Using multiple ontologies as background knowledge in ontology matching

Zharko Aleksovski[1], Warner ten Kate[1], and Frank van Harmelen[2]

[1] Philips Research, Eindhoven,
[2] Vrije Universiteit, Amsterdam

Abstract. Using ontology as a background knowledge in ontology matching is being actively investigated. Recently the idea attracted attention because of the growing number of available ontologies, which in turn opens up new opportunities, and reduces the problem of finding candidate background knowledge. Particularly interesting is the approach of using *multiple ontologies as background knowledge*, which we explore in this paper. We report on an experimental study conducted using real-life ontologies published online.

The first contribution of this paper is an exploration about how the matching performance behaves when multiple background ontologies are used cumulatively. As a second contribution, we analyze the impact that different types of background ontologies have to the matching performance. With respect to the precision and recall, more background knowledge monotonically increases the recall, while the precision depends on the quality of the added background ontology, with high quality tending to increase, and the low quality tending to decrease the precision.

1 Introduction

Ontology matching is regarded as one of the most urgent and most important problems in the Semantic Web. It is scientifically challenging and inherently very difficult problem [1–6]. It generated a lot of research in the past years, which resulted in many different solution methods proposed. Good surveys of the existing ontology matching methods can be found in [2, 3, 7]. According to [7] they can be divided into four categories: *terminological* that use lexical similarities between names, comments etc., *structural* that use the similarities in the structure of the matching ontologies, *instance-based* that use the classified instance-data in the ontologies, and *using background knowledge* that rely on external structured resources to find matching entities across different ontologies. In this paper, we focus on the last category - using ontologies as background knowledge in the matching.

Background knowledge in matching has been used in different ways [8–10]. In this study we use a very simple approach. We try to match each pair of concepts from the matching ontologies in two steps - *anchoring* and *deriving relations*. In the anchoring, we look if the matching concepts can be themselves matched to the background knowledge, and in the deriving relations we check if they match to background concepts which are related to one another. If they are, then we report that the testing pair of concepts are matched. This type of match we call an *indirect match* because it is being discovered indirectly through the background knowledge ontology.

In respect to the matching success, regardless of the choice, no background knowledge ontology is likely to provide all the matches we would like to find. Instead, it is reasonable to expect that matches missed by one background ontology can be found using some other. Hence, hoping to find more of the matches we desire to find, we can use *multiple ontologies as background knowledge*. The question we face now is how the characteristics of the background ontologies will impact the matching performance. As discussed in the study of [11], the landscape of the online published ontologies is very diverse.

We set to investigate the feasibility of the matching when multiple background ontologies are used. To stress the paradigm, we present the results of several experiments in which we set our objectives as follows: *(i)* the anchoring to the background knowledge is a simple lexical matching technique, i.e. we only use simple matching as needed to obtain relatively successful anchoring (see Section 3 for further explanation), *(ii)* the background knowledge candidates are relatively large sized ontologies[3], and *(iii)* there is lexical overlap between the matching ontologies and the background knowledge, that is, lexical match is possible between the matching ontologies and the background knowledge. We use multiple background knowledge ontologies by using each ontology separately, and then combining the sets of obtained matches. We are interested to see how do the background ontologies perform together as compared to how each of them performs alone.

Multiple ontologies as background knowledge have already been used [9]. Contribution of this paper is that we study the contribution of each background ontology individually as compared to their cumulative contribution, and we also study the effects of the different types of ontologies when used as background knowledge. All the test data was selected from online published ontologies, and it consisted of two ontologies which we matched to one another and six other that we used as background knowledge. Having selected our data from online published ontologies, our results reflect on the current state of the published ontologies.

The experiments revealed that using background ontologies published on the semantic web provide low but stable increase of recall as compared to a simple direct matching. Multiple background ontologies find almost disjoint sets of matches, and hence result in cumulative increase of recall. The precision of background-based matching mostly, but not entirely depends on the quality of the background ontologies. The low quality ontologies increase the recall, but they reduce the precision. The high-quality background knowledge ontologies

[3] Ontologies of size around 30 concepts are common as demonstration examples, however, they are trivial to analyze and do not provide well-grounded empirical insight. Hence, we focused our attention on ontologies of larger size with at least couple of hundreds of concepts as more interesting candidates.

also find wrong matches, but these are mainly caused by the different context of the knowledge, not by mistakes in the ontologies.

The rest of the paper is organized as follows: in Section 2 we will describe our approach to using background knowledge in detail, in Section 3 we will describe our case study with the experimental data and the results, in Section 4 we will discuss the findings of the experiments, and finally with Section 5 we will conclude the paper.

2 Using background knowledge in ontology matching

In our approach we match two ontologies while using a third one as background knowledge. We call the ontologies being matched the source and the target, however, this naming is not discriminative - the matching algorithm treats them equally, and swapping their places will only invert the result set[4]. As we mentioned in the introduction, the algorithm proceeds in two steps - anchoring and deriving relations. Its scheme is depicted on Figure 1.

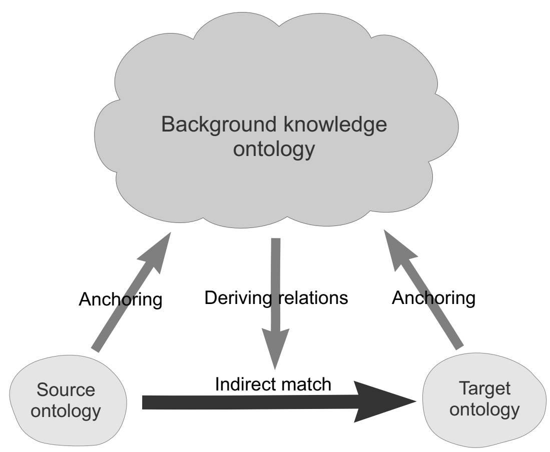

Fig. 1. Scheme of ontology matching using background knowledge.

Anchoring is matching the source and target concepts to the background knowledge. In general, this process can be performed by using any existing ontology matching technique. In our case we only use a simple lexical matching. Using other methods can make it difficult to explain the experimental results, because they may produce wrong matches, and simple technique while being rigid it is very precise and allows us to concentrate on the use of the background knowledge itself.

[4] The source and the target concept in each match on the result set will have their places swapped as well

Deriving relations is the process of discovering relations between source and target concepts by looking for relations between their anchored concepts in the background knowledge. Both the source and target concepts anchors are part of the background knowledge, and checking if they are related means using the reasoning service in the background knowledge ontology. Combining the anchor matches with the relations between the background knowledge concepts derives the match between source and target concepts, which is what we are looking for.

To explain this process in the context of life-sciences ontologies, we can see a realistic example on Figure 2: the source concept SRC: Brain is anchored to background knowledge concept BK: Brain, and the target concept TAR: Head is anchored to a background knowledge concept BK: Head. The background knowledge reveals a relation BK: Brain part-of BK: Head, and we derive a relation that source concept SRC: Brain has a narrower meaning than the target concept TAR: Head. Using background knowledge was crucial in this case; the match was not found by directly matching the source to the target ontology, SRC: Brain is classified under SRC: Central nervous system which is in no way related to the concept TAR: Head.

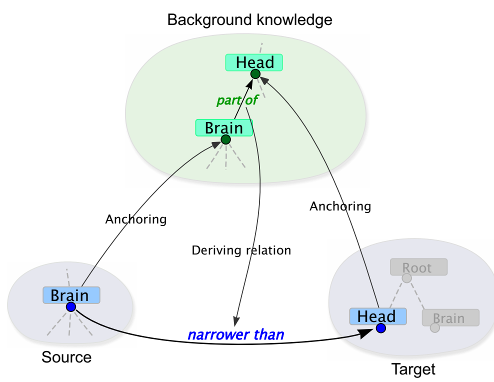

Fig. 2. Example using background knowledge in the matching process.

As suggested by the example above, of particular interest in our approach is exploiting the structure of the background knowledge ontology. It is done in the deriving relations step, when checking for relatedness between the anchored concepts in the background knowledge ontology. Before moving to the experimental part of the work, we will first introduce the formal definitions of all the components in this framework, which we will later use in the experimental part.

2.1 Formal framework

Concept is a class of things grouped together due to some shared property. It is named with a label, and sometimes with additional alternative names (synonyms). Besides the name(s), the meaning of a concept is determined by its semantic neighborhood, that is how it is related to the other concepts in the ontology. We will refer to concepts with capital italic letters X, Y..., with X^{ONT} to a concept from specific ontology, and we will also use the concept's label (in Sans Serif font), like Temporal lobe, or ONT: Temporal lobe, for the concept from the particular ontology.

Relation instance (also called just relation) is a triple $(X\ \mathit{relation}\ Y)$, where X and Y are concepts, and $\mathit{relation} \in \mathcal{T}$ is a relation type. $\mathcal{T}$ is the universal set of all relation types. The relation instance $(X\ \mathit{relation}\ Y)$ is interpreted as the concept X is related through the relation type *relation* to the concept Y. When clear from the context we will call the relation instances simply relations.

Ontology is a pair of sets: $\mathsf{ONT}(\mathcal{C}, \mathcal{R})$. $\mathcal{C}$ is a set of concepts, $\mathcal{R}$ is the set of relations among these concepts. We will refer to ontologies with their full name, like Foundational Model of Anatomy (or the name in italic *Foundational Model of Anatomy*), with short form of the name in Sans Serif font, like ONT for an arbitrary ontology, or FMA for the particular ontology *Foundational Model of Anatomy*

Ontology match between two ontologies S and T is a set of relation instances:

$$M \subseteq \mathcal{C}^{\mathsf{S}} \times \mathcal{T} \times \mathcal{C}^{\mathsf{T}} \tag{1}$$

Each element in this set $(X\ r\ Y) : X \in \mathcal{C}^{\mathsf{S}}, r \in \mathcal{T}, Y \in \mathcal{C}^{\mathsf{T}}$ we call a *match* between X and Y, or, X is matched to Y, through the relation type r. We will write it as $X \xrightarrow{r} Y$, or, $X \rightarrow Y$ when the relation type of the match is known from the context.

An ontology match is the result of any ontology matching technique. In practice, it plays the role of a bridge between different ontologies. Two specific ontology matches are of particular interest to our approach. They correspond to the two phases of the matching - anchoring and deriving relations.

2.2 Evaluation

To characterize the degree of success for matching we adopt two notions from the information retrieval field: *precision* and *recall*. In Information Retrieval (IR) the precision and recall are measures on performance of document retrieval [12]. They rely on a collection of documents and a query for which the relevancy of the document is known, assuming binary relevancy: a document is either relevant or non-relevant. In the ontology matching we define these measures through two sets - *desired* matches, and matches *found* by a matching method.

Precision is the proportion of desired and found matches, to all the found matches:

$$\textbf{Precision} = \frac{|\mathit{Desired} \cap \mathit{Found}|}{|\mathit{Found}|} \tag{2}$$

Recall is the proportion of desired and found matches, to all the desired matches:

$$\textsf{\textbf{Recall}} = \frac{|Desired \cap Found|}{|Desired|} \tag{3}$$

The precision represents the quality or the preciseness of the matches - what portion of the found matches are correct, and the recall represents the completeness of the matches - how many of the matches we want to find were actually found. The precision and recall have values between 0 and 1 inclusive. In practice they are often expressed in terms of percentage, ranging from 0% to 100%.

3 Case study

In our case study we matched two ontologies from the agricultural domain using six other ontologies as background knowledge. Motivated by the variety of ontologies that exist online, we decided to use background knowledge ontologies with varying origin. We investigated three different types of ontologies: different but related domain ontologies, general knowledge ontologies, and ontologies of an unknown origin. We set simple direct matching as a baseline to evaluate the matching performance, and we analyzed the matching performance by observing the precision and recall. All the test data was extracted in March 2007.

Matching ontologies The source ontology was NALT[5] and the target Agrovoc[6]. They both describe the domain of life sciences and agriculture. Agrovoc, as stated on the description provided on its homepage[7], I quote "is a multilingual, structured and controlled vocabulary designed to cover the terminology of all subject fields in agriculture, forestry, fisheries, food and related domains (e.g. environment)." NALT, as described on its homepage [8], I quote: "The NALT is primarily used for indexing and for improving retrieval of agricultural information. Currently, the NALT is the indexing vocabulary for NAL's bibliographic database of citations to agricultural resources, AGRICOLA[9]. The Food Safety Research Information Office[10] (FSRIO) and Agricultural Network Information Center[11] (AgNIC) also use the NALT as the indexing vocabulary for their information systems. In addition, the NALT is used as an aid for locating information on the ARS[12] and AgNIC web sites." In the experiments we used the versions of the OAEI 2006[13], which are publicly available. They contain 41,577 concepts NALT, and 28,174 concepts Agrovoc. Many of the concepts besides the labels are additionally described with synonyms.

---

5 `http://agclass.nal.usda.gov/agt`
6 `http://www.fao.org/agrovoc`
7 `http://www.fao.org/aims/ag_intro.htm`
8 `http://agclass.nal.usda.gov/about.shtml`
9 `http://agricola.nal.usda.gov/`
10 `http://www.nal.usda.gov/foodsafety/`
11 `http://www.agnic.org/`
12 `http://www.ars.usda.gov/`
13 Published on `http://www.few.vu.nl/\~wrvhage/oaei2006/`

Background knowledge We selected the background knowledge ontologies to faithfully represent the types of background knowledge we set to investigate. We used the Watson[14] ontology search engine to find them. We queried Watson for concept labels from the matching ontologies which are common English terms like *meat*, *animal*, *food*, etc. and selected six ontologies which frequently occurred in the retrieved results and also seemed like reasonable choice for the goal we set to analyze, that is exploring the different background knowledge types. Note that the choice of the search engine is not any special, in other studies different search engines have been successfully used for the same purpose, [9] used Swoogle to dynamically select background ontologies for an ontology matching task.

The selected six ontologies were the following: Economy which models a different but related domain as the matching ontologies; Mid-level, Sumo and Tap which are general knowledge ontologies; and A.com and Surrey which are ontologies of an unknown origin.

- *Background knowledge 1*: Economy ontology is described at `www.daml.org`, I quote: "is based on CIA World Fact Book (2002). Some industry concepts are based on the North American Classification System ('NAICS') - online at `http://www.census.gov/rpcd/www/naics.html`." As its name indicates, it intends to formally describe the domain of economy. It was engineered by Teknowledge Corporation[15] and submitted to the collection of ontologies gathered at `www.daml.org`. The size is 323 concepts.
- *Background knowledge 2*: Mid-level is constructed to play the role of bridge between the Sumo abstract level ontology, and the different varieties of Sumo domain-specific ontologies[16]. It is not domain-specific, and contains 1773 concepts.
- *Background knowledge 3*: Sumo (Suggested Upper Merged Ontology) is being created as part of the IEEE Standard Upper Ontology Working Group. It contains 576 concepts.
- *Background knowledge 4*: Tap as described in [13] is a shallow but broad knowledge base containing basic lexical and taxonomic information about a wide range of popular objects. It is claimed to be independent of a domain, however, a manual inspection indicated that it mainly covers the chemical, machine and electronic industry domains. It contains 5488 concepts.
- *Background knowledge 5*: A.com is an ontology with an unknown origin. By browsing it we got the impression that it has been produced as a result of merging several ontologies. In addition, noticeable are surprising relations such as:

  $\mathsf{Volume} \preceq \mathsf{Pollution}$

  which can be seen as an indication that some form of directory structure was the origin of the data. It seems to cover various domains, and its size is 5624 concepts.

[14] `http://watson.kmi.open.ac.uk/WatsonWUI/`

[15] `http://www.teknowledge.com/`

[16] `http://ontology.teknowledge.com/`

– *Background knowledge 6*: Surrey ontology, according to the Watson search engine, originates from the web site `www.surrey.co.uk`. In our analysis we did not manage to trace back its source, the download link does not work and on the web site the ontology is not available. Similarly as in the previous case, parts of its content gave the impression that it was created by transforming a directory structure into an ontology in a straight-forward way. Having no available documentation about how it was created, we treated it as an unknown origin ontology. Its size is 672 concepts.

Background knowledge ontology	Type of ontology	Size in number of concepts
BK_1: Economy	Different domain	323
BK_2: Mid-level	General knowledge	1773
BK_3: Sumo	General knowledge	576
BK_4: Tap	General knowledge	5488
BK_5: A.Com	Unknown origin	5624
BK_6: Surrey	Unknown origin	672

Fig. 3. Properties of the background knowledge ontologies

The six background knowledge ontologies with their properties are summarized on the table in Figure 3. With respect to the common ontology sizes found online [11], they are large sized ontologies.

Evaluation We manually evaluated the results of the matching experiments. As a reference use-case we set the task of document reclassification, which is realistic in this context because the matching ontologies are used for classifying books and articles.

3.1 Experiments

We performed seven experiments in which we matched NALT to Agrovoc. In the first experiment, which served as baseline, we matched the ontologies directly, and in the other six we matched them indirectly using the six previously described background ontologies, one per experiment.

Direct matching (Experiment 1) In the direct matching we combined lexical and structural matching. In the lexical phase the labels were normalized by discarding stop words (*the*, *and*, *an*, *a*) and interpunction, and then matched to one another accounting for different word order and plural/singular form of the words. As a result, the lexical phase produced list of pairs of equivalent concepts. In the structural phase the hierarchical structure of the ontologies was used to induce further matches. The direct matching algorithm is shown in Figure 4.

The set of direct matches is empty in the beginning
dmatches $:= \emptyset$

Lexical phase: find equivalent lexical matches
for every concept pair $X \in \mathcal{C}^{\mathsf{SRC}}$, $Y \in \mathcal{C}^{\mathsf{TAR}}$ **do**
  if $\textsc{FullLexMatch}(X, Y)$ **then**
    dmatches $\Leftarrow (X \xrightarrow{\equiv} Y)$
end for

Structural phase: use the structure to find more matches
for every two relations $(X_1 \preceq X_2) \in \mathcal{R}^{\mathsf{SRC}}$, $(Y_1 \preceq Y_2) \in \mathcal{R}^{\mathsf{TAR}}$
  if $(X_2 \xrightarrow{\equiv} Y_1) \in$ dmatches **then**
    dmatches $\Leftarrow (X_1 \xrightarrow{\preceq} Y_2)$
for every two relations $(X_1 \succeq X_2) \in \mathcal{R}^{\mathsf{SRC}}$, $(Y_1 \succeq Y_2) \in \mathcal{R}^{\mathsf{TAR}}$
  if $(X_2 \xrightarrow{\equiv} Y_1) \in$ dmatches **then**
    dmatches $\Leftarrow (X_1 \xrightarrow{\succeq} Y_2)$

Fig. 4. Algorithm for matching ontologies directly.

Even though the direct matching was done using such a simple and rigid technique (no edit distance, or other form of approximation), it produced 6,437 matches between NALT and Agrovoc. This number is comparable to the numbers obtained in the OAEI 2006 [5] on the same test data, where most of the participating matching systems produced between 5000 and 10,000 matches. Hence, our direct matching can be considered relatively successful, and a good base-line to measure the added value of the background knowledge.

Indirect matching (Experiments 2 - 7) In the indirect matching we lexically anchored the matching ontologies to the background knowledge, and then used the hierarchies of the background knowledge to induce the indirect matches. In other words, the indirect matching algorithm can be explained as follows: for two matching concepts we first find their equivalent concepts in the background knowledge (if possible), then check if these background concepts are hierarchically related, and if they are we report an indirect match between the matching concepts. The indirect matching algorithm is shown on Figure 5.

The table on Figure 6 summarizes the results of the anchoring phase showing the number of source and target anchors (NALT and Agrovoc, respectively) established to each background knowledge ontology. The Economy ontology has the highest number of anchors as compared to its size, roughly to about one third of its concepts there are anchors established from the matching ontologies. Contrary, ACom has much fewer anchors relatively to its size, roughly one out of each 90 concepts has anchor established to it. We can also observe from the table that this ratio is variable for the background ontologies of the same origin.

Generally, the number of anchors is much smaller than the sizes of the matching ontologies NALT and Agrovoc, which count in tens of thousands. However,

The set of indirect matches is empty in the beginning
$\mathsf{imatches} := \emptyset$

Anchoring phase: anchor SRC *and* TAR *to* BK *using direct matching*
$\mathsf{anch}^{S \mapsto B} := \textsc{MatchDirectly}(\mathsf{SRC}, \mathsf{BK})$
$\mathsf{anch}^{T \mapsto B} := \textsc{MatchDirectly}(\mathsf{TAR}, \mathsf{BK})$

Deriving relations phase: find indirect matches using the anchors and BK
for every two anchors $(X \overset{\preceq}{\longmapsto} Z_1) \in \mathsf{anch}^{S \mapsto B},\ (Y \overset{\succeq}{\longmapsto} Z_2) \in \mathsf{anch}^{T \mapsto B}$
  if $(Z_1 \preceq Z_2)$ **then**
    $\mathsf{imatches} \Leftarrow (X \xrightarrow{\preceq} Y)$
for every two anchors $(X \overset{\succeq}{\longmapsto} Z_1) \in \mathsf{anch}^{S \mapsto B},\ (Y \overset{\preceq}{\longmapsto} Z_2) \in \mathsf{anch}^{T \mapsto B}$
  if $(Z_1 \succeq Z_2)$ **then**
    $\mathsf{imatches} \Leftarrow (X \xrightarrow{\succeq} Y)$

Fig. 5. Algorithm for matching SRC to TAR indirectly through BK as a background knowledge.

Background knowledge	BK_i size	Source anchors	Target anchors
BK_1: Economy	323	121	106
BK_2: MidLevel	1773	330	271
BK_3: Sumo	576	79	72
BK_4: Tap	5488	367	227
BK_5: ACom	5624	66	69
BK_6: Surrey	672	102	95

Fig. 6. Overview of the anchoring results.

given the sizes of the background ontologies and the fact that they are not agriculture-specific, this anchoring result is not surprising.

The table in Figure 7 gives an overview on the indirect matching results. The third and fourth column show the number of indirect matches, and the number of additional indirect matches which were not found in the baseline direct matching. Each row in the table corresponds to one background knowledge ontology, except for the last one which shows the *cumulative* number of matches (union). Note that these cumulative numbers are not simple sums of the numbers above them, for example for the indirect matches the sum is 2287 and the cumulative number of matches is 2183. They are different because some of the matches are found by more than one background knowledge ontology. Similarly, the sum of the additional indirect matches is 1462 whereas the cumulative number is 1428. We see that the sum and the cumulative number are close to one another, which reveals very important and attractive behavior of using multiple background knowledge ontologies. Namely, different ontologies produce nearly disjoint sets of indirect matches. This means that the more ontologies we use - the more matches we will find. If we look at the cumulative matches, the additional indirect

Background ontology	BK_i size	Indirect matches	Additional matches on top of direct matches
BK_1: Economy	323	259	85
BK_2: MidLevel	1773	200	81
BK_3: Sumo	576	115	57
BK_4: Tap	5488	1003	625
BK_5: ACom	5624	87	71
BK_6: Surrey	672	623	543
Cumulatively all BK_i		2183	1428

Fig. 7. Overview of the indirect matching results, the number of matches established using each background ontology

matches represent 66% of all the indirect matches, which in turn means that an arbitrary indirect match has higher chances to be an addition to the baseline matches. However, these numbers say nothing about the quality of the matches, as a next step we will evaluate their correctness.

3.2 Evaluation

In order to get better insight in the matching process we decided to undertake the effort of manually assessing the matches. As a natural reference we choose the task of document reclassification: the obtained matches are expected to faithfully reclassify the documents from the source to the target ontology, ideally, in the same way as a human would do.

For the precision we did the evaluation as follows: each match was checked for validity, if the correctness was not obvious then Google was used as reference by querying for *define: label* to find the definition of the term *label*. The evaluation of the precision proceeded in two phases: first evaluate the direct and then the indirect matches. For the direct matching which produced more than 6000 matches, we choose the random sampling method. After drawing a random sample of 10% (640 matches), we manually assessed these matches as described above. For the indirect matches, which were in total little bit more than 2000, we took the effort to manually assess all of them.

The recall was hard to estimate because it requires all the correct matches between the matching ontologies available, which we don't have. Therefore we set to observe the change in recall between different experiments instead of estimating the achieved recall.

The evaluation revealed that the direct matching achieved 100% precision, i.e. all the matches in the evaluation sample were correct. The precision of the indirect matching and the change in recall are shown in the table on Figure 8.

Matching experiment	**Precision** indir. matches	**Precision** addit. matches	**ΔRecall**
Exp.2: BK_1: Economy	84.17%	51.76%	0.68%
Exp.3: BK_2: Mid-level	97.00%	92.59%	1.17%
Exp.4: BK_3: Sumo	76.52%	52.63%	0.47%
Exp.5: BK_4: Tap	57.23%	31.36%	3.04%
Exp.6: BK_5: A.Com	36.78%	22.54%	0.25%
Exp.7: BK_6: Surrey	35.63%	26.15%	2.21%
Cumulativly BK_1-BK_6	57.63%	35.22%	7.81%

Fig. 8. Performance of the indirect matching experiments

4 Analysis

First general observation on the matches (all the matches from all the seven experiments) is that they were established between a small subset of the matching concepts: 2241 in NALT, and 1757 concepts in Agrovoc participated in the matches, as compared to the size of NALT which is 41,577 concepts and Agrovoc 28,174 concepts. The number of concepts which participated in the matching results were in the order of about 5% of the size of the matching ontologies. But, this effect is not peculiarity of our experimental data, in other studies [14, 15] similar effect was noticed when matching the FMA and GALEN ontologies which model the human anatomy. These ontologies have 59,000 and 24,000 concepts respectively, and the number of matched concepts reported in the studies is in the order of 10% of the ontology sizes. It seems that this effect occurs when matching large ontologies even though they model the same domain. Most likely explanation for this is that for the general concepts there is much better naming agreement, while for the more specific ones, which represent the majority, there is almost no agreement. In such a situation the labeling problem is solved by using many words to name a single concept. As an example, in NALT there is a concept named Salmonella choleraesuis subsp.choleraesuis serovar Paratyphi A.

Precision and recall The table on Figure 8 shows the precision of the indirect matches, and the increase of recall with respect to the baseline direct matching. Each row corresponds to one background ontology, except for the last which shows the results for the cumulative use of all the background ontologies together.

All the indirect matches which were also found in the baseline were correct, incorrect matches only appeared when they were not found in the baseline matching. Hence, the precision of the additional indirect matches is lower than the precision of the indirect matches.

The Tap ontology resulted in 57.23% precision, however, a special situation had reduced the precision of this ontology. Many of the matches were wrongly established to the target concept called Node. The root concept in TAP is called Node, and the target concept anchored to it was found related to any source concept anchored in Tap. When these wrong matches are not taken into account, the precision of Tap is calculated to 92.13%. This example gave a very important insight, the indirect matching can be very sensitive to mistakes which are high in the background knowledge hierarchy. The fact that the root concept of Tap was named Node caused drastic change in the results when we used it as background knowledge.

The first four background ontologies which are expert-created exhibit high precision in the indirect matches (more than 75%), and relatively high precision in the additional indirect matches (more than 50%). On the other hand, the unknown-origin ontologies show lower precision which is not a surprising thing given the low quality of their content.

Observing the recall we see that Tap provides the highest increase of recall, shown in the third column, but the Surrey ontology is the second next to the Tap ontology in the recall increase. While the ontologies of an unknown origin might show low precision, that does not prevent the recall being increased considerably. We also see that Surrey is much smaller than Midlevel, Tap and ACom, which is an empirical proof that small size does not immediately imply low recall.

Causes of wrong matches For the first four background knowledge ontologies there were two main causes for wrong matches: contextual problems and relatively small mistakes. Examples of matches caused by contextual problems are the following:

$\mathsf{NALT}\colon \mathsf{Meat} \xrightarrow{\preceq} \mathsf{Agrovoc}\colon \mathsf{Product}$

$\mathsf{NALT}\colon \mathsf{Vehicle} \xrightarrow{\preceq} \mathsf{Agrovoc}\colon \mathsf{Product}$

$\mathsf{NALT}\colon \mathsf{Organism} \xrightarrow{\preceq} \mathsf{Agrovoc}\colon \mathsf{Agent}$

Meat can be seen as a kind of product in the domain of economy, however, for our matching task this was not a desirable match. These matches can be seen as relations establishing roles, meat and vehicles can have the role of a product, and organism can have the role of an agent. Such modeling is apparently good for the contexts of these background ontologies. For discussions related to the context issues in knowledge representation the reader is referred to the Cyc Knowledge Base [16] and the study of [17][17]. In addition to the context problems, few of the wrong matches were caused by relatively small mistakes, such examples are the matches:

$\mathsf{NALT}\colon \mathsf{Marine\ invertebrae} \xrightarrow{\preceq} \mathsf{Agrovoc}\colon \mathsf{Fish}$

$\mathsf{NALT}\colon \mathsf{Herbivore} \xrightarrow{\preceq} \mathsf{Agrovoc}\colon \mathsf{Mammals}$

Jellyfish are kind of Marine invertebrae but they are not fish, and some kinds of birds are herbivore but not mammals. These relations come close to generally accepted claims like "birds fly" while exceptions exist: "penguins are birds, and

[17] The study argues that the knowledge representation issues and the functionality of the system are intrinsically tied to one another, I quote: "Representation and reasoning are inextricably intertwined: we cannot talk about one without also, unavoidably, discussing the other. We argue as well that the attempt to deal with representation as knowledge content alone leads to an incomplete conception of the task of building an intelligent reasoner."

yet they do not fly". We stress here that there were no different causes for wrong matches between Economy and the other three general-knowledge ontologies. The high-quality ontologies, whether they model different domain or are general-knowledge, the same reasons caused them to produce wrong matches when they were applied as background knowledge.

For the last two ontologies, which have unknown origin, mistakes were the cause for the wrong matches. For example:

NALT: Gas $\xrightarrow{\preceq}$ Agrovoc: Turbines

NALT: Waste $\xrightarrow{\preceq}$ Agrovoc: Water

The concepts in these wrong matches are semantically related, however, no strict relation can be established. These matches are clearly wrong. This suggests that ACom and Surrey were obtained by straight-forward transformation of a directory structure into an ontology.

5 Conclusions

Based on the work presented in this paper, we conclude that using multiple ontologies as background knowledge in ontology matching is useful and practically feasable. Our experiments indicated the key factors that influence the matching performance. The recall increases monotonically with adding more background ontologies. This is an important property because the recall increase is seen as bigger challenge for the current matching systems. For the precision, the success primarially depends on the quality of the background ontologies.

Observing the precision, the expert-created ontologies such as Economy, Mid-level, Sumo and Tap resulted in relatively high precision (more than 75%), and the main causes of wrong matches were contextual differences with the matching ontologies and small mistakes. The ontologies of unknown origin like ACom and Surrey resulted in lower precision (less than 40%) and the main cause of wrong matches were mistakes. This makes the expert-created ontologies more trustworthy and clearly preferable background knowledge candidates over the unknown-origin ontologies with respect to the precision.

All the background ontologies together provided relatively small increase in the recall of about 8% in addition to the direct matching. However, they resulted in nearly disjoint sets of matches, which means that new ontologies are likely to provide new additional matches and further increase the recall.

Furthermore, the expert-created ontologies, regardless whether they modeled different domain from the matching ontologies (Economy) or they were general-knowledge (Mid-level, Sumo and Tap), they resulted in similar matching qualities. On our experimental data we could not discriminate by the precision or the recall increase, and all of them had the same causes of wrong matches.

Finally, the Tap ontology showed that the matching process can be very sensitive to mistakes high in the background knowledge hierarchy. Other mistakes also resulted in wrong matches, but the mistake in Tap with the root concept being labeled Node seriously affected the precision when applying this background ontology.

References

1. Rahm, E., Bernstein, P.A.: A Survey of Approaches to Automatic Schema Matching. VLDB Journal **10**(4) (2001)
2. Shvaiko, P., Euzenat, J.: A survey of schema-based matching approaches. Journal on Data Semantics **4** (2005) 146–171
3. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art. Knowl. Eng. Rev. **18**(1) (2003) 1–31
4. Euzenat, J., Isaac, A., Meilicke, C., Shvaiko, P., Stuckenschmidt, H., Svab, O., Svatek, V., van Hage, W.R., Yatskevich, M.: First results of the ontology alignment evaluation initiative 2007. In: Ontology Matching Workshop at International Semantic Web Conference (ISWC). (2007)
5. Euzenat, J., Mochol, M., Shvaiko, P., Stuckenschmidt, H., Svab, O., Svatek, V., van Hage, W.R., Yatskevich, M.: Results of the ontology alignment evaluation initiative 2006. In: Ontology Matching Workshop at International Semantic Web Conference (ISWC). (2006)
6. Euzenat, J., Stuckenschmidt, H., Yatskevich, M.: Introduction to the ontology alignment evaluation 2005. In: Integrating Ontologies. (2005)
7. Euzenat, J., Le Bach, T., Paolo, Debo, J., Kuntz, R.D., Ehrig, M., Hauswirth, M., Jarrar, M., Lara, R., Dianamaynard, Napoli, A., Stamou, G., Stuckenschmidt, H., Shvaiko, P., Sergiotessaris, Van Acker, S., Zaihrayeu, I.: State of the art on ontology alignment. deliverable d2.2.3 (2004)
8. Aleksovski, Z., Klein, M., ten Kate, W., van Harmelen, F.: Matching unstructured vocabularies using a background ontology. In: Proceedings of Knowledge Engineering and Knowledge Management (EKAW). (2006) 182–197
9. Sabou, M., dAquin, M., Motta, E.: Using the semantic web as background knowledge for ontology mapping. In: Proceedings of Ontology Matching Workshop. (2006)
10. Bouquet, P., Serafini, L., Zanobini, S.: Semantic coordination: A new approach and an application. In: International Semantic Web Conference (ISWC). (2003) 130–145
11. Wang, T.D., Parsia, B., Hendler, J.A.: A survey of the web ontology landscape. In Cruz, I.F., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L., eds.: International Semantic Web Conference. Volume 4273 of Lecture Notes in Computer Science., Springer (2006) 682–694
12. Baeza-Yates, R., Ribeiro-Neto, B., et al.: Modern information retrieval. Addison-Wesley Harlow, England (1999)
13. Guha, R., McCool, R.: Tap: A semantic web toolkit. Semantic Web Journal **1**(1) (2003) 81–87
14. Zhang, S., Bodenreider, O.: Aligning representations of anatomy using lexical and structural methods. AMIA Symposium (2003) 753–757
15. Mork, P., Pottinger, R.A., Bernstein, P.A.: Challenges in precisely aligning models of human anatomy using generic schema matching. Technical report, Microsoft Research (2004)
16. Lenat, D.: Cyc: a large-scale investment in knowledge infrastructure. Communications of the ACM **38**(11) (1995) 33–38
17. Davis, R., Shrobe, H., Szolovits, P.: What is a knowledge representation? AI Magazine **14**(1) (1993) 17–33