

An unsupervised approach for semantic relation interpretation

Emiliano Giovannetti

Istituto di Linguistica Computazionale - CNR

via G. Moruzzi 1, Pisa

E-mail: emiliano.giovannetti@ilc.cnr.it

Abstract

In this work we propose a hybrid unsupervised approach for semantic relation extraction from Italian and English texts. The system takes as input pairs of “distributionally similar” terms, possibly involved in a semantic relation. To validate and label the anonymous relations holding between the terms in input, the candidate pairs of terms are looked for on the Web in the context of reliable lexico-syntactic patterns. This paper focuses on the definition of the patterns, on the measures used to assess the reliability of the suggested specific semantic relation and on the evaluation of the implemented system.

1. Introduction

Semantic relation extraction is a crucial task in Ontology Learning from Text. In the literature, two main approaches can be identified. On the one hand, distributional approaches typically detect pairs of semantically related terms on the basis of their distribution in texts, on the other hand, pattern-based approaches look for semantically related terms appearing in lexico-syntactic patterns inside linguistically annotated texts.

Most state-of-the-art systems use either a distributional approach or a pattern-based approach. Distributional systems study co-occurrence distributions of words in order to calculate a semantic distance between the concepts represented by those words. This distance metric can be used for conceptual clustering (Faure & Nédellec, 1990), Formal Concept Analysis (Cimiano & Staab, 2004a), for classifying words inside existing ontologies (Pekar & Staab, 2003) and to learn concept hierarchies (Widdows, 2003). Concerning pattern-based systems, Hearst (1992) pioneered using patterns to extract hypernymy relations. Berland and Charniak (1999) applied the same technique concerning meronymy. More recently Girju et al. (2006) studied meronymic relations extraction while Turney (2008) proposed a uniform approach for the extraction of different kinds of relations from text. In (Ruiz-Casado et al., 2007) Wikipedia is used for the extraction of semantic relations to integrate inside the WordNet ontology. Some works make use of very large corpora, like the Web. Cimiano and Staab (2004b) describe a system that generates instances of lexico-syntactic patterns indicating specific relations and counts their occurrences in the WWW using the Google™ API. Pantel and Pennacchiotti (2006) propose a pattern matching algorithm to harvest semantic relations exploiting information redundancy of the Web to filter the matches of general patterns using reliable patterns.

In this work we propose a hybrid unsupervised approach for semantic relation extraction from Italian and English texts, where the concepts of “closed” and “open patterns” are being introduced. The system takes as input

pairs of “distributionally similar” terms, possibly involved in a semantic relation. To validate and label the anonymous relations holding between the terms in input, the candidate pairs of terms are looked for in a “support corpus” (in this case, the Web) in the context of reliable “low recall but high precision” lexico-syntactic patterns (RPs, in the following).

This work focuses on the definition and application of the lexico-syntactic patterns, on the measures used to assess the reliability of the specific semantic relation the system suggests and on the evaluation of the system. So far, the system is able to extract the following types of semantic relations: hyponymy, meronymy, and co-hyponymy. The approach can however be easily extended to manage other relations by defining the appropriate battery of RPs.

Accuracy of the RelEx system was found to be very promising, scoring 83.3% for hyponymy, 75% for meronymy and 72.2% for co-hyponymy extraction.

2. The approach

In the proposed approach two kinds of lexico-syntactic reliable patterns have been defined: “closed”, and “open” patterns, depending on the way they are applied to query the Web. In particular, Closed Patterns (CPs, in the following) are used for hyponymy and meronymy discovery and they contain both the candidate related words. Some examples of reliable CPs are shown in Table 1.

	Italian version	English version
h_1)	T_1 (e OR o) altri T_2	T_1 (and OR or) other T_2
h_2)	T_1 è [R] T_2	T_1 is [R] T_2
m_1)	T_1 è [R] parte [E A AE BAE] T_2	T_1 is [R RA RBA] part [E ER] T_2
m_2)	T_2 è (costituito OR costituita) da T_1	T_2 is made up of T_1

Table 1: Examples of Closed Patterns for hyponymy and meronymy relation extraction.

Patterns h_1 and h_2 are used to verify if T_1 is an hyponym of

T_2 . Pattern h_1 includes an “OR” operator: used to query Google™ it will match both “*cats and other felines*” and “*cats or other felines*” word sequences. Pattern h_2 includes a Part-Of-Speech placeholder, “R”, standing for “article”: h_2 will therefore match sequences like “*cat is a feline*” and “*cat is the feline*”.

Patterns m_1 and m_2 have the purpose of identifying if a meronymy relation holds between terms T_1 and T_2 (“E” stands for preposition, “A” for adjective and “B” for adverb). Examples of word sequences matching pattern m_1 may be: “the *nucleus* is the part of the *cell*” and “the *nucleus* is the most important part of the *cell*”.

2.1 Definition of patterns

The lexico-syntactic patterns have been defined at hand: the corpus (being the whole Web) is so large that a few patterns are enough to obtain a significant number of matches (see 2.2). Pattern-based relation extraction systems using smaller corpora (often linguistically annotated) need many more patterns to obtain relevant results: for this reason, a lot of work has been devoted to automatic pattern extraction (Casado et al., 2005; Mititelu, 2006).

The procedure we have adopted for the definition of RPs was inspired by Marti Hearst (1998):

- 1) decide on a semantic relation of interest,
- 2) decide a list of word pairs from WordNet in which this relation is known to hold,
- 3) extract sentences from the Web in which these words both occur, and record the lexical and syntactic context;
- 4) find the communalities among these contexts and hypothesize that the common ones yield patterns that indicate the relation of interest.

Experiments have shown that to isolate useful reliable lexico-syntactic patterns it is better to use non polysemic and domain-specific words, or “domain terms”.

2.2 The choice of the Web as the support corpus

Several experiments have been done concerning the choice of the most appropriate support corpus where RPs should have been applied on. As a basic assumption, we decided to use raw (not linguistically annotated) support corpora. We have taken this choice because, except for a few domains (such as the biomedical one), large annotated domain-specific corpora are not available and the greatest part of the them cannot be found for languages other than English. Since the objective of this work is to provide support in the process of ontology learning, domain-specific corpora are needed.

Since the initial design of the approach, it was evident that the most appropriate corpus would be the Web. However, some experiments have been conducted using large unannotated texts, also to provide an objective comparison with the WWW.

As expected, the application of RPs on raw corpora provided very poor results. Apart from the very few

matches obtained, the computational load required to search strings inside large texts was excessive, especially when using regular expressions. To overcome this problem, some experiments have been done by indexing the corpora using the Google Desktop™ freeware application: search time was cut to (almost) zero, but the application didn’t allow the use of wildcards inside the query, thus strongly limiting the application of RPs.

Apart from being, unfortunately, linguistically unannotated, the Web as a support corpus provides a lot of advantages:

- **size.** The (indexed) Web is several orders of magnitude larger than any other available collection of documents. The number of web sites indexed (by the most popular search engines) in the beginning of 2010 is estimated in more than 50 billions¹, each one containing a variable quantity of text.
- **languages.** Though most of web contents is in English (56.4% on the basis of a survey conducted in 2002²) it is possible to have access to millions of documents written in hundreds of different languages.
- **content.** The Web is composed of documents belonging to all existent domains of knowledge, incorporating an enormous variety of domain-specific (and very specific) corpora.
- **high redundancy.** The amount of repetition of information can represent a measure of its relevance: we cannot trust the information contained in an individual website, but we can give more confidence to a fact that is enounced by a considerable amount of possibly independent sources.
- **evolution.** The Web is not a static corpus: it evolves over time thanks to the daily contribution of millions of “text producing” users, in every domain of interest. For this reason, the Web is also the most up-to-date corpus, including, for example, neologisms as they are being produced.

3. The system

The implemented system, called **RelEx (Relation Extractor)**, takes as input a set of candidate related pairs of terms and classifies them with respect to the selected semantic relations (hyponymy, meronymy and co-hyponymy). The system makes use of two sub-components: the first one, called **CPM (Closed Patterns Module)**, uses CPs to verify if a relation of hyponymy or meronymy holds between the terms.

The second component, the **OPM (Open Patterns Module)**, is used to look for hypernyms of a specific term T . In this case, open lexico-syntactic patterns (OPs, in the following) are used. The nouns appearing in the word

¹ <http://www.worldwidewebsite.com/>

² <http://www.netz-tipp.de/languages.html>

sequences that result from the Web query and that immediately follow the pattern can be considered as hypernyms of T. It is possible to consider just the nouns or various combination of noun-adjective or adjective-noun.

For example, given the term “cyclosporine” and the OP “T is a” the Web can be queried using the pattern “cyclosporine is a”. Some of the candidate obtained hypernyms are: *drug*, *immunomodulator* and *medication*.

To verify if a co-hyponymy relation holds between two terms, the OPM is applied to both terms T_1 and T_2 . The two sets of candidate hypernyms are then compared: if the two terms share a significant number of hypernyms they can be considered co-hyponyms (see next paragraph).

The OPM could also be used to look for meronyms of a particular term: it can be accomplished by simply changing the battery of OPs with patterns such as “T is a part of”. Current work involves the upgrade of the system for co-meronymy extraction.

The RelEx system collects the information coming from the two sub-components and provides the most probable semantic relation (if present) holding between each pair of terms.

4. Definition of the scoring functions

Two different scoring functions have been defined, one for hypernymy and meronymy relations and the other for co-hyponymy. Concerning CPM, the score provides the number of matches for each relation. For example, given the terms “gatto” (*cat*) and “felino” (*feline*) and applying CPM using Italian versions of h_1 and m_1 results are 23 matches for h_1 and 0 matches for m_1 , thus indicating 100% confidence for an hyponymy relation holding between the two terms.

Accuracy of CPM is good: even when using just pattern h_1 (not as reliable as h_2) thanks to the high redundancy of the Web as a corpus it is possible to identify hyponymy or meronymy relations with very high precision.

Measures developed for OPM take into account the number of common hypernyms shared between the two terms. Let’s see some examples. Given the terms “tigers” and “lions”³ we applied the English pattern “T (and OR or) other” to both terms obtaining the results summarized in Table 2, taking the first 15 hypernyms ordered by frequency. Common hypernyms appear in bold.

hypernyms of “tiger”	freq	hypernyms of “lion”	freq
animal	28	predator	21
cat	24	animal	17
wildlife	8	cat	15
predator	6	wildlife	7
creature	6	carnivore	5
species	4	thought	2
team	4	country	1
mammal	4	thing	1

³ Using the open version of pattern h_1 we found out better results when using the plural forms of terms.

application	3	organization	1
beast	2	tale	1
kind	2	game	1
asian	1	source	1
thing	1	beasty	1
game	1	club	1
performer	1	individual	1

Table 2: Hypernyms of “tiger” and “lion” obtained by applying the open variant of pattern h_1 .

The correlation measure we have defined takes into account both the number of distinct common hypernyms and the relative frequency of each hypernym. In the example of Table 2, for instance, we have 6 distinct terms representing common hypernyms of “tigers” and “lions”. Furthermore, we weigh each term considering its relative frequency, thus giving more relevance, for example, to “animal” hypernym than to “wildlife”.

Given the set of common hypernyms $\{y_1, y_2, \dots, y_k\}$, n_1 and n_2 the total number of candidate hypernyms for, respectively, terms T_1 and T_2 , the relative frequency of term y_j for term T_1 defined as f_j^1 we can define the correlation measure for T_1 and T_2 as:

$$Corr(T_1, T_2) = \sum_{i=1}^k \left(\frac{f_1^i}{n_1} + \frac{f_2^i}{n_2} \right) \frac{f_1^i \cdot f_2^i}{n_1 \cdot n_2} * K$$

where K is a normalizing constant used to compare the score obtained for co-hyponymy with the other two scores (relative to hyponymy and meronymy).

The measure ranges from 0 (no common hypernyms) to $2 * K$ (every term is a shared hypernym). Terms “tigers” and “lions” get a score of $0.042 * K$ while “tigers” and “pines”, for example, get a score of $1.03 \cdot 10^{-4} * K$ and “tigers” and “chairs” get a score of $5.53 \cdot 10^{-6} * K$. Experiments conducted up to now suggest a score greater than $0.002 * K$ as an indication for positive co-hyponymy.

Table 3 reports some examples of term pairs extracted via the distributional system applied to a corpus in the History of Art domain and ranked with the aforementioned correlation measure. Constant K has been set to 1000, thus fixing the threshold to 2.

term pair (T_1, T_2)	Corr(T_1, T_2)
(door, road)	0.36700472
(<i>temple</i> , <i>vault</i>)	8.524048
(order, painting)	0.24807052
(loan, head)	0.08190192
(<i>sculpture</i> , <i>painting</i>)	17.0844
(<i>oratory</i> , <i>temple</i>)	5.7685953

Table 3: Examples of co-hyponymy correlation measures. Constant K set to 1000.

In the examples terms appearing in pairs (*temple*, *vault*), (*sculpture*, *painting*) and (*oratory*, *temple*) would be labeled as co-hyponyms, since the relative correlation measure is greater than 2.

Once the RelEx system has classified a term pair on the basis of each of the semantic relations of interest, in case of multiple positive results it has to choose one relation as the “winner”. For each term pair, five possible results can be obtained:

- *hypo*: the number of snippets where hyponymy RPs have matched
- *hyper*: the number of snippets where hypernymy RPs have matched
- *mero*: the number of snippets where meronymy RPs have matched
- *holo*: the number of snippets where holonymy RPs have matched
- *co-hypo*: the co-hyponymy correlation measure.

In brief, the greatest number is taken as the final result. In any case, a co-hypo lesser than 2 (the fixed threshold) is considered a negative result and would not be counted.

5. Evaluation

Before the final evaluation, a preliminary testing of the two sub-components composing the system has been carried out (Giovannetti et al., 2008), aimed at the definition of the most appropriate evaluation measures and to better understand which aspects of the approach could be refined. Those first experiments involved also the use of generic “high recall but low precision” lexico-syntactic patterns, as defined by Pantel and Pennacchiotti (2006).

5.1 Evaluating the system

In general, evaluating results of an automatic ontology learning methodology is a difficult task: either an expert opinion is needed to check the results manually or an ontological repository is required to perform any automatic evaluation.

The biggest and most widely used general purpose English and Italian repositories are WordNet (Miller, 1995) and ItalWordNet (Roventini et al., 2000). They offer a lexicon, a thesaurus and semantic linkage between terms. As detailed in the following, both human testing and WordNet have been used for the evaluation of the system.

Most of the systems working with lexico-syntactic patterns and described in the literature aim at evaluating the extraction of relation instances on the basis of precision. Marti Hearst, who pioneered the use of patterns, reported a 52% of precision in her works on hyponymy extraction (Hearst, 1992). A more recent variant of this technique was implemented by Alfonseca and Manandhar (2002) who compared the collocational patterns of words from *The Lord of the Rings* with those of words appearing WordNet, adding new nouns to WordNet with an accuracy of 28%. Cederberg and Widdows (2003) applied latent semantic analysis (LSA) to filter extracted hyponymy relations obtained with their system thus reducing the rate of error by 30%, achieving precision of 58%.

5.2 Combining human testing and ItalWordNet

We asked 3 subjects (raters) to consider a set of 100 Italian term pairs, randomly extracted from the output of the distributional system applied to an Italian corpus of 269,000 words in the History of Art domain. Each rater has been asked to assign a tag to each pair of terms (T_1 , T_2), choosing from the following predefined list:

- a) *hyponymy*: an hyponymy relation holds between T_1 and T_2
- b) *meronymy*: a meronymy relation holds between T_1 and T_2
- c) *co-hyponymy*: T_1 and T_2 are co-hyponyms
- d) none of the above relations holds between T_1 and T_2

To reduce the number of choices, hypernymy and holonymy relations (the inverse of hyponymy and meronymy) have been omitted. Indeed, once the RelEx system detects an hyponymy relation, it can establish which term is the hyponym and which is the hypernym with very high accuracy (except in presence of highly generic and ambiguous terms). The same thing happens concerning meronymy. In this perspective, it was sufficient for the rater to indicate that an hyponymy (or meronymy) relation held between the two terms to compare results. Of course, no such issues regard choices c) and d).

Raters had also to couple each relation with a label “direct” or “indirect” where:

- **direct**: a “direct” relation holds between the two terms, i.e. for hyponymy and meronymy no other terms may appear between T_1 and T_2 . For example, concerning hyponymy, it makes no sense to interpose a term between “Ministry” and “Ministry of the Interior”, or between “tree” and “acacia”. Similarly, “treetop” is a direct meronym of “tree”. We intend direct co-hyponymy if the two terms share a direct hypernym: for example, “leopard” and “jaguar” can be considered direct co-hyponyms of “feline”.
- **indirect**: an “indirect” relation holds between the two terms. In the case of hyponymy and meronymy, one or more terms can appear between T_1 and T_2 . Of course, a reasonably few number of terms may be interposed between the two terms: the term “cat”, for example, can be considered an indirect hyponym of “animal”, since a few other terms may be interposed between them (like “mammal”). In the case of co-hyponymy, it was necessary to pay attention, since every term can be considered co-hyponym of any other term: “ruby” and “chicken”, for instance, are co-hyponyms of the term “object”, but such a relation would be of no use. On the other hand, terms “wolf” and “chicken” can be labelled as indirect co-hyponym (for example, of “animal”).

Consensus among raters was good, with relevant differences just in the assignment of the direct/indirect label. Table 4 summarizes the agreement on tags, where “tag total agreement” stands for the percentage of cases of 3 raters on 3 assigning the same tag and “tag partial agreement” for 2 raters on 3.

	tag total agreement	tag partial agreement
hyponymy	79.2%	20.8%
meronymy	50%	50%
co-hyponymy	54.5%	45.5%

Table 4: Agreement among raters on tag assignment.

As it is evident from the table, greater agreement could be found about hyponymy recognition (see also Table 6).

5.2.1. Definition of the test set

Correct relations to be used as reference for the evaluation of the system were established following this criterion:

- if terms T_1 and T_2 are both present in ItalWordNet and one semantic relation among hyponymy, meronymy and co-hyponymy holds between them take it as the correct relation;
- if terms are not present in ItalWordNet, or they are present but no semantic relation holds between them, take as correct, by majority, the tag assigned by the human raters.

The obtained results have been compared with the ones automatically assigned to the same test set by the RelEx system. The accuracy of the RelEx system calculated on each single semantic relation is summarized in Table 5.

	Hypo	Mero	Co-Hypo	No Rel.
extracted	20	9	8	45
not extracted	4	3	3	8
total	24	12	11	53
accuracy	83.3%	75%	72.7%	84.9%

Table 5: Accuracy of the RelEx system.

We expected better results for co-hyponymy extraction. The obtained score was probably related to the presence of generic and highly ambiguous terms inside the test set, like, for example, “ricerca, lavoro” (“research, work”). The distributional system we have used to generate the Related Terms is currently being upgraded to incorporate contrastive domain techniques, through which it will be possible to filter out most of the generic terms.

5.2.2. Direct/indirect relations

Direct/indirect label assigned by the raters have been compared to each other. Just the 47 term pairs marked as semantically related (and used as reference in the evaluation) have been considered. Table 6 summarizes the agreement among raters concerning direct/indirect assignment, calculated similarly to tags of Table 4.

	d/i total agreement	d/i partial agreement
hyponymy	45.8%	54.2%
meronymy	33.3%	66.7%
co-hyponymy	36.3%	63.7%

Table 6: Agreement among raters direct/indirect label assignment.

Similarly to assignment of tags, consensus among raters was greater relatively to hyponymy, probably indicating hyponymy detection as an easier task.

We also investigated the possible correlations between direct/indirect relations assigned by the raters and the scores automatically assigned by the RelEx system to each term pair. Just hyponymy provided a sufficient number of total direct relation agreement (11 term pairs). Setting a threshold of $hypo = 6$ (where $hypo$ stands for the number of matched snippets found on the Web), the percentage of direct raters assignment was 72.7%. In other words, 72.7% of the term pairs labelled with the “hyponymy” tag by the RelEx system and with a $hypo$ value greater than 6, were tagged as “direct” by the raters, showing some kind of correlation between the scores assigned by the system and the direct/indirect raters assignment. Anyway, further investigations are needed to confirm this hypothesis.

5.3 Using RelEx to extend lexical ontologies

Novel semantic relation instances, not present in WordNet (and ItalWordNet for Italian) can be automatically detected and labelled, proving that the RelEx system can be used for concrete applications in the ontology learning from text domain.

In the experiments done for the evaluation, some of the semantically related term pairs that have been extracted could contribute to extend the ItalWordNet lexical ontology. For example:

- “mostra, retrospettiva” (in English, “*exhibition, retrospective*”): the word “retrospettiva” does not even appear in ItalWordNet. The RelEx system indicate it as co-hyponym of “mostra”. It is even possible to suggest the possible co-hypernym term, by considering the common hypernyms the system has found between the two terms. In this case, just one common hypernym is shared, the term “event”, found 12 times as hypernym of “mostra” and 3 times as hypernym of “retrospettiva”. Using “event” as co-hypernym it is easy to select the correct sense of “mostra” to be used (see next example).
- “padiglione, mostra” (in English, “*pavilion, exhibition*”): both terms appear in ItalWordNet, however, relatively to “mostra”, the sense of “place” is missing (while several dictionaries report it), being present just the senses of “act” (the act of showing) and “event”. In this case, the system would suggest to add a sense to the term “mostra” and to set it as holonym of “padiglione”.

6. Conclusions

In this paper we have introduced a hybrid unsupervised approach for semantic relation extraction from Italian and English texts. The implemented system, called RelEx, takes as input pairs of “distributionally similar” terms, possibly involved in a semantic relation. To validate and label the anonymous relations holding between the terms in input, the candidate pairs of terms are looked for on the Web in the context of reliable “low recall but high precision” lexico-syntactic patterns.

To evaluate the system, two different scoring

functions have been defined, one for hypernymy and meronymy relations and the other for co-hyponymy. Besides, a test set was defined by asking to 3 human raters to tag a set of 100 Italian term pairs, randomly extracted from the output of the distributional system applied to an Italian corpus in the History of Art domain.

Accuracy of the ReLEx system was found to be very promising, scoring 83.3% for hyponymy, 75% for meronymy and 72.2% for co-hyponymy extraction. Results appear to be strongly related to the quality of the term pair set in input: the more generic and ambiguous they are, the more difficult it is to correctly detect the semantic relation (if any) holding between the two terms. Current work includes the upgrade of the distributional system, where contrastive domain techniques will be applied to the terminology extraction component, thus reducing the number of generic terms appearing in the term pair set.

7. References

- Alfonseca, E., Manandhar, S. (2002). Extending a Lexical Ontology by a Combination of Distributional Semantics Signatures. In *Proceedings of EKAW'02*, pp. 1--7.
- Berland, M., Charniak, E. (1999). Finding parts in very large corpora. In *Proceedings of the 37th ACL*. University of Maryland, pp. 57--64.
- Casado, R. M., Alfonseca, E., Castells, P. (2005). Automatic Extraction of Semantic Relationships for WordNet by Means of Pattern Learning from Wikipedia. In *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems*, NLDB 2005. Alicante, Spain.
- Cederberg, S., Widdows, D. (2003). Using LSA and Noun Coordination Information to Improve the Precision and Recall of Automatic Hyponymy Extraction. In *Proceedings of CoNLL-2003*. Edmonton, Canada, pp. 111--118.
- Cimiano, P., Staab, S. (2004a). Clustering concept hierarchies from text. In *Proceedings of LREC-2004*. Lisbon, Portugal.
- Cimiano, P., Staab, S. (2004b). Learning by googling. *ACM SIGKDD Explorations Newsletter*, 6(2). pp. 24--33.
- Faure, D., Nédellec, C. (1998). A corpus-based conceptual clustering method for verb frames and ontology acquisition. In P. Velardi (Ed.), *Adapting lexical and corpus resources to sublanguages and applications*, Workshop of the 1st Intl. Conf. on Language Resources and Evaluation. Granada, Spain, pp. 1-8.
- Giovannetti, E., Marchi, S., and Montemagni, S. (2008). Combining Statistical Techniques and Lexico-syntactic Patterns for Semantic Relations Extraction from Text. In *Proceedings of the 5th Workshop on Semantic Web Applications and Perspectives (SWAP2008)*. Rome, Italy.
- Girju, R., Badulescu, A., Moldovan, D. (2006). Automatic Discovery of Part-Whole Relations. *Computational Linguistics*, 32(1), pp. 83--135.
- Hearst, M. A. (1992). Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the 14th International Conference on Computational Linguistics*. Nantes, France, pp. 539--545.
- Hearst, M. (1998). Automated Discovery of WordNet Relations. MIT Press, Cambridge MA.
- Miller, G. (1995). Wordnet: A lexical database. *Communication of the ACM*, 38, pp. 39--41.
- Mititelu, V.B. (2006). Automatic Extraction of Patterns Displaying Hyponym-Hypernym Co-Occurrence from Corpora. In *Proceedings of the First CESCL*. Budapest, Hungary.
- Pantel, P., Pennacchiotti, P. (2006). Espresso: Leveraging generic patterns for automatically harvesting semantic relations. In *Proceedings of ACL 2006*. Sydney, Australia, pp. 113--120.
- Pekar, V., Staab, S. (2003). Word classification based on combined measures of distributional and semantic similarity. In *Proceedings of EACL03*. Budapest.
- Roventini A., Alonge A., Calzolari N., Magnini B., Bertagna F. (2000). ItalWordNet: a Large Semantic Database for Italian. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation (LREC 2000)*. Athens, Greece, pp. 783--790.
- Ruiz-Casado, M., Alfonseca, E., Castells, P. (2007). Automatising the learning of lexical patterns: An application to the enrichment of WordNet by extracting semantic relationships from Wikipedia. *Data & Knowledge Engineering*, 61(3), pp. 484--499.
- Turney, P.D. (2008). A uniform approach to analogies, synonyms, antonyms, and associations. In *Proceedings of Coling 2008*. Manchester, UK, pp. 905--912.
- Widdows, D. (2003). Unsupervised methods for developing taxonomies by combining syntactic and statistical information. In *Proceedings of the HLT/NAACL 2003 - Vol. 1*. pp. 197-204.