Solving semantic ambiguity to improve semantic web based ontology matching

Conference or Workshop Item

Solving Semantic Ambiguity to Improve Semantic Web based Ontology Matching

Jorge Gracia[1], Vanessa López[2], Mathieu d'Aquin[2], Marta Sabou[2], Enrico Motta[2], and Eduardo Mena[1]

[1] IIS Department, University of Zaragoza, Spain
{jogracia,emena}@unizar.es
[2] Knowledge Media Institute (KMi), The Open University, United Kingdom
{v.lopez,m.daquin,r.m.sabou,e.motta}@open.ac.uk

Abstract. A new paradigm in Semantic Web research focuses on the development of a new generation of knowledge-based problem solvers, which can exploit the massive amounts of formally specified information available on the Web, to produce novel intelligent functionalities. An important example of this paradigm can be found in the area of Ontology Matching, where new algorithms, which derive mappings from an exploration of multiple and heterogeneous online ontologies, have been proposed. While these algorithms exhibit very good performance, they rely on merely syntactical techniques to anchor the terms to be matched to those found on the Semantic Web. As a result, their precision can be affected by ambiguous words. In this paper, we aim to solve these problems by introducing techniques from Word Sense Disambiguation, which validate the mappings by exploring the semantics of the ontological terms involved in the matching process. Specifically we discuss how two techniques, which exploit the ontological context of the matched and anchor terms, and the information provided by WordNet, can be used to filter out mappings resulting from the incorrect anchoring of ambiguous terms. Our experiments show that each of the proposed disambiguation techniques, and even more their combination, can lead to an important increase in precision, without having too negative an impact on recall.
Keywords: semantic web, ontology matching, semantic ambiguity.

1 Introduction

As result of the recent growing of the Semantic Web, a new generation of semantic applications are emerging, focused on exploiting the increasing amount of online semantic data available on the Web [5]. These applications need to handle the high semantic heterogeneity introduced by the increasing number of available online ontologies, that describe different domains from many different points of view and using different conceptualisations, thus leading to many ambiguity problems.

In this challenging context, a new paradigm, which uses the Semantic Web as background knowledge, has been proposed to perform automatic Ontology

Matching [8]. An initial evaluation of this method showed a 70% precision in obtaining mappings between ontologies [9]. These experiments have also shown that more than half of the invalid mappings are due to ambiguity problems in the *anchoring process* (see later Sections 2 and 3).

These ambiguity problems are shared by any other Ontology Matching system based on background knowledge. Indeed, they are shared by any other system which needs to find correspondences across heterogeneous sources. Nevertheless we focus on the above mentioned Semantic Web based matcher, because it deals with online ontologies, thus maximizing heterogeneity of sources (and ambiguity problems), and providing us a suitable scenario to develop our ideas.

In this paper we investigate the use of two different techniques from Word Sense Disambiguation. The objective is to improve the results of background knowledge based Ontology Matching, by detecting and solving the ambiguity problems inherent to the use of heterogeneous sources of knowledge. Our experiments, based on the system described in [8], confirm our prediction that precision can be improved by using the above mentioned semantic techniques, getting even better results by combining them.

The rest of this paper is as follows: Section 2 explains the paradigm of harvesting the Semantic Web to perform Ontology Matching. How semantic ambiguity hampers this method is explained in Section 3, whereas in Sections 4, 5, and 6 we show three different approaches to solve this problem. Our experimental results and some related work can be found in Sections 7 and 8, respectively. Finally conclusions and future work appear in Section 9.

2 Ontology Matching by Harvesting the Semantic Web

In [8] a *new paradigm to Ontology Matching* that builds on the Semantic Web vision is proposed: it derives semantic mappings by exploring multiple and heterogeneous online ontologies that are dynamically selected (using Swoogle[3] as semantic search engine), combined, and exploited. For example, when matching two concepts labelled *Researcher* and *AcademicStaff*, a matcher based on this paradigm would 1) identify, at run-time, online ontologies that can provide information about how these two concepts relate, and then 2) combine this information to infer the mapping. The mapping can be either provided by a single ontology (e.g., stating that *Researcher* $\sqsubseteq$ *AcademicStaff*), or by reasoning over information spread among several ontologies (e.g., that *Researcher* $\sqsubseteq$ *ResearchStaff* in one ontology and that *ResearchStaff* $\sqsubseteq$ *AcademicStaff* in another). The novelty of the paradigm is that the knowledge sources are not manually provided prior to the matching stage but dynamically selected from online available ontologies during the matching process itself.

Figure 1 illustrates the basic idea of Ontology Matching by harvesting the Semantic Web. A and B are the concepts to relate, and the first step is to find online ontologies containing concepts A' and B' equivalent to A and B. This

[3] http://swoogle.umbc.edu/

process is called *anchoring* and A' and B' are called the *anchor terms*. Based on the relations that link A' and B' in the retrieved ontologies, a mapping is then derived between A and B.

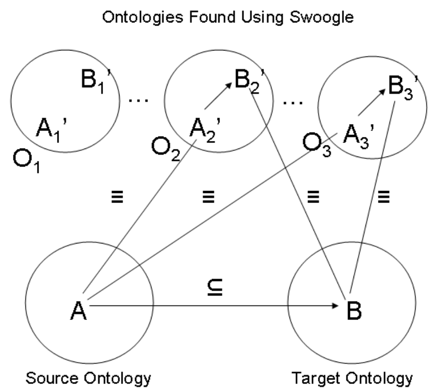

Fig. 1. Ontology Matching by harvesting the Semantic Web.

A baseline implementation of this technique has been evaluated [9] using two very large, real life thesauri that made up one of the test data sets in the 2006 Ontology Alignment Evaluation Initiative, AGROVOC and NALT[4]. A sample of 1000 mappings obtained thanks to this implementation has been manually validated, resulting in a promising 70% precision. However, a deeper analysis of the wrong mappings has shown that more than half of them (53%) were due to an incorrect anchoring: because of ambiguities, elements of the source ontology have been anchored to online ontologies using the considered terms with different senses. The employed naive anchoring mechanism is thus clearly insufficient, as it fails to distinguish words having several different senses and so, to handle ambiguity. Our hypothesis is that integrating techniques from Word Sense Disambiguation to complement the anchoring mechanism would lead to an important increase in precision.

3 Sense Disambiguation to Improve Anchoring

We have devised an improved way to perform Ontology Matching based on background knowledge, using techniques that take into account the *semantics* of the compared terms to validate the anchoring process.

For a better insight, let us see an example. The matcher described in Section 2 retrieved the following matching between two terms from the AGROVOC and NALT ontologies: $game \sqsupseteq sports$. "Game" is a "wild animal" in AGROVOC while "sports" appears in NALT as a "leisure, recreation and tourism" activity.

[4] http://www.few.vu.nl/ wrvhage/oaei2006/

The reason why this invalid mapping was derived is because "game" has been anchored in a background ontology[5], where it is defined as subclass of "Recreation or Exercise", and as superclass of "sport". This problem can be solved with an appropriate technique which deals with the *ambiguity* of the terms, being able to determine that "game" in the AGROVOC ontology (an animal) and "game" in the background ontology (a contest) are *different* concepts, thus avoiding their anchoring. Thus, our approach to handle semantic ambiguity is twofold:

First, we have considered the system proposed in [11]. Its goal is to disambiguate user keywords in order to translate them into semantic queries. In this context a *semantic similarity measure* has been defined to provide a *synonymy degree* between two terms from different ontologies, by exploring both their lexical and structural context. A configurable threshold allows the system to determine whether two ontological terms are considered or not the same (see Section 4 for more details).

Second, we have explored the use of a WordNet[6] based technique to perform a similar task. We reused parts of PoweMap [4], a hybrid knowledge-based matching algorithm, comprising terminological and structural techniques, and used in the context of multiontology question answering. Details of how PowerMap is used to filter semantically sound ontological mappings are given in Section 5.

In the following, we discuss the experiments we have conducted on the use of these two techniques, and on their combination, to improve Semantic Web based Ontology Matching.

4 Improving Anchoring by Exploring Ontological Context

In [2, 11] a system to discover the possible meanings of a set of user keywords by consulting a pool of online available ontologies is presented. First it proposes a set of possible ontological senses for each keyword, integrating the ones that are considered similar enough. Then these merged senses are used as input for a disambiguation process to find the most probable meaning of each keyword, to use them, finally, in the construction of semantic queries. These queries must represent the intended meaning of the initial user keywords.

Here we focus on the first step of the system, where an *ontological context* based similarity measure is applied to decide whether the semantics of two ontological terms represent the same sense or not.

4.1 Synonymy degree estimation

A detailed description of the above mentioned similarity measure is out of the scope of this paper, but we summarize here the key characteristics:

1. The algorithm receives two terms A, B from two different ontologies as input. Their ontological contexts are extracted (hypernyms, hyponyms, descriptions, properties,...).

[5] http://lists.w3.org/Archives/Public/www-rdf-logic/2003Apr/att-0009/SUMO.daml
[6] http://wordnet.princeton.edu/

2. An initial computation uses linguistic similarity between terms, considering labels as strings.
3. A subsequent recursive computation uses structural similarity, exploiting the ontological context of a term until a given *depth*. Vector Space Models are employed in the comparisons among *sampled* sets of terms extracted from the ontological contexts.
4. The different contributions (structural similarity, linguistic similarity, ...) are weighted, and a final *synonymy degree* between A, B is provided.

Therefore, this ontological context based similarity (let us call it $sim_{ont}(A, B)$) gets an estimated *synonymy degree* in $[0, 1]$ for a given *depth* of exploration (number of levels in the hierarchy that we explore).

4.2 Improved anchoring technique

Let us call, for the rest of the paper, A and B a particular pair of terms belonging respectively to the ontologies O_A and O_B to be aligned. We denote A' and B' their respective anchor terms in background ontologies, and O'_A and O'_B the respective background ontologies where they appear (sometimes $O'_A = O'_B$). Finally we denote as $\langle A, B, r, l\rangle$ a mapping between terms A and B, r representing the relation between them and l the level of confidence of the mapping.

Here is our first approach to take into account the semantics of the involved anchored terms in the matching process:

Scheme 1 ("filtering candidate mappings by exploring ontological context"). In this first approach, the validity of the anchoring is evaluated, a posteriori, on the mappings derived by the method explained in Section 2. The similarity between the ontological terms and their respective anchor terms is measured by analysing their ontological context up to a certain depth[7]: $sim_{ont}(A, A')$ and $sim_{ont}(B, B')$.

To qualify the mapping as a *valid* one, validity on each side of the mapping is required, hence both confidence degrees obtained must be above the required threshold. We compute the *confidence level* for the mapping $\langle A, B, r, l\rangle$ as:

$$l = min(sim_{ont}(A, A'), sim_{ont}(B, B')) \tag{1}$$

If $l > threshold$ then the mapping is accepted, otherwise is rejected.

The expected effect of this approach is an improvement in the precision, as many results erroneously mapped due to bad anchoring can be detected and filtered. Recalling the example discussed in Section 3: for the mapping $\langle game, sports, \sqsupseteq, l\rangle$ between AGROVOC and NALT ontologies, a value of $l = 0.269$ is computed. Then, if we have set up a threshold with a higher value, this erroneous mapping due to bad anchoring will be filtered out.

On the other hand, this approach is unable to improve the overall recall of the results (as it is unable to add new valid mappings). Indeed, we cannot discard a

[7] In this and subsequent experiments we compute sim_{ont} using $depth = 2$.

potential negative effect on recall, as some good mappings could also be filtered out if the computed similarities are not high enough (for example because of a poor description of the terms in ontologies).

5 Improving Anchoring by Exploring WordNet

As a complementary way, we have explored the use of a WordNet based algorithm implemented as part of PowerMap [4]. This makes possible to establish comparisons with the technique proposed in Section 4 and, eventually, to identify a combined use of both.

PowerMap is the solution adopted by PowerAqua, a multiontology-based Question Answering platform [4], to map user terminology into ontology-compliant terminology distributed across ontologies. The PowerMap algorithm first uses syntactic techniques to identify possible ontology matches, likely to provide the information requested by the user's query. WordNet based methods are then used to elicit the sense of candidate concepts by looking at the ontology hierarchy, and to check the semantic validity of those syntactic mappings, which originate from distinct ontologies, with respect the user's query terms.

5.1 PowerMap based method for the semantic relevance analysis

The PowerMap WordNet-based algorithm is adapted and used here to determine the validity of the mappings provided by the system described in Section 2. In this approach we do not perform similarity computation between terms and anchored terms, as we did in Schemes 1. Instead, similarity is computed directly between the matched ontology terms A and B.

Note that, here, *similarity* has a broader meaning than *synonymy*. We say that **two words are semantically similar** if they have a synset(s) in common (synonymy), or there exists an allowable IS-A path (in the hypernym/hyponym WordNet taxonomy) connecting a synset associated with each word. The rationale of this point is based on the two criteria of similarity between concepts established by Resnik in [7], where semantic similarity is determined as a function of the path distance between the terms, and the extent to which they share information in common. Formally, in the IS-A hierarchy of WordNet, similarity is given by the Wu and Palmer's formula described in [13].

5.2 Improved anchoring technique

In the following we explain how we apply this WordNet based method to determine the validity of mappings.

Scheme 2 ("filtering candidate mappings by exploring WordNet"). We compute the WordNet based confidence level $l = sim_{WN}(A, B)$ for the matching $\langle A, B, r, l \rangle$ as follows. Given the two ontological terms A and B, let $S_{B,A}$ be the set of those synsets of B for which there exists a semantically similar synset of A (according to Wu and Palmer's formula). If $S_{B,A}$ is empty, the mapping B is discarded because the intended meaning of A is not the same as that of the concept B. Finally, the true senses of B are determined by its place in the hierarchy of the ontology. That is, S_B^H consists only of those synsets of B that are similar to at least one synset of its ancestors in the ontology. We then obtain the valid senses as the intersection of the senses in S_B^H, with the senses obtained in our previous step, $\mathrm{S}_{B,A}$. Note that by intersection we mean the synsets that are semantically similar, even if they are not exactly the same synset. In case the intersection is empty it means that the sense of the concept in the hierarchy is different from the sense that we thought it might have in the previous step, and therefore that mapping pair should be discarded. The same process is repeated for the term A and its mapped term B.

The obtained confidence level l is in $\{0, 1\}$. This is a binary filtering, which only estimates whether there is semantic similarity between the mapped terms or not. The ontology mapping pair will be selected ($l = 1$) only if there is similarity between at least one pair of synsets from the set of valid synsets for A-B and the set of valid synsets for B-A. Otherwise, the mapping is rejected ($l = 0$)

Note that this method is not appropriate to evaluate disjoint mappings, producing unpredictable results. Also it is affected if the terms has no representation in WordNet. Therefore if $r = \perp$ or one of the terms to be mapped is not found in WordNet (i.e "zebrafish"), we left the value l as *undetermined*. Otherwise we compute the WordNet based *confidence level* for the mapping $\langle A, B, r, l \rangle$ as:

$$l = sim_{WN}(A, B) \tag{2}$$

Different strategies can be applied in case $l = undetermined$. By default we will not apply the filtering in these cases, thus assigning $l = 1$.

6 Combined Approach to Improve Anchoring

Finally, we propose a last strategy to improve anchoring: the combined use of the filtering schemes presented in Sections 4.2 and 5.2. We argue that, due to the different nature of these approaches, some of the false positives not filtered by one method could be detected by the other as inappropriate mappings, and vice versa. As an example, let us remind that the WordNet based method cannot evaluate disjoint mappings, thus this type of relations could be assisted by the other method. On the contrary if the internal structure of background ontologies is not rich enough, the ontological context based method could not filter properly, while the WordNet based one can.

Scheme 3 ("filtering candidate mappings by combining WordNet and Ontological Context based techniques"). Let us call l_{ont} the confidence level based on ontological context, computed with Equation 1 and l_{WN} the WordNet based confidence level obtained from Equation 2. We have identified two ways of combining both measures in an unified one:

Scheme 3.1: Promoting precision. As reported in Section 5.2, l_{WN} cannot be always computed. In such cases ($l_{WN} = undetermined$) we assign $l = l_{ont}$. Otherwise we compute the *confidence level* for the mapping $\langle A, B, r, l \rangle$ as:

$$l = min(l_{ont}, l_{WN}) \tag{3}$$

Criterion of minimizing the confidence degree optimizes precision (but penalizes recall), because the resultant filtering criteria are much more exigent: only mappings that both methods estimate as valid can pass the filter.

Scheme 3.2: Promoting recall. If $l_{WN} = undetermined$ then $l = l_{ont}$, else:

$$l = max(l_{ont}, l_{WN}) \tag{4}$$

This alternative scheme, that maximizes the confidence degree, can be used if our primary target is to obtain as many potentially good mappings as possible (among the total of valid ones), thus promoting recall instead of precision.

7 Experimental Results

Our experiments have been conducted to verify the feasibility of the proposed methods to improve the Semantic Web based Ontology Matching method. We have tested a basic implementation of Schemes 1, 2 and 3. The results confirm our initial hypothesis (the precision is increased by solving ambiguity problems) thus proving the value of the approach.

We applied our different filtering mechanisms to a sample of 354 evaluated mappings, out of the total set of data provided by the initial matching experiment mentioned in Section 2 (which lead to a baseline precision of 70%).

We have measured *precision* as the number of retrieved valid mappings out of the total which pass the filtering. Nevertheless, the filtering also rejects a number of valid mappings. In order to assess this we would need a recall measure but, due to the nature of the experiment, we are not able to provide it (our starting point, the experiment mentioned in Section 2, did not consider recall). Nevertheless we can estimate the *effect* that the filtering of mappings causes on recall (even if we do not know it), by using this expression:

$$effect\ on\ recall = \frac{number\ of\ retrieved\ valid\ mappings}{number\ of\ initial\ valid\ mappings}$$

This is a value to be multiplied by the recall of the initial matching process, to obtain the final recall. We consider as *initial valid mappings* those out of the utilized sample that are valid according to human evaluation.

7.1 Experiment 1: filtering by using Ontological Context

We have run our first experiment by applying the filtering mechanism discussed in Section 4.2. In Figure 2 we show (Scheme 1), the precision achieved by the prototype in the experiment. The worst value coincides with the baseline (70%), with minimum threshold. As we increase it, we reject more invalid mappings than valid ones, as reflects the increase of precision, which reaches soon values above **80%**. At some point (thresholds between 0.33 and 0.38) the precision moderates its ascending trend, fluctuating around 87%. This value is the predicted precision one can reach due to the anchoring improvement according to [9].

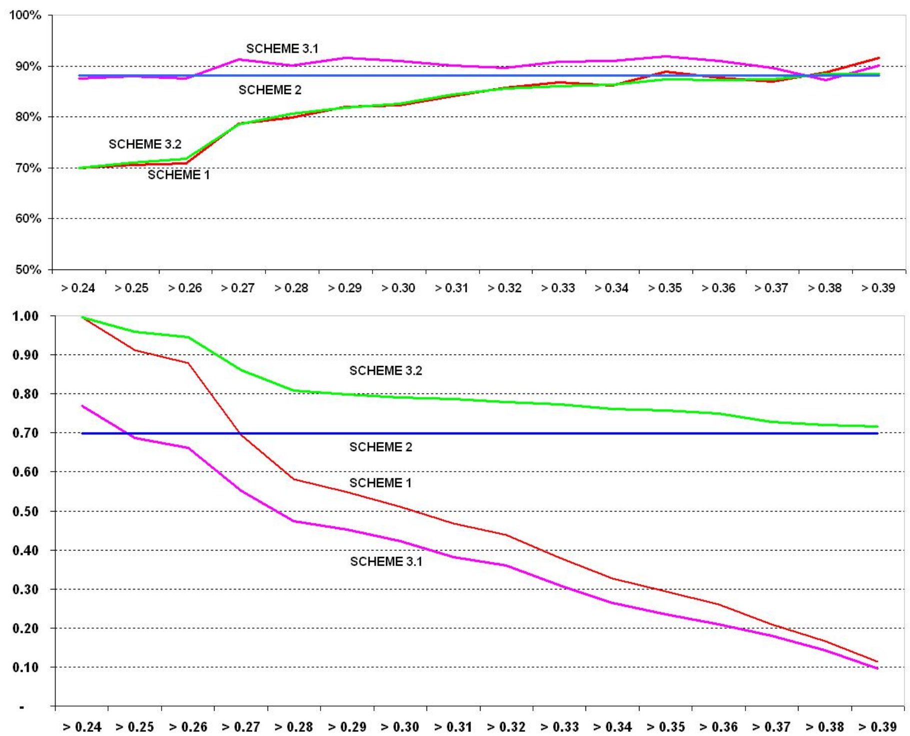

Fig. 2. *Precision* (upper) and *effect on recall* (lower). Baseline precision is 70%.

Figure 2 shows the effect on recall due to the filtering of mappings (Scheme 1). As we can expect, with the lowest threshold, no valid mappings are removed, so the recall is not influenced (effect on recall=1). As the threshold is raised the effect on recall decreases because more valid mappings are filtered out.

After first analysis of the mappings that hamper the method, we have discovered many ontological terms which are poorly described in background ontologies, and some other problems that we discuss later in Section 7.3.

Furthermore, we have run the experiment with a smaller set of randomly selected mappings (50), achieving almost identical effect on recall. This shows the feasibility of a training mechanism to obtain an optimal threshold with a small training set, to be reused later on the whole dataset.

7.2 Experiment 2: filtering by exploring WordNet

We analyse the results obtained from the same sample studied in Experiment 1. The WordNet based algorithm evaluated as correct 70% of valid mappings and 22% of invalid ones, leading to a *precision* of **88%** and an *effect on recall* of **0.70**. This sample help us to analyse the drawbacks of exclusively relying on sense information provided by WordNet to compute semantic similarity on ontology concepts. Those drawbacks are:

1. *Ontologies classes frequently use compound names without representation in WordNet.* Some compounds are not found in WordNet as such, i.e. "sugar substitutes" corresponds to two WordNet lemmas ("sugar", "substitutes"). Therefore, in many occasions the meaning can be misleading or incomplete.
2. *Synsets not related in the WordNet IS-A taxonomy.* Some terms considered similar from an ontology point of view, are not connected through a relevant IS-A path, i.e. for the term "sweeteners" and its ontological parent "food additive" (AGROVOC), therefore the taxonomical sense is unknown.
3. *The excessive fine-grainedness of WordNet sense distinctions.* For instance, the synsets of "crayfish" (AGROVOC) considering its parent "shellfish" are (1) "lobster-like crustacean..."; and (2) "warm-water lobsters without claws", but while considering its mapped term "animal" (NALT) the synset is (3) "small fresh water crustacean that resembles a lobster". This valid mapping is discarded as there is no relevant IS-A path connecting (3) with (1) or (2).
4. *Computing semantic similarity applying Resnik criteria to IS-A WordNet does not always produce good semantic mappings.* For instance the best synset obtained for, when computing the similarity between "Berries" and its parent "Plant" is "Chuck Berry – (United States rock singer)".

7.3 Experiment 3: combined approach

Here we have tested the behaviour of the improved anchoring schemes proposed in Section 6. In Figure 2 we can see the results (Schemes 3.1 and 3.2), and establish comparisons among all studied schemes.

As we predicted, Scheme 3.1 promotes precision. This combined approach slightly increases the precision achieved by Scheme 2, reaching a **92%** for a threshold of 0.285. Nevertheless we reduce recall almost to one half for this threshold. On the other hand Scheme 3.2 shows almost the same improvement in precision than Scheme 1, but with a very good behaviour in recall.

A precision of 92% obtained with Scheme 3.1 is the maximum we can reach combining both methods. At this point the system filters out most mappings considered invalid between AGROVOC and NALT, i.e. $\langle Fruit, Dessert, \perp, 0.25\rangle$ or $\langle Dehydration, drying, \sqsubseteq, 0\rangle$. Exploring the invalid mappings that pass our filters (particularly the ones that cause a slightly decrease in precision for high thresholds) we have found that the number of negative mappings due to bad anchoring is negligible, having found other types of errors that hamper our method, as bad modelling of relationships (using for example sumbsumption instead of part-of relation) i.e. $\langle East\_Asia, Asia, \sqsubseteq, 0.389\rangle$. Moreover, the meaning of an ontological concept must be precisely defined in the ontology: both similarity measures need to get the direct parents of the involved terms, but often the ancestor is Resource[8], and therefore the taxonomical meaning cannot be obtained, which introduces certain degree of uncertainty in the results.

8 Related Work

The anchoring process, where ambiguity problems can be present, is inherent to any Ontology Matching system based on background knowledge. Nevertheless most of them rely on merely syntactical techniques [1, 12]. Others, as S-Match [3], explore structural information of the term to anchor the right meaning, however it only accesses to WordNet as background knowledge source.

In some cases ambiguity in anchoring is a minor problem, because matched and background ontologies share the same domain [1], so it is expected that most polysemous terms have a well defined meaning. On the contrary, in our case the online ontologies constitute an open and heterogeneous scenario where, consequently, ambiguity becomes relevant.

Regarding the techniques we use from Word Sense Disambiguation, many others could be applied (see [10, 6] for example). Nevertheless we have selected the synonymy measure used in [11] to perform our disambiguation tasks because it has some convenient properties: it is not domain-dependent, it does not depend on a particular lexical resource, and it was conceived to deal with online ontologies. We also included the PowerMap based technique [4] to take advantage of the high quality description and coverage that WordNet provides, and because it combines in a clever way some well founded ideas from traditional Word Sense Disambiguation [7, 13].

9 Conclusions and Future Work

In this paper, we have presented different strategies to improve the precision of background knowledge based Ontology Matching systems, by considering the semantics of the terms to be anchored in order to deal with possible ambiguities during the anchoring process. We have explored the application of two similarity measures: one based on the ontological context of the terms, and another based on WordNet. A final strategy has been conceived by combining both measures.

In order to apply our ideas we have focused on a matcher that uses the Semantic Web as source of background knowledge. Our experimental results show that all filtering strategies we have designed improve the precision of the system (initially 70%). For example our Scheme 3.2 can reach a precision of 87%, affecting the overall recall in only a factor of 0.76.

Our experimental results encourage us to tackle further improvements and tests to our matching techniques. For example, a more advanced prototype will

[8] http://www.w3.org/2000/01/rdf-schema#Resource

be developed, which fully integrates the Semantic Web based Ontology Matcher with the filtering schemes that we have tested here. Also we will explore new ways to exploit semantics *during* the anchoring process, not only after it (as we currently do in our filtering schemes).

Acknowledgments. This work was founded by the OpenKnowledge IST-FF6-027253, NeOn IST-FF6-027595 and Spanish CICYT TIN2004-07999-C02-02 projects, and by the Government of Aragon-CAI grant IT11/06.

References

1. Z. Aleksovski, M. C. A. Klein, W. ten Kate, and F. van Harmelen. Matching unstructured vocabularies using a background ontology. In *EKAW*, volume 4248 of *LNSC*, pages 182–197, 2006.
2. M. Espinoza, R. Trillo, J. Gracia, and E. Mena. Discovering and merging keyword senses using ontology matching. In *1st International Workshop on Ontology Matching (OM-2006) at ISWC-2006, Athens, Georgia (USA)*, November 2006.
3. F. Giunchiglia, P. Shvaiko, and M. Yatskevich. Semantic schema matching. In *OTM Conferences (1)*, volume 3760 of *LNCS*, pages 347–365. Springer, 2005.
4. V. Lopez, M. Sabou, and E. Motta. Powermap: Mapping the real semantic web on the fly. In *Proc. of 5th International Semantic Web Conference, Athens, GA, USA, November 5-9, 2006, LNCS 4273*, 2006.
5. E. Motta and M. Sabou. Next generation semantic web applications. In *1st Asian Semantic Web Conference*, volume 4185 of *Lecture Notes in Computer Science*, pages 24–29. Springer, 2006.
6. R. Navigli and P. Velardi. Structural semantic interconnections: A knowledge-based approach to word sense disambiguation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(7):1075–1086, 2005.
7. P. Resnik. Disambiguating noun groupings with respect to Wordnet senses. In *Proceedings of the Third Workshop on Very Large Corpora*, pages 54–68. Association for Computational Linguistics, 1995.
8. M. Sabou, M. d'Aquin, and E. Motta. Using the semantic web as background knowledge for ontology mapping. In *1st International Workshop on Ontology Matching (OM-2006) at ISWC-2006, Athens, Georgia (USA)*, November 2006.
9. M. Sabou, M. d'Aquin, and E. Motta. Using the semantic web as background knowledge for ontology matching: A large-scale evaluation. *Submitted for peer review*, 2007.
10. S. P. Ted Pedersen, Satanjeev Banerjee. Maximizing semantic relatedness to perform word sense disambiguation, 2005. University of Minnesota Supercomputing Institute Research Report UMSI 2005/25.
11. R. Trillo, J. Gracia, M. Espinoza, and E. Mena. Discovering the semantics of user keywords. *Journal on Universal Computer Science. Special Issue: Ontologies and their Applications*, November 2007.
12. W. R. van Hage, S. Katrenko, and G. Schreiber. A method to combine linguistic ontology-mapping techniques. In *International Semantic Web Conference*, 2005.
13. Z. Wu and M. Palmer. Verb semantics and lexical selection. In *32nd. Annual Meeting of the Association for Computational Linguistics*, pages 133 –138, New Mexico State University, Las Cruces, New Mexico, 1994.