Automated Translation of Semantic Relationships

Dmitry Davidov
ICNC
Hebrew University of Jerusalem

Ari Rappoport
Institute of Computer Science
Hebrew University of Jerusalem
`arir@cs.huji.ac.il`

Abstract

We present a method for translating semantic relationships between languages where relationships are defined as pattern clusters. Given a pattern set which represents a semantic relationship, we use the web to extract sample term pairs of this relationship. We automatically translate the obtained term pairs using multilingual dictionaries and disambiguate the translated pairs using web counts. Finally we discover the set of most relevant target language patterns for the given relationship. The obtained pattern set can be utilized for extraction of new relationship examples for the target language.

We evaluate our method on 11 diverse target languages. To assess the quality of the discovered relationships, we use an automatically generated cross-lingual SAT analogy test, WordNet relationships, and concept-specific relationships, achieving high precision. The proposed framework allows fully automated cross-lingual relationship mining and construction of multilingual pattern dictionaries without relying on parallel corpora.

1 Introduction

Acquiring and understanding semantic relationships is crucial for many NLP applications. In many cases, we would like to know if a given term pair participates in a specified semantic relationship or if two different term pairs encode the same (possibly unspecified) type of relationship. Beyond the well-known major relationship types such as hyponymy (is-a) and meronymy (part-of), there is a huge number of other relationships between objects and concepts. Examples include general relations such as larger-than, contained-in, liked-by and domain specific ones such as country-language, product-manufacturer, product-seller, drug-disease etc.

The vast majority of NLP research is done in a few languages for which extensive corpora (including the web) are available. As a result, most relationship retrieval studies and lexical database compilation efforts target only a few languages. However, due to the substantial growth of the multilingual web[1] and a growing demand for NLP application coverage for less common languages, there is a need for relationship data in many less studied languages.

In this paper we address the task of translating relationships between languages, which has two obvious benefits. First, it can directly help applications such as machine translation, cross-lingual information retrieval, cross-lingual web mining and the construction and enrichment of semantic databases. Second, it can assist applications in a single language, especially when compensating for a relative scarcity of resources in that language. We focus on relations between two entities, which are the most common type.

When discussing the translation of relationships, it is important to define how these are represented and in what way the task differs from MT. While there is wide agreement on the definition and representation of major relationship types such as hypernymy and (to a lesser extent) meronymy, there is no single accepted method (or

---

[1]http://www.internetworldstats.com/stats7.htm

resources) for other less common relationships. Among the methods that have been proposed for specifying lexical relationships are natural language description and rules (Girju et al., 2007), distributional means (Turney, 2005), sample term pairs (Pasca et al, 2006), relationship instances (Banko et al., 2007) and pattern clusters (Davidov and Rappoport, 2008a).

In this paper we utilize the last definition. Following (Davidov and Rappoport, 2008a) each semantic relationship can be defined and represented by a set of lexical patterns such that the represented relation holds between entities filling the patterns' slots. We focus on pattern clusters relationship definition due to several reasons. First, as opposed to natural language descriptions, pattern clusters are formal. Second, as opposed to the other methods above, pattern clusters provide a 'generative' model for the represented relationship – it is possible to obtain from them relationship instances and term pairs, as we indeed utilize in this paper. Third, pattern clusters can be mined in a fully unsupervised manner, or in a focused manner when the relationship desired is known. Finally, pattern methods have proven to be highly efficient and effective for lexical acquisition tasks (Pantel et al, 2004; Davidov and Rappoport, 2006).

The proposed framework comprises the following stages. First, given a set of patterns defining a relationship in a source language, we obtain from the web a set of corresponding term pairs. Next, for each of the terms in the obtained term pairs, we retrieve sets of their translations to the target language using available multilingual dictionaries. Now that we have a set of translations for each term in each pair, we retrieve search engine snippets with the translated term pairs. We then select appropriate word senses using web counts, and extract a set of patterns which connect these disambiguated terms. As a result we get a set of relation-specific target language patterns, effectively obtaining the desired relationship definition. We can optionally use the retrieved pattern sets to obtain term pairs of target language relationships from the web.

We performed a thorough evaluation for various relationships involving 11 languages. We tested our framework on major relationships like meronymy, specific relationships like country-capital and unspecified unsupervisedly discovered English relationships. The obtained relationships were manually verified by human judges using cross-lingual SAT analogy questions, and a few specific factual relationships were evaluated using a gold standard.

Our main contribution is a novel framework for automated relationship translation across languages, where relationships are defined as pattern clusters or as term pairs. This framework allows fully automated cross-lingual relationship mining and construction of multilingual pattern dictionaries without relying on parallel corpora.

In Section 2 we discuss related work. Section 3 details the algorithm. Section 4 describes the evaluation, and Section 5 concludes.

2 Related work

Recently, with the development of practical applications which utilize WN-like databases in dozens of languages, great effort has been made to manually construct and interconnect such databases for different languages (Pease et al, 2008; Charoenporn et al., 2007). Some studies (e.g., (Amasyali, 2005)) use semi-automated methods based on language-specific heuristics and dictionaries.

At the same time, much work has been done on automated lexical acquisition for a single language, and in particular, on the web-based acquisition of various types of semantic relationships. There is a substantial amount of related studies which deal with the discovery of various relationship types represented in useful resources such as WordNet, including hypernymy (Pantel et al, 2004; Snow et al., 2006), synonymy (Davidov and Rappoport, 2006; Widdows and Dorow, 2002) and meronymy (Berland and Charniak, 1999; Girju et al, 2006). Since named entities are very important in NLP, many studies define and discover relations between named entities (Hassan et al., 2006). Work was also done on relations between verbs (Chklovski and Pantel, 2004). There is growing research on relations between nominals (Girju et al., 2007).

While the majority of studies focus on extracting pre-specified semantic relationships, several

recent studies were done on the automated discovery of unspecified relationship types. Thus Turney (2006) provided a pattern distance measure that allows a fully unsupervised measurement of relational similarity between two pairs of words on the same language. Banko et al. (2007) and Rosenfeld and Feldman (2007) find relationship instances where the relationships are not specified in advance. (Davidov and Rappoport, 2008a) introduced the idea that salient semantic relationships can be defined as pattern clusters, confirming it with SAT analogy test. As explained above, we use this definition in the present study. We also use pattern clusters given by (Davidov and Rappoport, 2008a) as input in our evaluation.

Most of the relationship acquisition studies were done in a single language. Those that experiment in several languages usually treat each language separately, while we extract a relationship definition for one language using the provided definition for the other language.

Our study is related to cross-language information retrieval (CLIR) frameworks. Both deal with multilingual information extracted from the Web. However, the majority of CLIR studies pursue different targets. Thus, one of the main CLIR goals is the retrieval of *documents* based on explicit queries, when the document language is not the query language (Volk and Buitelaar, 2002). These frameworks usually develop language-specific tools and algorithms including parsers, taggers and morphology analyzers in order to integrate multilingual *queries* and *documents* (Jagarlamudi and Kumaran, 2007). Our goal is to develop and evaluate a *language-independent* algorithm for the *cross-lingual translation of relationship-defining structures*. While our targets are different from those of CLIR, CLIR systems can greatly benefit from our framework, since we can translate the relationships in CLIR queries and subsequently check if the same relationships are present in the retrieved documents.

Another field indirectly related to our research is Machine translation (MT). Many MT tasks require automated creation or improvement of dictionaries (Koehn and Knight, 2001). However, MT mainly deals with translation and disambiguation of words at the sentence or document level,

while we translate relationship structures as a set of patterns, defined independently of contexts. We also perform pattern-set to pattern-set translation rather than the pattern-to-pattern or pair-to-pair translation commonly explored in MT studies. This makes it difficult to perform meaningful comparison to existing MT frameworks. However, the MT studies benefit from the proposed framework by enhancement and verification of translated relationship instances.

In (Davidov and Rappoport, 2009), we proposed a framework for automated cross-lingual concept mining. We incorporate several principles from this study including concept extension and disambiguation of query language (See Section 3.3). However our goals here are different since we target cross-lingual acquisition of relationship structures rather then concept term lists.

3 Relationship Translation Framework

Our framework has the following stages: (1) given a set of patterns in a source language defining some lexical relationship, we use the web to obtain source language term pairs participating in this relationship; (2) we automatically translate the obtained terms in each pair to the target language using available multilingual dictionaries; (3) we retrieve web snippets where these translations co-appear, disambiguating translations with web counts and extracting the corresponding patterns. As an optional final stage, the translated pattern cluster can be used to extract and extend a set of target language term pairs. Now we describe each of these stages in detail.

3.1 Acquisition of representative term pairs

We are provided with a pattern cluster, a set of patterns representing a specific lexical relationship in some language. The goal of the first stage is to discover the most representative term pairs for this cluster and language from the web. If the relationship is already specified by a representative set of term pairs, we skip this stage and continue to the next stage. Note that the method described below can also be used at the final stage to obtain representative *target* language term pairs once we obtain a *target* language pattern cluster.

The input lexical patterns are surface patterns

which include several fixed words or punctuation symbols and two slots for content words, e.g. *"the [X] of the [Y],"*. Given a cluster of patterns defining a semantic relationship, we would like to obtain from the web the most representative and frequent examples of the represented relationship. In order to do that we construct search engine queries[2] from the given patterns using wildcard symbols to represent pattern slots. For example, given a pattern *"the [X] of the [Y],"* we construct queries such as *"the * of the"; "the * * of the"*[3]. We collect all the retrieved search engine snippets and extract the appropriate term pairs found in these snippets.

Now we would like to select the most useful of the extracted pairs. Since the obtained pairs are only useful if we can translate them into the target language, we dismiss all pairs in which one or both terms have no translations to the target language in our dictionaries (see Section 3.2). Since each particular pattern can be ambiguous, we also dismiss pairs which were found for only a single pattern in the given cluster.

For the remaining term pairs we would like to estimate their specificity for the given pattern cluster. For each pattern, we retrieve and use two web hit counts: $F_{terms}(p, T1, T2)$, a hit count for co-appearance of the pair in a way similar to that in the pattern, and $F_{all}(p, T1, T2)$, the hit count of the full pattern instance.

For example, if for the pattern p=*"the * of the"* we obtain a term pair (CEO, company), then $F_{all}(p)$=*Hits("the CEO of the company")* and $F_{terms}(CEO, company)$= *Hits("CEO * * company")*. Given a pattern cluster C with patterns $\{p_1 \ldots p_n\} \in C$, we estimate the specificity of a term pair $(T1, T2)$ using the following simple probabilistic metric, giving to all patterns in the cluster an equal weight:

$$Spec(T1, T2) = \frac{1}{n} \sum_{p_i \in C} \frac{F_{all}(p_i, T1, T2)}{F_{terms}(p_i, T1, T2)}$$

We select the top 15 pairs with the highest specificity and use them in the next stage.

---

[2]We use Yahoo! Boss.

[3]Since the search engine API doesn't allow punctuation, we omit the punctuation in queries, but require a proper punctuation when processing the obtained snippet data.

3.2 Translation of the term pairs

After the previous stage we have a good representative set of term pairs for the desired source language relationship. Now we would like to translate the words in these pairs to the target language. In order to do that we use an extensive set of 1067 multilingual dictionaries developed for Star-Dict[4], including Wikipedia cross-language links and Wiktionary. For each term we obtain a set of its translations to the target language. If we get more than five different translations, we select the five having the highest number of dictionaries where this translation appears.

As discussed in Section 3.1, we dismissed terms for which no translation was found in any of the available dictionaries, so each term in each of the obtained pairs has at least a single translation to the target language. However, in many cases the available translations represent the wrong word sense, since both the source terms and their translations can be ambiguous. Thus at this stage many of the obtained term translations are irrelevant for the given relationship and require disambiguation.

3.3 Web mining for translation contexts

For this stage, we need to restrict web mining to specific target languages. This restriction is straightforward if the alphabet or term translations are language-specific or if the search API supports restriction to this language. In case where there is no such natural restrictions, we attempt to detect and add to our queries a few language-specific frequent words. Following (Davidov and Rappoport, 2009), we use our dictionaries to find 1–3 of the 15 most frequent words in a desired language[5] that are unique to that language and 'and' them with the queries to ensure proper language selection. This allows applying our algorithm to more than 60 diverse languages. The only data required for each language is at least a partial coverage of the obtained term pairs by some available dictionary.

Given a term pair $(T1, T2)$ we obtain a set of translations $(T1'_{i \in 1 \ldots n}, T2'_{j \in 1 \ldots m})$. For each combination $T1'_i, T2'_j$ of the obtained term translations, we construct and execute the following

---

[4]http://stardict.sourceforge.net/

[5]We estimated the word frequencies from text available in the corresponding multilingual dictionaries.

four queries: $\{$"$T1'_i \; * \; T2'_j$", "$T2'_j \; * \; T1'_i$", "$T1'_i \; * \; * \; T2'_j$", "$T2'_j \; * \; * \; T1'_i$"$\}$[6]. Since $Yahoo!Boss$ allows retrieval of up to the 1000 first results, we can collect up to four thousand snippets for each combination. However, the majority of these combinations return no snippets at all, effectively generating an average of a dozen snippets per query.

3.4 Pattern extraction

Now for each pair of term translations we would like to extract from the snippets all surface patterns which connect the terms in this pair. We use the basic two-slot meta-pattern type:

$$[Prefix] \; X \; [Infix] \; Y \; [Postfix]$$

X and Y should be the translated terms, Infix may contain punctuation, spaces, and up to four words (or up to eight symbols in languages without space-separated words like Chinese). Prefix and Postfix are limited to contain one or zero punctuation characters and/or up to two words. We do not allow empty Infix, Prefix of Postfix. If there are several possible combinations of Prefix and Postfix we generate a pattern set for all possible combinations (e.g., if we retrieve a snippet ...", *consider using [plexiglass] for [kitchen].*"..., we create patterns "using X for Y.", "consider using X for Y." and ", consider using X for Y.").

Now we would like to find the patterns representing the relationship in the target language. We do this in two stages. First we would like to detect the most common patterns for the given relationship. Let S_k be the union set of all patterns obtained for all combinations of the extracted translations for a specific source language term pair $k \in 1 \ldots K$. Let $Salience(p) = \frac{1}{K}|\{k|p \in S_k\}|$ be the portion of source language term pairs which lead to detection of the target language pattern p. We compute salience for each pattern, and select a subset of *salient patterns*, defined to be those whose Salience exceeds a predefined threshold (we used $1/3$). If one salient pattern is a substring of another salient pattern, we only select the longer one.

---

[6]These are Yahoo! queries where enclosing words in "" means searching for an exact phrase and "*" means a wildcard for exactly one arbitrary word.

In our salience estimation we mix data from all combinations of translations including incorrect senses and wrong translations of ambiguous terms. Now we would like to select a single correct target language pair for each source language pair in order to find more refined relationship representing patterns. For each source language term pair, we select the target language translated pair which captured the highest number of salient patterns. In case there are several pairs with the same number of salient patterns, we select a pair with the greatest web hit count. We drop term pairs with zero salient patterns.

Finally we would like to enhance the obtained set of salient patterns with more precise and representative relationship-specific patterns. Since we disambiguated the translated pairs, target language patterns captured by the remaining term pairs should be more trusted. We compare the target language pattern sets obtained for different remaining term pairs, and collect all patterns that were captured by at least three different term pairs. As before, if one pattern is a substring of another we retain only the longer one. As a result we get a comprehensive target language pattern cluster for the desired relationship.

3.5 Retrieval of target language term pairs

As an optional final stage, we can utilize the retrieved target language pattern clusters in order to discover target language term pairs for the desired relationship. We do this by utilizing the strategy described in Section 3.1 on the obtained target language pattern clusters. We do not dismiss obtained terms having no available dictionary translations, and we do not limit our search to the 15 terms with highest specificity. Instead we either select N term pairs with top specificity (where N is provided by user as in our evaluation), or we select all term pairs with specificity above some threshold.

4 Evaluation

In order to test the quality of the translated pattern clusters and the corresponding translated term pairs, we need to check both flexibility and correctness. Flexibility measures how well the retrieval works well *across languages* and for *many*

types of semantic relationships. To do that, we tested our framework on both generic and specific relationships for 11 languages. Correctness verifies that the retrieved set of target language patterns and the corresponding term pairs represent *the same semantic relationship* as the given set of source language term pairs or patterns. To do that, we used both manual cross-lingual analogy-based correctness evaluation and evaluation based of factual data.

4.1 Languages and relationships

One of the main goals in this research was to provide a fully automated and flexible framework, which requires minimal modifications when applied to different languages and relationships.

We examined an extensive set of target languages using English as a source language. Table 1 shows 11 languages used in our experiments. We included west European languages, Slavic languages like Russian, Semitic languages like Hebrew, and Asian languages such as Chinese. We developed a set of tools for automatic off-line access to an extensive set of 1067 multilingual dictionaries created for the StarDict platform. These dictionaries include recent dumps of Wikipedia cross-language links and Wiktionary data.

In our experiments we used three sets of relationships: **(1) Generic:** 15 unsupervisedly discovered English pattern clusters representing various generic relationships. **(2) H-M-C:** The three most studied relationships: hypernymy, meronymy and co-hyponymy.**(3) Specific:** Three factual relationships: country-capital, country-language and dog breed-origin. Below we describe the evaluation of each of these sets in detail. Note that our framework allows two ways of specifying a source language relationship – a pattern cluster and a set of term pairs.

4.2 Evaluation of generic pattern clusters

In our **Generic** evaluation setting, we utilized as input a random sample of 15 automatically discovered relationship definitions. We started from a set of 508 English pattern clusters, unsupervisedly discovered using the method of (Davidov and Rappoport, 2008a). Each of these clusters is assumed to represent a distinct semantic rela-

tionship. We randomly selected 15 pattern clusters from this set and executed our framework on these clusters to obtain the corresponding target language pattern clusters for each of the 11 tested languages. An example of a partial set of patterns in a cluster is: *'this [X] was kept in [Y],'; 'the X that he kept in [Y],'; 'the [X] in the [Y] and'; 'the [Y] containing the [X]'.*...

We then used the term pair selection algorithm described in Section 3.1 to select the most specific term pair for each of the 15 source language clusters and 10 pairs for each of the corresponding translated target language clusters. Thus for each of the 15 pattern clusters and for each of the 11 languages we produced a single source language term pair and up to 10 corresponding target language term pairs.

In order to check the correctness of translation of an unspecified semantic relationship we need to compare source and target language relationships. Comparison of relationships is a challenging task, since there are no relationship resources for most relationship types even in a single language, and certainly so for their translations across languages. Thus various studies define and split generic relationships differently even when describing relatively restricted relationship domains (e.g., relationships holding between parts of noun phrases (Nastase and Szpakowicz, 2003; Moldovan et al., 2004)). In order to compare generic relationships we used a manual cross-lingual SAT-like analogy human judgment evaluation[7]. This allowed us to assess the quality of the translated pattern clusters, in a similar way as (Davidov and Rappoport, 2008a) did for testing clusters in a single language.

For each of the 15 clusters we constructed a cross-lingual analogy question in the following manner. The header of the question was a term pair obtained for the source language pattern cluster. The six multiple choice items included: (1) one of the 10 discovered translated term pairs of the same cluster (the 'correct' answer)[8]; (2) three

---

[7]Using Amazon's Mechanical Turk.

[8]We avoid selection of the target language pairs which were obtained through direct translation of the source language pair given at the header of the question. This is crucial so that subjects will not judge correctness of translation but correctness of the relationship.

of the translated pairs of the other clusters among the 15; (3) a pair constructed by randomly selecting terms from different translated clusters; (4) the 6th option states that either the given options include broken words or incorrect language, or none of the presented pairs even remotely exemplifies the relationship in question. An example question for English-Italian:

The English pair: *(kennel, dog)*; (1) "correct" pair: *(acquario, pesce)*; (2)-(4) "wrong" pairs: *(topo, orecchio), (mela, rossa), (occhio, grande)*; (5) "random": *(scodella, scatola)*; (6) Pairs comprise non-Italian/broken words or no pair exemplifies the relationship

In order to check the English proficiency of the subjects we added 5 "easy" monolingual English SAT analogy questions. We also added a single hand-crafted cross-lingual question of an obvious analogy case, making a total of 16 cross-lingual questions. Subjects who failed more than one of the easy English SAT questions or failed the obvious cross-lingual question were rejected from the evaluation. Finally we have three subjects for each of the tested languages. We also asked the subjects to assign a confidence score from 0 (worst) to 10 (best) to express how well the selected term pair represents the source language relationship in question.

Language	P	% 6th	$Score_c$	$Score_w$
Chinese	71	9	9.1	1.8
Czech	73	9	8.3	2.0
French	80	10	8.4	1.9
German	68	9	8.3	1.5
Greek	72	11	8.7	2.0
Hebrew	69	11	9.0	2.5
Hindi	62	12	7.4	1.9
Italian	70	10	8.5	1.5
Russian	75	8	9.0	1.6
Turkish	61	13	9.1	2.0
Ukrainian	73	11	9.3	2.3
Average	70	10	9.1	1.9

Table 1: Averaged results for manual evaluation of 15 pattern clusters. P: precision (% of correct answers); % 6th: percentage of 6th selection; $Score_c$: averaged confidence score for correct selections; $Score_w$: confidence score for wrong selections.

We computed accuracy and agreement for the given answers (Table 1). We can see that for all languages above 61% of the choices were correct (comparing to 75% reported by (Davidov and Rappoport, 2008a) for a similar *monolingual* analogy test for the same set of pattern clusters). While the results are obviously lower than the corresponding single-language test, they are significantly above the random baseline of 20%[9]. Also note that as reported in (Turney, 2006), an average single-language highschool SAT grade is 57, which is lower than the scores obtained for our cross-lingual test. We can also see that for the correctly selected pairs the confidence score was very high, while the score for wrongly selected pairs was significantly lower.

4.3 Evaluation of the H-M-C relationships

In order to test how well our algorithm performs on the most common and useful relationships, hypernymy, meronymy and co-hyponymy, we automatically sampled from WordNet a set of 10 source language term pairs for each of these relationships and applied our framework to extract up to 100 target language term pairs for each of the three relationships as done above.

For each of the tested languages we presented to three human subjects for each language a short English definition of hypernymy, meronymy and co-hyponymy, along with the corresponding randomly selected 10 of 100 extracted pairs, and asked them to rank how well (0 (worst) to 10 (best)) each pair represents the described relationship. In order to reduce possible bias, we mixed in each set 3 randomly selected term pairs obtained for the other two relationships. Table 2 shows the average scores for this task.

Language	Hypernymy	Meronymy	Co-hyponymy	Random
Chinese	8.0	7.1	8.1	1.9
Czech	8.4	7.0	8.5	2.3
French	8.1	7.5	8.4	1.8
German	8.4	7.1	8.6	2.4
Greek	8.7	7.5	8.6	1.8
Hebrew	8.6	7.9	8.3	1.6
Hindi	7.5	7.1	7.8	2.2
Italian	7.9	7.8	8.2	1.5
Russian	8.6	8.1	8.9	1.7
Turkish	8.3	7.2	8.6	1.7
Ukrainian	8.2	7.7	8.2	1.7
Average	8.3	7.5	8.4	1.9

Table 2: Averaged results for hypernymy, meronymy and co-hyponymy translations. The three first columns show average scores for hypernymy, meronymy and co-hyponymy relationships. The last column shows scores for the random baseline.

We can see that our algorithm successfully detects the common relationships, achieving high scores. Also the results indicate that the patterns

---

[9]A reasonable random baseline omits the 6th option.

are sufficiently precise to extract at least 100 of the instances for the given salient relationships.

4.4 Evaluation of the specific relationships

To check how well our algorithm performs on some specific relationships, we examined its performance on three specific relationships explored in previous studies. We provided it with 10 source language (English) term pair examples for each of the (country, capital), (country, language) and (dog breed, origin) relationships. For each of these relationships we have factual information for every tested target language available through Wikipedia list articles. This allows us to perform an unbiased automated evaluation of the quality of the obtained target language data.

We applied our framework on these examples and generated 30 target language pairs with highest specificity for each of these relationships and languages. We compared the retrieved pairs to the factual data. Table 3 shows the precision of the results obtained for these patterns.

Language	Capital	Language	Dog breed
Chinese	0.87	0.83	0.8
Czech	0.93	0.83	0.77
French	0.97	0.9	0.87
German	0.93	0.9	0.83
Greek	0.87	0.83	0.77
Hebrew	0.83	0.8	0.8
Hindi	0.83	0.8	0.77
Italian	0.93	0.87	0.83
Russian	0.97	0.9	0.87
Turkish	0.87	0.83	0.83
Ukrainian	0.93	0.87	0.8
Average	0.9	0.85	0.81

Table 3: Precision for three specific relationship types: (country, capital), (country, language) and (dog breed,origin).

The precision observed for this task is comparable to precision obtained for Country-Capital and Country-Language in a previous single-language acquisition study (Davidov et al., 2007)[10]. The high precision observed for this task indicates that the obtained translated patterns are sufficiently good as a seed for pattern-based mining of specific relationships.

---

[10]It should be noted however that unlike previous work, we only examine the first 30 pairs and we do not use additional disambiguating words as input.

5 Conclusion

We proposed a framework which given a set of patterns defining a semantic relationship in a specific source language uses multilingual dictionaries and the web to discover a corresponding pattern cluster for a target language. In the evaluation we confirmed the applicability of our method for different languages and relationships.

The obtained set of target language pattern clusters can be used for acquisition of relationship instances as shown in our evaluation. An interesting direction for future work is to use the discovered target language pattern clusters in NLP tasks like textual entailment which require distinguishing between semantic relationships.

Applying our framework to the set of unsupervisedly discovered relationships allows a fully automated construction of a relationship dictionary, where pattern clusters in one language correspond to patten clusters in many other languages. Unlike the majority of existing machine translation systems, construction of this dictionary does not require parallel corpora. Such a dictionary can be useful for machine translation, cross-lingual textual entailment and query translation, to name just a few applications. In the future we plan to create a multilingual pattern cluster dictionary which interconnects pattern clusters from many languages and allows cross-lingual definition of lexical relationships.

References

Amasyali Fatih, 2005. Automatic Construction of Turkish Wordnet. *Signal Processing and Communications Applications Conference*.

Mishele Banko, Michael Cafarella , Stephen Soderland, Matt Broadhead, Oren Etzioni, 2007. Open information extraction from the Web. *IJCAI '07*.

Matthew Berland, Eugene Charniak, 1999. Finding parts in very large corpora. *ACL '99*.

Thatsanee Charoenporn, Virach Sornlertlamvanich, Chumpol Mokarat, and Hitoshi Isahara, 2008. Semi-automatic Compilation of Asian WordNet. *Proceedings of the 14th NLP-2008, University of Tokyo, Komaba Campus, Japan.*

Timothy Chklovski, Patrick Pantel, 2004. VerbOcean:

mining the web for fine-grained semantic verb relations. *EMNLP '04.*

Dmitry Davidov, Ari Rappoport, 2006. Efficient unsupervised discovery of word categories using symmetric patterns and high frequency words. *COLING-ACL '06.*

Dmitry Davidov, Ari Rappoport and Moshe Koppel, 2007. Fully Unsupervised Discovery of Concept-Specific Relationships by Web Mining. *ACL '07.*

Dmitry Davidov, Ari Rappoport. 2008a. Unsupervised Discovery of Generic Relationships Using Pattern Clusters and its Evaluation by Automatically Generated SAT Analogy Questions. *ACL '08.*

Dmitry Davidov and Ari Rappoport, 2008b. Classification of relationships between nominals using pattern clusters. *ACL '08.*

Dmitry Davidov and Ari Rappoport, 2009. Translation and Extension of Concepts Across Languages. *EACL '09.*

Roxana Girju, Adriana Badulescu, and Dan Moldovan, 2006. Automatic discovery of part-whole relations. *Computational Linguistics*, 32(1).

Roxana Girju, Marthy Hearst, Preslav Nakov, Vivi Nastase, Stan Szpakowicz, Peter Turney and Yuret, D., 2007. Task 04: Classification of semantic relations between nominal at SemEval 2007. *4th Intl. Workshop on Semantic Evaluations (SemEval '07), in ACL '07.*

Hany Hassan, Ahmed Hassan and Ossama Emam, 2006. Unsupervised information extraction approach using graph mutual reinforcement. *EMNLP '06.*

Jagadeesh Jagarlamudi, A Kumaran, 2007 Cross-Lingual Information Retrieval System for Indian Languages *Working Notes for the CLEF 2007 Workshop.*

Philipp Koehn and Kevin Knight. 2001. Knowledge sources for word-level translation models. *EMNLP '01.*

Dan Moldovan, Adriana Badulescu, Marta Tatu, Daniel Antohe, and Roxana Girju, 2004. Models for the semantic classification of noun phrases. *HLT-NAACL '04 Workshop on Computational Lexical Semantics.*

Vivi Nastase, Stan Szpakowicz, 2003. Exploring noun-modifier semantic relations. *In Fifth Intl. Workshop on Computational Semantics (IWCS-5).*

Patrick Pantel, Deepak Ravichandran, Eduard Hovy, 2004. Towards terascale knowledge acquisition. *COLING '04.*

Marius Pasca, Dekang Lin, Jeffrey Bigham, Andrei Lifchits, Alpa Jain, 2006. Names and similarities on the web: fact extraction in the fast lane. *COLING-ACL 06.*

Adam Pease, Christiane Fellbaum, Piek Vossen, 2008. Building the Global WordNet Grid. *CIL18.*

Benjamin Rosenfeld , Ronen Feldman, 2007. Clustering for unsupervised relation identification. *CIKM '07.*

Rion Snow, Daniel Jurafsky, Andrew Ng, 2006. Semantic taxonomy induction from heterogeneous evidence. *COLING-ACL '06.*

Peter Turney, 2005. Measuring semantic similarity by latent relational analysis, *IJCAI '05.*

Peter Turney, 2006. Expressing implicit semantic relations without supervision. *COLING-ACL '06.*

Martin Volk, Paul Buitelaar, 2002 A Systematic Evaluation of Concept-Based Cross-Language Information Retrieval in the Medical Domain. *In: Proc. of 3rd Dutch-Belgian Information Retrieval Workshop.* Leuven.

Dominic Widdows, Beate Dorow, 2002. A graph model for unsupervised Lexical acquisition. *COLING '02.*