AlignmentVis: Visual Analytics for Ontology Matching

Jillian Aurisano*
EVL, University of Illinois at Chicago

Amruta Nanavaty†
ADVIS Lab, University of Illinois at Chicago

Isabel F. Cruz‡
ADVIS Lab, University of Illinois at Chicago

Abstract

We present AlignmentVis, a visualization tool that supports users in the process of interactive ontology matching, which is a central data integration component. AlignmentVis enables tasks such as the exploration of the resulting mappings, the assessment of the performance of the matching algorithms, and the identification of sets of ontology entities that share similar matching characteristics with the purpose of facilitating the diagnosis of matching errors and of algorithm optimization. We offer a user-centric application design to enable these tasks through multiple coordinated views.

1 Introduction

Ontologies specify a conceptualization of a domain in terms of entities and their attributes, which is mainly hierarchical [5]. A central component of data integration consists of mapping semantically related entities between a *source ontology* and a *target ontology*. The resulting set of mappings is called an *alignment*. A variety of automatic matching algorithms or *matchers* use various criteria including lexical, syntactic, and structural. The final alignment is obtained by combining the results obtained from the various algorithms.

In spite of the success of automatic algorithms, the need to support user feedback as a component of the matching process has been recognized [12] and is part of a more general trend in information integration systems [1]. We explore the use of visualization to understand and evaluate the results from different matchers, as implemented in AlignmentVis, a prototype visual application. AlignmentVis uses coordinated visualizations to provide complementary views of the ontology matching process. For quick prototyping, we have implemented AlignmentVis using Processing.

AlignmentVis works with the AgreementMaker ontology matching system [3]. AgreementMaker has ranked first in the most important tracks of the Ontology Evaluation Alignment Initiative (OAEI) [8, 9] and is recognized as one of the top ontology matching systems [12]. It is also one of the most widely used ontology matching systems, distributed to more than one hundred users. It implements a large variety of matchers and is particularly effective in the combination of their results [4].

While the field of ontology matching is booming, the use of visualization has been sparse even if there are several worthwhile contributions that include PROMPT+COGZ [10], AgreementMaker [6], and iMerge [7]. The richer in number of views is PROMPT+COGZ, while iMerge uses different complementary views for the understanding of a matcher. However, neither compares the results provided by the different matchers like AgreementMaker does, a capability also supported by AlignmentVis. In addition, AlignmentVis supports a rich palette of views like PROMPT+COGZ and tightly links multiple views like iMerge.

*e-mail: jillian.aurisano@gmail.com
†e-mail: aybgr8@gmail.com
‡e-mail: isabelcfcruz@gmail.com

2 AlignmentVis

2.1 Task Description

AlignmentVis supports user tasks, according to a typology [2] gathered through discussions with two users who are ontology matching experts. They wish to *evaluate* the performance both of individual algorithms as well as their combination. Moreover, they would like to *diagnose* and characterize the error types in the output of a matcher, such as missed mappings or false mappings. Once the errors are *identified*, users need to diagnose their potential causes. This diagnosis includes *locating* details about individual entities and *browsing* to see commonalities such as identical properties or performance characteristics shared by false mappings. In addition, the diagnosis may involve analysis of the performance of each matcher. For instance, one matcher may perform poorly because it produces a high percentage of false mappings. *Exploration* and *comparison* tasks allow for evaluation and diagnosis, particularly in a scenario where a *reference alignment* (gold standard) is unavailable. In this case, mappings with different confidence scores across different matchers may be targets for exploration. The rapid exploration of the ontologies and of their alignments together with the comparison of results across matchers is essential to refining the matching approach.

2.2 Visualization Prototype

We have implemented four coordinated views as follows.

Matcher Output–Grid View shows a matrix of the mapping confidence scores for the output of a matcher in the interval [0,1]. The source entities are listed vertically and the target entities are listed horizontally. Each cell represents one mapping color coded from black (0) to bright blue (1). Green indicates a correct mapping, which is present in both the current alignment and in the reference alignment. Red indicates a missed mapping, which is present in the reference alignment but not in the current alignment. Orange indicates a false mapping, which is present in the current alignment but not in the reference alignment. Users can *brush* a cell to see the labels of the source and target entities and the mapping confidence score. This view is coordinated with the other views, such that brushing in another view highlights entities in this view and vice versa. Several sorting algorithms (by entity name, mean and standard deviation of the confidence score) can be applied to the source and target entity lists to detect patterns. Missed, false, and correct mappings can be displayed in groups thus allowing for the identification of sets of entities with similar algorithm results.

Entity Mapping Characteristics–Scatter Plot View displays matching performance statistics associated with entities in the source and target ontologies. Options for the axes include *mean confidence score*, *standard deviation score*, and *mapping correctness* (missed, false, or correct). Points are colored according to whether they are from the source or target, thus allowing for the identification of different characteristics in the source and target ontologies. This view is coordinated with other views through brushing. Users can also switch between the *x* and *y* axes.

Ontology Hierarchy–Tree View allows for users to explore the hierarchical relationship between mapped entities in the source and target ontologies. On hovering over a section of the tree, the particular section under the mouse is expanded, and the mappings above a fixed threshold are displayed by a colored line between entities

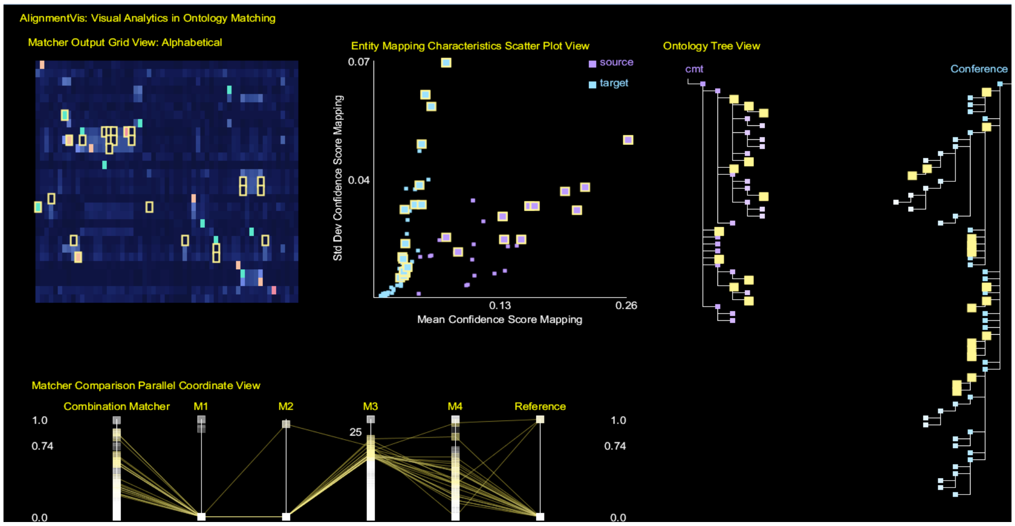

Figure 1: AlignmentVis user interface.

in the two trees. The color scheme is the same as in the grid view. Mappings are displayed on demand to avoid information overload. When users hover over a mapping, it is also displayed in other views through brushing.

Matcher Comparison–Parallel Coordinate View enables the comparative analysis of the performance of all the matchers involved (including the combination matcher) and the reference alignment. Each axis depicts a matcher, with mappings displayed as points that are positioned relative to their confidence score. Lines are drawn between corresponding mappings in different matchers. This view allows for an overview of the distribution of the mappings for each matcher, such as whether the confidence score of many mappings fall at either 0 or 1, as in the binary string matcher, or whether many fall in the middle. On hovering over an axis, the mappings in that area are highlighted and lines are shown. Users can therefore evaluate whether a high confidence score for a matcher corresponds to a high or low confidence score in the other matchers and how it compares with the reference alignment. By presenting the data in this view, users are able to rapidly identify differences in performance between the matchers. In addition, by linking this view to other views through brushing, users can observe whether high or low confidence scores for one matcher correspond to entities that have particular characteristics in the other views.

Figure 1 shows the user interface of AlignmentVis with three individual matcher views (in this case for the combination matcher) and the parallel coordinate view.

3 Conclusions and Future Work

An initial assessment based on the OAEI conference dataset indicates that the proposed interactive interface with coordinated views is useful in the evaluation, diagnosis, comparison, and exploration of the resulting alignments. With these capabilities, AlignmentVis goes beyond the state of the art of visualization approaches for ontology matching [11]. Future work will concentrate on additional techniques to sort and group elements in the existing views, as well as work to support evaluation of mappings for larger ontologies. Further evaluation of the benefits of this visual tool in the ontology matching process will also be conducted.

Acknowledgements

This research was partially supported by NSF Awards IIS-1143926, IIS-1213013, and CCF-1331800 .

References

[1] K. Belhajjame, N. W. Paton, A. A. A. Fernandes, C. Hedeler, and S. M. Embury. User Feedback as a First Class Citizen in Information Integration Systems. In *CIDR Conference on Innovative Data Systems Research*, pages 175–183, 2011.

[2] M. Brehmer and T. Munzner. A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.

[3] I. F. Cruz, F. Palandri Antonelli, and C. Stroe. AgreementMaker: Efficient Matching for Large Real-World Schemas and Ontologies. *PVLDB*, 2(2):1586–1589, 2009.

[4] I. F. Cruz, F. Palandri Antonelli, and C. Stroe. Efficient Selection of Mappings and Automatic Quality-driven Combination of Matching Methods. In *ISWC International Workshop on Ontology Matching (OM)*, volume 551 of *CEUR Workshop Proceedings*, pages 49–60, 2009.

[5] I. F. Cruz and A. Rajendran. Semantic Data Integration in Hierarchical Domains. *IEEE Intelligent Systems*, 18:66–73, 2003.

[6] I. F. Cruz, C. Stroe, and M. Palmonari. Interactive User Feedback in Ontology Matching Using Signature Vectors. In *IEEE International Conference on Data Engineering (ICDE)*, pages 1321–1324, 2012.

[7] Z. El Jerroudi and J. Ziegler. iMERGE: Interactive Ontology Merging (poster). In *International Conference on Knowledge Engineering and Knowledge Management Knowledge Patterns (EKAW)*, 2008.

[8] J. Euzenat, A. Ferrara, L. Hollink, A. Isaac, C. Joslyn, V. Malaisé, C. Meilicke, A. Nikolov, J. Pane, M. Sabou, F. Scharffe, P. Shvaiko, V. Spiliopoulos, H. Stuckenschmidt, O. Sváb-Zamazal, V. Svátek, C. T. dos Santos, G. A. Vouros, and S. Wang. Results of the Ontology Alignment Evaluation Initiative 2009. In *ISWC International Workshop on Ontology Matching (OM)*, volume 551 of *CEUR Workshop Proceedings*, pages 73–126, 2009.

[9] J. Euzenat, A. Ferrara, C. Meilicke, J. Pane, F. Scharffe, P. Shvaiko, H. Stuckenschmidt, O. Sváb-Zamazal, V. Svátek, and C. T. dos Santos. Results of the Ontology Alignment Evaluation Initiative 2010. In *ISWC International Workshop on Ontology Matching (OM)*, volume 689 of *CEUR Workshop Proceedings*, pages 85–117, 2010.

[10] S. M. Falconer and M.-A. D. Storey. A Cognitive Support Framework for Ontology Mapping. In *International Semantic Web Conference/Asian Semantic Web Conference (ISWC/ASWC)*, volume 4825 of *Lecture Notes in Computer Science*, pages 114–127. Springer, 2007.

[11] M. Granitzer, V. Sabol, K. W. Onn, D. Lukose, and K. Tochtermann. Ontology Alignment—A Survey with Focus on Visually Supported Semi-Automatic Techniques. *Future Internet*, 2(3):238–258, 2010.

[12] P. Shvaiko and J. Euzenat. Ontology Matching: State of the Art and Future Challenges. *IEEE TKDE*, 25(1):158–176, 2013.