Matching Domain and Top-level ontologies exploring Word Sense Disambiguation and Word Embedding

Daniela Schmidt*, Rafael Basso*, Cassia Trojahn†, Renata Vieira*

*Pontifical Catholic University of Rio Grande do Sul (Brazil)
{daniela.schmidt, rafael.basso}@acad.pucrs.br, renata.vieira@pucrs.br
†Université de Toulouse 2 & IRIT (France)
cassia.trojahn@irit.fr

Abstract. Top-level ontologies play an important role in the construction and integration of domain ontologies, providing a well-founded reference model that can be shared across domains. While most efforts in ontology matching have been particularly dedicated to domain ontologies, the problem of matching domain and top-level ontologies has been addressed to a lesser extent. This is a challenging task in the field, specially due to the different levels of abstraction of these ontologies. This paper addresses this problem by proposing an approach that relies on existing alignments between WordNet and top-level ontologies. Our approach explores word sense disambiguation and word embedding models. We evaluate our approach in the task of matching DOLCE and SUMO top-level ontologies to ontologies from three different domains.

1 Introduction

Guarino [7] classifies ontologies according to their "level of generality", in particular: *(i) top-level ontologies* describe very general concepts (e.g., space, time, object, etc.), which are independent of a particular problem or domain. These ontologies, also named upper or foundational ontologies [33], are equipped with a rich axiomatic layer; *(ii) domain ontologies* that describe the entities and other information related to a generic domain (e.g., biology or aeronautic). While the rich semantics and formalization of top-level ontologies are important requirements for ontology design [18], they act as well as semantic bridges supporting very broad semantic interoperability between ontologies [15,16]. In that sense, they play a key role in ontology matching.

Whereas the area of ontology matching [3] has developed fully in the last decades, matching ontologies from different levels of abstraction as domain and top-level ontologies is still an early tackled challenge. This is a complex task, even manually, that requires the deep identification of the semantic context of concepts and, in particular, the identification of subsumption relations. The latter is largely neglected by most state-of-the-art matchers. The main problem of matching top-level and domain ontologies using these matching systems is that, despite the variety of approaches, most of them typically rely on string-based techniques as an initial estimate of the likelihood that two elements refer to the same real world phenomenon, hence the found correspondences represent equivalences with concepts that are equally or similarly written. However, in many cases, this correspondence is wrong [32]. In fact, when having different levels of abstraction it might be the case that the matching process is rather capable of identify subsumption correspondences than equivalence, since the top ontology has concepts at a higher level. Approaches dealing with this task are mostly based on manual matching [1,18].

This paper proposes an approach to match domain and top-level ontologies that exploits existing alignments between top-level ontologies and WordNet [20]. These alignments act as bridges for aligning domain and top-level ontologies. The notion of *context* of concepts is used for disambiguating the senses that better express the meaning of domain ontology entities in this external resource. Contexts are constructed from the available terminological information about a domain ontology entity (e.g., entity naming, annotations, and information on the neighbours of entities[1]). Here, we exploit two similarity measures for synset disambiguation : (1) an adaptation of the Lesk measure [13] and (2) word embeddings [19]. Once the domain synset has been selected, we exploit the relation between this synset and a top-level concept via existing alignments between WordNet and the top-level ontologies. Most strategies we apply here, in particular indirect matching [9,8,39], WordNet-based matching [14,38], the classical notion of context [37,30,2] and word-sense disambiguation [21], have been already exploited in different ways in the field. However, we argue that the novelty of our approach relies on their combination, which remains unexplored in the specific task of matching top-level and domain ontologies. The use of word embedding for the matching task is, however, less studied [40,36]. Here, we focus on DOLCE and SUMO top-level ontologies and on their alignments to WordNet [6,23]. This choice is motivated by the fact that they are the most used top-level ontologies and serve as a reference model for the modelling and integration of ontologies [24]. We align them to ontologies from three domains (SSN[2], CORA [28], and OAEI Conference ontologies[3]).

The main contributions of our paper can be summarised as follows : (i) to the best of our knowledge, our approach is the first attempt to automatically match domain and top-level ontologies; (ii) we provide an evaluation of our approach and compare how state-of-the-art matching results can be improved by exploiting existing alignments between WordNet and top-level ontologies; and (iii) our results may form a baseline for an OAEI task since there is no current track involving this kind of challenge.

The rest of the paper is organised as follows. §2 introduces top-level ontologies, WordNet and existing alignments to WordNet. §3 discusses the main related work. §4 presents our matching approach. §5 presents the experiments and discusses the results. Finally, §6 concludes the paper and presents future work.

2 Background

2.1 Top-level ontologies

A top-level ontology is a high-level and domain independent ontology. The concepts expressed are intended to be basic and universal to ensure generality and expressiveness for a wide range of domains. It is often characterized as representing common

[1] Here, we do not exploit restrictions and other axioms (e.g., disjointness)
[2] https://www.w3.org/TR/vocab-ssn/
[3] http://oaei.ontologymatching.org/2017/conference/index.html

sense concepts and is limited to concepts which are meta, generic, abstract and philosophical. Several top-level ontologies have been proposed in the literature. The reader can refer to [15] for a review of them. Here, we briefly introduce DOLCE and SUMO, which are further used in our experiments. DOLCE [5,17] (Descriptive Ontology for Linguistic and Cognitive Engineering) was designed to include the most reusable and widely applicable upper-level categories, rigorous in terms of axiomatization and extensively researched and documented. It is an ontology of particulars which has four top-level concepts: *endurant*, *perdurant*, *quality*, and *abstract*. *Endurants* represent objects or substances while *perdurants* correspond to events or processes. The main relation between *endurants* and *perdurants* is that of participation, e.g., a person which is an *endurant*, may participate in a discussion, which is a *perdurant*. *Qualities* can be seen as the basic entities that we can perceive or measure, e.g., shapes, colors, sizes, etc. *Abstracts* do not have spatial or temporal qualities, and they are not qualities themselves. DOLCE has many variations, such as DOLCE-Lite [6], which is an OWL-DL fragment of DOLCE. DOLCE-Lite has been extended in modules for representing information, communication, plans, and with some domain information for representing e.g. legal, biomedical notions. The combination of DOLCE-Lite and the mentioned additional modules is called DOLCE-Lite-Plus[4].

SUMO [22] (Suggested Upper Merged Ontology) provides definitions for general-purpose terms and acts as a foundation for more specific ontologies. It is being used for research and applications in search, linguistics and reasoning. It is an ontology of particulars and universals which has two top-level concepts: *physical* and *abstract*. *Physical* represent an entity that has a location in space-time. An *abstract* can be said to exist in the same sense as mathematical objects such as sets and relations, but they cannot exist at a particular place and time without some physical encoding or embodiment.

2.2 WordNet and its alignments to top-level ontologies

WordNet [20] is a general-purpose large lexical database of English frequently adopted as an external resource in automatic ontology matching between domain ontologies [38,37,30]. In the following, we discuss its alignments to top-level ontologies.

DOLCE to WordNet alignment (OntoWordNet) Gangemi et al. [6] developed the OntoWordNet, a resource which expresses the alignment between WordNet 1.6 version and DOLCE-Lite-Plus. The authors assume that the hyponymy relation could be aligned to the subsumption relation and the synset notion could be aligned to the notion of concept. In OntoWordNet, the named concepts were normalized to obtain one distinct name for each synset. Hence, if a synset had a unique noun phrase, it is used as a concept name (*e.g.* Document__Written_Document__Papers). If the noun phrase was polysemous, the concept was numbered (*e.g.* Writting_1, Writting_2). Figure 1 presents a fragment of WordNet synsets (as concepts) linked to DOLCE-Lite-Plus concepts. The first-level concepts (in lower case) correspond to a DOLCE-Lite-Plus concept. The upper case concepts represent WordNet synsets. Each concept in OntoWordNet is associated to an annotation containing the corresponding gloss of the synset in WordNet.

[4] http://www.loa.istc.cnr.it/old/ontologies/DLP_397.owl

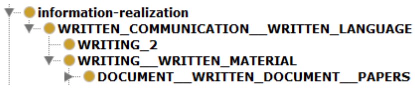

Fig. 1. Example of WordNet synsets linked to DOLCE.

SUMO to WordNet alignment Niles and Pease [23] construct an alignment between SUMO top-level ontology and WordNet 1.6. For each identified correspondence, the synset of WordNet is augmented with three information : *(i)* a prefix (&%) that indicates that the term is taken from SUMO; *(ii)* the SUMO concept; and *(iii)* a suffix indicating the kind of relation. The suffix '=' indicates that the correspondence relation is synonymy. '+' indicates that the concept is a hypernym of the associated synset. The instantiation relation is indicated by the suffix '@'. An example of the structure of SUMO and WordNet correspondence representing a synonymy relation can be seen below. In the example, *"02761392 06 n 03 automaton 0 robot 0 golem"* corresponds to the synset. The gloss is defined as *"a mechanism that can move automatically"*, the prefix *"&%"* indicates that the term is taken from SUMO. *"Device"* corresponds to the SUMO concept and the signal *"+"* is the suffix indicating the hyponymy relation.

02761392 06 n 03 automaton 0 robot 0 golem — a mechanism that can move automatically
&%Device+

There are other efforts that provide alignments of WordNet to top-level ontologies (as Cyc and BFO). The reader can be refer to [29,34] for details.

3 Related work

This section discusses works on aligning domain and top-level ontologies, Wordnet as background knowledge in the matching task, and word embeddings.

Domain and top-level ontology matching. We see a growing importance of aligning domain and top-level ontologies. In [26], correspondences between DBPedia ontology and DOLCE-Zero [4] are used to identify inconsistent statements in DBPedia. In that sense, in [18], a domain ontology describing web services (OWL-S) is manually aligned to DOLCE-Lite-Plus, in order to overcome conceptual ambiguity, poor axiomatization, loose design and narrow scope of the domain ontology. In [35] an alignment between an upper ontology (BFO) and a biomedical ontology (GO) is used for filtering out correspondences at domain level that relate two different kinds of ontology entities. Analysing the impact of using top ontologies as semantic bridges has been done in [16], where a set of algorithms exploiting such bridges are applied and the circumstances under which upper ontologies improve matching approaches are studied. A close approach to ours in terms of data set has been proposed by [25], where OAEI Conference ontologies were manually aligned to UFO, adopting a set of patterns grounded by UFO ontology. There are also works concerning alignment between different top ontologies. In [10,11], the ROMULUS repository aims at improving semantic interoperability between foundational ontologies (DOLCE, BFO and GFO), which are aligned with each other in a semi-automatic way using available matching tools, whose results have been manually evaluated. While these proposals mainly generate manual alignments between top level and domain ontologies, here we propose an approach to automatise this task. A preliminar study is presented in [31].

WordNet as a resource to ontology matching and contexts. Background knowledge from resources such as WordNet has been largely exploited in ontology matching. In [12], a lexical measure considers aggregating sets including names of ontology entities and WordNet synset's words (including hypernyms and meronyms relations). In [38], a set of twelve element-level matchers using WordNet as background knowledge is proposed. The use of WordNet is frequently coupled with the notion of context. In [30], virtual documents (context) represent the meaning of ontology entities and WordNet entries and entities are coupled according to their document similarities. The notion of context has also been exploited in [37], where *semantic description documents* refer to the information about concept hierarchies, related properties and instances, or in [2] where a bag of words describing a concept is exploited within a mining approach. On the other hand, the use of context is very common in the Word Sense Disambiguation, which can be carried out using a diversity of approaches [21]. Here, we adapt the [13] Word Sense Disambiguation to the task of synset disambiguation.

Word embeddings in ontology matching. Word embedding has been largely adopted in several tasks of NLP [19]. It is an umbrella name for a set of NLP language modelling and feature learning techniques which represent words as vectors in a semantic space. Models are trained to produce a vector space and reconstruct the linguistic contexts of words. Each unique word in the corpus is assigned a corresponding vector in the space. Word vectors are positioned in the vector space such that words that share common contexts in the corpus are located in close proximity to one another in the space. The similarity between words is calculated using functions as the cosine similarity, Euclidean distance. Such approach represents an alternative to WordNet similarities, which may fail due to the low WordNet coverage of specific domains. To the best of our knowledge, few works have exploited word embeddings in ontology matching [40,36]. In [40], a hybrid approach combines word embeddings and lexical similarities. The performance of edit distance, WordNet, Latent Semantic Analysis (LSA), word embeddings (using Wikipedia Word2Vec trained model) and the hybrid method were compared, showing that the performance of the hybrid method outperforms the others. In [36], the approach relies on word-to-word similarities exploiting the GloVe model. The hypothesis is that two entities can be matched based on the words in their names using the word-to-word similarity provided by the model. Close to [40], but for a different task, we combine WordNet and word embeddings.

4 Our approach

Our matching approach has two main steps. The first step disambiguates the domain concept, selecting the most appropriated WordNet synset; and the second matches the domain concept to the top-level concept via existing correspondences between WordNet and the top-level ontologies, as detailed below.

4.1 Synset disambiguation

In order to select the synset that better expresses the meaning behind the ontology concept, we adopt the notion of *context*. A *context* is constructed from all information available about an ontology entity, including entity naming (ID), annotation properties (usually labels and comments) and information on the neighbours (super and sub-concepts). Given $Sup(e)$ and $Sub(e)$, the sets of terms denoting the super-concepts and sub-concepts of the entity e, and $Ann(e)$ the set of terms from its annotations, a naive strategy for building a *context* (*context*) considers these sets as a *bag of words* :

$$context(e) = \{e, w | w \in Sup(e) \cup w \in Sub(e) \cup w \in Ann(e)\}$$

This context is used to find the closer synset using two strategies, as above.

Lesk measure The Lesk measure for word sense disambiguation [13] relies on the calculation of the word overlap between the sense definitions of two or more target words. Given a word w, it identifies the sense of w whose textual definition has the highest overlap with the words in the context of w :

$$score_{Lesk}(S) = |context_{Lesk}(w) \cap gloss(S)|$$

where $context_{Lesk}(w)$ is the bag of all content words in a context window around the target word w. Here, we overlap the $context_e$ with the context of each WordNet synset :

$$context(synset) = \{w | w \in Terms(synset) \cap w \in Gloss(synset)\}$$

where $Terms(synset)$ the set of terms in a *synset* and $Gloss(synset)$ the corresponding set of terms from the gloss (i.e, textual description containing definitions and examples) associated to the synset. We hence retrieve the highest overlap between $context(e)$ and $context(synset)$

$$score'_{Lesk}(e) = |context(e) \cap context(synset)|$$

Word embeddings The second similarity measure compares contexts of entities $context(e)$ and of WordNet synsets $context(synset)$ (represented as vectors of words). The comparison is based on the distance of contexts in vector spaces. This method adopts the cosine distance between two words generated by the word embedding model to identify the similarity between them. We retrieve the similarity between $context(e)$ and $context(synset)$, then we calculate the average similarity. After calculating this average to all elements of the context, we calculate the average of the context, considering the context length. The synset with the higher average is selected.

4.2 Identification of correspondences to top-level ontologies

In this step we perform the identification of the top concept. This step relies on the representation of the given existing alignments.

DOLCE correspondence identification This step uses existing alignments between DOLCE-Lite-Plus and WordNet 1.6. For each concept of the domain ontology, we use the selected synset (step 1) to identify the corresponding concept in OntoWordNet. To select the concept in OntoWordNet we compare the WordNet synset with each concept c in OntoWordNet (recall that concepts are represented by the concatenation of words). A bag of words for the OntoWordNet concept is created from the concatenated words and gloss, i.e., $context(c)$. Then, we overlap the synset and c.

$$score'_{Lesk}(c) = |context(c) \cap context(synset)|$$

After finding the OntoWordNet concept c corresponding to the synset, the higher level concept h^c of c is retrieved,. h^c corresponds to the DOLCE concept (Figure 1).

SUMO correspondence identification Similarly to the correspondence identification in DOLCE, this step uses existing alignments between SUMO and WordNet 1.6, in order to identify the domain and top concepts correspondences. As SUMO-WordNet alignment is a file containing the synset ID, terms, gloss, and the alignment to top concept (§2.2), we search for the domain selected synset in this file and, if the synset is found, we match the domain concept with the top-level concept related to the synset.

As described above, our approach depends on the availability of alignments between the background knowledge resource (here, WordNet) and the top-level ontologies. Hence, we are able to exploit other top-level ontologies in case such alignments exist. This leads also to the question on the maintenance of these alignments with the evolution of the ontologies and the given resource, which is out of the scope of this work.

5 Experiments

5.1 Material and methods

Domain ontologies We consider a set of ontologies from three different domains. First, SSN (W3C Semantic Sensor Network Ontology) describes sensors, devices, observations, measurements and other terms, enabling reasoning of individual sensors and the connection of them. A recent version of SSN includes a lightweight core ontology called SOSA (Sensor, Observation, Sample, and Actuators). SSN is aligned to SOSA and both ontologies are aligned to DOLCE Ultralite (DUL). SSN is composed of 18 first level concepts, from those, 8 concepts are aligned to the top ontology DUL. CORA (Core Ontology for Robotic and Automation) [28] specifies the main concepts, relations, and axioms of robotics and automation domains. Second, CORA is aligned with SUMO top ontology. CORA ontology, considering all its modules (CoraX, Cora, RParts, and POS) is composed of 34 first level concepts, from which 29 of them are aligned to SUMO. Finally, seven ontologies from the OAEI Conference data set[5] have been used (Cmt, ConfTool, Edas, Ekaw, Iasted, Sigkdd, SofSem). These ontologies are involved in reference alignments. These ontologies sum up 501 concepts, however, we consider in our

[5] http://oaei.ontologymatching.org/2017/conference/index.html

experiments the first-level concept of the hierarchies, what corresponds to 70 concepts (assuming that the other concepts will inherit their alignment with top ontologies from their roots). The choice for these ontologies is motivated by the fact that they are either widely adopted in real world scenarios or in experiments regarding automatic ontology matching approaches.

WordNet top-level alignments We use DOLCE, SUMO, and existing WordNet to top-level ontology alignments (§2). These previous alignments have been developed by specialists, hence if the selected synset is correct, the top-level concept (aligned as super-concept of that synset) is assumed to be a super-concept of the domain concept.

Word embedding models We used pre-trained models, GloVe [27] and GoogleNews[6]. GloVe is an *unsupervised learning algorithm to obtain vector representations for words*[7]. The training phase uses the Wikipedia 2014 and Gigaword5 corpora. It has 6 billions tokens, 400 thousand vocabulary size and neural network dimension of 200. The GoogleNews model is trained on part of Google News dataset (about 100 billion words). The model contains 300-dimensional vectors for 3 million words and phrases.

OAEI 2017 tools Our baseline corresponds to the results of a set of matching tools participating in OAEI 2017, with exception only of those specialised in instance matching (Legato, I-match and njuLink) and one specialised in the bio domain (Yam-bio). The matchers that were tested in our experiment are: ALIN, AML, CroLOM, KEPLER, LogMap, LogMap-Lite, ONTMAT, POMap, SANOM, WikiV3, WikiMatch and XMap. The reader can refer to OAEI papers[8] for a detailed description of them. All tools were run with their default configuration settings. All generated correspondences are available in `https://github.com/danielasch/top-match`.

5.2 Results and discussion

We run our system with the Lesk similarity (*lesk*) and word embedding models (*WE-GloVe and WE-GloogleNews*) and the OAEI tools for 16 matching tasks (SSN and DLP, CORA and SUMO, and 7 Conference ontologies with DLP and SUMO). All alignments generated by our approach are available online[9]. They have been evaluated in terms of precision and recall. With respect to the reference alignments, for the pairs involving SSN and CORA, given that these ontologies are already aligned to the top ontologies, we adopt these existing alignments as reference. We note that SSN is originally aligned with a different version of DOLCE. We hence consider the results in an interpreted way which consists at looking each generated correspondence and identify if they are the exact correspondence or related to the previous alignment via a subsumption relation. In the same way, we observe that some found correspondences from CORA and SUMO, were not exact the same of the adopted reference, however, they are hierarchically related, hence, we also adopted the interpreted evaluation.

For the Conference data set, which is not equipped with reference alignments to DOLCE and SUMO, the generated correspondences were manually evaluated by three

[6] https://code.google.com/archive/p/word2vec/
[7] https://nlp.stanford.edu/projects/glove/
[8] http://www.om2017.ontologymatching.org/#ap
[9] https://github.com/danielasch/top-match

specialists. Firstly, one evaluator analysed each correspondence, after, the results were discussed with all evaluators, maintaining or changing the initial analysis. For this data set, we made the hypothesis that, for each top domain concept, a corresponding WordNet synset exists. Hence, we are able to compute both precision and recall.

As shown in Figure 2, the best results were obtained for the conference domain, with .80 of F-measure with WE-GlooglеNews. We observe that overall WE-GloogleNews performs better than Lesk and WE-GloVe. However, looking at the SSN and CORA domain ontologies, the obtained results are lower than for Conference. Our hypothesis is that concepts from the conference ontology are more general (common sense) than these other domains. Note that the selected word embedding models were trained with general domain texts. The better performance obtained with the WE-GoogleNews model over the WE-GloVe model could be explained by the larger coverage of the first with respect to the training set.

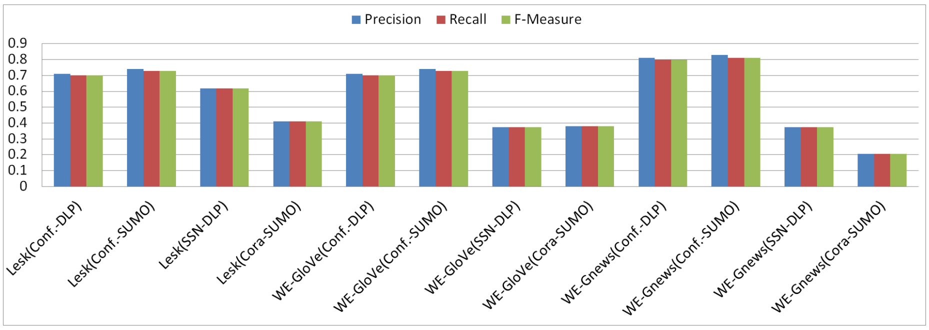

Fig. 2. Precision, recall and F-measure for each synset disambiguation strategy.

Regarding the number of correspondences, our approach was able to find 69 out of 70 correspondences from the Conference ontologies (we were not able to find the correspondences for 1 concept, for which there is no entry in WordNet) Considering Lesk and WE-GloVe, 51 correct correspondences were found when aligned to SUMO and 49 correct with DOLCE. This number increased up when using WE-GloogleNews (57 and 56, respectively). For SSN-DOLCE, we have 5 correct correspondences out of 8 considering Lesk, and 3 correct with WE-GloVe and WE-GloogleNews. For CORA-SUMO, 12 correct in a total of 29 correspondences considering Lesk, 11 correct with WE-GloVe and 6 correct with WE-GloogleNews.

Although our approach was able to found a high number of correspondences for the three domains, in some cases, the generated correspondences were wrong. First, as we adopt the context of concepts, this seems not to be enough to disambiguate the sense of the domain concept (Conference domain ontologies are not equipped of comments and labels). This can be improved by enriching the terminological layer. Second, we can observe that word embedding based on Google News model contributes to the disambiguation step, mainly with the Conference ontologies. However, for SSN and CORA it is still not able to retrieve the right synsets. In order to overcome this weaknesses, one direction is to use domain-specific embedding models. Third, the word sense disambiguation here is still based on the overlapping of words, and word sense disambiguation techniques could be used instead.

OAEI 2017 matching tools Only 4 tools (AML, LogMap, LogMapLite, and POMap) were able to find correspondences for 6 pairs of ontologies. Considering the correspondences found by these tools, 13 domain concepts from conference (out of 70) were aligned. Regarding the number of correspondences, *AML* was able to find 12 correspondences, and 7 of them were correct. *POMap* found 7 correspondences, and 6 were correct. *LogMap* and *LogMapLite* found 6 correspondences respectively, and 5 of them were correct. Figure 3 presents precision, recall and F-measure for each tool (including our evaluated techniques). Related to CORA, 1 correspondence was correctly found by *POMap*. As shown, our approach outperforms all system in terms of Recall and F-measure. Looking at *WE-GloogleNews*, the results are quite similar in terms of precision and better than all in terms of recall and F-measure. As somehow expected, while the tools perform well in terms of precision, they retrieve a limited number of correspondences.

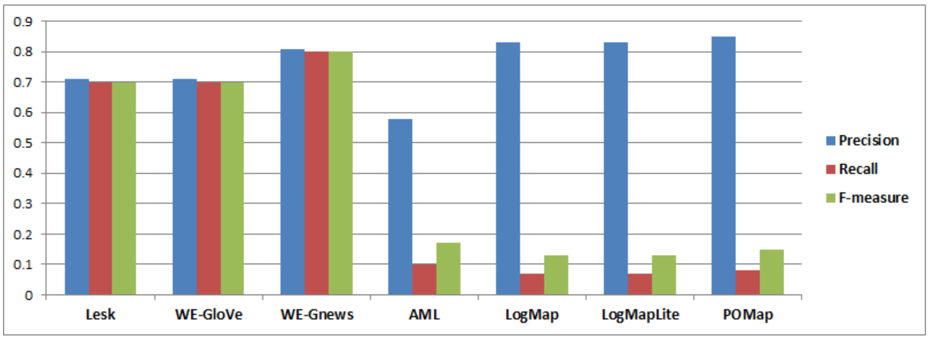

Fig. 3. Precision, recall and F-measure from each matching tool.

6 Concluding remarks and future work

This paper presented an approach to match domain and top-level ontologies, exploiting alignments between WordNet and top ontologies. Our evaluation was based on ontologies from three domains with DOLCE and SUMO top-level ontologies. Overall, we consider that existing top-level and WordNet alignments is a valuable resource for the task, at least for certain general domains. For most of the concepts from the domain ontologies we found a correspondence with the top ontology. We have evaluated OAEI matching tools in this task and, as expected, our approach outperforms all of them. Even though they were not exactly developed for that purpose, their results were the only available for comparison, and we set that as a baseline. To the best of our knowledge, our approach is the first attempt towards automatizing the process of aligning top and domain ontologies. As future work, we plan to provide a reference alignment involving the OAEI Conference dataset and DOLCE and SUMO ontologies with the aim of proposing a OAEI track for this task involving top and domain ontologies. We plan as well to combine Wordnet measures with other distributional semantics approaches and adopt other background knowledge resources as BabelNet.

This work will be published as part of the book "Emerging Topics in Semantic Technologies". ISWC 2018 Satellite Events. E. Demidova, A.J. Zaveri, E. Simperl (Eds.), ISBN: 978-3-89838-736-1, 2018, AKA Verlag Berlin.

References

1. Brodaric, B., Probst, F.: DOLCE ROCKS: Integrating Geoscience Ontologies with DOLCE. In: Semantic Scientific Knowledge Integration. pp. 3–8 (2008)
2. David, J.: AROMA results for OAEI 2011. In: Proc. of the 6th Intern. Workshop on Ontology Matching. pp. 122–125 (2011)
3. Euzenat, J., Shvaiko, P.: Ontology Matching. Springer-Verlag, 1st edn. (2007)
4. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A.: Sweetening WORDNET with DOLCE. AI Magazine 24(3), 13–24 (2003)
5. Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., Schneider, L.: Sweetening Ontologies with DOLCE. In: Proc. of the 13th Intern. Conf. on Knowledge Engineering and Knowledge Management. pp. 166–181 (2002)
6. Gangemi, A., Navigli, R., Velardi, P.: The OntoWordNet Project: Extension and Axiomatization of Conceptual Relations in WordNet, pp. 820–838. Springer Berlin Heidelberg (2003)
7. Guarino, N.: Formal Ontology in Information Systems: Proc. of the 1st Intern. Conference. IOS Press, 1st edn. (1998)
8. Jung, J.J., Håkansson, A., Hartung, R.: Indirect Alignment between Multilingual Ontologies: A Case Study of Korean and Swedish Ontologies. In: Agent and Multi-Agent Systems: Technologies and Applications: 3rd KES Intern. Symposium. pp. 233–241 (2009)
9. Kachroudi, M., Chelbi, A., Souid, H., Ben Yahia, S.: Fiona: A framework for indirect ontology alignment. In: Foundations of Intelligent Systems. pp. 224–229 (2015)
10. Khan, Z., Keet, C.M.: The Foundational Ontology Library ROMULUS. In: Proc. of the 3rd Conf. on Model and Data Engineering. pp. 200–211 (2013)
11. Khan, Z., Keet, C.M.: Toward semantic interoperability with aligned foundational ontologies in ROMULUS. In: Proc. of the 7th Conf. on Knowledge Capture. pp. 23–26 (2013)
12. Kwak, J., Yong, H.: Ontology Matching based on Hypernym, Hyponym, Holonym and Meronym Sets in WordNet. Web and Semantic Technology 1(2), 1–14 (2010)
13. Lesk, M.: Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from an ice cream cone. In: Proc. of the 5th Intern. Conf. on Systems Documentation. pp. 24–26 (1986)
14. Lin, F., Sandkuhl, K.: A Survey of Exploiting WordNet in Ontology Matching. In: Proc. of the 20th Intern. Federation for Information Processing, pp. 341–350 (2008)
15. Mascardi, V., Cordì, V., Rosso, P.: A Comparison of Upper Ontologies. In: Proc. of the 8th AI*IA/TABOO Joint Workshop on Agents and Industry. pp. 55–64 (2007)
16. Mascardi, V., Locoro, A., Rosso, P.: Automatic Ontology Matching via Upper Ontologies: A Systematic Evaluation. IEEE Trans. on Knowl. and Data Eng. 22(5), 609–623 (2010)
17. Masolo, C., Borgo, S., Gangemi, A., Guarino, N., Oltramari, A.: WonderWeb deliverable D18 ontology library (final). Tech. rep., IST Project 2001-33052 WonderWeb: Ontology Infrastructure for the Semantic Web (2003)
18. Mika, P., Oberle, D., Gangemi, A., Sabou, M.: Foundations for Service Ontologies: Aligning OWL-S to Dolce. In: Proc. of the 13th Conf. on World Wide Web. pp. 563–572 (2004)

19. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. CoRR (2013)
20. Miller, G.A.: WordNet: A Lexical Database for English. Commun. ACM 38(11), 39–41 (1995)
21. Navigli, R.: Word Sense Disambiguation: A Survey. ACM Comput. Surv. 41(2), 10:1–10:69 (2009)
22. Niles, I., Pease, A.: Towards a Standard Upper Ontology. In: Proc. of the Intern. Conf. on Formal Ontology in Information Systems. pp. 2–9 (2001)
23. Niles, I., Pease, A.: Linking Lexicons and Ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In: Proc. of the Intern. Conf. on Information and Knowledge Engineering. pp. 412–416 (2003)
24. Oberle, D., Ankolekar, A., Hitzler, P., et al.: DOLCE ergo SUMO: On foundational and domain models in the SmartWeb Integrated Ontology (SWIntO). Web Semantics: Science, Services and Agents on the World Wide Web 5(3) (2007)
25. Padilha, N.F., Baião, F., Revoredo, K.: Alignment Patterns based on Unified Foundational Ontology. In: Proc. of the Brazilian Ontology Research Seminar. pp. 48–59 (2012)
26. Paulheim, H., Gangemi, A.: Serving DBpedia with DOLCE - More than Just Adding a Cherry on Top. In: The Semantic Web, pp. 180–196 (2015)
27. Pennington, J., Socher, R., Manning, C.D.: Glove: Global vectors for word representation. In: Empirical Methods in Natural Language Processing. pp. 1532–1543 (2014)
28. Prestes, E., Carbonera, J.L., Fiorini, S.R., et al.: Towards a core ontology for robotics and automation. Robotics and Autonomous Systems 61(11), 1193 – 1204 (2013)
29. Reed, S., Lenat, D.: Mapping Ontologies into Cyc. In: Proc. of the Workshop on Ontologies For The Semantic Web. pp. 1–6 (2002)
30. Schadd, F.C., Roos, N.: Coupling of Wordnet Entries for Ontology Mapping Using Virtual Documents. In: Proc. of the 7th Conf. on Ontology Matching. pp. 25–36 (2012)
31. Schmidt, D., Basso, R., Trojahn, C., Vieira, R.: Matching Domain and Top-level Ontologies via OntoWordNet. In: Workshop on Ontology Matching. pp. 1–2 (2017)
32. Schmidt, D., Trojahn, C., Vieira, R.: Analysing Top-level and Domain Ontology Alignments from Matching Systems. In: Workshop on Ontology Matching. pp. 1–12 (2016)
33. Semy, S., Pulvermacher, M., Obrst, L.: Toward the use of an upper ontology for U.S. government and U.S. military domains: An evaluation. Tech. rep., Sub. to Workshop on Information Integration on the Web, in conjunction with VLDB-2004 (2004)
34. Seppala, S.: Mapping WordNet to Basic Formal Ontology using the KYOTO ontology. In: Proc. of the Intern. Conf. on Biomedical Ontology. pp. 1–2 (2015)
35. Silva, V., Campos, M., Silva, J., Cavalcanti, M.: An Approach for the Alignment of Biomedical Ontologies based on Foundational Ontologies. Information and Data Management 2(3), 557–572 (2011)
36. Vieira, R., Revoredo, K.: Using word semantics on entity names for correspondence set generation. In: Workshop on Ontology Matching. pp. 223–224 (2017)
37. Wang, P.: Lily Results on SEALS Platform for OAEI 2011. In: Proc. of the 6th Intern. Workshop on Ontology Matching. pp. 156–162 (2011)
38. Yatskevich, M., Giunchiglia, F.: Element level semantic matching using WordNet. In: Meaning Coordination and Negotiation Workshop. pp. 37–48 (2004)
39. Zhang, S., Bodenreider, O.: Alignment of multiple ontologies of anatomy: Deriving indirect mappings from direct mappings to a reference. In: American Medical Informatics Association Annual Symposium (2005)
40. Zhang, Y., Wang, X., Lai, S., He, S., Liu, K., Zhao, J., Lv, X.: Ontology matching with word embeddings. In: Sun, M., Liu, Y., Zhao, J. (eds.) Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data. pp. 34–45 (2014)