Measuring the quality of ontology mappings: A multifaceted approach.

Jennifer Sampson

Department of Computer and Information Science
Norwegian University of Science and Technology
Trondheim, Norway
Jennifer.Sampson@idi.ntnu.no

Abstract. While there is considerable research towards developing solutions and methodologies for ontology mapping, there is very little research towards evaluating the quality of both the tools and the outputs of the mapping process. Therefore the goal for this research is to develop a methodology and formal quality measures for evaluating ontology mapping results. This research will also require the definition of experimental protocols with fixed controls that will make the results of different evaluations comparable. Initial quality goals to be formally defined are syntactic, semantic, taxonomic and procedural quality.

1 Introduction

Ontologies are to provide a shared understanding of common domains, and provide a means for data exchange at a syntactic and a semantic level. Recently several ontology mapping techniques and approaches have been proposed as one step towards achieving interoperable systems. A mapping between ontologies means that for each concept (relation, property) in one ontology we try to locate a corresponding concept in the second ontology, with the same or closest meaning.

Problems with interoperating systems stem from the fact that the developers of information systems use different vocabularies to express the information contained in the systems even when describing the same domain. In many cases applications require information from multiple information sources. However, an application can only use information from two systems completely and accurately if it is able to establish the precise semantic correspondences among the concepts in the two systems. This is a very difficult problem to solve and current approaches often result in missing correspondences – whereby both matches and mismatches can produce errors of omission [10]. The implications of using incorrectly matched information in decision making are vast. [10] comments that, “for information sources using diverse vocabularies, an application that interoperates between the sources needs precise and complete articulations of the ontologies that contain the specification of those vocabularies” (p.53).

Ontology mapping is where we relate similar concepts or relations from different ontologies to each other by an equivalence relation. Ontology alignment is closely related but is where we bring together two or more ontologies into mutual agreement, with the aim of making them consistent and coherent. There are many tools which can, to differing degrees, achieve a matching via either term-based matching, structure-based matching or instance-based matching.

2 Research Purpose

The research goal is to develop a methodology and formal quality measures for evaluating ontology mapping results. One sub goal is to evaluate ontology mapping tools by determining the quality of the mapping outputs. Another research sub-goal is to define formal metrics for comparing the distance between ontologies. This work will also require the definition of experimental protocols with fixed controls that will make the results of different evaluations comparable.

To determine the quality of mapping results we may also need to determine the quality of individual ontologies before the mapping/merging process occurs. Initial quality goals to be defined are syntactic, semantic, taxonomic and procedural. In order to evaluate the quality of an ontology within these quality goals we need to know what kinds of things are there and can there be, in an ontology. For determining semantic and taxonomic quality we need to be able to formally define the correctness or otherwise of the taxonomic structure. [4] have defined Ontoclean as a methodology for ontological analysis. They propose four fundamental notions of formal ontology: identity, unity, essence and dependence, and attempt to show how they can be used as foundation of a methodology for conceptual modeling and ontology modeling. However [5] claims the formal statements of their proposals as logical schemata are flawed in many ways. Moreover [5] argues that [4] use an inconsistent notation making the intended semantics of the logic unclear, and [4] make false claims of logical consequence, and definitions that provably result in the triviality of some of their property features.

2.1 Research questions

The main research questions are:
"How can we measure the quality of an ontology mapping given we have diverse tools, and diverse user requirements which impact on the degree or type of mapping?"
How do we define practical criteria (quality characteristics) for evaluating the quality of ontologies and ontology-mappings?

2.2 Sub-questions

What is ontology-mapping and how can we classify the different approaches to mapping?
What parts of an ontology can be semantically interconnected?

Why evaluate the quality of ontology-mappings? Is it possible to evaluate the quality of ontology-mappings? Do we need to evaluate the quality of individual ontologies as a prerequisite for evaluating the quality of the ontology-mapping result?
What current tools are available for ontology-mapping, and what kinds of mapping techniques are supported by the ontology-mapping tool?
How can we measure the distance between two or more ontologies?
How can we measure how similar two ontologies are? Is this a more useful indicator than one which tells us how different two ontologies are?

2.3 Contribution

The research will define quality goals and quality metrics for evaluating ontology mapping results. To assist with defining these quality goals we are currently undertaking an indepth analysis of the ontology mapping and merging tools Chimaera and Prompt, to establish what are the existing problems ontology engineers face when using such tools. Once we have established an understanding of the problem we will consider formal theoretical frameworks and approaches as candidates for defining the quality measures and goals (e.g. psychologically motivated theories versus philosophically motivated theories e.g. BWW).

Once we have established the quality metrics we will conduct an empirical evaluation of current available tools (Chimaera [8], Prompt[12], Glue [1], OntoMerge[2], ONION [10]). These experiments will be to compare ontology mapping results using the proposed quality goals and metrics. The main contributions will be the development of a comprehensive methodology for ontology-mapping which describes how the proposed quality measures can be applied. We also hope to provide recommendations for future enhancements to existing tools and/or extension of an existing ontology tool which provides an automated mechanism for evaluating ontology-mappings.

In addition we would like to further test the quality metrics using a prototype agent-based ontology-mapping system. While we recognise that fully automatic ontology mapping is still impossible, we believe that software agents can assist humans in several aspects of ontology mapping for example finding useful concepts and relations among concepts from original data sources. Parts of ontologies, such as consistent concepts in data or concurrence of two concepts, can be discovered by software agents. The results of such discoveries however should be subjected to review of domain experts, or possibly other software agents. Therefore we claim that, in principle, it is possible for software agents to participate in the ontology mapping exercise. In addition we note that the ontological approach for the semantic web forces a specific approach namely taxonomic hierarchies, where the ontology describes what things are and not necessarily what they are used for. We need to consider whether taxonomic hierarchies are appropriate for meeting specialist objectives and relevant to users goals. We may consider the use of different types of hierarchies for web services for example – hierarchies that are more specific to application or workflow.

2.4 Algebraic semiotics

Algebraic semiotics [3] studies sign systems and morphisms in the framework of category theory. [3] has developed a mathematically precise theory of semiotics, called algebraic semiotics and has applied the mathematical fomalism to user interface design. In his research, [3] describes the notion of semiotic morphisms in the context of sign systems. A semiotic morphism transforms such a sign system (the source) into another (the target) supposed more suited to a particular use. Our claim is that ontology mapping can be seen as a semiotic morphism from one system of concepts to another. Semiotic morphisms are a way to describe the mapping, translation, interpretation and representation of concepts in one ontology to concepts in another ontology.

If we consider two ontologies O_1 and O_2 which are first mapped and then merged into one ontology O_G we can check the quality of the resulting merged ontology by investigating if the structure of the individual ontologies is preserved. [3] claims that "A good semiotic morphism should preserve as much of the structure in its source sign system as possible", but recognises that in many real world examples, it is not possible to preserve everything. [3] suggests that the structure that is preserved by semiotic morphisms provides an important way to compare their quality. Obviously a semiotic morphism $M : O_1 \rightarrow O_G$ need not be total, some concepts in O_1 may have no representation in O_G, we also know that some of the axioms, properties and relations of O_1 could be lost. We propose the use of semiotic morphisms as a way to measure the quality of a ontology merging process as available in the Prompt ontology merging environment [12]. Using algebraic semiotics it is possible to define and then determine a value for expressing the quality of two semiotic morphisms $M : O_1 \rightarrow O_G$ and $M' : O_2 \rightarrow O_G$.

3. Preliminary findings

This research is partly motivated by the fact that existing measures for assessing ontology pairs, for example Precision and Recall evaluation are insufficient or non existent. [11] describe tools for mapping, aligning, and merging ontologies as "tools that help users find similarities and differences between source ontologies". They claim that mapping tools either identify potential correspondences automatically or provide the environment for the users to find and define these correspondences, or both. However [11] point out that even though theories on how to evaluate either type of tools are not well established, there are already several frameworks for evaluating ontology development tools. [11] also examined the aspects that these frameworks compare: interoperability with other tools and the ability to import and export ontologies in different representation languages; expressiveness of the knowledge model; scalability and extensibility; availability and capabilities of inference services; usability of the tools. In our research we will consider and extend these aspects to develop a more formal, comprehensive approach.

3.1 Understanding quality measures for ontology mapping

While we recognise the importance of evaluating ontology tools our research is to more formally define specific quality goals for evaluating the results of an ontology mapping. Firstly we need to establish a consensus on what the desirable characteristics are required of an ontology-mapping. In order to create a meaningful set of evaluation criteria it is essential to determine the purpose, or purposes, the ontology-mapping is intended to serve. (We refer to conceptual model quality as a reference for ontology model quality). Until the publication of Lindland et al.'s paper in 1994 [7], research into determining quality in conceptual models had focused almost exclusively on the quality of the end product and resulted in a number of lists of desirable features and properties. [6] observe that previous lists of desirable properties for conceptual models have not provided a "systematic structure for evaluating them" and consequently propose a framework that not only "identifies major quality goals but gives the means for achieving them" (p.43).

The framework of [7] is primarily based on the linguistic concepts of syntax, semantics and pragmatics, in recognition that "modelling is essentially making statements in some language" and it is thus able to subsume all the characteristics they had listed from previous studies. They observed several significant trends:

· Many definitions are vague, complicated, or in some cases, even lacking...
· The list [of characteristics] is unstructured and the properties are partly overlapping...
· Specification properties are mixed with language and method properties...
· Some properties presuppose the existence of a design and even an implementation...
· Some goals are unrealistic, even impossible to reach" (ibid. p.43).

Their framework is intended to deal with those issues as well as making a clear distinction between the goals and the means, i.e. "by separating what you are trying to achieve...from how to achieve it" (ibid. p.42). The framework of [6] extends that of [7] by incorporating one further concept, participant knowledge of the domain, and two further dimensions, perceived semantics and social agreement.

In the context of ontology-mapping we have preliminary descriptions for quality assessment:

Syntactic quality: The relationship between the ontology mapping and the language used for the ontology modeling (syntactic correctness).

Semantic quality: The relationship between the ontology mapping and the ontology domain(s) (semantic correctness and semantic validity).

Taxonomic quality: The correctness of the semantic relations in the ontology. Due to the way in which ontologies are constructed we must consider that all ontologies represent valid viewpoints. Nonetheless, while there is no, one notion of a "correct" classification or taxonomy, we recognise that some taxonomic structures are more correct than others. Taxonomic quality may include the evaluation of the subtyping behaviour for each type rank (taxonomic correctness and taxonomic richness) .

[4] describe the ontological well-foundedness of the is-a partial-ordering relation. They discuss techniques based on the philosophical notions of identity, unity, essence, and dependence, which have been adapted to the needs of information systems design. They claim that they have demonstrated the "effectiveness of these techniques by taking real examples of poorly structured taxonomies, and revealing cases

of invalid generalization. The result of the analysis is a cleaner taxonomy that clarifies the modeler's ontological commitments"[4]. However like [5], we argue that the mathematical execution of their formalisation is problematic. In our research taxonomic quality will include determining the correctness or otherwise of 'is-a' and 'part-of' relations. We recognise that the 'is-a' construct is misused in the design of many ontologies, which potentially has large consequences for mapping ontologies based on incorrect classification systems.
Pragmatic quality: The relationship between the ontology mapping and the ontology designer's interpretation of the ontology mapping.
Procedural quality: The appropriateness of the approach followed to achieve the mapping.

3.2 Transformation – information loss or gain?

We can also examine the quality of an ontology mapping from the perspective of the information lost in the transformation or the information gained from the transformation. Missikoff [9] has defined possible interoperability clashes caused by differences in the conceptual schemas of two applications attempting to cooperate. The possible clashes are lossless clashes and lossy clashes:

- Lossless clashes, which can be solved with no loss of information. A lossless clash is where there exists a transformation that, given an information input produces an information output with a different representation but with the same semantic content. [9] comments that these clashes are quite intuitive and essentially of a syntactic nature. Examples include naming clashes, when the same information is represented by different labels; structural clashes, when information elements are grouped in a different way; and unit clashes, when a scalar value (typically an amount of money, or a distance) is represented with different units of measure.
- Lossy clashes, which include the clashes for which any conceivable transformations (in either direction) will cause a loss of information. [9] comments that typical cases are information represented at different levels of granularity, refinement, or precision. Another case of information loss is the presence, locally, of information that is not represented at all in the ontology [9].

3.3 Difficulties with evaluating ontology-mapping tools

Noy and Musen [11] point out that comparing and evaluating ontology development tools is very different from the comparison and evaluation of ontology mapping tools. While ontology-development tools have very similar inputs and desired outputs [11] ontology-mapping tools vary with respect to the precise task that they perform. While we can broadly say that ontology mapping tools are designed to find similarities and differences between source ontologies, [11] have identified three obstacles we need to consider when evaluating ontology-mapping tools. First they claim that tasks for which the mapping tools are designed, greatly. From the user's point of view the tools greatly in what tasks this analysis of similarities and differences supports. Second the different mapping tools rely on different inputs (e.g. only class hierarchies of the sources, classes and slots and value restrictions, existence of instances in each of the source ontologies, instances and that the source ontologies share a set of instances. Third, since the tasks that the mapping tools support differ considerably, the interaction between a user and a tool is very different from one tool to another [11].

4. Experimental Research Design

We will use controlled experiments to evaluate existing ontology mapping tools. Moreover we recognise that a controlled experiment provides the most effective way to evaluate the quality of ontology-mappings because:
It allows direct comparisons to be made between different ontology mapping results under controlled conditions through manipulation of experimental treatments.
It enables the ontology-mapping to be evaluated using objective and quantitative data.
It enables the ontology-mapping to be evaluated using independent participants.

4.1 Identifying ontology-mapping results for comparison

The experiments will only include those ontology-mappings that can be obtained by using ontology mapping tools which are freely accessible to researchers and practitioners alike. In addition due to the diversity of types of ontology mapping tools we need to establish certain groups of tools which are comparable. [11] follow the OntoWeb initiative who suggest a pragmatic classification of tools based on the users task requirements. These pragmatic selection criteria are: input requirements, level of user interaction, type of output and content of output. Initial tools we are considering for empirical evaluation are: Chimaera [8], Glue [1], ONION [10], Prompt (Protégé 2000 [12] plugin) and OntoMerge [2].

Conclusion

The overall research goal presented in this paper is to develop a methodology and a formal quality measures for evaluating ontologies and ontology mapping results. In this paper we have introduced the research project and described some of our preliminary ideas towards achieving the research goal. The aim is to empirically investigate the problems associated with ontology mapping and evaluate the existing solutions according to a solid theoretical framework.

Acknowledgments

This work is partially supported by the Norwegian Research Foundation in the framework of Information and Communication Technology (IKT-2010) program – the ADIS project.

References

1. Doan, A., Madhavan, J., Domingos, P., and Halevy A.: Learning to map between ontologies on the semantic web. In The Eleventh International WWW Conference, Hawaii, (2002).
2. Dou, D, McDermott, D and Qi, P.: Ontology Translation on the Semantic Web. In Proc. Int'l Conf. on Ontologies, Databases and Applications of SEmantics (ODBASE2003). LNCS 2888, pp. 952-969, (2003).
3. Goguen, J.: An Introduction to Algebraic Semiotics, with Application to User Interface Design. In: Computation for Metaphor, Analogy and Agents, edited by Nehaniv, C. Springer Lecture Notes in Artificial Intelligence, volume 1562, pages 242-291 (1999).
4. Guarino, N. and C. Welty, C.: Ontological analysis of taxonomic relationships. In A. Laender and V. Storey, editors, ER-2000: The 19th International Conference on Conceptual Modeling. Springer-Verlag, 2000.
5. Kaplan, A.: Towards a Consistent Logical Framework for Ontological Analysis. In Proceedings of the International Conference on Formal Ontology in Information Systems, Ogunquit, Maine, October, (2001).
6. Krogstie, J., Lindland O.I and Sindre, G.: Towards a deeper understanding of Quality in Requirements Engineering, Proceedings of 7th CaiSE, Jyvaskyla, Finland. June (1995).
7. Lindland, O.I., Sindre, G and Sølvberg, A.: Understanding Quality in Conceptual Modelling, IEEE Software, March: 42-49 (1994).
8. McGuinness, D., Fikes, R., Rice, J. and Wilder, S.: An environment for merging and testing large ontologies. In A. G. Cohn, F. Giunchiglia, and B. Selman, editors, Principles of Knowledge Representation and Reasoning: Proceedings of the Seventh International Conference (KR2000). Morgan Kaufmann Publishers, San Francisco, CA, 2000.
9. Missikoff, M.: Harmonise: An Ontology-Based Approach for Semantic Interoperability. ERCIM News, No. 51, October (2002) (Online edition).
10. Mitra, P.: An Algebraic Framework for the Interoperation of Ontologies. PhD Dissertation, Stanford University, August (2004).
11. Noy, N. and Musen, M.: Evaluating Ontology Mapping Tools: Requirement and Experience, In the Proceedings of the Workshop on Evaluation of Ontology Tools at EKAW'02 (EON2002), Siguenza, Spain, October (2002).
12. Noy, N. and Musen, M.: PROMPT: Algorithm and tool for automated ontology merging and alignment. In Seventeenth National Conference on Artificial Intelligence (AAAI-2000), Austin, TX, (2000).
13. Protege. The Prot´eg´e project, http://protege.stanford.edu, 2002.
14. Pustejovsky, J.: Type Construction and the Logic of Constructs, In: The Syntax of Word Meaning. Bouillon, P. and Busa F. (eds.) Cambridge University Press (2001).