STUDIA UNIV. BABEŞ–BOLYAI, INFORMATICA, Volume **LV**, Number 2, 2010

REASONING INTROSPECTION AND VISUALISATION FRAMEWORK FOR ONTOLOGY MAPPING ON THE SEMANTIC WEB

MIKLOS NAGY, MARIA VARGAS-VERA

Abstract. Cognitive support for ontology mapping systems become more and more important because the size and complexity of the results increases due to the availability of more and more ontologies on the Semantic Web. This support is especially required for reasoning and result introspection since the results need to be presented in a way that users can easily understand them. User understanding is crucial because the end users are the only one who can actually judge if a certain reasoning process is flawed or not. As such the quality and usability of the system is directly dependent on these kind of supports. In this paper we present a representation framework that can be used by different systems in order to store and visualise the reasoning behind ontology mapping.

1. Introduction

To date the quality of the ontology mapping was considered to be an important factor for systems that need to produce mappings between different ontologies. However, evaluation of ontology mapping systems has demonstrated that even if systems use a wide variety techniques, it is difficult to push the mapping quality beyond certain limits. It has also been recognised [1] that in order to gain better user acceptance, systems need to introduce cognitive support for the users i.e. reduce the difficulty of understanding the presented mappings. Further in order to improve the quality of the mapping systems these intermediary details need to be exposed to the users who can actually judge if the certain reasoning process is flawed or not. This important feedback or the ability to introspect can then be exploited by the system designers or

Received by the editors: December 6, 2009.

2010 *Mathematics Subject Classification.* 68T30, 68U35.

1998 *CR Categories and Descriptors.* I.2.4 [**Artificial Intelligence**]: Knowledge Representation Formalisms and Methods – *Representation languages*;

Key words and phrases. Artificial intelligence, Semantic Web, Ontology mapping.

This paper has been presented at the International Conference Interdisciplinarity in Engineering (INTER-ENG 2009), Târgu-Mureş, Romania, November 12–13, 2009.

ultimately the system itself through improving the reasoning processes, which is carried out behind the scenes in order to produce the end results. This ability to introspect the internal reasoning steps is a fundamental component of how human beings reason, learn and adapt. However, many existing ontology mapping systems that use different forms of reasoning exclude the possibility of introspection because their design does not allow a representation of their own reasoning procedures as data. Using a model of reasoning based on observable effect it is possible to test the ability of any given data structure to represent reasoning. Through such a model we present a minimal data structure necessary to record a computable reasoning process and define the operations that can be performed on this representation to facilitate computer reasoning. This model facilitates the introduction and development of basic operations, which perform reasoning tasks using data recorded in this format. It is necessary that we define a formal description of the structures and operations to facilitate reasoning on the application of stored reasoning procedures. By the help of such framework provable assertions about the nature and the limits of numerical reasoning can be made.

Our main objective is to establish a standard framework for ontology mapping systems that allows mapping systems to detect reasoning failures and to refine the function of reasoning mechanisms in order to improve system performance and avoid future reasoning problems. To achieve this goal, the users have the possibility to introspectively monitor the end result of the reasoning process and determine the possible causes of its failures and possibly perform actions that affect future reasoning processes. This is important aspect of the need to visualise the mapping stems from the fact that even though ontology mapping tools converge towards automatic mapping processes, users always play an important role for validating these results. Therefore, it is important that mappings are presented in a way that can easily be understood by end users and not just by specialised domain experts.

The main contribution of this paper is a reasoning introspection and visual mapping representation framework for ontology mapping that goes beyond the end results and includes the intermediary reasoning steps, which can be exploited by both end users, system designers or ontology engineers.

The paper is organised as follows. In section 2 we describe the human reasoning process, the modelled reasoning workflow and the reasoning steps that is supported by our introspection framework. In section 3 we present the visualisation components for both mapping and reasoning and discuss the level of cognitive support that can be achieved by 3D modelling. In section 4 we present our experiments with the OAEI benchmarks and in section 5 we summarise the advantages and current limitations of our framework. Section 6 presents the related work and in section 7 we draw our conclusions.

2. Representation model

2.1. **Human reasoning process.** When humans create mappings between two ontologies they rely heavily on their past experiences or existing knowledge about the domain. Experts can follow different processes due to their personal preferences, however we have considered a generalised mapping process in our scenario. First they select an initial set of terms from both ontologies that they believe can correspond to each other. At this step the candidate mappings are selected from the whole results. Candidate mappings are not more than hypothesises for the mapping that needs to be proved correct. Naturally the selected hypothesis is associated with a great deal of uncertainty, which stems from the lack of information about the context of these terms. Ideally more experts are involved in this process at the same time from probably different domains. Each expert with its own knowledge and experience selects candidate mappings from both ontologies. Once the candidate mappings are selected based on evidences that support their hypothesis they need to combine their subjective opinions into a more coherent view. This procedure ideally results in a consensus where the best mappings are selected. The process can be summed up in 5 steps:

(1) Select candidate mappings.
(2) Build hypotheses for possible mappings.
(3) Find evidence for proving that our hypothesis is true.
(4) Eliminate the terms that do not support our initial belief.
(5) Combine different beliefs into a more coherent ones i.e. reach consensus over the selected mappings.

The before mentioned process is perfectly modelled by existing evidential reasoning approaches e.g. the Dempster-Shafer theory (DS theory) of evidence. This model includes all levels of sub attributes, with the possibility of different frame of discernment. Further it is possible to derive the expected utility values directly from the combined experts' belief distributions.

When human experts create mappings across different domains they usually base the end result on some sort of consensus between different experts. Each expert examines a subset of the terms from both ontologies and using their background knowledge and experience they gradually eliminate terms from the subset till they reach the final result. Each expert goes through the same reasoning process and finally they discuss the results explaining why they have selected a particular mapping. Our proposed reasoning inspection framework intends to support ontology mapping systems that model the above-mentioned human reasoning process.

2.2. **Modelled reasoning workflow for ontology mapping.** Ontology mapping systems carry out several iterative steps (Fig. 1) to select the final mappings from a number of candidate mappings. During this process background knowledge is consulted in order to extend the original variables with ones that can possible describe the concepts that need to be matched. The overall process can be described as follows:

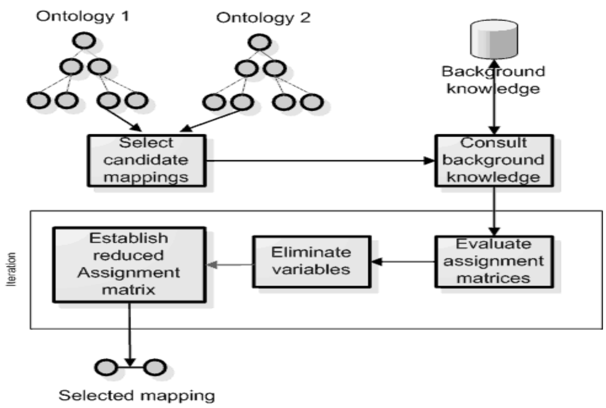

FIGURE 1. Reasoning process

(1) Select mapping candidates: In this step the system takes candidate mappings from both ontologies. The main objective of this step is to create an initial set of concepts that need to be compared to each other. Different systems can use various methods to select these candidates e.g. string similarities can be used to pre-filter concepts or a certain number of concepts are selected from both source and target ontologies.
(2) Consult background knowledge: The main objective of this step is to determine the possible meaning of the selected terms from the ontologies. Different background knowledge can be used e.g. WordNet or the Semantic Web itself. In this step the system selects a pre-defined number of additional terms that will be added to the candidate mapping sets. The system creates the preliminary assignment matrix.
(3) Evaluate assignment matrices: Using a wide variety of methods e.g. string similarity, graph structure or probabilistic information the system evaluates the initial similarities between the source and target ontologies. Different systems can use a single method or different methods, which need to be combined into a more coherent view.

(4) Eliminate variables: In this step the system selects the terms, which either above a pre-defined threshold in terms of similarity or most likely to be a match based on probabilistic information.
(5) Establish reduced assignment matrix: The system reduces the number of variables and if no clear matching is found it consults the background knowledge and starts the assignments again from step 3.

The above mentioned process is a general one. Different systems can implement differently each step however the main characteristics of these steps remain the same across different systems. Therefore in our representation framework we foresee to support any system that implements the before mentioned process.

3. Visualisation of the mapping

Visualising ontology mappings involves situations where the number of items that are concurrently displayed on the screen increases, which in turn worsens the graphical perception of the scene and complicates spotting details. If the amount of visualization space needed to represent all the mappings within the result set outnumbers the space available on the screen, a few options remain available: to scale down the whole image to the detriment of readability, to present on the screen just a portion of it and allow its navigation or to summarize the information in a condensed graph and provide means for exploration and expansion. As the effectiveness of these options depends on the task the users need to carry out, a combined usage of them offers a suitable approach. However combining these approaches using 3D space can considerably enhance[2] the productivity of the users.

Therefore in order to visualise the mappings we propose two different panes (each ontology has different pane) in 3D where the selected concepts are visible but their connections initially are not. The idea is to hide initially the network of connections as this potentially distracts the user from the details. Once the user selects a concept the system can reveal the connections to the selected concepts from the second ontology. In case the user wants to proceed for the reasoning the system reveals the reasoning states. During 3D visualization, the mappings are graphically encoded and displayed together with the different term's relationships, such as the belief in their similarities and differences. The strengths of the similarities can be represented differently using positions and several retinal variables: colour and orientation. For example once the user selects a term from the first ontology the mapped terms can be displayed where the most probable correspondences are displayed opposite to the selected term and the least probably terms are farther from the centre. Because of the data can be projected onto topological spaces with 3 orthogonal axis one can transpose these relations into a perspective that spontaneously highlights the magnitude of the semantic relations involved between the terms. As a result the position of a term on the side of each representation can directly give its amplitude of similarity in comparison to the other ones in the group.

Recreating and visualising the behaviour of complex reasoning procedures can be accomplished via two tasks. Primarily the amount of data available to a reasoning system, which records its own actions for further use is overwhelmingly large. Thus, some process must select data that needs to be considered for the visualisation process. The second portion of this process of visualising reasoning is the formation of a mapping between the chosen data and the recent history of events. These processes are complex, however, through the development of a formal data model we hope to facilitate the development of algorithms, which accomplish these tasks as well as establishing limitations on their operations. Further the task of managing and accessing large information spaces like reasoning is a problem of large scale cognition. This is because it is hard to visualise the large number of terms and how these terms are related during the reasoning process. As we have discussed in the previous sections the gradually decreasing size of the reasoning space can be perceived as hierarchical (pyramid like structure in 3D) spaces. Our basic idea is to visualise the reasoning based on conetree [4, 5]. A cone tree is a 3D representation of a tree structure, i.e. a standard $G = (V, E)$ graph with vertices and edges. In a cone tree representation, the root of a tree (represented by a cube, a sphere or some other appropriate object) is located at the tip of a transparent cone. The children of the root node are arranged around the base of the cone. Each child can be the root node of a subtree, which is represented in a recursive fashion by a cone whose tip is located at the object representing the child. Cone tree visualisation can be improved, particularly for very large datasets, by techniques such as usage-based filtering, animated zooming, hand-coupled rotation, coalescing of distant nodes, texturing, effective use of colour for depth cueing. For representing the reasoning space for term comparisons like ontology mapping the top of the hierarchy represents the result mapping pair of terms and situated in the apex of the cone with its children placed evenly spaced along its base. The next layer of nodes is drawn below the first, with their children in cones. Cone base level diameters are reduced at each level, which ensures that the shape will form a pyramid like structure(Fig. 2). This representation allows the user to easily navigate between the layers as directional movements ensure that nothing blocks the view of cones behind the user's point of view.

Our reasoning representation model is converted into a 3D model by our visualisation framework. This model can be viewed by a virtual reality viewer

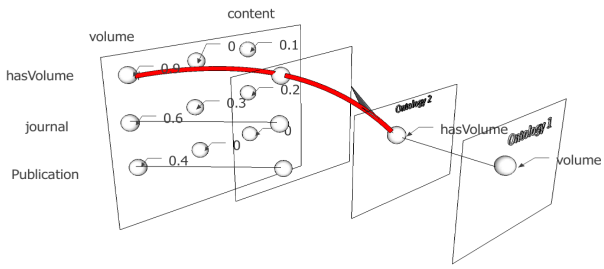

FIGURE 2. Reasoning with two steps

such as jReality [1] or different 3D modellers like Google SketchUp [2]. Therefore the user has the possibility to zoom in, out or move around the mapping model in order to discover the mappings and the reasoning model.

3.1. **Pluggable framework overview.** One of our main objective when designing the introspection and visualisation framework was that it should easily be integrated into other ontology mapping systems or 3D modellers. The conceptual overview of the systems is depicted on Fig. 3.

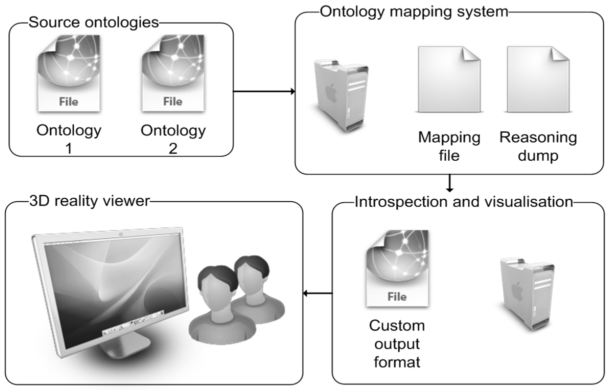

FIGURE 3. Visualisation framework overview

[1]http://www3.math.tu-berlin.de/jreality/
[2]http://sketchup.google.com/

Ontology mapping systems work with source ontologies and produce

(1) Intermediary dump files that contain the reasoning steps for the retained mappings.
(2) The result mapping file in e.g. in OAEI format.

Without visualisation this is the only output that the systems produce to the users. This approach works fine with small mapping files, however quickly become unmanageable for the users once the size of these files increase. Our introspection and visualisation framework takes these files as its input and convert it to custom 3D modells based on the representation described in section 2 . These model files can be opened with different 3D modelling tools by the user, which provides the functionality for zooming and moving around the mappings. An example interface using Google SketchUp is depicted on Fig 4.

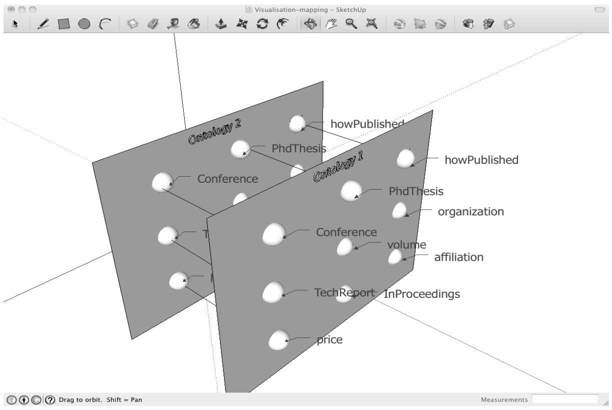

FIGURE 4. Example Google SketchUp interface

4. EXPERIMENTS WITH THE OAEI BENCHMARKS

We have carried out experiments with the benchmark ontologies of the Ontology Alignment Evaluation Initiative(OAEI) [3], which is an international initiative that has been set up for evaluating ontology matching algorithms. The experiments were carried out in order to determine the average level of layers, terms that need to be stored and visualised to the user. This is important as the more reasoning steps we need to visualise and store the more difficult is to manage in terms of computational complexity. Our main objective was to evaluate the correlation between concepts, properties in the

[3] http://oaei.ontologymatching.org/

ontology and the number of steps(iteration) necessary to select a candidate mapping to the result set. The OAEI benchmark contains tests, which were systematically generated starting from some reference ontology and discarding a number of information in order to evaluate how the algorithm behave when this information is lacking. The bibliographic reference ontology (different classifications of publications) contained 33 named classes, 24 object properties, 40 data properties. Our results are depicted in Table 1.

Table 1. Number of iterations to be presented

Reasoning steps	1	2	3	4	5	6	7	8
Classes	1972	58	79	68	82	113	91	689
Properties	2722	22	46	53	55	35	30	1532

We have measured the number of cases that belong to a particular iteration number for the whole dataset. It is worth to note that the used ontology mapping system [7] limits the number of iterations per mapping selection to a maximum of 8. Therefore table 1 contains all the possible number of iterations i.e. steps that need to be presented to the user from 1 to 8. Experiments have shown that the majority of cases(1972 for classes, 2722 for properties) belong to the single step reasoning i.e. the system selects the best mapping from the first assigned similarity assignment matrix. On the second place the maximum number of iterations 8 needs to be displayed. The rest of the cases, which constitutes only a small proportion of the total cases are divided between 2-7 reasoning steps. The initial results are encouraging as it demonstrates that the reasoning and mapping visualisation needs to present a one step reasoning for the majority of the cases. However these results also show that there is a need to investigate how to give the possibility to the user to navigate through complex reasoning steps.

5. Advantages and limitations

Our proposed framework has several advantages. First of all our framework has been designed as a pluggable component that can be fitted to both ontology matching systems and industry standard 3D visualisation framework. This is an important aspect as the usability of our visualisation framework does not depend on certain tools. Therefore the users can use a wide variety of mapping tools and virtual reality viewers. Secondly the structure, concepts and relations between concepts of mapping or reasoning can be presented to the user in a graphical form to facilitate a better understanding of the results. This 3D view can then be represented in a space where shapes, sizes and locations are governed by the sizes, overlaps and other properties of the different shapes, giving the user an intuitive feel of where the bulk of the information in the space exists. Additionally, one can manipulate the mapped view to show only those parts that appear to be the interest of the user. This view should provide a clearer picture of the relations between the resulting concepts, properties and instances.

The existing limitation of our proposed framework can be grouped into two distinct categories i.e. limitations on the storage of reasoning steps and the visualisation components. Concerning how the reasoning steps are stored our proposed framework fits well to most ontology mapping systems that use internal similarity matrixes for producing mappings. However not all mapping systems can be integrated directly with our framework. In these cases conversion needs to be done first that might not always be feasible. Concerning the visualisation components there is considerable room for improvements. Currently using different 3D modelling frameworks users can zoom and move around the 3D graph structures but the framework does not allow the interaction with the 3D structure itself i.e. users cannot change the presented structure. This can pose difficulties if the users wants to investigate a particular part of the reasoning space. Our primary research goal for the future is to find appropriate visualisation framework that could accommodate this need. Further currently it is not possible to use different filters or sorting for the mapping result set therefore out initial prototype is more suitable for mappings where the result mapping set is not too large. Nevertheless it is our future objective to investigate how different filtering and sorting possibilities can improve the visualisation of the mappings.

6. Related work

Several ontology editing tools [8, 9] provide visual interface for visualising ontologies including various plug-ins that aim to support different aspects of the ontology management lifecycle including, creation, checking and visualisation. Different tools and plugins usually implement different approached for editing or visualising these ontologies. However the most common visualisation model is 2D graph.

It has been acknowledged that interfaces will play an important role on the Semantic Web in order to gain better user acceptance. However most of the initial visualisation work [10] was carried out on representing ontologies for knowledge engineers. These techniques include tree [8] , network [11] or probably the most commonly used graph [9] representations. Most techniques display information in 2D nevertheless 3D representations [12] become popular as large scale ontologies can be presented more comprehensively for the users. Popular 3D ontology representation technique is to project the ontology network on the sphere [12] where the most relevant element appear bigger than the irrelevant ones on the peripheries of the sphere. Other 3D technique for information visualisation is called cone or cam trees [3] where hierarchical information is arranged in circles on different levels. Levels in the tree corresponds to visual depth. This technique was proposed well before the semantic web has been conceived. Based on the ontology visualisation techniques further research has been carried out on representing ontology mappings as well. CogZ [1] is a tree based technique, which proposes various cognitive support for the decision making processes(interaction, analysis, representation) used in the mapping task. Its main objective is to reduce the cognitive load experienced by users. Other solution employs [2] a three dimensional tree based visualisation that allows for the selection of multiple class, exclusion of classes, and saving the merged classes to an OWL ontology.

Several techniques have been proposed to visualise Semantic Web Data using both 2D and 3D. Nevertheless the dominant approaches are based on graph visualisation[13, 14]. 2D solutions in particular involve certain limitations when visualizing complex networks therefore many researchers have studied different graph visualization in 3D.

7. Conclusion

The process of developing algorithms to support numerical and introspective reasoning and visualisation for ontology mapping systems requires a great deal of understanding of these domains. By specifying standard data types and approaches to these methods, it is our hope that the development of these algorithms can be further investigated. Further introspective reasoning requires a domain independent approach to reasoning technique. The model presented here demonstrates one method of achieving this domain independence in a way, which is designed to allow the development of future algorithms. Based on our proposed framework we hope that future introspective mapping systems can be developed, which contributes to a better acceptance of these systems. Our initial visualisation framework can be improved in several ways. Firstly the visual interface does not give the possibility to the users to filter the result set. These aspects of our work need to be investigated further. Secondly the users cannot change the layout of the presented mappings. In order to achieve higher level of interaction our future objective is to investigate how the possibility of changing the layout can contribute to the objective to provide a better an easier understanding of the mappings. From the contribution point of view our framework proposes a standard way of storing the states of the system during the reasoning process. This is an important aspect of the mapping, which was not investigated so for in the context of ontology mapping. Our 3D visualisation framework build that presents the reasoning steps and the mapping itself can help the end users to navigate easily between the mappings without being overwhelmed by the complexities inherent to any 2D graph representation model.

References

[1] S.M. Falconer, M.A.D. Storey, *A Cognitive Support Framework for Ontology Mapping*, Proceedings of 6th International Semantic Web Conference (ISWC2007), 2007, pp. 114–127.

[2] D. Brokenshire, P. Lougheed, *3D Visualization – Ontology Visualization for Mapping*, Technical Report, School of Interactive Arts and Technology, Simon Fraser University, 2008.

[3] G.G. Robertson, J. D Mackinlay, S. K. Card, *Cone trees: Animated Cone trees: 3D visualizations of hierarchical information*, Proceedings of ACM SIGCHI '91 Conference on Human Factors in Computing Systems, 1991, pp. 189–194.

[4] G.G. Robertson, J. D Mackinlay, S. K. Card, *Cone Trees: animated 3D visualizations of hierarchical information*, CHI 91: Proceedings of the SIGCHI conference on Human factors in computing systems, 1991, pp. 189–194.

[5] A. P. Calitz, D. Munro, *Representation of hierarchical structures in 3D space*, AFRIGRAPH 01: Proceedings of the 1st international conference on Computer graphics, virtual reality and visualisation, 2001, pp. 59–64.

[6] B. Shneiderman, *Tree visualization with tree-maps: 2-d space-filling approach*, ACM Trans. Graph., 1/11 (1992), pp. 92–99.

[7] M. Nagy, M. Vargas-Vera, P. Stolarski, *DSSim results for OAEI 2008*, The 3rd International Workshop on Ontology Matching, 2008.

[8] Stanford Center Biomedical Informatics Research, *Protege : http://protege.stanford.edu/*, 2009.

[9] Michael Sintek, *OntoViz : http://protegewiki.stanford.edu/index.php/OntoViz*, 2007.

[10] A. Katifori, C. Halatsis, G. Lepouras, C. Vassilakis, E. Giannopoulou, *Ontology visualization methods-a survey*, ACM Comput. Surv., 4 (2007).

[11] P. W. Eklund, N. Roberts, S. P. Green, *OntoRama: Browsing an RDF Ontology using a Hyperbolic-like Browser*, The First International Symposium on CyberWorlds (CW2002), 2002, pp. 405–411.

[12] A. Bosca, D. Bonino, P. Pellegrino, *OntoSphere: more than a 3D ontology visualization tool*, CEUR Workshop Proceedings, 166 (2005).

[13] P. Mutton, J. Golbeck, *Visualization of Semantic Metadata and Ontologies*, IV '03: Proceedings of the Seventh International Conference on Information Visualization, 2003.

[14] I. Herman, G. Melanon, M. S. Marshall, *Graph Visualization and Navigation in Information Visualization: a Survey*, IEEE Transactions on Visualization and Computer Graphics, 6 (2000), pp. 24–43.

The Open University, United Kingdom
E-mail address: M.Nagy@open.ac.uk, M.Vargas-Vera@open.ac.uk