Ontology Engineering Assignment 1

Adriaan Louw (53031377)

May 1, 2018

Question 1

An ontology can be defined as "An explicit specialization of a conceptualization" Gruber (1993)

Genesereth and Nielson (1987) claims that "A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose". We can see it as a mental model of the world. It is not yet expressed formally. Genesereth and Nielson (1987) also more formally define a conceptualization as, *a tuple,(D,* $\boldsymbol{R}$*) where*

- *D is a set called the universe of discourse*
- *$\boldsymbol{R}$ is a set of relation on D*

In a conceptualization of an insect classification system (insect taxonomy) the elements of D would be things like the insects themselves, what the insects eat, insect body parts ect. $\boldsymbol{R}$ would be relationships like insect A eats insect B or Unitary relationships like insect C can fly.

According to Guarino, Oberle, and Staab (2009) we need a language (formal or informal) to express our conceptualization. We say that the language commits to a conceptualization and that once we commit we only admit models that are intended(Guarino et al., 2009, p.8). In other words only the models that fit our conceptualization. In our insect example the relationship "to eat" could be interpreted in many ways. Does it mean that insect A can or cant eat and insect of the same type? Also if insect A can eat insect B and insect B can eat insect C, can insect A eat insect C.

The language that is used need to only have the relations with the meaning intended in the conceptualization. There should be no ambiguity.

Conceptualizations can be specified in 2 ways: *extensionally* and *intentionally*(Guarino et al., 2009, p.8). To *extensionally* specify our example we have to list every possible relationship in R, which is impossible Guarino et al. (2009). We can only partially specify our world. In contrast a we can specify our conceptualization *intentionally* by fixing "a language want to use to talk of it, and to constrain the interpretations of such a language in and *intentional* way." (Guarino et al., 2009, p.8) This can be done by *meaning postulates* Carnap (1956). For example we define that "to eat" in our previous example is reflexive and transitive. In other words an insect can eat an insect of the same type i.e. cannibalism. Also, if insect A can eat insect B, which can eat insect C, then insect A can eat insect C. Thus by *intentionally* specifying the conceptualization we have created an *approximate* specification Guarino et al. (2009), in contrast to the explicit specification we would get from specifying the conceptualization explicitly. Thus by specifying the conceptualization explicitly all assumptions will be made explicit.

References

Carnap, R. (1956). Meaning and Necessity. A Study in Semantics and Modal Logic. *The University of Chicago Press, second edition.*

Genesereth, M. R., & Nielson, N. (1987). *Logical Foundations of Artificial Intelligence.*

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220.

Guarino, N., Oberle, D., & Staab, S. (2009). *Handbook on Ontologies.*

Question 2

Today's content on the web is mostly human readable and not machine readable. This it difficult for machines to independently, form a human intelligence, find and deduce information. Current systems can only to text searches, but cannot deduce meaning. Hence the development of the Semantic Web. "The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users". (Berners-Lee, Hendler, & Lassila, 2001) Examples of services the Semantic Web will enable are information brokers, search agents and information filters. (Decker et al., 2000)

Agents will be able to deduce the meaning of the data and infer new knowledge from what it already knows. Because the agents will be using the semantics embedded in the web pages, we will not need AI with a level of sophistication of a human being. (Berners-Lee et al., 2001)

The semantic information has to be encoded into web pages using a machine readable language for these agents to function because semantics of data is rarely explicit. (Heiler, 1995) With this in mind, there are a couple of languages that have come forward in which we can do this encoding. RDF and RDF SCHEMA were the initial focus of the but we found to be lacking in expressive power. The World Wide Web Consortium (W3C) created the Web Ontology Group to develop an Ontology language for the Semantic Web. The relationship between terms are to be provided by this ontology in a structured vocabulary. Making it easy for agent to interpret unambiguously. (Horrocks, Patel-Schneider, & van Harmelen, 2003) The language developed by the Web Ontology Group is called Web Ontology Langyage or OWL.

Using common languages, allows agents to search through different web pages to find relevant data. But using the same language, like OWL, does not guarantee that the meaning of particular terms are the same across different web pages or domains. We can guarantee *semantic interoperability* (meaning of the data) by using the same standard ontology. (van Diggelen & Dignum, 2007). In such a case the meaning of all terms and relationships are the same by definition. The definition of the ontology.

But there are very many websites and databases that do not use the same ontologies. Different domains use different ontologies because each ontology focuses on a different aspect of reality.(Chandrasekaran, Johnson, & Benjamins, 1999) Different domains do not need to express the fine details of other domains that they will never use. Even within a domain there might be multiple ontologies describing the same domain. For example Grosjean, Soualmia, Bouarech, Jonquet, and Darmoni (2014) mentions multiple portals to search multiple ontologies in the Biomedical domain.

It might be of interest that we can distinguish between *Internal Interoperability* and *External Interoperability. Internal Interoperability* relates to interoperability between elements of a system. *External Interoperability* refers to the interoperability between systems. (García-Castro & Gómez-Pérez, 2011) We will be ocussing on *External Interoperability*

If it is not possible for systems to use the same ontology, some interoperability can be achieved by using the same "upper-level" ontology. "A top-level ontology describes very general concepts that are the same for every domain and are not dependent on task and purpose." (van Diggelen & Dignum, 2007)

Another option is *Ontological Alignment.* This involves creating mappings between the entities and relationships of the ontologies. The problem with this approach is the sheer number of mappings that need to be made. (van Diggelen & Dignum, 2007) An intermediate ontology can be used to reduce the number of mappings.(Ciocoiu, Gruninger, & Nau, 2000; van Diggelen & Dignum, 2007)

When some form of *semantic interoperability* is achieved between systems, these systems can be used to answer queries. This can be done through *Ontology based data access* (OBDA). "Ontology-based data access (OBDA) is a recent paradigm that proposes the use of an ontology as a conceptual, reconciled view of the information stored in a set of existing data sources." (Bienvenu & Rosati, 2015) In other words the queries are done through the ontology to the data. This data could be anything form a relational database, tripple stores or datalog engines.(Rodríguez-Muro, Kontchakov, & Zakharyaschev, 2013)

These queries can be formally expressed using Descriptive Logics which is a family of formal knowledge representation languages. There are different types queries. Some types include *Conjunctive Queries* and *Instance Queries.* Bienvenu and Rosati (2015)

These queries in Descriptive Logics format then needs to be translated into the language of the target system. For instance if the target systems use a relational database, the query needs to be translated to a SQL query. (Kontchakov, Rodríguez-Muro, & Zakharyaschev, 2013)

The Semantic Web is an extension of the the current Web, Berners-Lee et al. (2001), and allows for much more comprehensive searches than than traditional text based searches (Google,Bing). Allowing the user to specify not only the objects that they are searching for but the relationships these objects have.

References

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Sci. Am.*, *284*(5), 34–43.

Bienvenu, M., & Rosati, R. (2015). Query-based comparison of OBDA specifications. In *Ceur workshop proceedings* (Vol. 1350).

Chandrasekaran, B., Johnson, T. R., & Benjamins, V. R. (1999). Ontologies: what are they? why do we need them? *IEEE Intelligent Systems and Their Applications*, *14*, 20–26.

Ciocoiu, M., Gruninger, M., & Nau, D. S. (2000). Ontologies for Integrating Engineering Applications. *Journal of Computing and Information Science in Engineering*, *1*(1), 12–22.

Decker, S., Melnik, S., Harmelen, F. V., Fensel, D., Klein, M., Broekstra, J., … Horrocks, I. (2000). THE SEMANTIC WEB: The Roles of XML and RDF. *IEEE Internet Comput.*, *4*(5), 63–74.

García-Castro, R., & Gómez-Pérez, A. (2011). Perspectives in Semantic Interoperability. In *The semanticweb research and applications 8th extended semanticweb conference* (p. 16).

Grosjean, J., Soualmia, L. F., Bouarech, K., Jonquet, C., & Darmoni, S. J. (2014). An Approach to Compare Bio-Ontologies Portals. In *Studies in health technology and informatics* (Vol. 205, pp. 1008–1012).

Heiler, S. (1995). Semantic Interoperability. *ACM Computing Surveys*, *27*(2), 271–273.

Horrocks, I., Patel-Schneider, P., & van Harmelen, F. (2003). From SHIQ and RDF to OWL:The Making of a Web Ontology Language. *J. Web Semant.*, *1*(1), 7–26.

Kontchakov, R., Rodríguez-Muro, M., & Zakharyaschev, M. (2013). Ontology-based data access with databases: A short course. In (Vol. 8067 LNAI, pp. 194–229).

Rodríguez-Muro, M., Kontchakov, R., & Zakharyaschev, M. (2013). Ontology-based data access: Ontop of databases. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, *8218 LNCS*(PART 1), 558–573.

van Diggelen, J., & Dignum, F. (2007). Developing semantically interoperable e-commerce systems. In *Proceedings of the ninth international conference on electronic commerce* (pp. 117–126). New York, NY, USA: ACM.

Question 3

Web Ontology Language (OWL) was created as a standard by the Web Ontology Working group, which in turn was formed by the World Wide Web Consortium (W3C). Grau et al. (2008) OWL was created because of a number of limitations in expressiveness was found in RDF and RDFS. The reason RDFS was not simply extended, was because of "... the trade-off between expressive power and efficient reasoning" Antoniou and Van Harmelen (2009). In other words, if the RDFS was made more expressive it would have had to be at the expense of being able to write efficient reasoners for it.

There are 3 language variants of OWL. Owl Lite, OWL DL and OWL Full. Where OWL DL has more expressive power than OWL Lite and OWL Full in turn has more expressive power than OWL DL. (Consortium, 2004)

OWL uses various commands from RDFS. For instance *rdfs::comment* is used to store comments on ontology elements, where as *rdfs::label* is used to add human readable names to elements in the ontology. (Horridge et al., 2011) Additionally *rdfs::subClassOf* is used to define whether one class is a subclass of another class. (Antoniou & Van Harmelen, 2009)

There are a number of useful features that cannot be expressed using RDF and RDFS. For instance:

- Whether 2 classes are disjoint or mutually exclusive. This can be accomplished using *owl:disjointWith* in OWL.

- Cardinality cannot be restricted. For example we cannot say that a store has only one owner. OWL uses *owl::cardinality* to express this concept. OWL also has *owl::minCardinality* and *owl::maxCardinality* to express a minimum or a maximum cardinality when the cardinality encompasses a range. For example lecturer must have a minimum of 5 students and a maximum of 400.

- New classes cannot be expressed as the union, intersection and/or complement of other classes. *owl::complementOf* can be used to declare a class to be the compliment of another class. *owl::unionOf* is used to declare a class the uniion of a set of classes and *owl:intersectionOf* is used to declare a class to be the intersection of a set of classes.

- The scope of properties cannot be made local. The operation rdfs:range defines the scope of a property for all classes. In OWL, *owl::Restriction* in conjunction with *owl::allValuesFromm* *owl::someValuesFrom* and *owl::hasValue* can be used to place various restrictions on classes.

- The RDFS cannot express whether a property is transitive or symmetric.
  This is solved in OWL using *owl::TransitiveProperty* and *owl::SymmetricProperty*.

- RDFS cannot express when one property is the inverse of another property.
  Owl uses *owl::inverseOf* for this functionality.

(Antoniou & Van Harmelen, 2009)

In conclusion, we can see that OWL was created to provide features that RDFS could not or it would not have been feasible to provide.

References

Antoniou, G., & Van Harmelen, F. (2009). Handbook on Ontologies. In S. Staab & R. Struder (Eds.), *Handbook on ontologies* (Second Edition ed., pp. 91–110). Springer. Retrieved from `http://link.springer.com/10.1007/978-3-540-92673-3` doi: 10.1007/978-3-540-92673-3

Consortium, W. W. W. (2004). *Owl web ontology language semantics and abstract syntax* (Vol. 10).

Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P. F., & Sattler, U. (2008). OWL2: The next step for OWL. *J. Web Semant.*, *6*(4), 309–322.

Horridge, M., Knublauch, H., Rector, A., Stevens, R., Wroe, C., Jupp, S., … Brandt, S. (2011). A Practical Guide To Building OWL Ontologies Using Protégé 4 and CO-ODE Tools Edition 1.3.