopology: A theory of parts and boundaries. *Data and Knowledge Engineering*, *20*(3), 287–303. Retrieved from http://ontology.buffalo.edu/smith/articles/Mereotopology.pdf doi: 10.1016/S0169-023X(96)00015-8

Sowa, J. F. (2000). *Knowledge representation: logical, philosophical, and computational foundations* (Vol. 13). Brooks/Cole Pacific Grove.