Validating Top-level and Domain Ontology Alignments using WordNet

Daniela Schmidt*, Cassia Trojahn†, Renata Vieira*, Mouna Kamel†

*Pontifical Catholic University of Rio Grande do Sul (Brazil)
daniela.schmidt@acad.pucrs.br, renata.vieira@pucrs.br
† Institut de Recherche en Informatique de Toulouse (France)
{cassia.trojahn, mouna.kamel}@irit.fr

Abstract. *Matching domain and top-level ontologies is an important task but still an open problem in the ontology matching field, particularly due to their different levels of abstraction. Beyond that, validating candidate alignments is crucial before exploiting them within ontology construction and integration processes involving such kinds of ontologies. This paper concerns the automatic validation of candidate alignments between top and domain ontologies, exploiting WordNet background knowledge. We apply our approach for validating alignments generated by a set of matching tools and discuss the results against manual validation. Our experiments are based on one domain ontology and three well-known top-level ontologies.*

1. Introduction

According to their "level of generality", ontologies can be classified in different categories [Guarino 1998]. *Top-level ontologies* describe very general concepts (e.g., space, time, object, etc.), which are independent of a particular domain, and are usually equipped with a rich axiomatic layer; *domain ontologies* describe the entities related to a domain (e.g., biology or aeronautic). While the rich semantics and formalization of top-level ontologies are important requirements for ontology building [Mika et al. 2004], they act as well as semantic bridges supporting very broad semantic interoperability between domain ontologies [Mascardi et al. 2010]. For grounding domain concepts, one is required to find relationships with concepts from a top-level ontology, i.e., to match them.

There have been many proposals to the matching task in the literature [Euzenat and Shvaiko 2007], which exploit in different ways the knowledge encoded within each ontology when identifying correspondences between their features or structures. However, most efforts have been particularly dedicated to domain ontologies and the problem of matching domain and top-level ontologies has been addressed to a lesser extent. Matching domain and top-level ontologies poses different challenges, in particular due to their different levels of abstraction. This is a complex task, even manually, that requires to deeply identify the semantic context of concepts. It involves going beyond the frontiers of the knowledge encoded in the ontologies and, in particular, the identification of subsumption relations. The latter is in fact largely neglect by most matchers. Beyond that, the validation of alignments is a crucial step before exploiting them within ontology building and integration processes. While manual validation is a time-consuming task requiring expert judges, automatic ones generally rely on gold standards, which are not always available or may suffer of imperfections or low coverage. Exploiting background knowledge can be an alternative. However, knowledge on top level ontologies is highly specialized and requires resources covering different levels of abstraction.

This paper concerns the automatic validation of candidate alignments between top-level and domain ontologies, relying on background knowledge from WordNet [Miller 1995] and on the notion of *context*, which has been exploited in different ways in ontology matching [Maedche and Staab 2002, Wang 2011, Schadd and Roos 2012, Djeddi and Khadir 2014]. Contexts are constructed from all available information about an ontology entity, including entity naming (ID), annotation properties (labels and comments) and information on the neighbors of entities (super and sub-concepts). They are used for disambiguating the senses that better express the meaning of ontology entities, before looking for relations between these senses in WordNet in order to validate or not a correspondence between those entities. Although our approach can be applied for validating any kind of alignment (i.e., domain ontology alignments), it is particularly suitable for validating alignments which are rich in subsumption relations, what is typically the case when matching domain and top-level ontologies, since it takes advantage of the network of conceptual-semantic and lexical relations in WordNet (in particular, hypernym relations). Furthermore, although this approach can be seen as a matching approach itself, our motivation here is to automatically validate alignments in the lack of gold standards. We apply our approach for validating alignments generated by a set of matching tools (see more details in [Schmidt et al. 2016]) and discuss the results against a manual validation. Our experiments are based on three well-known top-level ontologies (DOLCE, GFO, and SUMO) and one domain ontology from one of the most used data set in the Ontology Alignment Evaluation Initiative (OAEI) campaigns (the Conference data set). The rest of the paper is organised as follows. §2 introduces WordNet structure, top-level ontologies and ontology matching. §3 describes the validation approach. §4 presents the experiments. §5 discusses related work and, finally, §6 concludes the paper.

2. Background

2.1. WordNet

WordNet [Miller 1995] is a general-purpose large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept. Synsets are interlinked by means of conceptual-semantic and lexical relations. A synset denotes a concept or a sense of a group of terms. WordNet also provides textual descriptions of the concepts (gloss) containing definitions and examples. For instance, for the concept "Poster", one of the associated WordNet synsets (SID-06793426-N) groups the synonyms "poster, posting, placard, notice, bill, card", together with a gloss "a sign posted in a public place as an advertisement; a poster advertised the coming attractions".

2.2. 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 expressivity for a wide range of domains. It is often characterized as representing common sense concepts and concerns concepts which are meta, generic, abstract and philosophical. There are two approaches for the use of top-level ontologies. The top-down approach uses the ontology as a foundation for deriving concepts in the domain ontology, taking advantage of the knowledge and experience already expressed in the top-level ontology. In a bottom-up approach, one usually matches a new or existing domain ontology to a top-level ontology. This paper concerns bottom-up approaches. Several top-level ontologies have been proposed in the literature. The reader can refer to [Mascardi et al. 2007] for a review of them. Here, we briefly introduce some well-known and largely used top-level ontologies which are used further in our experiments:

- DOLCE [Gangemi et al. 2002]: Descriptive Ontology for Linguistic and Cognitive Engineering has been proposed at LOA (Laboratory for Applied Ontology). Its focus is to grasp the underlying categories of human cognitive tasks and the socio-cultural environment. It is an ontology of particulars and includes concepts such as abstract quality, abstract region, physical object, process, and so on. Here, we consider the version DOLCE Lite[1], which is composed by 37 concepts and 70 object properties.
- GFO [Herre et al. 2007]: General Formal Ontology is a top-level ontology for conceptual modeling that has been proposed by the Onto-Med Research Group. It includes elaborations of categories such as objects, processes, time and space, properties, relations, roles, functions, facts, and situations. The version we consider here, GFO Basic[2], is composed by 45 Concepts and 41 Object properties.
- SUMO [Niles and Pease 2001]: Suggested Upper Merged Ontology has been proposed by The IEEE Standard Upper Ontology Working Group. It provides definitions for general-purpose terms and acts as a foundation for more specific domain ontologies and is used for research and applications in search, linguistics and reasoning. We consider the version SUMO[3] which is composed by 4.558 concepts and 778 object properties.

2.3. Ontology matching

The process of finding correspondences between ontology entities is known as ontology matching. It takes as input two ontologies o_s (source) and o_t (target) and an (possibly empty) alignment A to be completed, and determines as output an alignment A', i.e., a set of correspondences. Here, we borrow the definition of correspondence from [Euzenat and Shvaiko 2007]:

Definition 1 (Correspondence) *A correspondence can be defined as* $<e_s,e_t,r,n>$*, such that:* e_s *and* e_t *are entities (e.g., concepts, properties, instances) of* o_s *and* o_t*, respectively; r is a relation holding between two entities* e_s *and* e_t*, (for instance,* equivalence, subsumption, disjointness, overlapping*); and n is a confidence measure number in the [0;1] range. The confidence assigns a degree of trust on the correspondence from the matcher.*

Different matching approaches have emerged from the literature [Kalfoglou and Schorlemmer 2003, Euzenat and Shvaiko 2007]. While *terminological* methods lexically compare strings (tokens or n-grams) used in naming entities (or in the labels and comments of entities), *semantic* methods utilize model-theoretic semantics to determine whether or not a correspondence exists between 2 entities. Approaches may consider the *internal* ontological structure, such as the range of their properties, their cardinality, and their transitivity and/or symmetry features, or alternatively the *external*

[1] http://www.loa.istc.cnr.it/old/DOLCE.html
[2] http://onto.eva.mpg.de/gfo-bio/gfo-bio.owl
[3] http://www.adampease.org/OP/SUMO.owl

ontological structure, such as the position of the 2 entities within the ontological hierarchy. The instances of concepts can also be compared using statistical, probabilistic and linguistic approaches. Moreover, ontology matching systems rely on several approaches.

3. Validation approach

Our validation approach is based on the intuition that a correspondence c from an alignment A, expressing a relation r (in particular, *equivalence* or *sumbumption*) between the entities e_d and e_t of a domain ontology o_d and a top-level ontology o_t, respectively, may exist as a semantic relation in WordNet (synonymy or hypernymy). Given e_d and e_t concepts in the ontologies, we have first to retrieve the most appropriated WordNet synsets and then check if there exists a relation r between such synsets. Our approach thus is based on two main steps : (i) *synset disambiguation* and (ii) *synset relation identification*.

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 neighbors (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* (*ctx*) considers these sets as a *bag of words*. In a pre-processing phase, the annotations are tokenized and stop words and special characters are removed. Hence, the context ctx_e of an entity e is a bag of words from the pre-processed sets Sup_e, Sub_e and Ann_e, including e itself (i.e., its ID) :

$$ctx_e = \{e, w \mid w \in Sup_e \cup w \in Sub_e \cup w \in Ann_e\}$$

We take ctx_e for retrieving the closer WordNet synset with respect to the meaning of e. For that, we overlap ctx_e with the context of each retrieved $synset_i^e$ when looking for e in WordNet, and select the set with higher intersection. We compute the number of words occurring in both sets, based on the exact match between words. Given $Terms_i^e$ the set of terms in a $synset_i^e$ and $Gloss_i^e$ the corresponding set of terms from the gloss (i.e, textual description containing definitions and examples) associated to the synset, the context $ctx_{synset_i^e}$ of the $synset_i^e$ is defined as follows:

$$ctx_{synset_i^e} = \{e, w \mid w \in Terms_i^e \cup w \in Gloss_i^e\}$$

The overlap between ctx_e and $ctx_{synset_i^e}$ is calculated as follows (where n is the number of synsets of e):

$$overlap(ctx_e, ctx_{synset_i^e}) = \max_{1 \leq i \leq n} |ctx_e \cap ctx_{synset_i^e}|$$

Synset relation identification. Having selected $synset_i^{e_d}$ and $synset_i^{e_t}$, the closest WordNet synsets expressing, respectively, the meaning of the entities e_d and e_t, we verify if $synset_i^{e_d} = synset_i^{e_t}$. In this case, we assume that e_d and e_t are equivalent (r = equivalence). Otherwise, we construct the set $Super\_synset_i^{e_d}$ which is composed of all the hypernym synsets of $synset_i^{e_d}$ of rank k (k being the maximum length of the path from $synset_i^{e_d}$ to one of its hypernym synsets in WordNet, based on a depth-first search strategy). The intuition here is that if the concept in the top-level ontology belongs to the set of hypernyms of the domain concept, then a subsumption relation exists between them. Hence, $synset_i^{e_d}$ subsumes $synset_i^{e_t}$ if $synset_i^{e_t}$ belongs to the set $Super\_synset_i^{e_d}$ (r = subsumption).

Example. Given <Presentation$_{Conference}$,Event$_{GFO}$ >, one correspondence to be validated (taken from the ontologies used in our experiments) and the sets:

- $Sup_{Presentation}$ = {conference, document, contribution, invited, talk} (from the original set {Conference_document, Conference_contribution, Invited_talk});
- $Sub_{Presentation}$ = ∅ (the domain concept has no sub-concepts);
- $Ann_{Presentation}$ = ∅ (the domain concept has no annotations);
- Sup_{Event} = {individual, concrete, processual, structure, occurrent} (from the original set {individual, concrete, processual_structure, occurrent});
- Sub_{Event} = ∅ (the domain concept has no sub-concepts);
- Ann_{Event} = {event, processual, structure, process} (from the annotation "Events are processual structures comprising a process ...");

The following *contexts* can be constructed:

- $ctx_{presentation}$ = {presentation, conference, document, contribution, invited, talk}
- ctx_{event} = {event, individual, concrete, processual, structure, occurrent, process}

Overlapping the context of the concepts with all their WordNet synsets, the closer synsets are = {SID-01048697-N (presentation)} and {SID-00029378-N (event)}. As they are different synsets, the equivalence between them can not be verified. Hence, we look for the *Super_synset* of the concept presentation (from domain ontology) which is = {SID-01048697-N (presentation), SID-01027379-N (ceremony), SID-00407535-N (activity), SID-00030358-N (act, deed, human_action, human_activity), SID-00029378-N (event)}. As there is an overlap between this *Super_synset* and the synset for event (SID-00029378-N (event)) then we assume that Presentation$_{Conference}$ subsumes Event$_{GFO}$.

4. Experiments

4.1. Data set and matchers

OAEI Conference data set. The OAEI Conference set[4] contains 16 ontologies covering the domain of conference organization. We have chosen this data set because it contains expressive ontologies and is one of the most popular data set in the field [Cheatham and Hitzler 2014]. In our experiments, we have used only the Conference ontology[5]. This ontology has 60 concepts, 46 object properties and 18 data properties.

Ontology matching tools. A set of tools, publicly available, from previous OAEI campaigns (not limited to Conference track top participants), and implementing different matching strategies was selected. Even though they are not exhaustive and were not exactly developed for that purpose, their output might help us to investigate the problem of aligning domain and top-level ontologies. Aroma[6] is a hybrid tool based on association rules; Falcon-AO[7] applies linguistic and structural approaches, as Lily[8], which includes debugging strategies; LogMap[9] applies logical reasoning and repair strategies and its variant LogMap-Lite is essentially based on string similarities; MaasMatch adopts a similar-

[4]http://oaei.ontologymatching.org
[5]http://oaei.ontologymatching.org/2015/conference/data/Conference.owl
[6]https://exmo.inrialpes.fr/software/aroma/
[7]http://ws.nju.edu.cn/falcon-ao/
[8]http://cse.seu.edu.cn/people/pwang/lily.htm
[9]https://www.cs.ox.ac.uk/isg/tools/LogMap/

ity cube and a disambiguation phase as described in [Schadd and Roos 2012]; WeSeE-Match[10] uses web search results for improving similarity measures; WikiMatch[11] uses Wikipedia as external knowledge source and YAM++[12] applies linguistic, graph-based and machine learning approaches. MaasMatch and YAM++ use WordNet as background knowledge (see §5). All the tools were run with their default configuration settings.

4.2. Results and discussion

Manual analysis. We first qualitatively analysed how the available matching tools perform when aligning domain with top-level ontologies, based on a manual validation of the generated alignments, by 3 evaluators. For our experiments, we ran each of the above mentioned systems for the pairs composed by the *Conference* ontology against each top-level ontology. We then merge the alignments generated by the matchers, resulting in 28 correspondences, and submitted the resulting merge to the analysis of three evaluators. The evaluators are researchers that have common-sense knowledge about the domain of conferences, with a strong background in Computer Science and well-familiarised with ontology matching. The evaluators analysed each correspondence and selected one type of relation (Equivalent, Subsumption or None), according to the relation they judged as correct, based on the concept description provided in the top-level ontology (see more details in [Schmidt et al. 2016]). Regarding the evaluators judgement, there was a total agreement among them in 21 (out of 28) correspondences. For 15 of them, no relation has been identified so that more than a half of the automatically aligned concepts were considered neither equivalent or subsumed. In 3 cases there was total agreement regarding subsumption, and in 3 cases total agreement for equivalence. For the 7 pairs, a partial agreement was achieved as subsumption (see Table 2, column 'Manual evaluation'). In the following, we quantitatively analyse the results obtained by the matchers, using precision and considering 4 different sets:

- P_1 considers the cases of total agreement among the judges (21 correspondences), where the correspondence has been marked either as equivalent or subsumed;
- P_2 considers both total and partial agreements (for equivalent or subsumed);
- P_3 considers total agreement for equivalences (matchers generated only equivalences);
- P_4 considers both total and partial agreements only for equivalences.

Table 1. Precision of systems for their complete alignments with the 3 top-level ontologies. Systems marked with * use WordNet in their matching strategies.

System	P_1	P_2	P_3	P_4
Aroma	0	.33	0	0
Falcon-AO	0	0	0	0
Lily	-	-	-	-
LopMap	.33	.55	.33	.27
LogMapLite	.33	.33	.33	.33
MaasMatch*	.30	.42	0	.08
WeSeE-Match	-	-	-	-
WikiMatch	-	-	-	-
Yam++*	.27	.38	.18	.15
Total	.29	.39	.19	.18

[10]http://www.ke.tu-darmstadt.de/resources/ontology-matching/wesee-match
[11]http://www.ke.tu-darmstadt.de/resources/ontology-matching/wikimatch
[12]http://www.lirmm.fr/yam-plus-plus/

Table 1 shows the matchers results for these 4 sets, in terms of precision. From 9 matchers, 2 of them (MaasMatch and YAM++) have used WordNet in their matching approaches (in particular, MassMatch with a disambiguation phase). As expected, the best results were obtained for P_2 for both systems. However, the best results for this set have been obtained by LogMap. Although, ideally, avoiding a bias in the evaluation is highly desirable and the background resources used in the validation process should not be used for validating correspondences generated on basis of them, our evaluation is not limited to matchers relying on WordNet. We have used a set of available matchers using different matching strategies and tried to maximize the testing set by merging their output. We can in fact observe that MassMatch and YAM++ were able to generate more correspondences than the other systems (with LogMap and its variant coming just behind). Contrary to what would be expected, these systems (and all others, in fact) were not able to generate subsumption (even though some have been designed to). They generated only equivalences, even when they were in fact subsumptions. This is actually an important distinction that our validation tries to identify.

Automatic validation. We have submitted the 28 automatically generated correspondences to the automatic validation and compared the results with the manual evaluation (Table 2). Although, by definition, top-level ontologies should be able to cover most domain concepts, the matchers were not able to generate correspondences for most of the concepts (42 in a total of 60). A total of 8 pairs were automatically validated. From these pairs, 5 were correspondences validated manually by the evaluators. While the automatic validation distinguishes relations of equivalence and subsumption, it is still not able to distinguish as much as the manual analysis. The comparison of the automatic with the manual validation is given in terms of precision, recall, f-measure, and accuracy, in Table 3. Since the matchers do not distinguish equivalences and subsumptions, when evaluating our automatic approach, P_3 and P_4 have not been considered.

Looking at the results in detail (Table 2), we see 5 pairs with agreement between automatic and manual validation:

- Equivalence: *Poster* and *Poster (SUMO)*, *Publisher* and *Publisher (SUMO)*;
- Subsumption: *Organizer* and *Organism (SUMO)*, *Presentation* and *Event (GFO)*, and *Person* and *Individual (GFO)*.

For the first 4 pairs above, the evaluators totally agreed on equivalence or subsumption relations and for the fifth pair, they partially agreed on subsumption. One of them has been generated by MassMatch (pair *Presentation* and *Event(GFO)*) and the other 4 by YAM++. In the automatically validated relations, the pairs involving an equivalence (*Poster* and *Publisher*), the same WordNet synset for both top-level and domain concepts was selected in the disambiguation step – respectively, SID-06793426-N (poster, posting, placard, notice, bill, card) and SID-08062623-N (publisher, publishing_house, publishing_firm, publishing_company). For subsumption relations, for the pair *Organizer* and *Organism(SUMO)*, the selected synsets were not the same – respectively, SID-10383237-N (organizer, organiser, arranger) and SID-00004475-N (organism, being) but the domain synset is related to the top-level synset via a hypernym relation (*Organizer* is-a *Organism*). The same was observed for the pairs *Presentation* and *Event*, and *Person* and *Individual*.

For the 3 pairs of concepts automatically validated but manually annotated as 'None', the concept *Abstract* appears in 2 pairs, and the third pair is related to the concept

Table 2. List of automatically generated correspondences and their manual and automatic validations. (PA) indicates a partial agreement between judges. In bold, the automatically validated pairs.

Domain concept	Top-level concept	Manual evaluation	Automatic evaluation
Poster	**Poster (SUMO)**	**Equivalence**	**Equivalence**
Publisher	**Publisher (SUMO)**	**Equivalence**	**Equivalence**
Organization	Organization (SUMO)	Equivalence	None
Organizer	**Organism (SUMO)**	**Subsumption**	**Subsumption**
Presentation	**Event (GFO)**	**Subsumption**	**Subsumption**
Registered_applicant	Entity (GFO)	Subsumption	No WN entry
Person	**Individual (GFO)**	**Subsumption (PA)**	**Subsumption**
Presentation	Proposition (DOLCE)	Subsumption (PA)	None
Topic	Concept (GFO)	Subsumption (PA)	None
Conference_part	Relevant-part (DOLCE)	Subsumption (PA)	No WN entry
Conference_proceedings	Process (DOLCE)	Subsumption (PA)	No WN entry
Conference_proceedings	Process (GFO)	Subsumption (PA)	No WN entry
Conference_volume	Particular (DOLCE)	Subsumption (PA)	No WN entry
Abstract	Abstract (GFO)	None	Equivalence
Abstract	Abstract (SUMO)	None	Equivalence
Workshop	Workshop (SUMO)	None	Equivalence
Abstract	Abstract (DOLCE)	None	None
Chair	Chair (SUMO)	None	None
Organizer	Organization (SUMO)	None	None
Paper	Paper (SUMO)	None	None
Poster	Property (GFO)	None	None
Topic	Time-interval (DOLCE)	None	None
Conference_applicant	Non-physical-object (DOLCE)	None	No WN entry
Conference_document	Particular (DOLCE)	None	No WN entry
Conference_part	Region (DOLCE)	None	No WN entry
Extended_abstract	Abstract-region (DOLCE)	None	No WN entry
Registered_applicant	Physical-object (DOLCE)	None	No WN entry
Workshop	Spatial-location_q (DOLCE)	None	No WN entry

Table 3. Precision, Recall, and F-measure of automatically validated pairs.

	P_1	P_2
Precision	.57	.63
Recall	.67	.38
F-Measure	.62	.48
Accuracy	.76	.61

Workshop. For the first pair involving *Abstract* (GFO), 2 contexts intersect both domain and top-level concept contexts with the same number of words. However, the wrong synset was selected for the domain concept. While *Abstract* in the domain ontology corresponds to the sense SID-06468951-N (outline, synopsis, abstract, precis) and *Abstract* in the top-level ontology to the sense SID-05854150-N (abstraction,abstract), the selected synset for the domain concept was SID-05854150-N (abstraction,abstract). This is due to the fact that the domain concept context available was not enough for the disambiguation step and intersects with both SID-05854150-N and SID-06468951-N, with an intersection set containing the same number of elements. A similar behaviour was observed for the second pair involving *Abstract(SUMO)* and again for the pair involving the concept *Workshop*. For the latter, the same synset was selected for both domain and top-level ontology, SID-04603081-N (workshop,shop). The concept *Abstract* appears once more in another correspondence, for DOLCE, but it was not validated automatically nor manually. For 11 pairs (out the 28), very specific terms are used to express either domain or top-level concepts, and there are no respective entries in WordNet. As a general purpose resource, it does not cover as well domain specialized vocabularies. These specialized terms are usually expressed by noun compounds and represent 39% of correspondences that could not be validated. These limitations could be addressed by extending the validation to additional external resources.

Among the 13 pairs annotated as correct by the evaluators, 5 of them were validated automatically; 5 others belong to very specific terms for which there are no entries in WordNet, and 3 were found in WordNet, but not automatically validated: *Organization* and *Organization (SUMO)*, *Presentation* and *Proposition (DOLCE)*, and *Topic* and *Concept (GFO)*. For *Organization* in the domain ontology, the synset SID-08008335-N (organization,organisation) was selected, which represents a group of people working together, but for the top-level, a wrong synset was selected, SID-08164585-N (administration, governance, governing_body, establishment, brass, organization, organisation). For the pair *Presentation* and *Proposition*, the context was not enough to identify the correct synset. While for the domain concept the synset SID-01048697-N (presentation) was selected, for the top-level the synset SID-06750804-N (proposition) has been chosen and there was no hierarchical relation between them. For the pair *Topic* and *Concept*, the synset SID-06599788-N (subject,topic,theme) and SID-05835747-N (concept,conception,construct) were selected and, no relation was found between them in WordNet.

We could also observe that the domain ontology is not terminologically rich enough (i.e., entities have no labels nor comments) for helping the disambiguation and selection of the appropriate synsets. In some cases, only the identifier of the concept is available, once it is a root node in the hierarchy. For instance, the domain concept *Organization*, which has 7 WordNet synsets, has no annotations and no neighbours in the ontology hierarchy. Furthermore, the inconsistencies introduced in the resulting alignments (merge) should be taken into account. For instance, it seems contradictory that *Conference_applicant* matches *Non_physical_object* and *Registered_applicant* to *Physical_object*, considered that the latter is a subclass of *Conference_applicant* in the domain ontology. Finally, we could enrich the generated set of alignments by introducing simple hierarchical reasoning (if *Conference_applicant* is a *Physical_object*, all its subclasses are as well). Thus, covering most domain concepts.

5. Related work

Background knowledge from external resources such as WordNet has been largely exploited in matching domain ontologies, as a way for improving similarity measures [Lin and Sandkuhl 2008]. In [Kwak and Yong 2010], a lexical similarity for ontology matching considers the rate of including names of ontology entities and WordNet synset's words in an aggregated set (including hypernyms, hyponyms, holonyms and meronyms relations). Similar to us, they overlap word sets built from the available information on ontology entities and synsets. However, we exploit the context of ontology entities, including their sub-concepts, super-concepts and annotations but are limited to hypernym relations. In [Yatskevich and Giunchiglia 2004], a set of twelve element-level matchers using WordNet as background knowledge is proposed. The authors distinguish knowledge-based matchers (which exploit the structural properties of WordNet, often combined with statistics from large scale corpora) and gloss-based matchers (which exploit WordNet synset descriptions). We also exploit the structural properties of WordNet (but focusing on hypernym relations) and the glosses of synsets. However, they do not exploit the notion of context of ontology entities. As in our case, both works distinguish equivalence and subsumption relations. None of the matchers here indicated subsumption (§4.2).

A closer approach to ours is proposed in [Schadd and Roos 2012], where virtual documents represent the meaning of ontology entities and WordNet entries and entities are coupled according to their document similarities. A document is constructed from entities naming and annotation properties and information on the neighbours of the ontology entity, as we do in this paper. However, we exploit the WordNet hypernym relations in another way (not for constituting the virtual document associated to a synset), but in the step of identifying the relation between synsets. As in their approach, ours is related to the field of Word Sense Disambiguation, which can be carried out using several different techniques [Navigli 2009]. Although it is possible to tackle this problem from various angles and relying on different similarity metrics, here we adopt a naive method based on the exact match of words in a context. The notion of context has also been exploited in [Wang 2011], where *semantic description documents* (SDD) refer to the information about concept hierarchies, related properties and instances, or in [David 2011] where a bag of words describing a concept is exploited within a mining approach. Other matching approaches exploiting WordNet include YAM++ [Ngo and Bellahsene 2013] and XMAP [Djeddi and Khadir 2014] systems. YAM++ applies diverse strategies coming from machine learning, information retrieval and graph matching. For terminological similarity measures, YAM++ includes Wordnet-based metrics as Lin and Wu-Palmer [Budanitsky and Hirst 2006]. XMAP defines a semantic similarity measure, defined using a domain-specific resource (UMLS) and WordNet. In order to deal with lexical ambiguity, the notion of *scope* is introduced, which represents the context where a concept occurs (i.e., the set of those concepts that are close to a given concept in therms of shortest path in a semantic network). They exploit super-concepts and sub-concepts in a hierarchy, as we did, but they do not take into account entity annotations as comments or definitions.

Our approach is inspired from those WordNet-based matching and sense disambiguation approaches for automatically confirming or not generated alignments that rely on different strategies. There exist different ways to evaluate alignments generated by matchers [Hong-Hai et al. 2002]. One classical way consists of comparing alignments to reference ones (gold standard). However, constructing such references is a time-consuming task. They are not always available or may suffer of imperfections or low coverage. In the lack of such resources, alternatives include manual labeling on sample alignments, alignment coherence measurements [Meilicke 2011], and correspondence patterns mining. Furthermore, the quality of a matcher can be assessed regarding its suitability for a specific task or application [Isaac et al. 2008, Hollink et al. 2008]. Finally, alternative approaches for validating alignments consider the generation of natural language questions to support end-users in the validation task [Abacha et al. 2014] or validation of correspondences using graph-based algorithms in a semi-automatic way [Serpeloni et al. 2011].

6. Final remarks and future work

This paper presented an approach for validating alignments between top-level and domain ontologies, using WordNet. We presented preliminary results that point to the strengthen and weaknesses of our approach. In particular, we do not mainly treat terms (but single words); the appropriate synset may not be selected if a wrong synset overlaps with the context of a concept, with the same number of elements in the overlapping set; sometimes the disambiguation process does not help due to the lack of context (cases where the ontologies have not enough terminological information or are structurally flat); we can not validate correspondences involving concepts with no entries in the resource; we do not take into account logical inconsistency nor simple hierarchical reasoning for filtering out and enhancing the set of correspondences. For overcoming these weaknesses, we plan to enrich the contexts by adding the contexts of neighbours, use distributional resources within the disambiguation process for better measuring the overlap of contexts, apply alternative similarity metrics for measuring the overlap between contexts, and deal with logical reasoning. We also plan to combine different background resources, including indirect alignments between WordNet and top-level ontologies and the lexico-semantic resource BabelNet, which integrates several lexicons and encyclopedias, and explore the superior layer of concepts in domain ontology to retrieve the more appropriate synset.

References

Abacha, A., Dos Reid, J., and Mrabet, Y. (2014). Question Generation for the Validation of Mapping Adaptation. In *6th International Symposium on Semantic Mining in Biomedicine*, pages 27–31.

Budanitsky, A. and Hirst, G. (2006). Evaluating WordNet-based Measures of Lexical Semantic Relatedness. *Comput. Linguistics*, 32(1):13–47.

Cheatham, M. and Hitzler, P. (2014). Conference v2.0: An Uncertain Version of the OAEI Conference Benchmark. In *13th Semantic Web Conference*, pages 33–48.

David, J. (2011). AROMA results for OAEI 2011. In *6th Workshop on Ontology Matching*, pages 122–125.

Djeddi, W. E. and Khadir, M. T. (2014). A Novel Approach Using Context-Based Measure for Matching Large Scale Ontologies. In *16th Conf. on Data Warehousing and Knowledge Discovery*, pages 320–331.

Euzenat, J. and Shvaiko, P. (2007). *Ontology Matching*. Springer-Verlag, 1st edition.

Gangemi, A., Guarino, N., Masolo, C., Oltramari, A., and Schneider, L. (2002). Sweetening Ontologies with DOLCE. In *13th Conf. on Knowledge Engineering and Knowledge Management*, pages 166–181.

Guarino, N. (1998). *Formal Ontology in Information Systems: Proceedings of the 1st Conference*. IOS Press, 1st edition.

Herre, H., Heller, B., Burek, P., Hoehndorf, R., Loebe, F., and Michalek, H. (2007). General Formal Ontology (GFO): A Foundational Ontology Integrating Objects and Processes. In *Basic Principles, Research Group Ontologies in Medicine).*

Hollink, L., Van Assem, M., Wang, S., Isaac, A., and Schreiber, G. (2008). Two Variations on Ontology Alignment Evaluation: Methodological Issues. In *5th European Semantic Web Conference*, pages 388–401.

Hong-Hai, D., Melnik, S., and Rahm, E. (2002). Comparison of schema matching evaluations. In *Workshop on Web, Web-Services, and Database Systems*, pages 221–237.

Isaac, A., Matthezing, H., van der Meij, L., Schlobach, S., Wang, S., and Zinn, C. (2008). Putting Ontology Alignment in Context: Usage Scenarios, Deployment and Evaluation in a Library Case. In *5th European Semantic Web Conference*, pages 402–417.

Kalfoglou, Y. and Schorlemmer, M. (2003). Ontology mapping: the state of the art. *Knowledge Engineering Review*, 18(1):1–31.

Kwak, J. and Yong, H. (2010). Ontology Matching based on Hypernym, Hyponym, Holonym and Meronym Sets in WordNet. *Web and Semantic Technology*, 1(2):1–14.

Lin, F. and Sandkuhl, K. (2008). A Survey of Exploiting WordNet in Ontology Matching. In *IFIP 20th World Computer Congress*, pages 341–350.

Maedche, A. and Staab, S. (2002). Measuring Similarity between Ontologies. In *13th Conf. on Knowledge Engineering and Knowledge Management*, pages 251–263.

Mascardi, V., Cordì, V., and Rosso, P. (2007). A Comparison of Upper Ontologies. In *8th AI*IA/TABOO Joint Workshop on Agents and Industry*, pages 55–64.

Mascardi, V., Locoro, A., and Rosso, P. (2010). Automatic Ontology Matching via Upper Ontologies: A Systematic Evaluation. *IEEE Trans. on Knowl. and Data Eng.*, 22(5):609–623.

Meilicke, C. (2011). *Alignment incoherence in ontology matching*. PhD thesis, University of Mannheim.

Mika, P., Oberle, D., Gangemi, A., and Sabou, M. (2004). Foundations for Service Ontologies: Aligning OWL-S to Dolce. In *13th Conf. on World Wide Web*, pages 563–572.

Miller, G. A. (1995). WordNet: A Lexical Database for English. *Commun. ACM*, 38(11):39–41.

Navigli, R. (2009). Word Sense Disambiguation: A Survey. *ACM Comput. Surv.*, 41(2):10:1–10:69.

Ngo, D. and Bellahsene, Z. (2013). YAM++ - Results for OAEI 2013. In *8th Workshop on Ontology Matching*, pages 211–218.

Niles, I. and Pease, A. (2001). Towards a Standard Upper Ontology. In *Conf. on Formal Ontology in Information Systems*, pages 2–9.

Schadd, F. C. and Roos, N. (2012). Coupling of Wordnet Entries for Ontology Mapping Using Virtual Documents. In *7th Workshop on Ontology Matching*, pages 25–36.

Schmidt, D., Trojahn, C., and Vieira, R. (2016). Analysing Top-level and Domain Ontology Alignments from Matching Systems. In *11th Workshop on Ontology Matching*, pages 1–12.

Serpeloni, F., Moraes, R., and Bonacin, R. (2011). Ontology mapping validation. *International Journal of Web Portals*, 3(3):1–11.

Wang, P. (2011). Lily Results on SEALS Platform for OAEI 2011. In *6th Workshop on Ontology Matching*, pages 156–162.

Yatskevich, M. and Giunchiglia, F. (2004). Element level semantic matching using WordNet. In *Meaning Coordination and Negotiation Workshop*, pages 37–48.