Combining statistical and semantic approaches to the translation of ontologies and taxonomies

John McCrae
AG Semantic Computing
Universität Bielefeld
Bielefeld, Germany
jmccrae@cit-ec.uni-bielefeld.de

Mauricio Espinoza
Universidad de Cuenca
Cuenca, Ecuador
mauricio.espinoza@ucuenca.edu.ec

Elena Montiel-Ponsoda, Guadalupe Aguado-de-Cea
Ontology Engineering Group
Universidad Politécnica de Madrid
Madrid, Spain
{emontiel, lupe}@fi.upm.es

Philipp Cimiano
AG Semantic Computing
Universität Bielefeld
Bielefeld, Germany
cimiano@cit-ec.uni-bielefeld.de

Abstract

Ontologies and taxonomies are widely used to organize concepts providing the basis for activities such as indexing, and as background knowledge for NLP tasks. As such, translation of these resources would prove useful to adapt these systems to new languages. However, we show that the nature of these resources is significantly different from the "free-text" paradigm used to train most statistical machine translation systems. In particular, we see significant differences in the linguistic nature of these resources and such resources have rich additional semantics. We demonstrate that as a result of these linguistic differences, standard SMT methods, in particular evaluation metrics, can produce poor performance. We then look to the task of leveraging these semantics for translation, which we approach in three ways: by adapting the translation system to the domain of the resource; by examining if semantics can help to predict the syntactic structure used in translation; and by evaluating if we can use existing translated taxonomies to disambiguate translations. We present some early results from these experiments, which shed light on the degree of success we may have with each approach.

1 Introduction

Taxonomies and ontologies are data structures that organise conceptual information by establishing relations among concepts, hierarchical and partitive relations being the most important ones. Nowadays, ontologies have a wide range of uses in many domains, for example, finance (International Accounting Standards Board, 2007), bio-medicine (Collier et al., 2008) (Ashburner et al., 2000) and libraries (Mischo, 1982). These resources normally attach labels in natural language to the concepts and relations that define their structure, and these labels can be used for a number of purposes, such as providing user interface localization (McCrae et al., 2010), multilingual data access (Declerck et al., 2010), information extraction (Müller et al., 2004) and natural language generation (Bontcheva, 2005). It seems natural that for applications that use such ontologies and taxonomies, translation of the natural language descriptions associated with them is required in order to adapt these methods to new languages. Currently, there has been some work on this in the context of ontology localisation, such as Espinoza et al. (2008) and (2009), Cimiano et al. (2010), Fu et al. (2010) and Navigli and Penzetto (2010). However, this work has focused on the case in which exact or partial translations are found in other similar resources such as bilingual lexica. Instead, in this paper we look at how we may gain an adequate translation using statistical machine translation approaches that also utilise the semantic information beyond the label or term describing the concept, that is relations among the concepts in the ontology, as well as the attributes or properties that describe concepts, as will be explained in more detail in section 2.

Current work in machine translation has shown that word sense disambiguation can play an important role by using the surrounding words as context to disambiguate terms (Carpuat and Wu, 2007) (Apidianaki, 2009). Such techniques have

Proceedings of SSST-5, Fifth Workshop on Syntax, Semantics and Structure in Statistical Translation, pages 116–125,
ACL HLT 2011, Portland, Oregon, USA, June 2011.

been extrapolated to the translation of taxonomies and ontologies, in which the "context" of a taxonomy or ontology label corresponds to the *ontology structure* that surrounds the label in question. This structure, which is made up of the lexical information provided by labels and the semantic information provided by the ontology structure, defines the sense of the concept and can be exploited in the disambiguation process (Espinoza et al., 2008).

2 Definition of Taxonomy and Ontology Translation

2.1 Formal Definition

We define a taxonomy as a set of concepts, C, with equivalence (synonymy) links, S, subsumption (hypernymy) links, H, and a labelling function l that maps each concept to a single label from a language Σ^*. Formally we define a taxonomy, T, as a set of tuples (C, S, H, l) such that $S \subseteq \mathcal{P}(C \times C)$ and $H \subseteq \mathcal{P}(C \times C)$ and l is a function in $C \rightarrow \Sigma^*$. We also require that S is a transitive, symmetric and reflexive relation, and H is transitive. While we note here that this abstraction does not come close to capturing the full expressive power of many ontologies (or even taxonomies), it is sufficient for this paper to focus on the use of only equivalence and subsumption relationships for translation.

2.2 Analysis of ontology labels

Another important issue to note here is that the kind of language used within ontologies and taxonomies is significantly different from that found within free text. In particular, we observe that the terms used to designate concepts are frequently just noun phrases and are significantly shorter than a usual sentence. In the case of the relations between concepts (dubbed *object properties*) and attributes of concepts (*data type properties*), these are occasionally labelled by means of verbal phrases. We demonstrate this by looking at three widely used ontologies/taxonomies.

1. **Friend of a friend**: The Friend of a Friend (FOAF) ontology is used to describe social networks on the Semantic Web (Brickley and Miller, 2010). It is a small taxonomy with very short labels. Labels for concepts are compound words made up of up to three words.
2. **Gene Ontology**: The Gene Ontology (Ashburner et al., 2000) is a very large database of terminology related to genetics. We note that while some of the terms are technical and do not require translation, e.g., *ESCRT-I*, the majority do, e.g., *cytokinesis by cell plate formation*.
3. **IFRS 2009**: The IFRS taxonomy (International Accounting Standards Board, 2007) is used for providing electronic financial reports for auditing. The terms contained within this taxonomy are frequently long and are entirely noun phrases.

We applied tokenization and manual phrase analysis to the labels in these resources and the results are summarized in table 1. As can be observed, the variety of types of labels we may come across when linguistically analysing and translating ontology and taxonomy labels is quite large. We can identify the two following properties that may influence the translation process of taxonomy and ontology labels. Firstly, the length of terms ranges from single words to highly complex compound phrases, but is still generally shorter than a sentence. Secondly, terms are frequently about highly specialized domains of knowledge.

For properties in the ontology we also identify terms which consist of:

- Noun phrases identifying concepts.
- Verbal phrases that are only made up of the verb with an optional preposition.
- Complex verbal phrases that include the predicate.
- Noun phrases that indicate possession of a particular characteristic (e.g., *interest* meaning *X has an interest in Y*).

3 Creation of a corpus for taxonomy and ontology translation

For the purpose of training systems to work on the translation of ontologies and taxonomies, it is necessary to create a corpus that has similar linguistic structure to that found in ontologies and taxonomies. We used the titles of Wikipedia[1] for the following

[1] http://www.wikipedia.org

	Size	Mean tokens per label	Noun Phrases	Verb Phrases
FOAF	79	1.57	94.9%	8.9%
Gene Ontology	33795	4.45	100.0%	0.0%
IFRS 2009	2757	8.39	100.0%	0.0%

Table 1: Lexical Analysis of labels

	Link	Direct	Fragment	Broken
German	487372	484314	1735	1323
Spanish	347953	346941	330	682

Table 2: Number of translation for pages in Wikipedia

reasons:

- Links to articles in different languages can be viewed as translations of the page titles.
- The titles of articles have similar properties to the ontologies labels mentioned above with an average of 2.46 tokens.
- There are a very large number of labels. In fact we found that there were 5,941,890[2] articles of which 3,515,640 were content pages (i.e., not special pages such as category pages)

We included non-content pages (in particular, category pages) in the corpus as they were generally useful for translation, especially the titles of category pages. In table 2 we see the number of translations, which we further grouped according to whether they actually corresponded to pages in the other languages, as it is also possible that the translations links pointed to subsections of an article or to missing pages.

Wikipedia also includes redirect links that allow for alternative titles to be mapped to a given concept. These can be useful as they contain synonyms, but also introduce a lot more noise into the corpus as they also include misspelled and foreign terms. To evaluate the effectiveness of including these data for creating a machine translation corpus, we took a random sample of 100 pages which at least one page redirects to (there are 1,244,647 of these pages in total). We found that these pages had a total of 242 extra titles from the redirect page of which 204 (84.3%) where true synonyms, 19 (7.9%) were misspellings, 8 (3.3%) were foreign names for concepts (e.g., the French name for "Zeebrugge"), and 11 (4.5%) were unrelated. As such, we conclude that these extra titles were useful for constructing the corpus, increasing the size of the corpus by approximately 50% across all languages. There are several advantages to deriving a corpus from Wikipedia, for example it is possible to provide some hierarchical links by the use of the category that a page belongs to, such as has been performed by the DBpedia project (Auer et al., 2007).

4 Evaluation metrics for taxonomy and ontology translation

Given the linguistic differences in taxonomy and ontology labels, it seems necessary to investigate the effectiveness of various metrics for the evaluation of translation quality. There are a number of metrics that are widely used for evaluating translation. Here we will focus on some of the most widely used, namely BLEU (Papineni et al., 2002), NIST (Doddington, 2002), METEOR (Banerjee and Lavie, 2005) and WER (McCowan et al., 2004). However, it is not clear which of these methods correlate best with human evaluation, particularly for the ontologies with short labels. To evaluate this we collected a mixture of ontologies with short labels on the topics of human diseases, agriculture, geometry and project management, producing 437 labels. These were translated with web translation services from English to Spanish, in particular Google Translate[3], Yahoo! BabelFish[4] and SDL FreeTranslation[5]. Having obtained translations for each label in the ontology we calculated the evaluation scores using the four metrics mentioned above. We found that the source ontologies had an average

[2] All statistics are based on the dump on 17th March 2011

[3] http://translate.google.com

[4] http://babelfish.yahoo.com

[5] http://www.freetranslation.com

	BLEU	NIST	METEOR	WER
Evaluator 1, Fluency	0.108	0.036	0.134	0.122
Evaluator 1, Adequacy	0.209	0.214	0.303	0.169
Evaluator 2, Fluency	0.183	0.062	0.266	0.164
Evaluator 2, Adequacy	0.177	0.111	0.251	0.194
Evaluator 3, Fluency	0.151	0.067	0.210	0.204
Evaluator 3, Adequacy	0.143	0.129	0.221	0.120

Table 3: Correlation between manual evaluation results and automatic evaluation scores

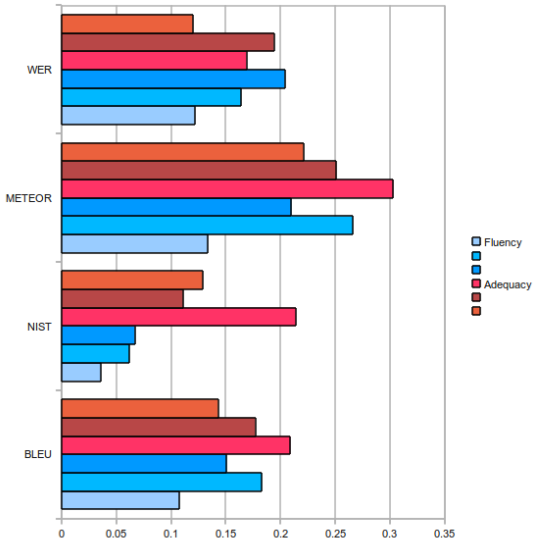

Figure 1: Correlation between manual evaluation results and automatic evaluation scores

label length of 2.45 tokens and the translations generated had an average length of 2.16 tokens. We then created a data set by mixing the translations from the web translation services with a number of translations from the source ontologies, to act as a control. We then gave these translations to 3 evaluators, who scored them for adequacy and fluency as described in Koehn (2010). Finally, we calculated the Pearson correlation coefficient between the automatic scores and the manual scores obtained. These are presented in table 3 and figure 1.

As we can see from these results, one metric, namely METEOR, seems to perform best in evaluating the quality of the translations. In fact this is not surprising as there is a clear mathematical deficiency that both NIST and BLEU have for evaluating translations for very short labels like the ones we have here. To illustrate this, we recall the formulation of BLEU as given in (Papineni et al., 2002):

$$\text{BLEU} = BP \cdot \exp(\sum_{n=1}^{N} w_n \log p_n)$$

Where BP is a brevity penalty, w_n a weight value and p_n represents the n-gram precision, indicating how many times a particular n-gram in the source text is found among the target translations. We note, however, that for very short labels it is highly likely that p_n will be zero. This creates a significant issue, as from the equation above, if any of the values of p_n are zero, the overall score, BLEU, will also be zero.

For the results above we chose $N = 2$, and corrected for single-word labels. However, the scores were still significantly worse, similar problems affect the NIST metric. As such, for the taxonomy and ontology translation task we do not recommend using BLEU or NIST as an evaluation metric. We note that METEOR is a more sophisticated method than WER and, as expected, performs better.

5 Approaches for taxonomy and ontology translation

5.1 Domain adaptation

It is generally the case that many ontologies and taxonomies focus on only a very specific domain, thus it seems likely that adaptation of translation systems by use of an in-domain corpus may improve translation quality. This is particularly valid in the case of ontologies which frequently contain "subject" annotations[6] for not only the whole data structure but often individual elements. To demonstrate this we tried to translate the IFRS 2009 taxonomy using the Moses Decoder (Koehn et al., 2007), which we trained on the EuroParl corpus (Koehn, 2005), translating from Spanish to English. As the IFRS taxonomy is on the topic of finance and accounting, we

[6]For example from the Dublin Core vocabulary: see `http://dublincore.org/`

	Baseline	With domain adaptation
WER*	0.135	0.138
METEOR	0.324	0.335
NIST	1.229	1.278
BLEU	0.090	0.116

Table 4: Results of domain-adapted translation. *Lower WER scores are better

	$P(syn\|s)$	$P(syn\|p)$	$P(syn\|n)$
English	0.147	0.012	0.001
Dutch	0.137	0.011	0.001
German	0.125	0.007	0.001
Spanish	0.126	0.012	0.001

Table 5: Probability of syntactic relationship given a semantic relationship in IFRS labels

chose all terms from our Wikipedia corpus which belonged to categories containing the words: "finance", "financial", "accounting", "accountancy", "bank", "banking", "economy", "economic", "investment", "insurance"and "actuarial" and as such we had a domain corpus of approximately 5000 terms. We then proceeded to recompute the phrase table using the methodology as described in Wu et al, (2008), computing the probabilities as follows for some weighting factor $0 < \lambda < 1$:

$$p(e|f) = \lambda p_1(e|f) + (1 - \lambda)p_d(e|f)$$

Where p_1 is the EuroParl trained probability and p_d the scores on our domain subset. The evaluation for these metrics is given in table 4. As can be seen with the exception of the WER metric, the domain adaption does seem to help in translation, which corroborates the results obtained by other authors.

5.2 Syntactic Analysis

One key question to figure out is: if we have a semantic model can this be used to predict the syntactic structure of the translation to a significant degree? As an example of this we consider the taxonomic term "statement", which is translated by Google Translate[7] to German as "Erklärung", whereas the term "annual statement" is translated as "Jahresabschluss". However, if the taxonomy contains a subsumption (hypernymy) relationship between these terms we can deduce that the translation "Erklärung" is not correct and the translation "Abschluss" should be preferred. We chose to evaluate this idea on the IFRS taxonomy as the labels it contains are much longer and more structured than some of the other resources. Furthermore, in this taxonomy the original English labels have been translated into ten languages, so that it is already a multilingual resource that can be used as gold standard. Regarding the syntax of labels, it is often the case that one term is derived from another by addition of a complementary phrase. For example the following terms all exist in the taxonomy:

1. Minimum finance lease payments receivable
2. Minimum finance lease payments receivable, at present value
3. Minimum finance lease payments receivable, at present value, end of period not later than one year
4. Minimum finance lease payments receivable, at present value, end of period later than one year and not later than five years

A high-quality translation of these terms would ideally preserve this same syntactic structure in the target language.We attempt to answer how useful ontological structure is by trying to deduce if there is a semantic relationship between terms then is it more likely that there is a syntactic relationship. We started by simplifying the idea of syntactic dependency to the following: we say that two terms are syntactically related if one label is a sub-string of another, so that in the example above the first label is syntactically related to the other three and the second is related to the last two. For English, we found that there were 3744 syntactically related terms according to this criteria, corresponding to 0.1% of all label pairs within the taxonomy, for all languages. For ontology structure we used the number of relations indicated in the taxonomy, of which there are 1070 indicating a subsumption relationship and 987 indicating a partitive relationship[8]. This means that

[7]Translations results obtained 8th March 2011

[8]IFRS includes links for calculating certain values, i.e., that "Total Assets" is a sum of values such as "Total Assets in Prop-

$e \rightarrow f$	$P(syn_f \mid syn_e, s)$	$P(syn_f \mid syn_e, p)$	$P(syn_f \mid syn_e, n)$
English $\rightarrow$ Spanish	0.813 ± 0.059	0.750 ± 0.205	0.835 ± 0.013
English $\rightarrow$ German	0.835 ± 0.062	0.417 ± 0.212	0.790 ± 0.013
English $\rightarrow$ Dutch	0.875 ± 0.063	0.833 ± 0.226	0.898 ± 0.013
Average	0.841 ± 0.035	0.665 ± 0.101	0.841 ± 0.008

Table 6: Probability of cross-lingual preservation of syntax given semantic relationship in IFRS. Note here s refers to the source language and t to the target language. Error values are 95% of standard deviation.

0.08% of label pairs were semantically related. We then examined if the semantic relation could predict whether there was a syntactic relationship between the terms in a single language. We define N_s as the number of label pairs with a subsumption relationship and similarly define N_p, N_n and N_{syn} for partitive, semantically unrelated and syntactically related pairs. We also define $N_{s \wedge syn}$, $N_{p \wedge syn}$ and $N_{n \wedge syn}$ for label pairs with both subsumption, partitive or no semantic relation and a syntactic relationships. As such we define the following values

$$P(syn|s) = \frac{N_{s \wedge syn}}{N_s}$$

Similarly we define $P(syn|p)$ and $P(syn|n)$ and present these values in table 5 for four languages.

As we can see from these results, it seems that both subsumption and partitive relationships are strongly indicative of syntactic relationships as we might expect. The second question is: is it more likely that we see a syntactic dependency in translation if we have a semantic relationship, i.e., is the syntax more likely to be preserved if these terms are semantically related. We define N_{syn_e} as the value of N_{syn} for a language e, e.g., $N_{syn_{en}}$ is the number of syntactically related English label pairs in the taxonomy. As each label has exactly one translation we can also define $N_{syn_e \wedge syn_f \wedge s}$ as the number of concepts whose labels are syntactically related in both language e and f and are semantically related by a subsumption relationship; similarly we define $N_{syn_e \wedge syn_f \wedge p}$ and $N_{syn_e \wedge syn_f \wedge n}$. Hence we can define

$$P(syn_f|syn_e, s) = \frac{N_{syn_f \wedge syn_e \wedge s}}{N_{syn_e \wedge s}}$$

erty, Plant and Equipment", we view such a relationship as semantically indicative that one term is part of another, i.e., as partitive or meronymic

And similarly define $P(syn_f|syn_e, p)$ and $P(syn_f|syn_e, n)$. We calculated these values on the IFRS taxonomies, the results of which are represented in table 6.

The partitive data was very sparse, due to the fact that only 15 concepts in the source taxonomy had a partitive relationship and were syntactically related, so we cannot draw any strong conclusions from it. For the subsumption relationship we have a clearer result and in fact averaged across all language pairs we found that the likelihood of the syntax being preserved in the translation was nearly exactly the same for semantically related and semantically unrelated concepts. From this result we can conclude that the probability of syntax given either subsumptive or partitive relationship is not very large, at least from the reduced syntactic model we used here. While our model reduces syntax to n-gram overlap, we believe that if there was a stronger correlation using a more sophisticated syntactic model, we would still see some noticable effect here as we did monolingually. We also note that we applied this to only one taxonomy and it is possible that the result may be different in a different resource. Furthermore, we note there is a strong relationship between semantics and syntax in a mono-lingual context and as such adaption of a language model to incorporate this bias may improve the translation of ontologies and taxonomies.

5.3 Comparison of ontology structure

Our third intuition in approaching ontology translation is that the comparison of ontology or taxonomy structures containing source and target labels may help in the disambiguation process of translation candidates. A prerequisite in this sense is the availability of equivalent (or similar) ontology structures to be compared.

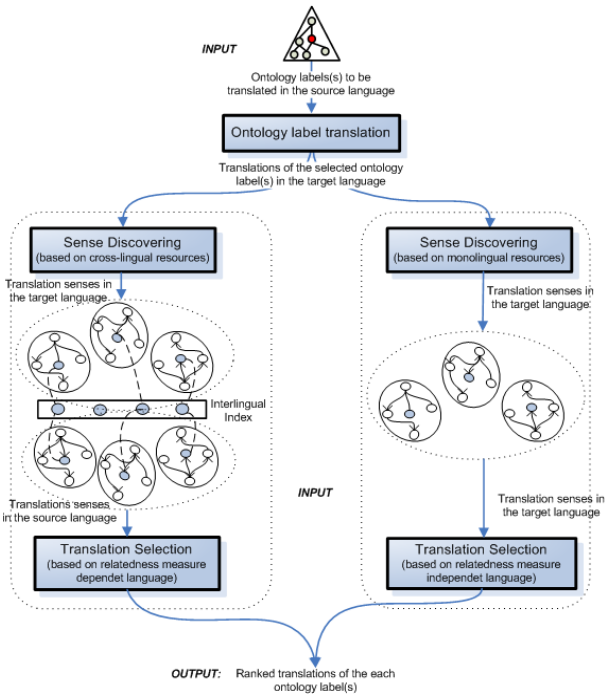

Figure 2: Two approaches to translate ontology labels.

From a technical point of view, we consider the translation task as a word sense disambiguation task. We identify two methods for comparing ontology structures, which are illustrated in Figure 2.

The first method relies on a multilingual resource, i.e., a multilingual ontology or taxonomy. The ontology represented on the left-hand side of the figure consists of several monolingual conceptualizations related to each other by means of an interlingual index, as is the case in the EuroWordNet lexicon (Vossen, 1999). For example, if the original label is *chair* for seat in English, several translations for it are obtained in Spanish such as: *silla* (for seat), *cátedra* (for university position), *presidente* (for person leading a meeting). Each of these correspond to a sense in the English WordNet, and hence each translation selects a hierachical structure with English labels. The next step is to compare the input structure of the original ontology containing *chair* against the three different structures in English representing the several senses of chair and obtain the corresponding label in Spanish.

The second method relies on a monolingual resource, i.e., on monolingual ontologies in the target language, which means that we need to compare structures documented with labels in different languages. As such we obtain a separate translated ontologies for each combination of label translations suggested by the baseline system. Selecting the correct translations is then clearly a hard optimization problem.

For the time being, we have only experimented with the first approach using EuroWordNet. Several solutions have been proposed in the context of ontology matching in a monolingual scenario (see (Shvaiko and Euzenat, 2005) or (Giunchiglia et al., 2006)). The ranking method we use to compare structures relies on an *equivalence probability measure* between two candidate structures, as proposed in (Trillo et al., 2007).

We assume that we have a taxonomy or ontology entity o_1 and we wish to deduce if it is similar to another taxonomy or ontology entity o_2 from a reference taxonomy or ontology (i.e., EuroWordNet) in the same language. We shall make a simplifying assumption that each ontology entity is associated with a unique label, e.g., l_{o_1}. As such we wish to deduce if o_1 represents the same concept as o_2 and hence if l_{o_2} is a translation for l_{o_1}. Our model relies on the Vector Space Model (Raghavan and Wong, 1986) to calculate the similarity between different labels, which essentially involves calculating a vector from the bag of words contained within each labels and then calculating the cosine similarity between these vectors. We shall denotes this as $v(o_1, o_2)$. We then use four main features in the calculation of the similarity

- The VSM-similarity between the labels of entities, o_1, o_2.
- The VSM-similarity between any glosses (descriptions) that may exist in the source or reference taxonomy/ontology.
- The hypernym similarity given to a fixed depth d, given that set of hypernyms of an entity o_i is given as a set

$$h^O(o_i) = \{h|(o_i, h) \in H\}$$

  Then we calculate the similarity for $d > 1$ recursively as

$$s_h(o_1, o_2, d) = \frac{\sum_{h_1 \in h^O(o_1), h_2 \in h^O(o_2)} \sigma(h_1, h_2, d)}{|h^O(o_1)||h^O(o_2)|}$$

$$\sigma(h_1, h_2, d) = \alpha v(h_1, h_2) + (1-\alpha) s_h(h_1, h_2, d-1)$$

And for $d = 1$ it is given as

$$s_h(o_1, o_2, 1) = \frac{\sum_{h_1 \in h^O(o_1), h_2 \in h^O(o_2)} v(h_1, h_2)}{|h^O(o_1)||h^O(o_2)|}$$

- The hyponym similarity, calculated as the hypernym similarity but using the hyponym set given by

$$H^O(o_i) = \{h | (h, o_i) \in H\}$$

We then incorporate these factors into a vector $\mathbf{x}$ and calculate the similarity of two entities as

$$s(o_1, o_2) = \mathbf{w}^T \mathbf{x}$$

Where $\mathbf{w}$ is a weight vector of non-negative reals and satisfies $||\mathbf{w}|| = 1$, which we set manually.

We then applied this to the FOAF ontology (Brickley and Miller, 2010), which was manually translated to give us a reference translation. After that, we collected a set of candidate translations obtained by using the web translation resources referenced in section 3, along with additional candidates found in our multilingual resource. Finally, we used EuroWordNet (Vossen, 1999) as the reference taxonomy and ranked the translations according to the score given by the metric above. In table 7, we present the results where our system selected the candidate translation with the highest similarity to our source ontology entity. In the case that we could not find a reference translation we split the label into tokens and found the translation by selecting the best token. We compared these results to a baseline method that selected one of the reference translations at random.

These results are in all cases significantly stronger than the baseline results showing that by comparing the structure of ontology elements it is possible to significantly improve the quality of translation. These results are encouraging and we believe that more research is needed in this sense. In particular, we would like to investigate the benefits of performing a cross-lingual ontology alignment in which we measure the semantic similarity of terms in different languages.

	Baseline	Best Translation
WER*	0.725	0.617
METEOR	0.089	0.157
NIST	0.070	0.139
BLEU	0.103	0.187

Table 7: Results of selecting translation by structural comparison. *Lower WER scores are better

6 Conclusion

In this paper we presented the problem of ontology and taxonomy translation as a special case of machine translation that has certain extra characteristics. Our examination of the problem showed that the main two differences are the presence of structured semantics and shorter, hence more ambiguous, labels. We demonstrated that as a result of this linguistic nature, some machine translation metrics do not perform as well as they do in free-text translations. We then presented the results of early investigations into how we may use the special features of taxonomy and ontology translation to improve quality of translation. The first of these was domain adaptation, which in line with other authors is useful for texts in a particular domain. We also investigated the possibility of using the link between syntactic similarity and semantic similarity to help, however although we find that mono-lingually there was a strong correspondence between syntax and semantics, this result did not seem to extend well to a cross-lingual setting. As such we believe there may only be slight benefits of using techniques, however further investigation is needed. Finally, we looked at using word sense disambiguation by comparing the structure of the input ontology to that of an already translated reference ontology. We found this method to be very effective in choosing the best translations. However it is dependent on the existence of a multilingual resource that already has such terms. As such, we view the topic of taxonomy and ontology translation as an interesting sub-problem of machine translation and believe there is still much fruitful work to be done to obtain a system that can correctly leverage the semantics present in these data structures in a way that improves translation quality.

References

Marianna Apidianaki. 2009. Data-driven semantic analysis for multilingual WSD and lexical selection in translation. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*.

Michael Ashburner, Catherine Ball, Judith Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan Davis, et al. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics*, 25(1):25–29.

Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak, and Zachary Ives. 2007. Dbpedia: A nucleus for a web of open data. *The Semantic Web*, 4825:722–735.

Satanjeev Banerjee and Alon Lavie. 2005. METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. *Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization*, page 65.

Kalina Bontcheva. 2005. Generating tailored textual summaries from ontologies. In *The Semantic Web: Research and Applications*, pages 531–545. Springer.

Dan Brickley and Libby Miller, 2010. *FOAF Vocabulary Specification 0.98*. Accessed 3 December 2010.

Marine Carpuat and Dekai Wu. 2007. Improving Statistical Machine Translation using Word Sense Disambiguation. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2007)*.

Philipp Cimiano, Elena Montiel-Ponsoda, Paul Buitelaar, Mauricio Espinoza, and Asunción Gómez-Pérez. 2010. A note on ontology localization. *Journal of Applied Ontology (JAO)*, 5:127–137.

Nigel Collier, Son Doan, Ai Kawazoe, Reiko Matsuda Goodwin, Mike Conway, Yoshio Tateno, Quoc-Hung Ngo, Dinh Dien, Asanee Kawtrakul, Koichi Takeuchi, Mika Shigematsu, and Kiyosu Taniguchi. 2008. BioCaster: detecting public health rumors with a Web-based text mining system. *Oxford Bioinformatics*, 24(24):2940–2941.

Thierry Declerck, Hans-Ullrich Krieger, Susan Marie Thomas, Paul Buitelaar, Sean O'Riain, Tobias Wunner, Gilles Maguet, John McCrae, Dennis Spohr, and Elena Montiel-Ponsoda. 2010. Ontology-based Multilingual Access to Financial Reports for Sharing Business Knowledge across Europe. In József Roóz and János Ivanyos, editors, *Internal Financial Control Assessment Applying Multilingual Ontology Framework*, pages 67–76. HVG Press Kft.

George Doddington. 2002. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of the second international conference on Human Language Technology Research*, pages 138–145. Morgan Kaufmann Publishers Inc.

Mauricio Espinoza, Asunción Gómez-Pérez, and Eduardo Mena. 2008. Enriching an Ontology with Multilingual Information. In *Proceedings of the 5th Annual of the European Semantic Web Conference (ESWC08)*, pages 333–347.

Mauricio Espinoza, Elena Montiel-Ponsoda, and Asunción Gómez-Pérez. 2009. Ontology Localization. In *Proceedings of the 5th International Conference on Knowledge Capture (KCAP09)*, pages 33–40.

Bo Fu, Rob Brennan, and Declan O'Sullivan. 2010. Cross-Lingual Ontology Mapping and Its Use on the Multilingual Semantic Web. In *Proceedings of the 1st Workshop on the Multilingual Semantic Web, at the 19th International World Wide Web Conference (WWW 2010)*.

Fausto Giunchiglia, Pavel Shvaiko, and Mikalai Yatskevich. 2006. Discovering missing background knowledge in ontology matching. In *Proceeding of the 17th European Conference on Artificial Intelligence*, pages 382–386.

International Accounting Standards Board, 2007. *International Financial Reporting Standards 2007 (including International Accounting Standards (IAS) and Interpretations as at 1 January 2007)*.

Philipp Koehn, Hieu Hoang, Alexandra Birch, Chris Callison-Burch, Marcello Federico, Nicola Bertoldi, Brooke Cowan, Wade Shen, Christine Moran, Richard Zens, et al. 2007. Moses: Open source toolkit for statistical machine translation. In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, pages 177–180.

Philipp Koehn. 2005. Europarl: A parallel corpus for statistical machine translation. In *Proceedings of the Tenth Machine Translation Summit*.

Philipp Koehn. 2010. *Statistical Machine Translation*. Cambridge University Press.

Iain McCowan, Darren Moore, John Dines, Daniel Gatica-Perez, Mike Flynn, Pierre Wellner, and Hervé Bourlard. 2004. On the use of information retrieval measures for speech recognition evaluation. Technical report, IDIAP.

John McCrae, Jesús Campana, and Philipp Cimiano. 2010. CLOVA: An Architecture for Cross-Language Semantic Data Querying. In *Proceedings of the First Mutlilingual Semantic Web Workshop*.

William Mischo. 1982. Library of Congress Subject Headings. *Cataloging & Classification Quarterly*, 1(2):105–124.

Hans-Michael Müller, Eimear E Kenny, and Paul W Sternberg. 2004. Textpresso: An ontology-based information retrieval and extraction system for biological literature. *PLoS Biol*, 2(11):e309.

Roberto Navigli and Simone Paolo Ponzetto. 2010. Babelnet: Building a very large multilingual semantic network. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 216–225.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting on association for computational linguistics*, pages 311–318. Association for Computational Linguistics.

V.Vijay Raghavan and S.K.M. Wong. 1986. A critical analysis of vector space model for information retrieval. *Journal of the American Society for Information Science*, 37(5):279–287.

Pavel Shvaiko and Jerome Euzenat. 2005. A survey of schema-based matching approaches. *Journal on Data Semantics IV*, pages 146–171.

Fabian Suchanek, Gjergji Kasneci, and Gerhard Weikum. 2007. Yago: a core of semantic knowledge. In *Proceedings of the 16th international conference on World Wide Web*, pages 697–706.

Raquel Trillo, Jorge Gracia, Mauricio Espinoza, and Eduardo Mena. 2007. Discovering the semantics of user keywords. *Journal of Universal Computer Science*, 13(12):1908–1935.

Piek Vossen. 1999. EuroWordNet a multilingual database with lexical semantic networks. *Computational Linguistics*, 25(4).

Hua Wu, Haifeng Wang, and Chengqing Zong. 2008. Domain adaptation for statistical machine translation with domain dictionary and monolingual corpora. In *Proceedings of the 22nd International Conference on Computational Linguistics-Volume 1*, pages 993–1000. Association for Computational Linguistics.