Using Concept Maps and Ontology Alignment for Learning Assessment (September 2012)

Alex Alves da Silva, Natalia Padilha, Sean Siqueira, Fernanda Baião and Kate Revoredo

Abstract—Concept Maps are used as a tool of the constructivist theory of Meaningful Learning. Just like domain ontologies, concept maps enable clarification of concepts and relationships of a domain. This article presents a learning assessment model supported by the use of concept maps and ontology alignment techniques, respectively for clarification of the organization of concepts in the student’s cognitive structure and for the automatic comparison of these maps with a reference ontology. An example illustrating this application is presented and discussed.

Index Terms—concept maps, learning assessment, ontologies, ontology matching

I. INTRODUCTION

ASSESSMENT in the educational context is the process of determining what a student knows. Concept maps, which are a tool of cognitive constructivist theory of Meaningful Learning, provide a means to capture student knowledge and are particularly effective for representing the organization that students see among the concepts. In this sense, they represent a valuable assessment tool [1].

Using concept maps as part of an assessment approach requires considering two issues: (1) how the maps will be constructed and (2) how they will be interpreted [1]. The approach described in this paper proposes the use of ontology alignment techniques for addressing the second issue. To enable this semi-automatic interpretation the maps will be constructed in a tool that automatically encodes the concept maps into ontology files supported by the ontology alignment systems.

Domain ontologies are artifacts that enable the modeling of specific knowledge about a domain of discourse, through the representation of the entities and their interrelationships. There is a growing use of ontologies in diverse areas of research, including in the educational field [2]. Ontologies are knowledge repositories that can support a variety of educational applications [3]. Specifically, the use of ontologies as a knowledge repository for use in the learning assessment has been explored in the literature [3].

The heterogeneity of ontologies covering an area of common knowledge motivated the development of several techniques for aligning ontologies, as well as tools to automate this task [4]. The process of aligning two ontologies includes the identification of corresponding elements (concepts, relations and instances) in these ontologies.

Finding elements with the same intended meaning is not a task of comparing their labels and identifying the identical ones. This simple method ignores, for example, the existence of synonymous terms and even minor differences in the label spelling. The scientific research in the area of ontology alignment resulted in several techniques of aligning concepts, considering the syntactic and structural characteristics of an element and even external resources, like lexical databases.

Considering the existence of a significant set of tools like Protégé Suite Prompt [5], COMA++ [6], FOAM [7], SemMatcher [8] and MEMO [9], which automates the alignment process, one of the motivations of this work is to minimize the overhead in the learning assessment with concept maps.

This work is organized into seven sections. The following section introduces the definition of concept maps and ontologies. Section III presents the process and techniques of ontology alignment. The proposed approach using concept maps and alignment of ontologies is described in section IV, supplemented by an example of use in Section V. The related works are presented in section VI. Finally, in Section VII, we present the final considerations.

II. CONCEPT MAPS AND DOMAIN ONTOLOGIES

Concept maps are a tool of cognitive constructivist theory of Meaningful Learning, proposed by David Paul Ausubel [10]. A concept map consists of concepts and propositions. Concepts are representations of objects, events, situations, or properties designated by a label. Propositions express relations between
these concepts, through the association between two or more concepts labeled by a connecting line.

Ontologies are an explicit specification of a conceptualization [11], [12]. When considered from the perspective of an engineering artifact and within a knowledge domain, ontologies consist of a formal structure of concepts and relations between domain concepts, and a set of axioms restricting the interpretation of this structure and allows the derivation of new knowledge from factual knowledge represented in the structure [13].

In [14] concept maps and ontologies are compared in a specific context. Table I reproduces some of the results from that work.

<table>
<thead>
<tr>
<th>TABLE I</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>COMPARISON BETWEEN CONCEPT MAPS AND ONTOLOGIES ADAPTED FROM [6]</strong></td>
</tr>
<tr>
<td>Concept Maps</td>
</tr>
<tr>
<td>Serves to clarify the concepts</td>
</tr>
<tr>
<td>Requires the use of formality</td>
</tr>
<tr>
<td>Free elaboration</td>
</tr>
<tr>
<td>Software support</td>
</tr>
</tbody>
</table>

According to them, concept maps and ontologies serve to clarify the concepts, but in different ways. While the process of building a concept map is free, the elaboration of ontology requires the use of formality. The ontology representation language currently recommended by the World Wide Web Consortium (W3C⁴) for the Semantic Web is the OWL Web Ontology Language, which enables the processing of content by computer applications.

Concept maps have traditionally been used as cognitive tools in the educational process while ontologies are often used to capture and share knowledge [3]. In our approach, the ontologies will be generated from concept maps, preventing the increased complexity of building a formal ontology in the process of representation of an educational content. After that, the ontologies will be aligned to a reference ontology.

### III. ONTOLOGY ALIGNMENT

The Semantic Web community faces a heterogeneity problem. Many individual ontologies representing the same contents have emerged recently. They were developed independently to meet specific purposes and communities [15].

The process of alignment of two or more ontologies means that for each entity (concept, relation or instance), we try to find a corresponding entity, which has the same intended meaning, in other ontology. For some entities no corresponding entity might exist [15]. Ontologies are supplied to the ontology alignment system and alignments are returned, as depicted in Figure 1.

Fig. 1. Ontology Alignment. From [21]

Several techniques have been used in the process of ontology alignment. According to [16], these techniques can be classified according to the granularity of the analysis in element-level and structure-level. Element-level techniques exploit the entities, not considering the relationships between these structures. Structure-level techniques, in turn, evaluate the composition of these entities in the structure. From a second perspective, these techniques are classified according to the resource used for the analysis of correspondence (syntactic, semantic or external). Syntactic techniques interpret the input according to its structure, with the application of a clearly defined algorithm. The semantic techniques are based on a formal semantics, such as model theory. Some techniques, in turn, rely on the use of an external resource such as thesauri or other ontologies.

Fig. 2. Element-level alignment techniques. Adapted from [16].

In Fig. 2 the element-level techniques are illustrated. **String-based** techniques compare the labels of the entities. **Language-based** ones consider names as words in a natural language such as Portuguese, Spanish or English. Some techniques are based on Natural Language Processing [17], exploiting morphological properties of words that are provided as input. **Constraint-based** techniques use algorithms that deal with internal restrictions applied to the definitions of entities, such as data type and cardinality of attributes. **Linguistic Resources** consider linguistic features such as domain-specific thesauri (Wordnet² - a lexical database for the English and UMLS³ - Unified Medical Language System are examples) which contains synonyms, generalizations and specializations. **Alignment reuse** techniques explore external resources as the results of previous alignment between ontologies. For example, if we have stored results of the alignment between p and p', and between p and p'' we can reuse these results in the alignment of

Fig. 3. Structure-level alignment techniques. Adapted from [16].

---

⁴ http://www.w3.org/
² http://wordnet.princeton.edu/
³ http://www.nlm.nih.gov/research/umls/
p’ and p”. Upper ontologies can be used as external sources of common knowledge and the domain specific ontologies can also be used as external resources, for using only terms with relevant meanings to the field and are not related to similar concepts in other domains [16], [18], [4], [19].

In Fig. 3 the structure-level techniques are shown. Data analysis and statistics take advantage of (preferably large) samples of a population in order to find regularities and discrepancies. This helps to group items or compute distances between them. Correspondence analysis and frequency distributions are examples of these types of techniques. Graph-based algorithms consider ontologies as labeled graphs. The analysis of similarity between a pair of nodes representing the concepts in two ontologies is based on their position within the graph, considering that if two nodes in two ontologies are similar, their adjacent nodes, should also present a certain degree of similarity. Taxonomy-based also consist of graph-based algorithms, which consider only the relation of specialization. Repositories of structures: ontology repositories store structures and their fragments, and similarity measures for pairs of entities. Model-based algorithms (or semantically based) manipulate entries based on its semantic interpretation. The idea behind this type of technique is that if two entities are similar, they share the same logical interpretation. Technical description logic inferences are examples of model-based techniques [16], [18], [4], [19].

These techniques are usually combined in the tools of ontology alignment. Fig. 4 illustrates an alignment example of two concepts: Object and Thing, Car and Automobile. The alignment of these terms is possible with the use of external linguistic resources and structure-level techniques, considering that the strings of the terms substantially differ from each other.

IV. LEARNING ASSESSMENT SUPPORTED BY ONTOLOGY MATCHING

The approach described in this paper involves the application of ontology alignment techniques for the semi-automatic comparison between the concept maps developed by the students and a reference one.

The process begins with a tutor (a teacher or a system), which presents the content to be learned by the student. In the assessment phase the student generates a concept map on the subject learned. See Table II and Fig. 5.
TABLE II
COMPARISON BETWEEN CONCEPT MAPS AND ONTOLOGIES

<table>
<thead>
<tr>
<th>Role</th>
<th>Responsibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tutor</td>
<td>Construct the reference concept map and interpret the results of alignment</td>
</tr>
<tr>
<td>Student</td>
<td>Construct concept map</td>
</tr>
<tr>
<td>COE</td>
<td>Convert concept map to an ontology in OWL format</td>
</tr>
<tr>
<td>Tool for ontology alignment</td>
<td>Align the concept maps generated</td>
</tr>
</tbody>
</table>

Since most of the ontology alignment tools adopt the standard representation language OWL, it is necessary that both the students and the reference concept maps are encoded in OWL. COE is a project whose goal is to develop an integrated suite of software tools for constructing, sharing and viewing OWL encoded ontologies based on CmapTools, a concept mapping software used in educational settings, training, and knowledge capturing. Concept maps provide a human-centered interface to display the structure, content, and scope of an ontology.

Then, the alignment result of an automated tool will be the basis for the evaluation of the student.

V. EXAMPLE OF USE

In this section we will discuss some aspects of the proposed approach using a practical example. The presented concept maps were built by students of a discipline of Knowledge Discovery in Databases (KDD).

The concept map taken as reference is reproduced in Fig 6. The reference concept map may be developed by the tutor or be an existing domain ontology of the subject. It is not expected the concept maps to be identical to the reference one. The reference concept map is not an unquestionable standard but if it is well designed and comprehensive in the field, it can be used as a good reference for comparison. The automatic alignment will facilitate this comparison indicating the corresponding entities identified by the tool, which are a good beginning for the tutor analysis.

The tool used to process the alignments was COMA++\(^6\), a schema and ontology alignment tool that extends previous prototype COMA utilizing a composite approach to combine different match algorithms [20]. This tool has a good interface for presenting the results, which facilitates the analyses of the found alignments.

Fig. 7 illustrates the concept map of one of the students and Fig. 8 shows the resulted alignment. The lines in this figure connect the entities aligned and the color indicates the strength of the alignments, the green ones for the best similarity values and the orange ones for lower coefficients. Fig. 9 shows the similarity values for each line in the Fig. 8.

\(^4\) http://www.ihmc.us/groups/coe/
\(^5\) http://cmap.ihmc.us/
\(^6\) http://dbs.uni-leipzig.de/Research/coma.html
Fig. 7. Concept map of student 1

Fig. 8. Student 1 x Reference

Fig. 9. Coefficients of the alignments in Fig. 8

Fig. 10. Summary of the alignment
As shown in the summary of the alignment in Fig. 10, the tool identified 12 correspondences, 5 of them with similarity value 1.0. This is the case when the entities have the same string, as in “association rules” and “clustering”.

However, the main motivation for the use of alignment techniques is the identification of entities with the same meaning, even when the strings are not identical. In the example above and in the following ones we have some alignments demonstrating this case. The alignments of “Preprocessing” and “pre-processing” and “Naive Bayes” and “bayesian networks” are examples in the Fig 8.

In the alignment shown in Fig. 12, the “business rules” concept in the student’s map (Fig. 11) and the “business rule mining” concept in the reference map were aligned by the tool. In the alignment of Fig. 14 we have a similar example with “Process” and “process mining” and “Text” and “text mining”. that the words “business rules”, “text” and “process” are connected to the word “mining” by a relation but they mean the
same idea in the reference concept map.

The example above shows a relation alignment between “can be” and “may be”.

In the following example (Fig. 15 and Fig. 16), the term “Knowledge Discovery Database” is aligned to the term “knowledge discovery from data”.

By evaluating all alignments we can observe that some concepts such as “association rules” appear in all concept maps. The comparison of the frequency of the terms in the concept maps configures a feedback for the tutor and points to the most and less learned concepts.

The motivation of the work presented in [21] is based on the assertion that the use of concept maps as a tool to support the learning process implies an increased workload for the teacher. It discussed the use of concept maps in assessment of learning and focused on the importance of an automated system to support the teacher in this task.

Caldas & Favero [22] proposed a tool for automatic evaluation of concept maps in ODL environments, which through Artificial Intelligence techniques (n-grams algorithm and KNN) performs an assessment both quantitatively (by a score given to the student concept map) and qualitative (through a "report-guide" provided to the student after the development of his concept map).

VI. FINAL CONSIDERATIONS

This paper proposes a learning assessment approach that combines the use of concept maps as a tool for explanation of the student’s knowledge about a subject, and alignment of ontologies for the automatic comparison of these maps with one of reference. This proposal differs from others identified in the literature because of the focus given to the alignment of ontologies as a support for automating the evaluation of concept maps. Ontology alignment techniques are subject of current research, with contributions from development tools from events like the Ontology Alignment Evaluation Initiative [14]. The automatic evaluation of concept maps supported by the use of these techniques can be positively influenced by developments in this area. The enrichment of the tools has enabled satisfactory results, with the identification of correspondences that are based not only on comparison of the elements of the string, but also on thesaurus such as WordNet, that enables the alignment of synonymous terms. Moreover, the proposed approach was able to consider the semantics embedded in the structure of ontologies.

The use of automated tools for evaluation of concept maps produced for students does not eliminate the role of the teacher, but speeded up the process, which is a fundamental requirement due to the great volume of assessments, common in distance education, for example.

The future work must evaluate how different techniques and algorithms for ontology alignment would be more appropriate in this context of application, bringing better results in terms of maximizing correct alignments found and minimizing wrong matches.

The proposed approach will also be evaluated in a formal case study.

ACKNOWLEDGMENT

We thank Sabine Massmann for giving us a COMMA ++ demo to accomplish the experiments and Masters students Alysson Gonçalves, Eduardo Fritzen, Marcos Afonso and Wladimir Pereira for their contributions.

REFERENCES

Alex Alves da Silva is Master Student in Information Systems at the Federal University of Rio de Janeiro (UNIRIO). Graduated in Computer Technology (2002) from the University Center of the City of Rio de Janeiro. Graduated in Computer Science (2009) from the University Center of the City of Rio de Janeiro and is a member of the Brazilian Computer Society.


Sean Siqueira is a Professor of the Department of Applied Informatics of the Federal University of the State of Rio de Janeiro (UNIRIO) since 2004. He received the Doctor of Science degree from the Pontifical Catholic University of Rio Janeiro (PUC-Rio) in 2005. He also worked as a professor at the Federal University of Goiás (UFG) in 2006. His current research interests include Semantic Web, ontologies, "technology-enhanced learning", "learning objects", data integration, information retrieval, knowledge management, data, text and web mining, social networking and web 2.0. He participates in research projects in those areas, with funding from several Brazilian government agencies, including CNPq, CAPES and FAPERJ. He participates in several program committees of national and international conferences and workshops, and is a member of ACM, IEEE and of the Brazilian Computer Society.

Fernanda Baião is a Professor of the Department of Applied Informatics of the Federal University of the State of Rio de Janeiro (UNIRIO) since 2004, where she is one of the coordinators of the NP2Tec research group. She received the Doctor of Science degree from the Federal University of Rio de Janeiro (COPPE/UFRJ) in 2001. During the year 2000 she worked as a visiting student at the University of Wisconsin, Madison (USA). Her current research interests include conceptual modelling, well-founded representation languages, information architecture, data management in scientific workflows. She participates in research projects in those areas, with funding from several Brazilian government agencies, including CNPq, CAPES and FAPERJ. She participates in several program committees of national and international conferences and workshops, and is a member of ACM and of the Brazilian Computer Society.

Kate Revoredo is doctor of Science from the Federal University of Rio de Janeiro (COPPE/UFRJ). During the year 2006 she worked as a visiting student at the University of Freiburg, Freiburg (Germany). She is a Professor of the Department of Applied Informatics of the Federal University of the State of Rio de Janeiro (UNIRIO) since 2009. Her current research interests include Data Mining, Relational Learning (inductive logic programming - ILP), Statistical Relational Learning (SRL), Theory Revision, Predicate invention, Social Network Analysis and Semantic Web. She participates in several program committees of national and international conferences and workshops, and is a member of the Brazilian Computer Society.