Towards visualizing the mapping incoherences in Bioportal

Catarina Martins [1], Ernesto Jimenez-Ruiz [2], Emanuel Santos [1] and Catia Pesquita [1]

[1]LaSIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa
[2]Department of Computer Science, University of Oxford

ABSTRACT

The integration of biomedical ontologies via their matching often results in logical errors, due to incorrect mappings or incompatible ontological models. To solve this issue, repair algorithms remove mappings, decreasing the size and possibly the coverage of the alignment. Understanding these errors is crucial to support an effective use of alignments, since different scenarios may favor coherence above completeness and vice-versa. This could be supported by visualization, however, the challenges in visualizing ontology alignments are further compounded by the need to provide an explanation to the logical conflict between the mappings.

We present a preliminary visualization tool that supports the identification of mapping incoherence, by displaying sets of mappings involved in logical conflicts between several BioPortal ontologies pairs, as well as the classes and axioms involved.

1 INTRODUCTION

Establishing meaningful correspondences between biomedical ontologies is crucial to effectively explore the knowledge they model in an articulated fashion. Creating these correspondences, or mappings, can be accomplished by ontology matching techniques (Euzenat *et al.*, 2007).

Bioportal (Whetzel *et al.*, 2011), a web portal that provides access to more than 400 biomedical ontologies, provides mappings between ontologies which are automatically generated or manually added by experts. However, the integration of the ontologies via these mappings can result in incoherences due either to erroneous mappings or incompatibilities between both ontologies (Meilicke and Stuckenschmidt, 2008). The example in Figure 1 illustrates this problem. Although individual mappings appear to be correct, their integration results in a logical conflict, since in NCI Thesaurus *Anatomic_Structure_System_or_Substance* is disjoint with *Gene_Product*, and *Fibrillar_Actin* cannot be a subclass of both. This is a result of the different domain models followed by Foundational Model of Anatomy (FMA) and National Cancer Institute Thesaurus (NCIT).

To address this issue, (Faria *et al.*, 2014) applied both AML (Faria *et al.*, 2013) and LogMap (Jiménez-Ruiz and Grau, 2011) to detect and repair the incoherences in 19 pairs of ontologies from Bioportal and their mappings, and discovered that 11 in 19 had logical errors involving in average 22% of the mappings. These algorithms aim at eliminating incoherences by removing or altering mappings, and although they can provide logically sound solutions, these may not always be correct, since the choice of which mappings to eliminate is based on a change minimization strategy.

In this context, a visualization tool to identify the incoherences caused by mappings between ontologies would support their identification and correction by expert users. Moreover, it would support the decision for or against using a repaired alignment, since coherence often sacrifices completeness, and depending on the
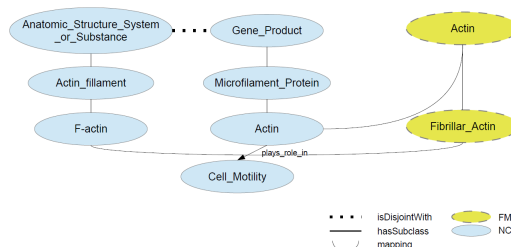

Fig. 1. Example of incoherent mappings between the FMA ontology and the NCIT
Taken with permission from Pesquita *et al.*, 2013

application, users may prefer one or the other (Pesquita *et al.*, 2013). Here, we present a preliminary version of a web tool[1] for visualizing sets of mappings involved in causing incoherences between several BioPortal ontologies pairs, as detected by AML using its repair algorithm (Santos *et al.*, 2013).

2 CHALLENGES IN VISUALIZING MAPPING INCOHERENCES

Biomedical ontologies present several visualization challenges due to their size, richness and complexity of vocabulary. There are two main paradigms to support the visualization of the ontologies and both have their drawbacks and benefits in the context of mapping visualization (Fu *et al.*, 2013). On the one hand, trees are appropriate when representing hierarchical relations but are confusing when representing multiple inheritance or several kinds of relations. Graphs, on the other hand, can handle both issues, but when the number of nodes is very high, visualization can be impaired. These aspects are relevant when visualizing alignments, and are further compounded by the representation of mappings between large ontologies (Pesquita *et al.*, 2014; Ivanova and Lambrix, 2014). On top of the challenges in visualizing ontology alignments, there are additional constraints when considering the visualization of mapping incoherence. The goal here is to give the user sufficient information to understand the reason behind the incoherence and allow him or her to evaluate the mappings and decided if a correction is needed, and how it should be performed.

To accomplish this we have identified the minimum set of information to show when displaying a set of conflicting mappings: (1) the classes involved in the mappings; (2) the mappings between classes; (3) the disjoint axiom involved in causing the incoherence; and (4) the relations between the mapped classes and the classes involved in the disjoint axiom.

---

[1] http://xldb.di.fc.ul.pt/biotools/vizrepair/

Ontology 1	Ontology 2	Total Mappings	Conflicting Mappings
BDO	NCIT	1636	1374
CCONT	NCIT	2097	1136
EFO	NCIT	2507	1541
EP	FMA	78489	109
EP	NCIT	2465	307
MA	FMA	961	22
OMIM	NCIT	5178	1078
UBERON	FMA	1932	121

Table 1. Total and conflicting mappings in the ontologies used. Bone Dysplasia Ontology (BDO), Cell Culture Ontology (CCONT), Experimental Factor Ontology (EFO), Cardiac Electrophysiology Ontology (EP), Foundational Model of Anatomy (FMA), Mouse Adult Gross Anatomy Ontology (MA), National Cancer Institute Thesaurus (NCIT), Online Mendelian Inheritance in Man (OMIM), Uber Anatomy Ontology (UBERON).

Our preliminary webtool represents this complexity by using graph-based visualization techniques. However, given the drawbacks of graphs when displaying a considerable number of nodes, our webtool focuses on displaying sets of conflicting mappings (i.e., the mappings that taken together cause an incoherence), rather than the whole alignment at once.

3 WEBTOOL

The backend of our tool is supported by a database that stores all the ontology and alignment information. This corresponds to the ontologies and alignments shown in Table 1, as well as a list of the conflict sets. Having data stored in a relational database allows for faster retrieval of the information to draw the graph.

The user interface allows users to select an alignment, and then browse the conflict sets in a table format. Each conflict set can be selected for graph visualization (supported by sigma.js). To allow users to understand the conflict we need to show the relations between the classes involved in mappings and the classes involved in the disjoint axiom(s), however in many cases this would result in showing the several classes that compose the path from the mapping to the disjoint axiom. To reduce this visual clutter, we compute the transitive closure between these classes and display them as directly linked (see Fig. 2).
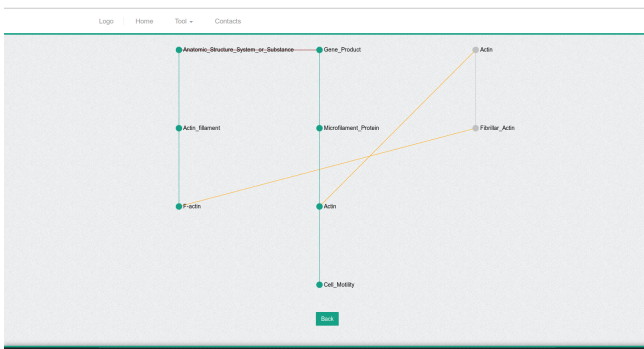

Fig. 2. Our webtool displaying the conflict set described in Figure 1. Green: NCIT; Gray: FMA; Red: disjoint axiom; Yellow: mappings.

4 CONCLUSION

Given the error rates of current repair algorithms (Pesquita *et al.*, 2013), supporting user involvement in alignment repair is key to providing coherent and correct alignments. Understanding the impact of incoherence in ontology alignments is crucial to support an effective use of alignments, which depending on the task at hand, should enforce or relax coherence. This can be of particular importance for biomedical ontologies where it is fairly common that ontologies covering the same domain are based on incompatible models. So for instance, if the task is to support the cross-references between ontologies, then coherence can be sacrificed to achieve a greater coverage, whereas in more complex tasks that depend on reasoning, such as querying support (Solimando *et al.*, 2014), ensuring coherence is paramount.

Our webtool supports users in this task by allowing them to visualize all classes, mappings and axioms involved in a logical conflict. We are currently extending the tool to permit users to manually solve conflicts, and export the repaired alignments. In future work we would like to link our tool to BioPortal, to support access to all ontologies and mappings it provides.

REFERENCES

Euzenat, J., Shvaiko, P., *et al.* (2007). *Ontology matching*, volume 18. Springer.

Faria, D., Pesquita, C., Santos, E., Palmonari, M., Cruz, I. F., and Couto, F. M. (2013). The Agreementmakerlight ontology matching system. In *On the Move to Meaningful Internet Systems: OTM 2013 Conferences*, pages 527–541. Springer.

Faria, D., Jiménez-Ruiz, E., Pesquita, C., Santos, E., and Couto, F. M. (2014). Towards annotating potential incoherences in BioPortal mappings. In *The Semantic Web–ISWC 2014*, pages 17–32. Springer.

Fu, B., Noy, N. F., and Storey, M.-A. (2013). Indented tree or graph? A usability study of ontology visualization techniques in the context of class mapping evaluation. In *The Semantic Web–ISWC 2013*, pages 117–134. Springer.

Ivanova, V. and Lambrix, P. (2014). User Involvement for Large-Scale Ontology Alignment. In *International Workshop on Visualizations and User Interfaces for Knowledge Engineering and Linked Data Analytics*, pages 34–47.

Jiménez-Ruiz, E. and Grau, B. C. (2011). Logmap: Logic-based and scalable ontology matching. In *The Semantic Web–ISWC 2011*, pages 273–288. Springer.

Meilicke, C. and Stuckenschmidt, H. (2008). Incoherence as a basis for measuring the quality of ontology mappings. In *In Proceedings of the 3rd ISWC international workshop on Ontology Matching*, pages 1–12.

Pesquita, C., Faria, D., Santos, E., and Couto, F. M. (2013). To repair or not to repair: reconciling correctness and coherence in ontology reference alignments. In *Ontolpgy Matching workshop at ISWC 2013*, pages 13–24.

Pesquita, C., Faria, D., Santos, E., Neefs, J.-M., and Couto, F. M. (2014). Towards visualizing the alignment of large biomedical ontologies. In *Data Integration in the Life Sciences*, pages 104–111. Springer.

Santos, E., Faria, D., Pesquita, C., and Couto, F. (2013). Ontology alignment repair through modularization and confidence-based heuristics. *arXiv preprint arXiv:1307.5322*.

Solimando, A., Jiménez-Ruiz, E., and Pinkel, C. (2014). Evaluating ontology alignment systems in query answering tasks. In *ISWC 2014 Posters and Demonstrations Track*.

Whetzel, P. L., Noy, N. F., Shah, N. H., Alexander, P. R., Nyulas, C., Tudorache, T., and Musen, M. A. (2011). BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. *Nucleic acids research*, **39**(suppl 2), W541–W545.