Demo: A community based approach for managing ontology alignments

Conference or Workshop Item

A Community Based Approach for Managing Ontology Alignments

Gianluca Correndo, Yannis Kalfoglou, Paul Smart, Harith Alani

University of Southampton,
Electronic and Computer Science Department
[gc3, y.kalfoglou, ps02v, h.alani]@ecs.soton.ac.uk,
WWW home page: http://ecs.soton.ac.uk
SO17 1BJ, United Kingdom

Abstract. The Semantic Web is rapidly becoming a defacto distributed repository for semantically represented data, thus leveraging on the added on value of the network effect. Various ontology mapping techniques and tools have been devised to facilitate the bridging and integration of distributed data repositories. Nevertheless, ontology mapping can benefit from human supervision to increase accuracy of results. The spread of Web 2.0 approaches demonstrate the possibility of using collaborative techniques for reaching consensus. While a number of prototypes for collaborative ontology construction are being developed, collaborative ontology mapping is not yet well investigated. In this paper, we describe a prototype that combines off-the-shelf ontology mapping tools with social software techniques to enable users to collaborate on mapping ontologies. Emphasis is put on the reuse of user generated mappings to improve the accuracy of automatically generated ones.

1 Introduction

The transformation of the Web from a mere collection of documents to a queryable Knowledge Base (KB) is one of the most prominent targets of Semantic Web (SW). To help reach this goal, knowledge repositories need to publish semantic representations of their data models to enable other machines to understand and query their content. To this end, much research and development has focused on building tools and capabilities for ontology and KB construction. However, support for distributed teams to remotely and continuously collaborate on building and updating ontologies and knowledge repositories is still underdeveloped. In this paper we describe an approach and present a prototype for facilitating ontology mapping by supporting social collaboration and reuse of mapping results for supporting data integration task. More specifically, our approach allows to: **align** local ontologies to shared ones; **exploit** social interaction and collaboration for improve alignment quality; **reuse** user ontology alignments for improving future automated alignments.

2 Collaboration for Knowledge Sharing

The need to make explicit, agree and publish data semantics is becoming increasingly central since more information systems are becoming largely decoupled and separately managed. To this end, the vision of the SW is moving towards a scenario where the task of creating and maintaining ontologies, that formalise data semantics, is going to be handed to the community that actually uses them [1]. Such vision requires that latent models shared by the community must emerge and tools and methodologies must be provided for fulfilling this task.

The rise of Web 2.0 has transformed the classical community of passive Web users into a community of active contributors. Leveraging upon this new perspective of web communities, several proposals have lately emerged to exploit users' contributions for supporting various knowledge tasks[2].

Collaborative Protègè [3] was recently developed as an extension to Protègè to support users to edit ontologies collaboratively, by providing them with services for proposing and tracking changes, casting votes, and discussing issues, thus infusing classical ontology editing with a number of popular social interaction features.

Other Web 2.0 inspired approaches rely on *lighter* ontologies, where the emphasis is put on sharing knowledge rather than creating an ontology. Some of these approaches use social tagging as the main driver for enacting collaborative lightweight ontology building [4]. Similarly, other tools are focussing on editing and sharing instance data, like OntoWiki [5] and DBin [6].

Most of the tools listed above focus on supporting users to collaboratively construct ontologies or to collaboratively populate an ontology with instance data. Unlike these tools, our proposed system, OntoMediate, extends the collaborative notion to support the task of *ontology mapping*, where users can collaborate and interact to map their existing ontologies and reuse mapping structures. A similar approach is the Zhadanova and Shvaiko [7] method. Focus of that work was on building such profiles to personalise reuse of ontology mappings. In OntoMediate though, we are exploring the use of collaborative features (discussions, voting, change proposals) to facilitate the curation, reuse and discussion of mappings by the community, and hence paving the way to integrate distributed knowledge bases.

3 OntoMediate System Description

In the OntoMediate [1] project we are studying how social interactions, collaboration and user feedback can be used in a community, in order to ease the alignment of ontologies and to share mapping results[2]. The implemented prototype is a Web application developed with J2EE and AJAX technologies. The ontologies are expected to be written in OWL and Jena API[3] is used for parsing documents. The system has been designed to be extended via its APIs and is composed of three main subsystems: ontologies and datasets manager (section 3.1); ontology alignment environment (section 3.2); social interaction environment (section 3.3).

[1] http://www.ecs.soton.ac.uk/research/projects/ontomediate

[2] This work was partially funded by a grant awarded to General Dynamics UK Ltd. and the University of Southampton as part of the Data and Information Fusion Defence Technology Centre (DIF DTC) initiative.

[3] http://jena.sourceforge.net

3.1 Ontologies and Datasets Manager

This part of the system allows users to register (as well as unregister) the datasets they intend to share with the community and the ontologies that describe their data vocabulary. An ontology browser allows then to inspect usual information about managed ontologies (i.e. hierarchy of concepts, labels, annotations, descriptions, properties and constraints). The ontologies that are loaded onto the system need to be aligned with one or more shared ontologies in order to enable querying of the published data by the community.

3.2 Ontology Alignment Environment

The full automation of ontology alignment is not an easy task [8]. The factors that affect the computation and accuracy of ontology alignments are so delicate that we can not afford not to take into account user input. Our system provides an API for automated ontology alignment tools to be plugged in and also maintains data structures to store parameters needed by a particular tool to execute (e.g. threshold values or available tool options). The API allows an easy integration of new alignment tools by means of wrappers (already integrated tools are: CROSI mapping system [9], INRIA Align [10] and Falcon OA [11]). These tools support the alignment task by proposing to the user some initial candidate mappings. The results from different tools can be merged using a weighted mean of each contribution and the decision of which combination of tools to use can be parameterised together with the configuration used to invoke each tool.

Once the automated mapping has been executed, the results are displayed in a dedicated interface for review and for searching further alignments. The interface has two view modalities: *hierarchical* and *detailed*. In the *hierarchical* view the two taxonomies are faced and mapped concepts are highlighted. The user can browse both taxonomies and create new mappings by dragging a source concept and dropping it into a destination concept. In the *detailed* view the description of two focused concepts are faced and the user can inspect the descriptions and map the properties using the same drag & drop facility used for mapping the concepts. The users can alternatively **accept** or **reject** some automatically proposed mappings. This choice will be recorded by the system and will be used to filter future mappings towards this target concept, thus increase future ontology alignment *precision*.

3.3 Social Interaction Environment

This functionality allows users of a community that deal with similar data - and therefore have a mutual interest to maintain good quality alignments - to socially interact with each other. The aim of the social interaction is to exploit community feedback in order to enhance the overall quality of the ontology alignment and achieve agreement on semantics of concepts by means of community acceptance. This subsystem proposes to the user three views: **Ontology View**; **User View** and **Forum View**. The **Ontology View** (see Figure 1 top-left corner) displays an enhanced taxonomy browser for the selected shared ontology. The enhancements concern the user activities affecting the shared concepts, visualising additional information (e.g. concepts that have some incoming mappings are

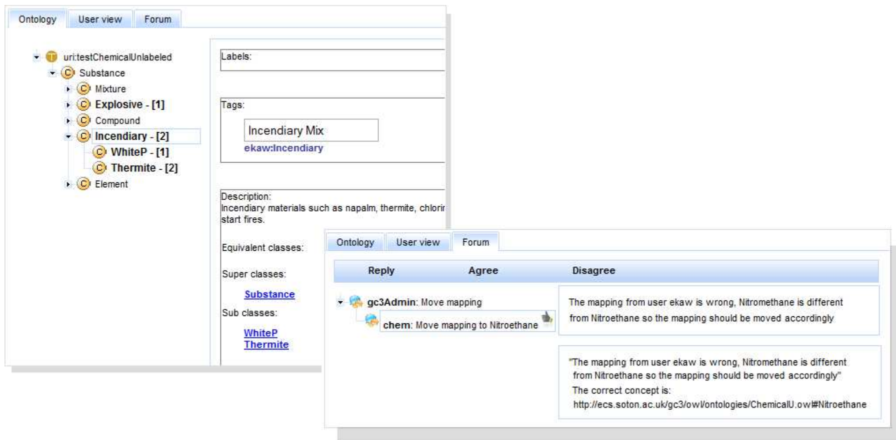

Fig. 1. Discussion environment - Ontology View

highlighted and the number of mappings is reported in brackets). Moreover, the interface allows to inspect the set of labels used for equivalent concepts (i.e. the ones provided with the alignments) in local ontologies (see the *Tags* text field in Figure 1). The user or administrator can edit such labels and add them to the shared concept to enrich the concept description with users' contributions. The new mapping, and the edited/added labels, will be logged in a database to be reused later to improve the *recall* of future ontology alignment tasks. When the user selects a concept that has some user mappings associated with it, he/she can switch to the **User View** that displays information about the local mappings for the focused concept. The user can then inspect a summarised description (i.e. subconcepts, superconcepts, properties etc.) of the local concepts and decide if they are relevant to the target concept or initiate a discussion thread in the forum (see Figure 1 bottom-right corner) in order to change them. Interacting within the forum users can debate the proposal, **reply** with a new one or simply **agree** or **disagree** with it. Relevant events are notified interested users (e.g. all the users that provided a mapping towards this target concept and all the others who explicitly asked to be informed).

3.4 Ontology Mapping Reuse

In OntoMediate system, one of the aims is to reuse user inputs in order to increase the quality of data integration and ease the ontological alignment task. Our approach for fulfilling this goals is twofold and involves user alignment results as an important input for increasing system performances and sharing achieved alignments.

The adoption of lexical information from local concept and properties for enriching target entities' description is just an example of how local contributions can help in building up a community tailored ontology. Such approach have shown, based on preliminary tests, to increase the performances of automated tools for successive alignment tasks.

Moreover, the system provides an additional functionality that allows to seamlessly share the agreed ontology alignments by means of POAF (Portable Ontology Aligned Fragments) fragments [12]. POAF uses existing alignments and OWL taxonomic reasoning to identify fragments can be reused as minimal information bundles for building an integrated data network.

4 Summary and Future Work

This paper presented a prototype for supporting ontology mapping with community interactions, where users can collaborate on aligning their ontologies. Our initial experiment showed an increase in both precision and recall in ontology mapping when reusing past mapping results. Next, we plan to run much larger experiments to further test the validity of the social approach, and the usability of the services and features provided. We will also implement services to allow users to submit and manage more complex mapping relationships.

References

1. Shadbolt, N., Berners-Lee, T., Hall, W.: The semantic web revisited. Intelligent Systems, IEEE **21**(3) (2006) 96–101
2. Correndo, G., Alani, H.: Survey of tools for collaborative knowledge construction and sharing. In: Workshop on Collective Intelligence on Semantic Web (CISW 2007). (November 2007)
3. Tudorache, T., Noy, N.: Collaborative Protégé. In: Workshop on Social and Collaborative Construction of Structured Knowledge (CKC 2007) at WWW 2007, Banff, Canada (2007)
4. Zacharias, V., Braun, S.: SOBOLEO – social bookmarking and lightweight engineering of ontologies. In: Proc. WWW 2007 Workshop on Social and Collaborative Construction of Structured Knowledge, Banff, Canada (May 2007)
5. Auer, S., Dietzold, S., Lehmann, J., Riechert, T.: OntoWiki: A tool for social, semantic collaboration. In: Workshop on Social and Collaborative Construction of Structured Knowledge (CKC) at WWW 2007, Banff, Canada (2007)
6. Tummarello, G., Morbidoni, C., Nucci, M.: Enabling semantic web communities with DBin: An overview. In: Proc. 5^{th} Int. Semantic Web Conf., ISWC 2006, Athens, GA, USA. (2006)
7. Zhdanova, A.V., Shvaiko, P.: Community-driven ontology matching. In: ESWC. (2006) 34–49
8. Kalfoglou, Y., Schorlemmer, M., Uschold, M., Sheth, A., Staab, S.: Semantic interoperability and integration. Seminar 04391 - executive summary, Schloss Dagstuhl - International Conference and Research Centre (September 2004)
9. Kalfoglou, Y., Hu, B., Reynolds, D., Shadbolt, N.: Capturing, representing and operationalising semantic integration (CROSI) project - final report (October 2005)
10. Euzenat, J.: An api for ontology alignment. In: Proc. 3^{rd} Int. Semantic Web Conf. (ISWC), Hiroshima ,Japan (2004)
11. Jian, N., Hu, W., Cheng, G., Qu, Y.: Falcon-AO: Aligning ontologies with falcon. In: Workshop on Integrating Ontologies (K-CAP 2005). (2005) 85–91
12. Kalfoglou, Y., Smart, P., Braines, D., Shadbolt, N.: POAF: Portable ontology aligned fragments. In: Proc. of the ESWC'08 Int. Workshop on Ontologies: Reasoning and Modularity (WORM'08), Tenerife, Spain. (2008)