Combining Statistical Techniques and Lexico-syntactic Patterns for Semantic Relations Extraction from Text

Emiliano Giovannetti, Simone Marchi, Simonetta Montemagni

Istituto di Linguistica Computazionale, Consiglio Nazionale delle Ricerche, via Moruzzi 1, 56100 Pisa, Italy
{Emiliano.Giovannetti, Simone.Marchi, Simonetta.Montemagni}@ilc.cnr.it

Abstract. We describe here a methodology to combine two different techniques for Semantic Relation Extraction from texts. On the one hand, generic lexico-syntactic patterns are applied to the linguistically analyzed corpus to detect a first set of pairs of co-occurring words, possibly involved in “syntagmatic” relations. On the other hand, a statistical unsupervised association system is used to obtain a second set of pairs of “distributionally similar” terms, that appear to occur in similar contexts, thus possibly involved in “paradigmatic” relations. The approach aims at learning ontological information by filtering the candidate relations obtained through generic lexico-syntactic patterns and by labelling the anonymous relations obtained through the statistical system. The resulting set of relations can be used to enrich existing ontologies and for semantic annotation of documents or web pages.

Keywords: Ontology Learning from Text, Semantic Relation Extraction, Lexico-syntactic Patterns, Distributional Similarity.

1 Introduction

Learning ontologies from text is a crucial task in the Semantic Web scenario: if data is semantically annotated with respect to an ontology, it can be shared between different parties on the basis of its meaning and it can be searched and retrieved in a more effective way. The state-of-the-art in ontology learning involves advanced Natural Language Processing (NLP) technologies and includes a series of tasks, starting from terminology extraction and concept definition to more complex ones such as learning taxonomic and non-taxonomic relations.

Concerning relation extraction, two main approaches can be distinguished, pattern based and distributional, both of which applied to linguistically analyzed documents.

In this paper we propose a methodology for Semantic Relation (SR) extraction from texts which combines these two complementary approaches by looking for both “syntagmatic” and “paradigmatic” relations inside textual corpora. In particular, generic “high recall but low precision” lexico-syntactic patterns are applied to the linguistically analyzed (i.e. shallow parsed) corpus to detect a first set of pairs of co-occurring words. These words, appearing close to each other inside a sentence, are involved in a “syntagmatic” relation, such as, for example, *steer* and *car* in the sentence “*steer is part of the car*”. On the other hand, a statistical unsupervised association system is used to obtain a second set of pairs of “distributionally similar” terms, that appear to occur in similar contexts, and possibly involved in “paradigmatic” relations (this is the case, for instance, of *car* and *motorcycle* in the sentences “*I drive my car*” and “*Bob drives his motorcycle*”.

Current research is aimed at learning ontological information by filtering the candidate relations obtained through generic lexico-syntactic patterns and by labelling the anonymous relations obtained through the statistical system. Among the possible solutions we envisage the use of candidate pairs of words inside reliable patterns (i.e. low recall but high precision patterns) to be projected onto the Web, possibly automatically learnt using machine learning technologies.

2 Related work

Automatic extraction of information from textual corpora for ontology learning is now a well-known field with many different applications. Concerning SR extraction, we may broadly classify current approaches in two groups:

Systems based on distributional properties of words. They consist in studying 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 example, for conceptual clustering ([16], [10]), Formal Concept Analysis ([8]), for classifying words inside existing ontologies ([20], [1]) and to learn concept hierarchies ([6], [28]). On the other hand, [17] learn association rules from dependency relations between words which, combined with heuristics, are used to extract non-taxonomic relations.

Systems based on pattern extraction and matching. These rely on lexico-syntactic patterns to discover SRs between words in unrestricted texts. [13] pioneered using patterns to extract hypernymy relations. [4] applied the same technique concerning meronymy. More recently [12] have studied meronymic relations extraction while [27] has proposed a uniform approach for the extraction of different kinds of relations from text. In [22] they use Wikipedia for the extraction of SRs to integrate inside the WordNet ontology. Some works on SR extraction make use of very large corpora, like the Web. [9] describes a system that generates instances of lexico-syntactic patterns indicating specific SRs and counts their occurrences in the WWW using the Google™ API. [19] proposes a pattern matching algorithm to harvest SRs. They exploit information redundancy of the Web to filter the matches of ***general*** patterns using ***reliable*** patterns. The latter two are the works most similar to the the methodology we introduce in this paper.

Concerning “pattern-based” approaches, several techniques aim at providing support for the automatic (or semi-automatic) definition of the patterns to be used for SR extraction. Marti Hearst [14] proposes to look for co-occurrences of word pairs appearing in a specific relation inside WordNet. [25] uses WordNet to extract relations from text, but requires initial seed patterns for each relation. In [18] patterns are extracted from the BNC corpus by looking for words appearing in a hypernymy relation inside WordNet. [26] presents an unsupervised learning algorithm that mines large text corpora for patterns expressing implicit SRs.

[11] provides a comparison between unsupervised and supervised techniques for SR extraction. Hybrid approaches combining unsupervised (statistical) and supervised (pattern-based) techniques have been proposed, as in [1], where WordNet has been extended with concepts extracted from The Lord of the Rings. In [7] Latent Semantic Analysis has been applied to improve pattern-based hyponymy relations learning. More recently, [21] propose an algorithm for IS-A relation extraction from the english Wikipedia.

3 Two Complementary Techniques

The integration of the pattern-based and the distributional similarity approaches for the extraction of candidate semantic relations from corpora represents the main novelty of the proposed methodology.

From a semiotic point of view, it is worth distinguishing between syntagmatic and paradigmatic SRs between words: Saussure claims that meaning arises from these two types of relations between words. These two dimensions of meaning are often presented as two orthogonal "axes": syntagmatic relations involve associations between linguistic expressions that exist "*in presentia*", whereas paradigmatic relations involve associations that exist "*in absentia*". In a sentence like "the cat eats", for instance, the association between "*cat*" and "*eats*" is realised through their co-occurrence within the same sentence; the semantic association between "*eats*" and "*sleeps*", on the other hand, exists even if it does not show up explicitly in the sentence. To put it in other words, syntagmatic relations hold intratextually within co-occurring words, whilst paradigmatic relations refer intertextually to words which are absent from the text but which can be substituted one for another [23].

In this perspective, we can say that the distributional approach recognizes that two words are semantically similar based on distributional similarity of the different contexts in which the two words occur. The distributional method identifies a somewhat loose notion of semantic similarity (sometimes addressed to as *semantic relatedness* [5]), such as between *company* and *government*, denoting the presence of a *paradigmatic* relation of some kind between the two words. On the other hand, the pattern-based approach is based on identifying joint occurrences of the two words within particular lexico-syntactic patterns, which typically indicate a *syntagmatic* relationship. The pattern-based approach tends to yield more accurate hyponymy and meronymy relations, but is less suited to acquire near-synonyms which only rarely co-occur within short patterns in texts.

There have been just a few attempts to combine the two described approaches (see section 2): in this paper we propose a methodology for integrating distributional similarity with the pattern-based approach.

4 The Proposed Methodology

The general idea is to first identify candidate relationships by applying the distributional approach and the general patterns on the chosen corpus, then filtering and classifying the results using reliable patterns applied to the Web.

In more details, generic "high recall but low precision" lexico-syntactic patterns (in the following, GPs, for Generic Patterns) are applied to the linguistically analyzed corpus to detect a first set of pairs of co-occurring words. These words, appearing close to each other inside a sentence, can be involved in a "syntagmatic" relation, such as the words "*cougar*" and "*mammal*" in the sentence "*the cougar is a mammal*". On the other hand, a statistical unsupervised association system is used to obtain a second set of pairs of "distributionally similar" terms, that appear to occur in similar contexts, and are possibly involved in "paradigmatic" relations (e.g. "*canteen*" and "*toilet*" in the sentences "*Bob goes to the canteen*" and "*Simon goes to the toilet*").

Once the two sets of pairs of words have been extracted, a set of reliable ("low recall but high precision") lexical patterns are instantiated with the words and projected onto the Web, thus filtering the candidate relations obtained through GPs and labelling the anonymous relations obtained through the statistical system. The overall process can be summarized in the following sequence of steps:

Input: a specific (raw) corpus, from which semantic relations between words are extracted;
Step 1: the sets of "generic" and "reliable" patterns are manually defined, each one corresponding to a specific SR;
Step 2: the corpus is linguistically analyzed using a battery of NLP tools;
Step 3: a distributionally–based algorithm for building classes of semantically–related words is used to obtain a first set of pairs denoting candidate (anonymous) relations between words;
Step 4: the GPs are applied to the syntactical analyzed corpus obtaining a second set of pairs denoting candidate (labelled) relationships between words;
Step 5: the reliable patterns are applied to the obtained sets of words, by instantiating each patterns with their morphological variations and in either order;
Output: a set of labelled word pairs representing SR instances.

Step 1: definition of generic and reliable patterns. The definition of generic and reliable patterns is at the moment done by hand. The algorithm we use is inspired by [14]: i) decide on a semantic relation of interest, ii) decide a list of word pairs from WordNet in which this relation is known to hold, iii) extract sentences from a corpus in which these words both occur, and record the lexical and syntactic context; iv) find the communalities among these contexts and hypothesize that the common ones yield patterns that indicate the relation of interest.

The relations we chose to focus on are *hyponymy*, *meronymy*, *co-hyponymy* and *near-synonymy*[1]. However, we envisage the application of this methodology for the

[1] We prefer to use the term (near-)synonymy since "there are very few absolute synonyms, if they exist at all. So-called dictionaries of synonyms actually contain near-synonyms" [15].

extraction of generic, associative relations between words. Future work will also include the application of semi-automatic techniques for the discovery of patterns.

A GP for hyponymy could be "X è un Y" (*X is a Y*) and "X di Y" (*X of Y*) for meronymy. It goes without saying that both exemplified patterns will produce a lot of noise in the results since they are used to express many other SR types, as noun compounds do for English. Once GPs are defined, we distinguish three kinds of reliable patterns: "close", "partially open" and "open" patterns, depending on the way we use them to query the Web. In particular, partially open patterns (POPs, in the following) are constructed by inserting wildcards (such as "*") inside the query and contain both words; "open patterns" (OPs) contain just one of the words and can contain wildcards; close patterns (CPs) don't have any wildcard, contain both words and they are completely enclosed in quotation marks. Here are some examples of reliable patterns (the first three being CPs, the last being a POP) involving words A and B, one for each of the four relations of interest:

- *Hyponymy*: "A e altri B" ("*A and other B*")
- *Meronymy*: "A è composto da B" ("*A is made up of B*")
- *Near-synonymy*: "A chiamato anche B" ("*A also called B*")
- *Co-Hyponymy*: "A e B sono" * ("*A and B are*" *)

Step 2: linguistic analysis of the corpus. The linguistic analysis of the corpus is carried out by AnIta [3], an Italian parsing system consisting of a suite of linguistic tools in charge of: tokenisation of the input text; morphological analysis (including lemmatisation); syntactic parsing, in its turn articulated in two different and incremental steps, namely "chunking", carried out simultaneously with morpho-syntactic disambiguation, and dependency analysis. For the specific concerns of this study, a central role is played by the syntactic analysis stages which are differentially exploited during the different extraction steps. Whereas Step 4 operates on texts annotated with basic syntactic structures ("chunks"), NLP requirements of Step 3 are more demanding, i.e. clusters of semantically related words are identified starting from the dependency–annotated text.

Step 3: extraction of anonymous relations through distributional similarity. Identification of clusters of semantically related terms is carried out on the basis of distributionally-based similarity measures. To this end, we used CLASS [2], a distributionally-based algorithm for building classes of semantically-related words. According to CLASS, two words are considered as semantically related if they can be used interchangeably in a statistically significant number of syntactic contexts. CLASS grounds its semantic generalizations on controlled distributional evidence, where not all contexts are equally relevant to an assessment of the semantic similarity between words (e.g. contexts with so-called light verbs, such as "to take" in "to take a shower" or "to have" in "to have a drink", play quite a marginal role in the assessment of semantic similarity).

Step 4: extraction of candidate relations through generic patterns. The set of GPs is mapped onto the syntactically chunked corpus. As expected, many matches are produced for each relation, the majority of which is represented by "noisy" word pairs. See 5 for some examples of obtained results.

Step 5: application of reliable patterns on the Web. For each obtained word pair, such as "bicicletta, ruota" (*bicycle*, *wheel*), the morphological variations, in either order, are generated, such as "biciclette, ruote" (*bicycles*, *wheels*), "bicicletta, ruote", "ruota, biciclette", and so on. The new word pairs are then used inside the reliable patterns created in Step 1 and applied to the Web. Though we envisage the use of the Google™ API in the system implementation step, we are currently applying the reliable patterns on the Web by just querying the Google™ search engine. Every reliable pattern will be manually instantiated, applied to the Web and evaluated.

5 Preliminary testing of the methodology

In this section we report a case study concerning some experiments we have done analysing a part of the Italian Wikipedia and looking for instances of the relations of hyponymy, meronymy, co-hyponymy and near-synonymy. The section is divided on the basis of the "five-steps" algorithm we have introduced.

Step 1: definition of generic and reliable patterns. GPs are defined just for hyponymy and meronymy, since near-synonymy and co-hyponymy are mainly identified on a paradigmatic basis in the distributional similarity step. As a matter of fact, it appears infrequent to find inside a corpus "intrasentence" co-occurrences of co-hyponyms and near-synonyms, except in rare cases, which can be found just by querying the Web as we do using the reliable patterns in step 5. For the sake of simplicity, let's imagine to have just the two GPs introduced in section 4, one for hypernymy and one for meronymy.

Step 2: linguistic analysis of the corpus. In this example, we have used a part of the Italian Wikipedia (908 articles, for a total of 788,000 words) as the corpus to be analyzed. For some details about the used NLP tools please refer to section 4.

Step 3: extraction of anonymous relations through distributional similarity. By applying the distributional system to the corpus we have obtained 4407 "semantically related" word pairs, containing correct and spurious results. Every relation instance is "anonymous": every pair must be applied inside the reliable patterns to verify the actual presence of one of the four relations we are looking for and to provide a correct classification.

Step 4: extraction of candidate relations through generic patterns. The application of the two GPs has produced 553 candidate pairs for hyponymy and 19099 candidate pairs for meronymy.

Step 5: application of reliable patterns on the Web. Every relation we are going to detect through reliable patterns needs a slightly different strategy. CPs, for example, seem particularly suited for hyponymy, while concerning meronymy the contribution of POPs improves the results. To label word pairs obtained at step 3, all the reliable patterns must be applied, since there's no way of knowing in advance the involved (if any) relation.

The preliminary evaluation has been conducted just for hyponymy and meronymy by manually building the reliable patterns relative to the word pairs obtained through the two GPs. Concerning hyponymy, 100 out of the 553 obtained word pairs (A, B) have been randomly selected and analysed: 66% of them were "positive" (A is a direct hyponym of B), 11% were "partially positive" (A is an indirect hyponym of B) while the remaining 24% were negative. Five reliable patterns have been applied to 10 word pairs randomly extracted from each class (positive, partially positive and negative). The results are shown in the following table:

	direct hyponymy accuracy	**direct and indirect accuracy**	**harvesting value**
pattern 1	0.8876	0.9704	44.55 %
pattern 2	0.9122	1.0000	22.64 %
pattern 3	0.8235	1.0000	4.06 %
pattern 4	0.9430	0.9929	21.04 %
pattern 5	0.9408	1.000	7.69 %

The first and second columns indicate the capability of each reliable pattern to recognize, respectively, direct hyponyms only and both direct and indirect hyponyms. Pattern 1, for example, recognizes exclusively direct hyponyms 88,76% of the times, direct and indirect hyponymy 97,04% of the times, but indicates a match in the remaining 2,96% of the cases when no hyponymy relation is really involved. The third column is relative to the proportion of matches produced by each pattern. It is interesting to notice a possible inverted proportionality between the harvesting value and the accuracy, though further investigations are necessary.

The following table shows the five reliable patterns we have used for hyponymy:

pattern 1	A è un B - "anatra è un tipo di animale" ("*duck is a type of animal*")
pattern 2	A ed altri B - "anatra e altri animali" ("*duck and other animals*")
pattern 3	A o altri B - "anatra o altri animali ("*duck or other animals*")
pattern 4	B come A - "animali come l'anatra ("*animals like the duck*")
pattern 5	B quali A - "animali quali l'anatra ("*animals such as the duck*")

Concerning meronymy, the GP produced a very noisy set of more than 19,000 word pairs. We manually selected 3630 pairs to evaluate the percentage of positive matches: 0.64% of them involved direct meronymy, 0.30% involved indirect meronymy while the remaining 99.06% were negative matches. We then extracted 5 word pairs from each class and manually applied each pair to the 10 reliable patterns we have constructed. It has been a quite long task, considering that each word pair required the manual construction (and consequent Google™ querying and validation) of 256 lexical patterns (about 25 variants for each pattern), for a total of 3840 patterns. We have ranked the word pairs on the basis of the number of matches they have obtained from the pattern application of the whole battery of patterns. The following table summarizes the results, ordered by rank:

word pair	rank	class	WN	IWN
ora, giorno (*hour, day*)	157	direct	yes	no
cella, memoria (*cell, memory*)	151	direct	no	no
elettrone, atomo (*electron, atom*)	112	indirect	-	-
rione, città (*quarter, town*)	19	direct	no	no
facciata, edificio (*front, building*)	17	direct	no	yes
squama, strobilo (*scale, strobile*)	10	direct	no	no
mese, inverno (*month, winter*)	10	negative	-	-
muro, castello (*wall, castle*)	7	indirect	-	-
esecuzione, musica (*playing, music*)	1	negative	-	-
osso, animale (*bone, animal*)	1	indirect	-	-
azione, piede (*action, foot*)	0	negative	-	-

The last two columns indicate if the meronimy relation between the two words is present, respectively, inside Wordnet and ItalWordNet. It is interesing to notice, for example, how the word pair (cella, memoria) is not present inside both WN and IWN even if scoring an high rank. In conclusion, we have seen that reliable patterns are very selective (they exclude negative matches) and, in some cases, can even lead to reconsider some asserted ontological relations, such as about “elettrone” and “atomo”, two words not involved in a (direct) meronimical relation inside WN and IWN.

The last two examples are relative to co-hyponymy and near-synonymy: the candidate word pairs have been obtained by the distributional system. Concerning co-hyponymy recognition, we envisage to adopt two strategies: i) to apply reliable “co-hyponymy” patterns involving words A and B (as usual) and ii) to look for hypernyms of A and B (using open patterns, OPs) to verify if A and B share the same hypernym. For example, “adenina, timina” is a word pair obtained by the distributional system. We can apply a pattern and then analyse the results:

“timina e adenina sono” (“*thymine and adenine are*”) → 1 match
- **timina e adenina sono** <u>basi azotate</u>... (*thymine and adenine are nucleobases...*)

On the basis of this result, the two words are probably co-hyponyms of “base azotata” (*nucleobase*). To strengthen this hypothesis we can look for hypernyms of the two words (separately) by applying OPs and by comparing the resulting candidate hypernyms. Some OPs instantiated with “timina” (and “adenina”) we have tried are: “la timina è” (197 matches), “come la timina” (54 matches), “le * sono la timina” (4120 matches), “timina e altre” (1 match). From every (positive) resulting snippet we have extracted the candidate hypernyms for both words:

candidate hypernyms for timina: {base azotata pirimidinica, componente, pirimidina, mutazione letifera, nucleotide, base, nome, base azotata }
candidate hypernyms for adenina: {base azotata purinica, purina, sostanza, molecola organica, acido, ligando naturale, base nucleotidica, base azotata, base purinica, nucleotide, segno, componente}

The intersection of the two sets is: {componente, nucleotide, base azotata}. If we combine this result with the previous one obtained with the co-hyponymy pattern we can assert that “timina” and “adenina” are co-hyponyms of “base azotata”.

Near-synonymy can be faced in a similar way. Let's consider the word pair "puma, coguaro" (*puma, cougar*). Starting from near-synonymy patterns we obtain:

"puma" * "chiamato anche coguaro" ("*puma*" * "*also called cougar*") → 8 matches
- Il **puma** (Puma concolor), **chiamato anche coguaro** o leone di montagna...
- Il **puma** (Puma concolor dal 1993 [...]), **chiamato anche coguaro**...
- vive nel continente americano [...] **chiamato anche coguaro**, ... I **puma**...

OPs for hypernymy and hyponymy can be applied and the resulting sets intersected to look for near-synonymy between the two words. The idea is that if words A and B have similar hypernyms and hyponyms they can be considered near-synonyms.

hypernyms for puma: {veicolo trasporto truppe, felino selvatico, telefilm, elicottero, felino, campione, veicolo, personaggio, animale selvaggio, animale}
hypernyms for coguaro: {puma americano, felino, carnivoro, incrocio, animale}
hyponyms for puma: {pantera della Florida, onza}
hyponyms for coguaro: {pantera della Florida}

Both the intersections of the two sets of hypernyms and the two sets of hyponyms are non-empty, thus reinforcing the near-synonymy hypothesis.

6 Conclusions and Future Works

The paper shows a methodology to combine two different techniques for Semantic Relation Extraction from texts, namely, lexico-syntactic patterns and statistical distributional systems. The preliminary evaluation we have conducted has shown very promising results, both for hyponymy and meronymy relation extraction.

We are currently working to implement a fully automatic semantic relation extraction system based on the proposed approach: as soon as it will be available, a more accurate and articulated evaluation will be possible.

References

1. Alfonseca, E., Manandhar, S.: Extending a Lexical Ontology by a Combination of Distributional Semantics Signatures. In: Proceedings of EKAW'02, pp. 1--7. (2002)
2. Allegrini, P., Montemagni, S., Pirrelli, V.: Example-Based Automatic Induction Of Semantic Classes Through Entropic Scores. Linguistica Computazionale XVI-XVII, pp.1-45. (2003)
3. Bartolini, R., Lenci, A., Montemagni, S., Pirrelli, V.: Hybrid Constrains for Robust Parsing: First Experiments and Evaluation. In Proceedings of LREC 2004, Lisbon, Portugal (2004).
4. Berland, M., Charniak, E.: Finding parts in very large corpora. In: Proceedings of the 37th ACL, pp. 57--64. University of Maryland (1999)
5. Budanitsky, A., Hirst, G.: Evaluating WordNet-based Measures of Lexical Semantic Relatedness. Computational Linguistics. Volume 32, Issue 1, pp. 13-47. MIT Press Cambridge, MA, USA (2006)

6. Caraballo, S.: Automatic acquisition of a hypernym-labeled noun hierarchy from text. In: Proceedings of the 37th Annual Meeting of the ACL. pp. 120-126. (1999)
7. Cederberg, S., Widdows, D.: Using LSA and Noun Coordination Information to Improve the Precision and Recall of Automatic Hyponymy Extraction. In CoNLL 2003, pp. 111--118. Edmonton, Canada (2003)
8. Cimiano, P., Staab, S.: Clustering concept hierarchies from text. In: Proceedings of LREC-2004. Lisbon, Portugal (2004)
9. Cimiano, P., Staab, S.: Learning by googling. ACM SIGKDD Explorations Newsletter. Volume 6, Issue 2. pp. 24--33. New York, NY, USA (2004)
10. Faure, D., N'edellec, C.: A corpus-based conceptual clustering method for verb frames and ontology acquisition. In: LREC workshop on Adapting lexical and corpus resources to sublanguages and applications. Granada, Spain (1998)
11. Finkelstein-Landau, M., Morin, E.: Extracting semantic relationships between terms: Supervised vs. unsupervised methods. In: Proceedings of the International Workshop on Ontological Engineering on the Global Information Infrastructure. Dagstuhl, Wadern, Germany (1999)
12. Girju, R., Badulescu, A., Moldovan, D.: Automatic Discovery of Part-Whole Relations. Computational Linguistics, Volume 32, No. 1, pp. 83--135. (2006)
13. Hearst, M. A.: Automatic acquisition of hyponyms from large text corpora. In: In Proceedings of the 14th International Conference on Computational Linguistics, pp. 539--545. Nantes, France (1992)
14. Hearst, M.,: Automated Discovery of WordNet Relations. MIT Press, Cambridge MA (1998)
15. Inkpen, D., Hirst, G.: Building and Using a Lexical Knowledge Base of Near-Synonym Differences. Computational Linguistics, 32(2), 223--262 (2006).
16. Lee, L.: Similarity-Based Approaches to Natural Language Processing, Ph.D. thesis. Technical Report TR-11-97, Harvard University (1997).
17. Maedche, A., Staab, S.: Discovering conceptual relations from text. In: Proceedings of ECAI 2000. Technical Report 425, University of Karlsruhe (2000)
18. Mititelu, V.B.: Automatic Extraction of Patterns Displaying Hyponym-Hypernym Co-Occurrence from Corpora. In: Proceedings of the First CESCL. Budapest, Hungary (2006)
19. Pantel, P., Pennacchiotti, P.: Espresso: Leveraging generic patterns for automatically harvesting semantic relations. In: Proceedings of ACL 2006, pp. 113--120. Sydney, Australia. (2006)
20. Pekar, V., Staab, S.: Word classification based on combined measures of distributional and semantic similarity. In: Proceedings of EACL03. Budapest (2003)
21. Ryu, P., Choi, K.: Automatic Acquisition of Ranked IS-A Relation from Unstructured Text. In: Proceedings of OntoLex07, pp. 67--77. Busan, Korea (2007)
22. Ruiz-Casado, M., Alfonseca, E., Castells, P.: Automatising the learning of lexical patterns: An application to the enrichment of WordNet by extracting semantic relationships from Wikipedia. Data & Knowledge Engineering, Volume 61, Issue 3, pp. 484--499 (2007)
23. Saussure, F. de: Course in General Linguistics (trans. Roy Harris). Duckworth, London (1983)
24. Silverman, K. : The Subject of Semiotics. Oxford University Press, New York (1983)
25. Stevenson, M.: An unsupervised WordNetbased algorithm for relation extraction. In: Proceedings of the "Beyond Named Entity" workshop at LREC-04. Lisbon, Portugal (2004)
26. Turney, P.D.: Expressing implicit semantic relations without supervision. In: Proceedings of Coling/ACL-06, pp. 313--320. Sydney, Australia (2006)
27. Turney, P.D.: A uniform approach to analogies, synonyms, antonyms, and associations. In: Proceedings of Coling 2008, pp. 905--912. Manchester, UK (2008)
28. Widdows, D.: Unsupervised methods for developing taxonomies by combining syntactic and statistical information. In Proc. of the HLT/NAACL 2003 - Vol. 1, 197-204 (2003).