An Evidence-Based, Contextual Approach to the Validation of Concept Alignments to Support Ontology Reuse

Marwa Abdelreheim[1], Friederike Klan[2], Taysir Hassan A. Soliman[1]

[1] Assuit University, Assuit, Egypt
[2] DLR Institute of Data Science, Jena, Germany
marwa.abdelrehem@fci.au.edu.eg, friederike.klan@dlr.de,
taysser.soliman@fci.au.edu.eg

Abstract. Reusing ontology components can save development costs, preserves knowledge and fosters interoperability. Existing software tools for ontology reuse are based on lexical matching techniques comparing input keywords and labels of the concept and often don't consider the contextual semantics encoded in the ontology. In this paper, we propose a novel context-based semantic matching approach that identifies pairs of equivalent concepts in terms of the context of their ontologies. The algorithm is part of a framework that supports ontology engineers and domain experts in their task of developing new ontologies by reusing parts of existing ontologies. The proposed algorithm takes the correspondences obtained by existing matchmakers and checks the correctness of those mappings based on the semantics encoded in the given ontologies. Our algorithm correctly detects true positive mappings and discards false positive mappings, while in some cases it incorrectly discards true positives as these mappings are not similarly encoded in their ontologies.

Keywords: Ontology Development, Ontology Reuse, Contextual Semantic Ontology Matching.

1 Introduction

Ontologies formally represent domain knowledge in terms of concepts and relations between those concepts. They are used to easily share and reuse domain knowledge. In conceptually diverse domains, developing ontologies is a complex process that requires an extensive professional exchange between domain experts and ontology engineers. It is not a one-step process but rather an iterative one that needs to be repeated until a mutually accepted version of the ontology is reached. Rather than designing ontologies from scratch, it is good practice to reuse existing ontologies, or parts of them. This can reduce development time and effort. It can also produce high quality ontologies as it reuses parts/fragments of existing ontologies that (ideally) were validated by domain experts and tested against inconsistencies. The process typically taken in order to build a new ontology based on existing ones [4, 12, 13, 15] involves the following five steps

(1) defining the scope of the target ontology, (2) extracting domain-related keywords, which are later used to (3) select candidate ontologies for reuse, (4) evaluating and filtering those candidates based on some criteria (e.g. encoding quality requirements and possibly rank them). Finally, (5) integrating/merging the selected ontologies to produce a new ontology.

The main challenge in the outlined ontology reuse process is how to select the appropriate ontologies that accurately represent the target domain. Existing ontology selection and reuse tools/approaches [1, 4, 6, 10, 11] typically consider input coverage as the most important criterion to assess the ontology relatedness to a particular target domain. They use lexical matching (i.e., exact or partial matching) between user input keywords, and possibly their synonyms, and the labels of ontology concepts. The larger the number of lexical matches in a source ontology is, the higher is the input coverage score, and the more relevant is the source ontology considered to be to the domain of the target ontology. Another challenge when reusing ontologies arises when an ontology engineer wants to build upon an already existing target ontology or a part of it. In this case, he needs to identify concepts from existing source ontologies that semantically match the concepts of the target ontology and reuse them to extend it. To this end, the ontology engineer can use existing ontology matching tools/systems to find correspondences between concepts in a source and a target ontology.

Ontology matching typically considers two input ontologies and attempts to find semantically equivalent entities (i.e. concepts, properties, and individuals). Several ontology matching systems exist [3, 7, 14] and use complex strategies to find accurate alignments between the input ontologies' entities. They use a combination of lexical and structural approaches to find alignments. Some approaches consider external knowledge sources such as WordNet[1] to identify correspondences. In fact, existing ontology matching tools can determine alignments between concepts that have similar/synonymous labels, but their limitation is the fact that if two concepts have identical labels this does not necessarily mean that those concepts are semantically equivalent. This includes homonyms, which are concepts that share the same or similar name but have different definition in different domains. In addition, concepts in some ontologies may not have a definition indicating their meaning, and an expert should be entrusted to determine whether, or not, the two concepts are equivalent. For example, consider the concept "*Property*" in two different ontologies. A "*Property*" in one ontology could represent the weight, mass, or length of an object, while in another ontology could represent taste, color, or texture of an object. In this case, although the two concepts "*Property*" have the same label, they do not have the same contextual semantics in their ontologies and should not be considered equivalent.

In this paper, we aim to find contextual semantic equivalent concepts (i.e., two concepts have the same contextual semantics if they have the same intensional meaning in their ontologies). In logic, intensional meaning of a concept is equivalent to specifying a broad definition of it, then define specific properties it must have to be counted as a term reference. We propose a new context-based semantic matching algorithm that extends an ontology matching system that takes a set of ontology alignments (e.g.

[1] WordNet. https://wordnet.princeton.edu/

generated by an ontology matching system) and examines the contextual semantic similarity (intensional meaning) for each candidate corresponding concepts. This method can exclude alignments that do not have evidence for contextual semantic similarity. This would improve the ontology reuse process by producing more accurate mappings and assist the ontology engineer in the creation of a new ontology, by discovering concepts/parts of existing source ontologies and proposing such sections to extend the target ontology.

The paper is organized as follows. A brief literature review of the existing ontology matching systems is presented in section 2. In section 3, we describe our proposed algorithm for identifying concepts likely to be semantically equivalent in the context of their ontologies. In section 4, we present and discuss the results of an empirical evaluation of our methodology. Finally, we conclude by identifying directions for future work.

2 Related Work

Ontology matching considers two input ontologies and aims to find correspondences between related entities (i.e., concepts, properties, and individuals) in both ontologies. Due to the complexity of ontologies, in terms of their size and structure, existing ontology matching algorithms adapt complex strategies to find accurate alignments between ontology entities. COMA++ [3] is a tool that applies diverse lexical and taxonomic matching algorithms to support matching different types of schemas and ontologies. AgreementMakerLight (AML) [7] is an ontology matching system that comprises several matching techniques and applies it on different levels for the components being matched (conceptual vs. structural). YAM++ [14] encodes the structural information of an ontology in a bitmap to accelerate calculating the similarity by comparing whole ontologies and avoiding pairwise comparisons of concepts. In addition, it discovers new mappings by applying similarity propagation methods on the initial list of mappings. Extensible Mapping (XMAP) [5] is a scalable ontology matching system aims to map very large ontologies efficiently (in terms of computation time) and effectively (in terms of precision and recall scores). It applies different matchers such as string, linguistic, and structure-bases and use mathematical methods to aggregate similarity scores resulted by each matcher.

There are several approaches addressing this. After getting an initial set of mappings by performing lexical and structural matching, an ontology reasoner can be used to recognize sets of correspondences that prove to be inconsistent, as in ASMOV[8], or to discover new mappings and repair basic mappings that are not satisfiable as done by LogMap [9]. Finally, STROMA [2] is a new ontology matching approach aiming to more accurately specifying the kind of semantic relationship between two concepts. In contrast to other ontology mapping approaches, its goal is to identify more expressive relationships between concepts, such as is-a, part-of, and equality relations between concepts.

3 Evidence-Based, Contextual Alignment Validation

Our contextual semantic matching algorithm aims to find concepts in two ontologies that are semantically equivalent and exclude concepts that are not proven to be semantically equivalent in the context of their ontologies even if they have similar labels. In order to achieve this, we take the alignments identified by a given ontology matching system and check if there is evidence in the ontologies that supports the equivalence of the concepts referred to in an alignment. Mappings without supporting evidence are filtered out. To achieve this, we take the alignments identified by a given ontology matching system and check if there is evidence in the ontologies that supports the equivalence of the concepts referred to in an alignment in terms of encoded knowledge.

As shown in Algorithm 1 below, we use the AgreementMakerLight matcher (AML) [7] as the base matcher of our algorithm to derive an initial set of candidate correspondences. The reasons for this are: (1) AML has maintained its position among the top systems in all areas of the Ontology Alignment Evaluation Initiative (OAEI) competition. (2) AML is available on GitHub as an open source code project[2] and thus can be reused and extended. (3) AML applies lexical and structural matching methods and uses WordNet as an external source of knowledge to find correspondent concepts. (4) AML can handle very large ontologies in in efficient computing time. But, as other ontology matching systems, it relies on comparing concept labels, even at the structural level, which alone is not sufficient for determining contextually semantic equivalent entities. Our evidence-based, contextual semantic matching algorithm is summarized in Algorithm 1.

Algorithm 1 Context-based Semantic Matching Algorithm

Input: AML candidate correspondences
Output: set of context-based semantically equivalent concepts (alignments)

```
For each candidate mapping (X, X*) do
    if (there exists at least one candidate mapping (Y, Y*)) such that
        case1: Y and Y* have superclass relation with X and X* resp., or
        case2: Y and Y* have subclasses relation with X and X* resp., or
        case3: Y and Y* have equivalence relation with X and X* resp., or
        case4: Y and Y* have sibling relation with X and X* resp.
        Then this is an indicator for X and X* being semantically equivalent.
    end if
end for
```

To illustrate this, suppose that there is a list of mappings between two ontologies. Let concept *X* from one ontology and concept *X** from the other ontology map to each other with acceptable AML similarity score (e.g. above a given threshold). Then, our algorithm will examine, for any pair of concepts (X, X*), whether one (or more) of the

[2] https://github.com/AgreementMakerLight/AML-Project

following additional evidences of contextual semantic equivalence is given. We arrange evidences according to their strength of indicating contextual equivalence.

1. Corresponding super-classes

For each candidate mapping (X, X*), if there exists another candidate mapping (Y, Y*), where Y and Y* have a superclass relation (direct or indirect) with X and X* respectively, then this is an indicator for X and X* being semantically equivalent in their context, see Fig. 1(a). On the contrary, if there is no such mapping (Y, Y*), then there is no evidence that X and X* are semantically equivalent.

Example. Let's assume we have a concept *Car,* which is subclass of concept *Automobile,* which in turn is a subclass of concept *Vehicle* in one ontology. And, Concept *Motor Vehicle* is a subclass of concept *Vehicle* in the other ontology. If there is a mapping between the concepts *Car* and *Motor Vehicle* and between the concepts *Vehicle* and *Vehicle* (their super classes), then the mapping (*Motor Vehicle, Car)* is supported by additional evidence, represented in Fig. 1(b) as double dashed mapping. Here, both concepts represent something related to a *Vehicle*.

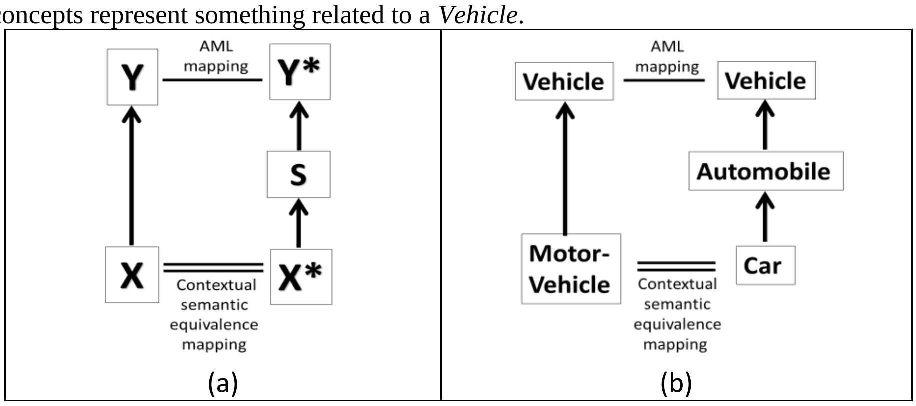


Fig. 1 a) Corresponding super-classes contextual semantic similarity evidence. b) An illustrative example for contextual semantic similarity mapping evidence based on super-class relationship.

2. Corresponding sub-classes

For each candidate mapping (X, X*), if there exists another candidate mapping (Y, Y*), where Y and Y* have a subclass relation with X and X* respectively, then this is an indicator for X and X* being semantically equivalent, see Fig. 2(a). The higher the number of candidate mappings which are sub-classes of (X, X*), the stronger the evidence that concepts (X, X*) have the same contextual semantics. On the other hand, if the candidate mapping (X, X*) does not have any corresponding sub-classes candidate mapping, then (X, X*) has no evidence for a contextual equivalent semantics.

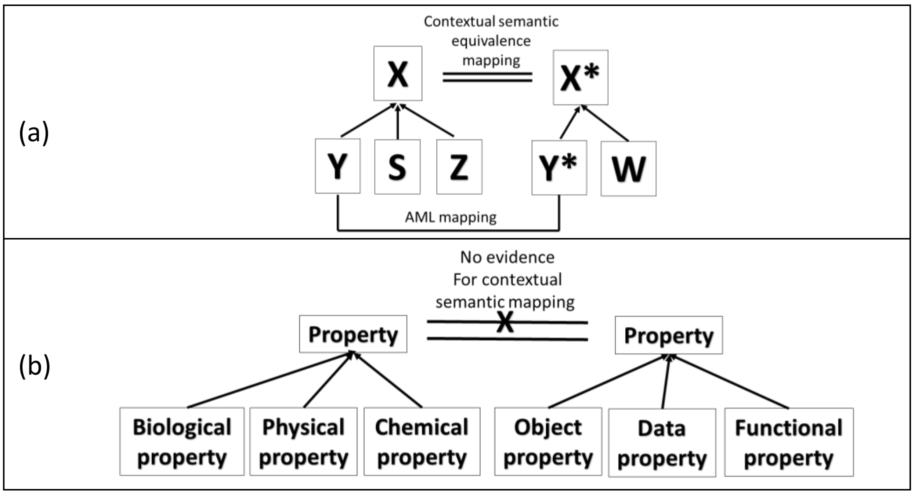


Fig. 2 a) Corresponding sub-classes semantic similarity evidence. b) An illustrative example where there is no evidence for contextual semantic similarity mapping based on sub-class relationship.

Example. In one ontology, concept *Property* has subclasses *Physical Property*, *Biological Property*, and *Chemical Property*. While, in the other ontology concept *Property* has subclasses *Object Property*, *Data Property*, and *Functional Property*. If the concepts *Property* in both ontologies map but none of their subclasses map to each other, this would indicate that the two concepts *Property* represent different things in the context of their ontologies (see Fig. 2(b)).

3. Corresponding equivalent concepts

For each candidate mapping (X, X*), see Fig. 3(a), if there exists another candidate mapping (Y, Y*), where Y and Y* have equivalence relation with X and X* respectively, then this is an indicator for X and X* being semantically equivalent.

Example. Let the concept *Physical Quality* be equivalent to the concept *Quality*. And, Concept *Physical Characteristics* be equivalent to the concept *Characteristics*, see Fig. 3(b). If the concepts *Physical Quality* and *Physical Characteristics* map to each other, and also the concepts *Quality* and *Characteristics* map to each other, according to the ontology matching algorithm, then this is an indicator that the concept *Physical Quality* in one ontology has the same contextual semantics as the concept *Physical Characteristics* in the other ontology. Here, the input mappings provide double support for the equivalence by resulting in mappings for both concepts and its equivalent concepts.

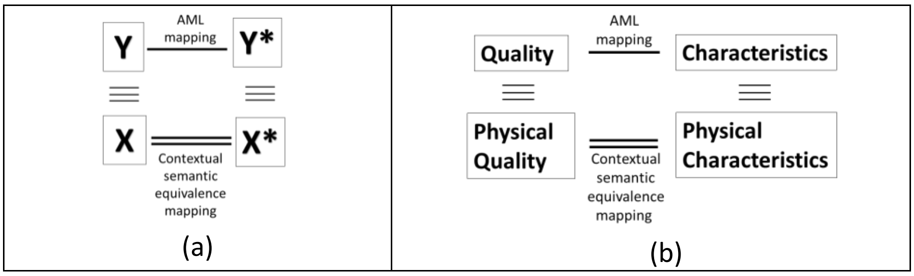
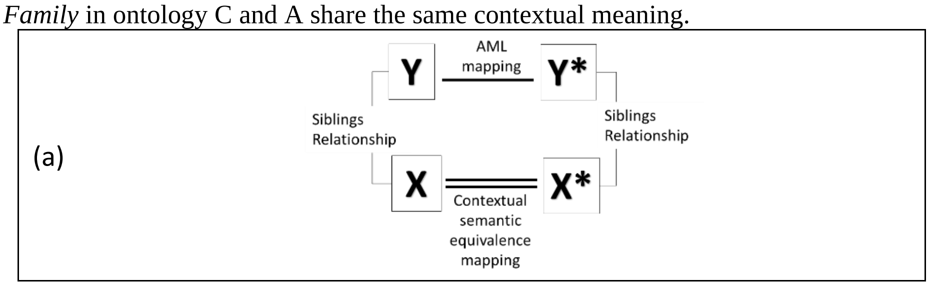


Fig. 3 a) Corresponding equivalent concepts semantic similarity evidence. b) An illustrative example for contextual semantic similarity evidence based on equivalent relationships.

4. Corresponding sibling concepts

For each candidate mapping (X, X*), if there exists another candidate mapping (Y, Y*), where Y and Y* have sibling relation with X and X* respectively, then X and X* are (not necessarily) semantically equivalent, but they share a similar context. The higher the number of candidate mappings which are siblings of (X, X*), the stronger the evidence that concepts (X, X*) have the same contextual semantics, represented as double dashed mapping in Fig. 4(a). On the contrary, if there are no candidate mappings among (X, X*)'s siblings, then there is no evidence that X and X* have a contextual equivalent semantics.

Example. In Fig. 4(b), the concept *Family* has sibling concepts *Cluster* and *Swarm in ontology A*. Now let's assume we have two other ontologies B and C. Ontology B, defines a concept *Family* with sibling concepts *Population of organisms*, *Collection of humans*, and *Species*, while in ontology C, the concept *Family* has sibling concepts *lifestyle*, *social status*, and *person*. If there are the mappings (*Collection of humans*, *Cluster*) and (*Population of organisms*, *Swarm*) in B and A, respectively. While there are no mappings between the concept *Family*'s sibling concepts in ontology C and A. Then, there is evidence that the concept *Family* in ontology B is contextually semantically equivalent to the concept *Family* in A, while there is no evidence that the concepts *Family* in ontology C and A share the same contextual meaning.



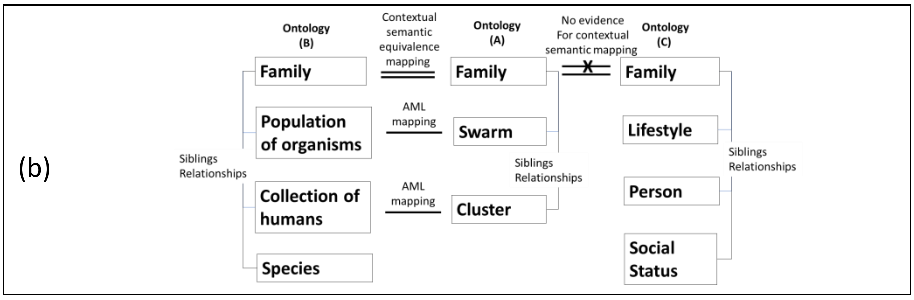


Fig. 4 a) Corresponding sibling concepts semantic similarity evidence. b) An illustrative example for contextual semantic similarity mapping evidence based on sibling relationships.

The algorithm aims to validate a set of input mappings (e.g. provided by AML) for two ontologies. Any pair of corresponding concepts that does not satisfy any of the conditions described above should be excluded from the final correspondences, since there is no evidence indicating that the two concepts referred to in the mapping share the same contextual meaning. This can improve the quality of any candidate set of corresponding concepts by excluding corresponding pairs that do not share the same contextual semantics. As mentioned above, an important application for this kind of algorithm is when suggesting ontology fragments for reuse. The algorithm provides ontology engineers with a set of validated mappings between a source and a target ontology (i.e., possible extension points). If a pair of corresponding concepts satisfies one or more of the conditions above, then the source concept is listed in the suggestions list, to extend the corresponding concept in the target ontology. In addition, the presented conditions are arranged from the strongest evidence for contextual semantic equivalence to the weakest, so the suggestions list of concepts from different source ontologies that could be used to extend the corresponding concept in the target ontology, could also be ranked. For, example, a concept is highly ranked in the suggestions list if it satisfies one of the strong evidences (e.g., 1 or 2), or if it satisfies more than one condition at the same time (e.g., 1 and 4).

4 Empirical Evaluation and Discussion

In this section we provide a quantitative and a qualitative analysis for our evidence-based approach, comparing it with the mapping results produced by the AML matchmaker in the context of the annual competitions organized by the Ontology Alignment Evaluation Intuitive (OAEI)[3]. Those competitions address different ontology matching systems to compare their performance based on benchmark matching tasks. OAEI has different tracks, each providing a specific set of ontologies, where ontology matching systems are tested on.

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

4.1 Evaluation Task

For our evaluation, we use the alignment tasks and ontologies of OAEI's Conference track. This track provides 16 ontologies about conferences developed by the OntoFarm project[4], popularly used as a benchmark dataset. The track consists of 21 alignment tasks. Each task comprises two ontologies for which a set of reference alignments is given which corresponds to the complete alignment space between 7 ontologies. To compare our algorithm's matching results with AML's results, we use the original reference alignments downloaded from the Conference track of (OAEI's 2018) edition of the competition[5]. We just consider the reference alignments for concepts (the competition also considers alignments between object properties, which our algorithm does not support). We run AML (version- 3.1) to get the initial set of mappings as input for our algorithm.

4.2 Results

After applying our evidence-based, contextual semantic matching algorithm to all 21 benchmark tasks, we compare AML ontology matching results and the results produced by our contextual semantic matching based on AML input, all results are included in Table 1. AML Results were downloaded from the results page of OAEI's 2018 conference track[6]. We compare the results in terms of precision (fraction of correct mappings returned by the algorithm), recall (fraction of the reference mappings returned by the algorithm), and F-measure (2*((precision * recall) / (precision + recall)). The best we can expect when applying our algorithm, is to improve precision by eliminating false positive alignments returned by AML, that do not represent the same context. The best we can achieve in terms of recall is to keep it the same value as AML's, as our algorithm cannot discover new true positive alignments, but only checks the contextual semantics of the alignments produced by the AML matchmaker. In Table 1, the dotted cells mean increasing in the score, white cells mean the score remains the same, and gray cells means decreasing in the score.

As shown in the Table 1, we can divide the table to four groups according to the increase/decrease in precision and recall:

- Group 1, precision increases and recall remains the same (no. 3 and 7): this means that our algorithm successfully eliminates AML mappings that are false positives (not in the reference alignments) without losing any true positives (true alignments). This happens when the two ontologies have similar contexts and describe their concepts in almost similar ways. These ontologies are typical candidates in an ontology reuse scenario, described in the introduction, as they are describing the conference domain from the same context.

[4] https://owl.vse.cz/ontofarm/
[5] http://oaei.ontologymatching.org/2018/results/conference/#crisp-ra
[6] http://oaei.ontologymatching.org/2018/conference/index.html

- Group 2, precision increases but at the same time recall decreases (no. 2, 8, 9, 10, 12, 13, 14, 16, and 17): this means that the algorithm successfully eliminates false positive mappings but also eliminates some true positive mappings. These are the majority of the cases; they happen when the two ontologies are generally having the same context, but they describe their concepts in slightly different ways.
- Group 3, precision decreases and recall decreases (no. 1, 4, 5, 6, 19, and 20): this means that the algorithm eliminates true positive mappings without discovering false positive mappings. That happens when the two ontologies are from the same domain, but they describe the concepts in different ways, or they are from interdisciplinary domains sharing common vocabularies.
- Group 4, precision and recall remain the same (no. 11 and 21), this means that AML mappings are all true positives and our algorithm successfully discovers all of them, and the two ontologies describe the same context.

Table 1. Comparison AMLs and our algorithms' precision, recall, and F-measure results for the alignment tasks of the Conference track.

		AML Results			Our Algorithm Results		
	Ontologies	**Precision**	**Recall**	**Fmeasure**	**Precision**	**Recall**	**Fmeasu**
1	cmt-conference	0.64	0.58	0.61	0.6	0.5	0.55
2	cmt-confOf	0.89	0.8	0.84	1	0.7	0.82
3	cmt-edas	0.89	1	0.94	1	1	1
4	cmt-ekaw	0.75	0.75	0.75	0.71	0.63	0.67
5	cmt-iasted	0.8	1	0.89	0.75	0.75	0.75
6	cmt-sigkdd:	0.9	0.9	0.9	0.89	0.8	0.84
7	conference-confOf:	0.73	0.8	0.76	0.8	0.8	0.8
8	conference-edas:	0.69	0.64	0.67	0.75	0.43	0.55
9	conference-ekaw:	0.78	0.78	0.78	0.84	0.7	0.76
10	conference-iasted:	0.83	0.38	0.53	1	0.23	0.38
11	conference-sigkdd:	0.83	0.83	0.83	0.83	0.83	0.83
12	confOf-edas:	0.9	0.6	0.72	1	0.47	0.64
13	confOf-ekaw:	0.94	0.8	0.86	1	0.75	0.86
14	confOf-iasted:	0.8	0.44	0.57	1	0.33	0.5
15	confOf-sigkdd:	1	0.83	0.91	1	0.33	0.5
16	edas-ekaw:	0.79	0.58	0.67	0.9	0.47	0.62
17	edas-iasted:	0.82	0.47	0.6	0.86	0.32	0.46
18	edas-sigkdd:	1	0.69	0.82	1	0.54	0.7
19	ekaw-iasted:	0.88	0.7	0.78	0.8	0.4	0.53
20	ekaw-sigkdd:	0.8	0.73	0.76	0.78	0.64	0.7
21	iasted-sigkdd:	0.81	0.87	0.84	0.81	0.87	0.84
		0.83	0.72	0.76	**0.87**	**0.59**	**0.68**

Dotted cells indicate an increase in the score
Gray cells indicate a decrease in the score
White cells indicate that the score remains the same

4.3 Discussion

In this section we discuss individual mapping results, in particular those where our algorithm mistakenly discards true positives, but also where it correctly discards false positives detected by AML.

Case 1. This is an example where our algorithm succeeds to confirm a true positive mapping identified by AML correctly. Consider the two ontologies Cmt and Conference (part of the competition tasks of the OAEI's Conference track), shown in Fig. 5. Consider the concept mappings (*Cmt:SubjectArea*, *Conference:Topic*) and (*Cmt:Chairman*, *Conference:Chair*) identified by AML.

Mapping (Cmt:SubjectArea, Conference:Topic): This mapping is supported by evidence of type 4 (having correspondent siblings) as defined in Sec.3. The mappings (*Cmt:Document*, *Conference:Conference_document*), (*Cmt:Person*, *Conference:Person*), and (*Cmt:Conference*, *Conference:Conferece*) referring to siblings of *Cmt:SubjectArea and*, *Conference:Topic*, provide evidence that the concepts *Subject-Area* and *Topic* share the same contextual meaning in the two input ontologies.

Mapping (Cmt:Chairman, Conference:Chair): This mapping is supported by evidence of type 1 (having correspondent super-classes). Since there is a mapping (*Cmt:Person*, *Conference:Person*), which refers to super-classes of *Cmt:Chairman and Conference:Chair* respectively, this is an evidence that the concepts *Chairman* and *Chair* share the same contextual meaning within the given input ontologies.

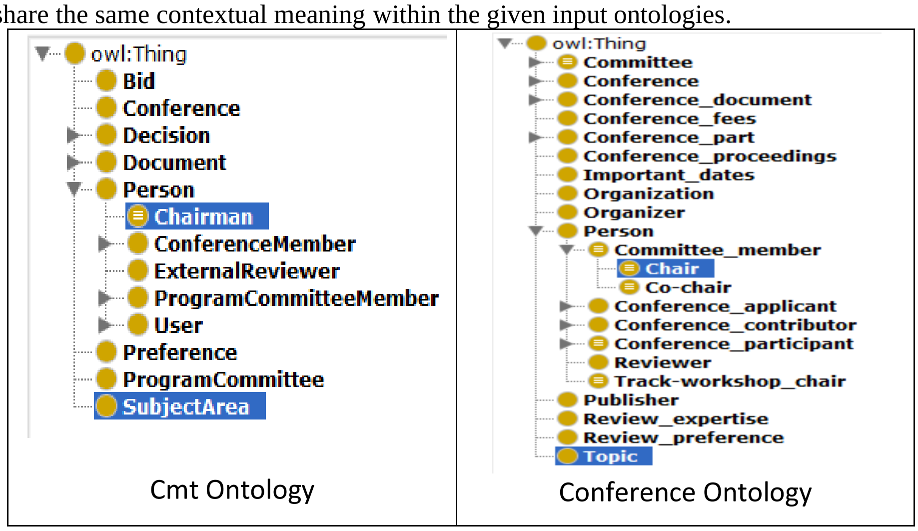


Fig. 5. Concepts *SubjectArea* and *Chairman* in Cmt ontology, and concepts *Topic* and *Chair* in Conference ontology

Case 2. This is as example where our algorithm succeeds to detect false positive mappings, which were mistakenly detected by AML as true mappings. Consider the ontologies Conference and Iasted (part of the competition tasks of the OAEI's Conference track), shown in Fig. 6., and ontologies Edas and Ekaw, shown in Fig. 7.

Mapping (Conference:Presentation, Iasted:Presentation): This mapping is not supported by any of the evidences defined in Sec.3. (there are no corresponding super-classes, sub classes, equivalent classes, or siblings). Thus, the algorithm discards this mapping which is a false positive one as it is not included in the reference alignments. Looking closely to both ontologies we can observe that concept *Presentation* is represented differently in each ontology, as in ontology Conference a presentation is a special type of conference documents while in Iasted ontology a presentation is a special type of a conference activity. To conclude, although the two concepts *Presentation* have exactly the same labels, they have different contextual semantics in both ontologies.

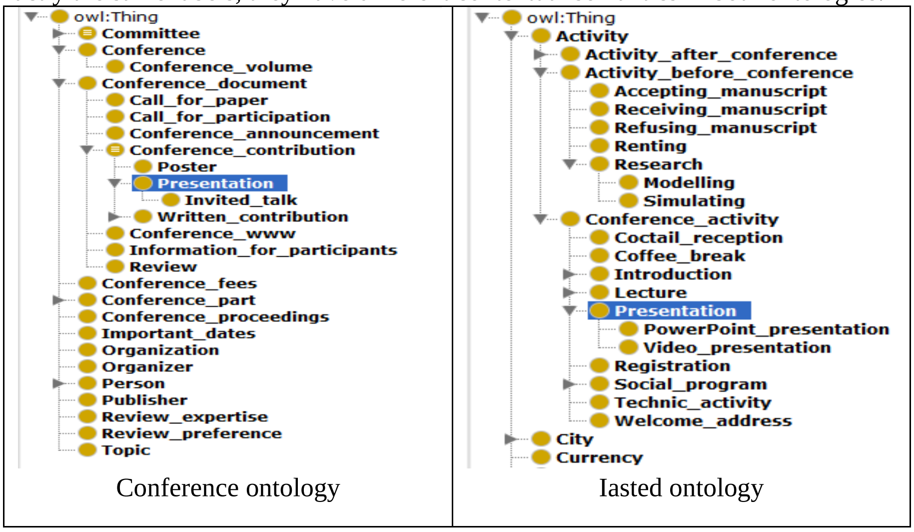


Fig. 6. Concepts *Presentation* in Conference and Iasted ontologies

Mapping (Edas:Conference, Ekaw:Conference): This mapping is not supported by any of the evidences defined in Sec.3. (there are no corresponding super-classes, sub classes, equivalent classes, or siblings). Thus, the algorithm discards this mapping which is a false positive one as it is not included in the reference alignments. Looking closely to both ontologies we can observe that concept *Conference* is represented differently in each ontology, as in ontology Edas the concept *Conference* is sub-class of a *Thing* and has sibling relations with concepts like *ConferenceEvent*, *Document*, *Person*, etc. On the other hand, a *Conference* in Ekaw ontology is represented differently as a special type of *Event*, namely *Scientific Event* and has sibling relations with the concepts *Workshop*, *Session*, *Track*, etc. Thus, the mapping (*Edas:Conference*, *Ekaw:Conference*), was discarded and considered as having different contextual semantics in both ontologies.

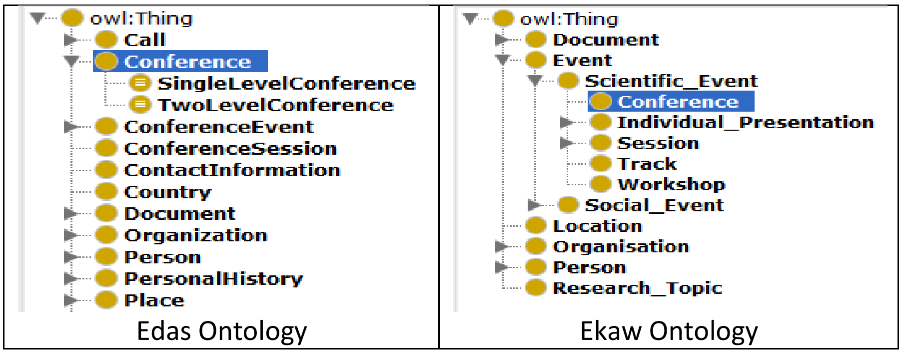


Fig. 7. Concepts *Conference* in Edas and Ekaw ontologies

Case 3. Finally, this is an example where our algorithm mistakenly discards mappings that are true positives. As it turns out, this is because the knowledge encoded in the ontologies does not provide evidence for this mapping. Consider the ontologies Cmt and Sigkdd (part of the competition tasks of the OAEI's Conference track), shown in Fig. 8.

Mapping (Cmt:ProgramCommittee, Sigkdd:Program_Committee): This mapping is not supported by any of the evidences defined in Sec.3. (there are no corresponding super-classes, sub classes, equivalent classes, or siblings). This mapping is a true positive one as it is included in the reference alignments, but our algorithm incorrectly discards it as it lacks evidence in the ontologies. In Fig. 8., concepts *Program Committee* are represented differently in the two ontologies, in Sigkdd ontology it is a type of committee and a committee is related to the conferences, while in Cmt ontology *Program Committee* is a concept having sibling relations with concepts conferences, documents, persons, etc.

Although our algorithm discards this type of mappings (the true positive ones), this does not mean that those mappings are not valid. It just means that they lack evidence in their ontologies and represented differently in the contexts of their ontologies (i.e., have different intensional meanings). In an ontology reuse scenario where we want to reuse existing ontologies to extend an input ontology, as part of the future work, the priority of suggested concepts will be given to concepts that are validated against evidence of contextual semantic equivalence. Other non-validated pairs (like this case) could be presented to the expert in a different list to evaluate their validity. This could be done using confidence score, which is part of the future work.

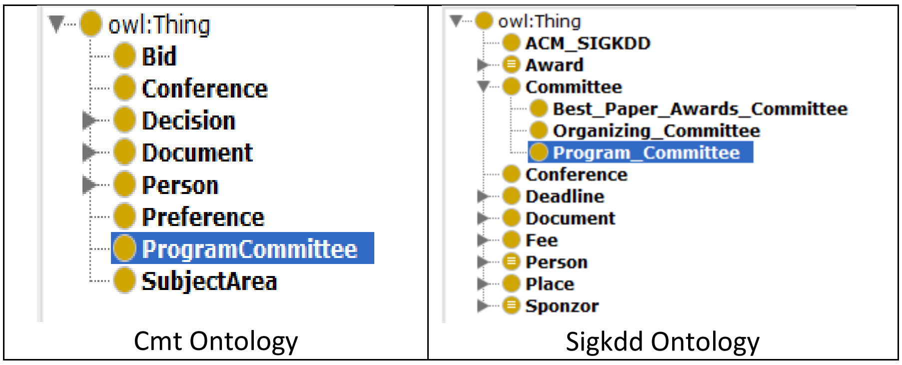


Fig. 8. Concepts *Program Committee* in Cmt and Sigkdd ontologies

5 Conclusion and Future Directions

Developing new ontologies by reusing ontology fragments or concepts from existing ontologies is considered a best practice in ontology engineering. To support this, we require methods that identify pairs of concepts in a given source ontology and target ontology that share the same contextual meaning (intensional meaning). This would help ontology engineers to choose appropriate concepts from source ontologies to extend their target ontology. Existing matchmakers do not distinguish different contextual meanings of concepts. In this paper, we propose a new evidence-based, contextual approach to validate matchings obtained from an ontology mapping system (e.g. AML) by checking whether they refer to concepts that share the same meaning within the context of their source ontologies. Our evaluation based on OAEI's competition tasks in the Conference track shows that our algorithm successfully confirms true positives and discards false negatives from the set of mappings returned by the AML matcher. In some cases, however, it failed to confirm true positives when concepts presented differently in their ontologies.

In the future we are intending to enhance our evidence-based, contextual matching algorithm to avoid losing the true positive mappings. This could be done by examining concepts' annotations (i.e., comments, definitions, etc.), if exists, and other external knowledge to verify contextual semantic equivalence of concepts. Then, we could build an ontology reuse system that aims to reuse existing ontologies to develop new ontologies by extending an input ontology in a target domain. Examining the validity of our algorithm in a real case scenario, the new developed ontology, would be assessed by domain expert.

References

1. Abdelreheim M., Klan F., Soliman T. (2017) Towards personalized support for ontology selection. In: the 29th GI-Workshop Grundlagen von Datenbanken. pp 13–18
2. Arnold P., Rahm E. (2014) Enriching ontology mappings with semantic relations. Data Knowl Eng 93:1–18. doi: https://doi.org/10.1016/j.datak.2014.07.001

3. Aumueller D., Do H-H, Massmann S., Rahm E. (2005) Schema and Ontology Matching with COMA++. In: Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data. ACM, New York, NY, USA, pp 906–908
4. Caldarola EG., Picariello A., Rinaldi AM. (2015) An Approach to Ontology Integration for Ontology Reuse in Knowledge Based Digital Ecosystems. In: Proceedings of the 7th International Conference on Management of Computational and Collective intElligence in Digital EcoSystems. ACM, New York, NY, USA.
5. Djeddi WE., Khadir MT. (2014) A Novel Approach Using Context-Based Measure for Matching Large Scale Ontologies. In: Data Warehousing and Knowledge Discovery. Springer International Publishing, Cham.
6. Faessler E., Klan F., Algergawy A., König-Ries B., Hahn U. (2017) Selecting and Tailoring Ontologies with JOYCE. In: Knowledge Engineering and Knowledge Management. Springer International Publishing, Cham, pp 114–118
7. Faria D., Pesquita C., Santos E., Palmonari M., Cruz IF., Couto FM. (2013) The AgreementMakerLight Ontology Matching System. In: On the Move to Meaningful Internet Systems: OTM 2013 Conferences. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 527–541
8. Jean-Mary YR., Shironoshita EP., Kabuka MR. (2009) Ontology matching with semantic verification. J Web Semant 7:235–251. doi: https://doi.org/10.1016/j.websem.2009.04.001
9. Jiménez-Ruiz E., Cuenca Grau B. (2011) LogMap: Logic-Based and Scalable Ontology Matching. In: The Semantic Web -- ISWC 2011. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 273–288
10. Martínez-Romero M., Jonquet C., O'Connor MJ., Graybeal J., Pazos A., Musen MA. (2017) NCBO Ontology Recommender 2.0: an enhanced approach for biomedical ontology recommendation. J Biomed Semantics 8:21.
11. Martínez-Romero M., Vázquez-Naya JM., Pereira J., Pazos A. (2014) BiOSS: A system for biomedical ontology selection. Comput Methods Programs Biomed 114:125–140. doi: https://doi.org/10.1016/j.cmpb.2014.01.020
12. Shah T., Rabhi F., Ray P., Taylor K. (2014) A Guiding Framework for Ontology Reuse in the Biomedical Domain. In: 2014 47th Hawaii International Conference on System Sciences. pp 2878–2887
13. Simperl E. (2009) Reusing ontologies on the Semantic Web: A feasibility study. Data Knowl Eng 68:905–925. doi: https://doi.org/10.1016/j.datak.2009.02.002
14. Tigrine AN., Bellahsene Z., Todorov K. (2015) Light-Weight Cross-Lingual Ontology Matching with LYAM++. In: On the Move to Meaningful Internet Systems: OTM 2015 Conferences. Springer International Publishing, Cham, pp 527–544
15. Zulkarnain NZ., Meziane F., Crofts G. (2016) A Methodology for Biomedical Ontology Reuse. In: Natural Language Processing and Information Systems. Springer International Publishing, Cham, pp 3–14