Alignment between versions of the same ontology

Ahmed ZAHAF

Evolutionary Engineering and Distributed Information Systems (EEDIS)
University of Djillali Liabès , Sidi BEL-ABBES, Algeria
zhmida@gmail.com

Abstract. Ontological change constitutes a knowledge source on the similarity and difference between versions. The usual algorithms of ontology matching do not take this knowledge into account. An ontology is a logical theory consisting of a pair of signature and axioms. The persistent signature is the focus of alignment between versions. We reformulate the alignment problem between versions as a problem of choosing among the elements of the signature persistent those who can form an ontology signature isomorphism. To ensure alignment coherence, we introduce a constraint on alignment semantics which we call changed meaning conservation. This constraint allows extending the computed alignment with correspondences for the remaining elements in a coherent manner. Regardless of the change, our approach identifies the persistent signature and provides an initial alignment between the elements of the two persistent signatures. Then it calculates the difference between versions to form the ontological change. The constraint of changed meaning conservation helps us on one side to revise the initial alignment to form an ontology signature isomorphism and on the other side to adjust the eliminated correspondences in the revision step to form a coherent alignment. Finally, we discuss the prototype implementation of our approach.

Keywords: ontological change, versions difference, alignment between versions, ontology signature morphism, changed meaning conservation, alignment revision, alignment Coherence.

1. Introduction

The alignment between versions of ontology facilitates the evolution of ontology based semantic systems by allowing them to continue to interoperate with each other.

The ontological change constitutes a knowledge source on the similarity and difference between versions. Usual algorithms of ontology matching [1] do not take this knowledge into account.

An ontology is a logical theory consisting of a pair of signature and axioms [2]. The signature is the vocabulary used to name the ontological entities (concept, property and individual). The axioms express intentional interpretation of this vocabulary.

The ontological change affects both the signature and axioms. The signature change is the set of added or deleted signature elements. The remaining elements form the persistent signature. The axiomatic change is the set of added or deleted axioms. The remaining axioms form the persistent axioms.

The persistent signature is the focus of alignment between versions. There are two types of such signature: the elements whose intentional interpretation, as specified by the axioms is not affected by the axiomatic change and those whose intentional interpretation is affected. We express the alignment between the elements of the first type as ontology signature isomorphism between the versions. Then, the alignment problem consists in finding the maximal ontology signature isomorphism. We introduce a constraint on alignment semantics which we call changed meaning conservation. This constraint ensures that the intentional interpretation of an element of the signature in a version is maintained vis-à-vis the knowledge propagation through alignment semantics. We then say, an alignment between versions that conserves changed meaning is a coherent alignment. The problem of alignment between the elements of the second type is therefore to extend the maximal ontology signature isomorphism with correspondences while respecting the constraint of changed meaning preservation.

Our approach to solve the problem of alignment as posed above spread over four steps. Regardless of the change, our approach identifies the persistent signature and provides an initial alignment between the elements of the two persistent signatures. Then it calculates the difference between the versions to form the ontological change in the second step. In the third step, the initial proposed alignment must be revised if a violation of changed meaning conservation constraint is detected. To support revision, we introduce a relevance relation on the elements of the signature of the axiom. This relation compares the degrees of intentional persistence of these elements. The last step is user driven to extend the revised alignment.

Section 2 is the foundation of our work. We present the problem statement of the alignment in Section 3, and then we present our approach in Section 4. Section 5 is reserved to describe the platform of the prototype implementation of our approach. We compare our approach with related works in Section 6 and we summarize our work in Section 7.

2. Preliminaries and notations

The concept of Ontology can be seen as a logical theory [2]. So it is a pair $(\boldsymbol{S}, \boldsymbol{A})$, where $\boldsymbol{S}$ is the signature - describing the vocabulary - and $\boldsymbol{A}$ is a set of axioms - specifying the intended interpretation of the vocabulary in a domain of discourse. The signature is the set $\boldsymbol{S} = \boldsymbol{C} \sqcup \boldsymbol{P} \sqcup \boldsymbol{I}$. $\boldsymbol{C}$ represents the set of vocabulary to designate concepts. $\boldsymbol{P}$ is the set of vocabulary to designate properties and $\boldsymbol{I}$ is the set of vocabulary to designate individuals. We distinguish between the origins axioms $\boldsymbol{A}$ and their logical consequences $\boldsymbol{A}^*$ (also called closure). Theory $(\boldsymbol{S}, \boldsymbol{A})$ is called the presentation of $(\boldsymbol{S}, \boldsymbol{A}^*)$. In this work, we limit ourselves only to $\boldsymbol{S} = \boldsymbol{C} \sqcup \boldsymbol{P}$ and we designate by ontological entity a concept or a property.

The ontological change affects both the signature and axioms. The signature change is the set of added or deleted signature elements. The remaining elements form the persistent signature. The axiomatic change is the set of added or deleted axioms. The remaining axioms form the persistent axioms.

We use the following notation: $\boldsymbol{S}_i^{\boldsymbol{p}}$ is persistent signature of a version i. $\boldsymbol{S}^-$ is the removed signature. $\boldsymbol{S}^+$ is the added signature. Similarly, $\boldsymbol{A}_i^{\boldsymbol{p}}$ is the set of persistent axioms of a version i. $\boldsymbol{A}^-$ is the set of deleted axioms. $\boldsymbol{A}^+$ is the set of added axioms.
We express the ontological change as the difference between versions.

Definition 1 (difference between versions). Given two versions of an ontology $\mathbf{v_1} = (\boldsymbol{S_1}, \boldsymbol{A_1})$ and $\mathbf{v_2} = (\boldsymbol{S_2}, \boldsymbol{A_2})$, the ontological change is the difference given by the set:

$$\boldsymbol{diff} = \{(\boldsymbol{S}^-, \boldsymbol{A}^-), (\boldsymbol{S}^+, \boldsymbol{A}^+)\}.$$

Ontology alignment is the task to detect links between elements from two ontologies. These links are referred to as correspondences and express a semantic relation. According to Euzenat and Shvaiko [1] we define a correspondence as follows and introduce an alignment as set of correspondences.

Definition 2 (Correspondence and Alignment). given ontologies $\mathbf{o_1}$ and $\mathbf{o_2}$, let $\boldsymbol{Q}$ be a function that defines sets of matchable elements $\boldsymbol{Q}(\boldsymbol{o_1})$ and $\boldsymbol{Q}(\boldsymbol{o_1})$. A correspondence between $\mathbf{o_1}$ and $\mathbf{o_2}$ is a 4-tuple $(\boldsymbol{e}, \boldsymbol{e'}, \boldsymbol{r}, \boldsymbol{n})$ such that $\boldsymbol{e} \in \boldsymbol{Q}(\boldsymbol{o_1})$, $\boldsymbol{e'} \in \boldsymbol{Q}(\boldsymbol{o_2})$, $\boldsymbol{r}$ is a semantic relation, and $\boldsymbol{n} \in [\boldsymbol{0};\ \boldsymbol{1}]$ is a confidence value. An alignment $\boldsymbol{M}$ between $\mathbf{o_1}$ and $\mathbf{o_2}$ is a set of correspondences between $\mathbf{o_1}$ and $\mathbf{o_2}$. We restrict r to be one of the semantic relations from the set $\{\sqsubseteq, \sqsupseteq, \equiv, \perp\}$.

In order to reason about alignment, two classes of approaches have been introduced. The first class is based on model theory. IDDL [7] and DDL [6] are two examples of approaches of this class. Based on an axiomatic approach, the second class called reductionist semantics [8] is to interpret correspondences of the alignment as axioms in some merged ontology. In this paper, we use an example of this semantic called natural semantic. It involves building a merged ontology through the union of the two ontologies to align and axioms obtained by translating relations of the alignment. We introduce this semantic through its merged ontology.

Definition 3 (Merged Ontology). given an alignment $\boldsymbol{M}$ between two ontologies $\mathbf{o_1}$ and $\mathbf{o_2}$ and $\boldsymbol{trans}$: $\boldsymbol{M} \longrightarrow \boldsymbol{A}$, a function that transforms a correspondence to an axiom. The merged ontology is defined by

$$\mathrm{o}_1 \cup_M \mathrm{o}_2 = \mathrm{o}_1 \cup \mathrm{o}_2 \cup trans(M).$$

Since an ontology is a logical theory, an ontology signature morphism (an important notion used as a foundation to our work) is a theory morphism. Theory morphism is a signature morphism whitch preserve the axioms [18].

Definition 4 (ontology signature morphism). given tow ontologies $\mathbf{o_1}$= $(\boldsymbol{S_1}, \boldsymbol{A_1})$ and $\mathbf{o_2}$= $(\boldsymbol{S_2}, \boldsymbol{A_2})$, an ontology signature morphism is a function $\boldsymbol{f}: \boldsymbol{S_1} \longrightarrow \boldsymbol{S_2}$ such that $\boldsymbol{A_2} \vDash \boldsymbol{f}(\boldsymbol{A_1})$, i.e., all models of $\boldsymbol{A_2}$ are models of the image of $\boldsymbol{A_1}$ by $\boldsymbol{f}$. The image of an axiom is obtained by systematically replacing signature elements of this axiom by their correspondents, according to the signature morphism $\boldsymbol{f}$.
When $\boldsymbol{f}$ is bijective, we say $\boldsymbol{f}$ is an ontology signature isomorphism.

3. Problem Statement

The persistent signature is the focus of alignment between versions. The objective is to establish semantic relations between elements of the two persistent signatures. The persistent signature includes two types of elements: element whose meaning as specified by the axioms is not affected by the axiomatic change and those whose meaning is affected. Thus, alignment must establish relations between elements of the first type so that their meanings are completely preserved. It therefore defines an ontology signature isomorphism (See Definition 4) between elements of this type. We call such condition, meaning preservation and we define it formally as follows,

Definition 5 (meaning preservation). given two versions of an ontology $\mathbf{v_1}$= $(\boldsymbol{S_1}, \boldsymbol{A_1})$ and $\mathbf{v_2}$= $(\boldsymbol{S_2}, \boldsymbol{A_2})$, an alignment $\boldsymbol{M}$ between $\mathbf{v_1}$ and $\mathbf{v_2}$ preserve meaning if and only if it define an ontology signature isomorphism $\boldsymbol{M}: \boldsymbol{S_1} \longrightarrow \boldsymbol{S_2}$ such that:

$$\boldsymbol{A_2} \vDash M(\boldsymbol{A_1}) \text{ and } \boldsymbol{A_1} \vDash M^{-}(\boldsymbol{A_2});$$

We can establish a variety of ontology signature isomorphism between versions depending on the number of correspondences established. The goal is to find the maximal one ($\boldsymbol{M_{max}}$).

The alignment is known to propagate knowledge from one version to another. If this propagation is not controlled, it can affect the meaning of elements of the second type. The control of knowledge propagation amounts to establish correspondences between the signature elements such that the changed meaning in one version is preserved. We call such condition, changed meaning conservation and we define it formally as follows,

Definition 6 (changed meaning conservation). an alignment $\boldsymbol{M}$ between two versions $\mathbf{v_1}$= $(\boldsymbol{S_1}, \boldsymbol{A_1})$ and $\mathbf{v_2}$= $(\boldsymbol{S_2}, \boldsymbol{A_2})$ conserve the changed meaning if and only if $\boldsymbol{M}$ verifies the following two properties:

$$\forall\, \delta \in A^{-} \;,\; \mathrm{v}_1 \cup_M \mathrm{v}_2 \nvDash M(\delta);$$

$$\forall\, \delta \in A^{+} \;,\; \mathrm{v}_1 \cup_M \mathrm{v}_2 \nvDash M^{-}(\delta).$$

We then say, an alignment between versions that conserves the changed meaning is a coherent alignment and it is incoherent otherwise.
The first property ensures the coherence of alignment with regard to the propagation by its natural semantics (see Definition 3) of deleted axioms. The second property ensures the coherence of alignment with regard to the propagation of the added axioms.
The problem of alignment between the elements of the two persistent signatures which their intentional interpretation is altered is therefore to extend $\boldsymbol{M_{max}}$ with correspondences between them so that the alignment is coherent.

4. Alignment Method

The objective of our alignment method is to compute an alignment between the elements of persistent signatures of different versions of the same ontology. This alignment must satisfy meaning preservation condition for the signature elements whose meaning is not altered by the ontological change and the conservation of the changed meaning vis-à-vis the propagation of knowledge by the semantics of the alignment for the other elements. Our method satisfies meaning preservation condition by establishing equivalence relations between signature elements of the first type. We can establish a variety of alignments of this type depending on the number of correspondences established. Our method tends to generate the maximal one in three steps: version matching, version difference and

alignment revision. Version matching step identifies persistent signature based on the comparison between the terminology elements of both signatures. Assuming no changed meaning had occurred for persistent elements, our method generates an initial alignment by establishment of equivalence relations between the elements of the two persistent signatures. The persistent signature serves as a guide to determine the ontological change as the difference between versions in the second step. The revision step of the initial alignment eliminates just the correspondences that are responsible for the incoherence of this alignment. The alignment result is the desired maximal alignment.
The alignment method is completed by the extension step. In this step, the eliminated correspondences must be reviewed by the user to establish the appropriate relations while respecting changed meaning conservation condition. This step is semi automatic. We describe in what follows only the first three steps.

4.1 Version Matching

The objective of this step is the identification of the persistent signature in both versions and expresses the correspondences between the elements of the two persistent signatures with equivalence relations to form the initial alignment. The identification of the persistent signature is based on the existence of a terminological matcher. The terminological matcher can be based on the syntax of terms to be compared or a relationship of synonymy from a thesaurus in the field of ontology versions. Formally defined,

$$\boldsymbol{s_1} \in \boldsymbol{S_1^p} \text{ and } \boldsymbol{s_2} \in \boldsymbol{S_2^p} \text{ if and only if there exists a matcher } \boldsymbol{M} \text{ such that } \boldsymbol{s_2} = \boldsymbol{M}(\boldsymbol{s_1});$$

4.2 version difference

The objective of this step is to compute the ontological change in the form of semantic difference between versions. First, our method computes the set of persistent axioms then use this set to deduce the sets of deleted and added axioms. An axiom in a version is considered as persistent if the other version contains its image. The image of an axiom is obtained by systematically replacing signature elements of this axiom by their correspondents, according to a matcher $\boldsymbol{M}$. Formally defined,

$$\delta_1 \in \boldsymbol{A_1^p} \text{ and } \boldsymbol{\delta_2} \in \boldsymbol{A_2^p} \text{ if and only if there exists a matcher } \boldsymbol{M} \text{ such that } \delta_2 = \boldsymbol{M}(\delta_1);$$

The following rules express the semantic difference:

$$A^- = A_1 - A_1^p; \text{(deleted axioms)}$$

$$A^+ = A_2 - A_2^p; \text{(added axioms)}$$

However, there may be exceptions to this, especially when considered as added or removed axioms can still be deducted. Therefore, we must refine the difference as follows,

$$\delta \in A^- \text{ and } \boldsymbol{v_2} \vDash M(\delta) \text{ then } A^- = A^- - \{\delta\}; \text{ (refined deleted axioms)}$$

$$\delta \in A^+ \text{ and } \boldsymbol{v_1} \vDash M^-(\delta) \text{ then } A^+ = A^+ - \{\delta\}\text{ ; (refined added axioms)}$$

4.3 Alignment Revision

In general, initial alignment cannot be coherent. Because, some correspondences propagate axioms from one version to another that violate the constraint of changed meaning conservation. The objective of this step is to identify these correspondences and provide a means to choose among them which must be eliminated. The identification of these correspondences is simply obtained by identifying the signature of the axiom propagated. To choose among correspondences, we introduce an order relation which we call relevance relation on the signature elements of the propagated axiom. The relevance relation (noted $<_{rel}$) compares the degrees of intentional persistence of these elements. The intentional persistence of an element signature s denoted ($intPersistance(s)$) is expressed as the ratio of the number of occurrences of this element in the persistent axioms set (denoted $nboccurrence(s, A_i^p)$ for a version i) on the total number of persistent axioms. Formally defined,

$$s_1 <_{rel} s_2 \text{ if and only if } \mathrm{intPersistance}(s_1) < intPersistance(s_2) \text{ and}$$

$$\mathrm{intPersistance}(s) = \mathrm{nboccurrence}(s, A_i^p)/|A_i^p|.$$

The signature element that has the less intentional persistent with respect to the relevance relation allows to choose the correspondence to eliminate from the initial alignment. When two of the signature elements have the same degree of persistence intentional, the choice is left to the user.

5. Implementation

The implementation of our method is at the time of this writing in an advanced stage. The first three phases of our approach is fully implemented. It remains to design and implement the extension phase. This phase requires a user-friendly interface to help the user to handle correspondences of the alignment with a flexible manner. Currently, the platform of our prototype is for OWL ontologies. We hope to extend it to other ontology languages in the near future. The platform is based on OWL API [14] and Align API [15]. The platform integrates pellet [16] as the main reasoning engine on OWL ontologies.

6. Related works

The problem of ontology matching has known the emergence of several approaches in recent years [1]. The main distinction between them is due to the nature of the knowledge encoded in the ontology, and how it is used in the identification of correspondences [9]. Terminological methods compare the lexicon used to designate ontological entities, while the semantic methods are based on model theory to determine the existence of a correspondence between two entities. Some approaches consider the internal structure of the ontology. Other approaches consider the external structure of the ontology. The ontology extension can also be used. The majority of the existing matching systems combine these techniques to cover different aspects of the ontology. With the exception of a few systems, such ASMOV [10] and S-Match [11], the alignment result is subject to logical contradictions. Other approaches [12] and [13] propose an additional component to revise the alignment. The revision is intended to ensure alignment coherence. Alignment coherence requires satisfiability preservation of ontological entities by alignment. None of these approaches considers the ontological change as a source of information about the similarity and difference between the versions. Meaning preservation and changed meaning conservation by alignment ensures alignment coherence between versions. In the case of alignment between versions, these two conditions are more general than satisfiability preservation of ontological entities.

The comparison between versions has been the subject of several approaches ([3], [4], [5], [17]). The purpose of the comparison is to calculate the difference between versions. Each approach is influenced by the underling representation of the ontology. For example, PromptDiff [3] consider ontology as a graph. Ontoview [4], SemVersion [5] and [17] consider ontology as a set of RDF triples. None of these approaches match our vision of ontology as a logical theory.

7. Conclusion and future works

We presented the problem of alignment between versions as a problem of establishing an ontology signature isomorphism between the persistent signature elements of versions. We introduced changed meaning conservation constraint both in building this isomorphism and its extension in a coherent manner. We also proposed an approach based on the concepts of meaning preservation and changed meaning conservation to build a coherent alignment between versions of the same ontology. We discussed the platform of the prototype of our approach and we hope to automate the extension step of our method and to evaluate the prototype in the near future.

References

1. Euzenat, J., Shvaiko, P.: Ontology Matching. Springer Verlag, 2007.
2. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art. The Knowledge Engineering Review 18(1), 1–31 (2003).
3. Noy, N.F., Musen, M.A.: Promptdiff: A fixed-point algorithm for comparing ontology versions. In National Conference on Artificial Intelligence and Fourteenth Conference on Innovative Applications of Artificial Intelligence, pages 744–750, Edmonton, Alberta, Canada, July 2002.
4. Michel, C., Klein, A.: Change Management for Distributed Ontologies. PhD thesis, Vrije Universiteit, Amsterdam), 2004.
5. VÂlolkel, M.: D2.3.3.v2 SemVersion Versioning RDF and Ontologies. Technical report,University of Karlsruhe, January 2006.
6. Borgida, A., Serafini, L.: Distributed description logics: Assimilating information from peer sources. Journal on Data Semantics, 2003.
7. Zimmermann, A., Le Duc, C.: Reasoning with a network of aligned ontologies. Proceeding of the 2nd International Conference on Web Reasoning and Rule systems (RR2008), 2008.
8. Meilicke, C., Stuckenschmidt, H.: An Efficient Method for Computing Alignment Diagnoses. Proceedings of the Third International Conference on Web Reasoning and Rule Systems (RR-09), Chantilly, Virginia, USA, 2009.
9. Euzenat, J. Meilicke, C. Stuckenschmidt, H. Shvaiko, P. Trojahn,C.: Ontology Alignment Evaluation Initiative: six years of experience. Journal on Data Semantics (JoDS), XV, pp. 158-192, 2011.
10. Jean-Mary, Y.R., Shironoshita, E.P., Kabuka, M.R.: Ontology matching with semantic verification. Journal of Web Semantics 7(3), 235–251 (2009).
11. Giunchiglia, F., Shvaiko, P., Yatskevich, M.: Semantic schema matching. In: Proceedings of the 13rd International Conference on Cooperative Information Systems (CoopIS), Agia Napa (CY). LNCS, vol. 3761, pp. 347–365 (November 2005).
12. Meilicke, C., Stuckenschmidt, H.: Incoherence as a basis for measuring the quality of ontology mappings. In: Proceedings of the ISWC 2008 Workshop on Ontology Matching, Karlsruhe, DE pp. 1–12 (October 2008).
13. Qi, G., Haase, P. and Ji, Q. A Conflict-based Operator for Mapping Revision--Theory and Implementation, In Proceedings of the 8th International Semantic Web Conference (ISWC'09), 2009.
14. Horridge, M., Bechhofer,S.: The OWL API: A Java API for Working with OWL 2 Ontologies. OWLED 2009, 6th OWL Experienced and Directions Workshop, Chantilly, Virginia, October 2009.
15. Euzenat, J.: An API for ontology alignment. In:McIlraith, S.A., Plexousakis, D., van Harmelen,F. (eds.) ISWC 2004. LNCS, vol. 3298, pp. 698–712. Springer, Heidelberg (2004).
16. Sirin, E., Parsia, B., Grau, B.C, Kalyanpur, A., Katz, Y.: Pellet: A practical OWL-DL reasoner. Journal of Web Semantics, 5(2), 2007.
17. Papavassiliou,V.Flouris,G., Fundulaki,I. Kotzinos,D. Christophides,V.: On Detecting High-Level Changes in RDF/S KBs. Proc. ISWC, 2009.
18. Meseguer,J.: General logics. In Logic Colloquium '87, pages 275{329. North Holland, 1989.