Semantically-Driven Extraction of Relations between Named Entities

Caroline Brun and Caroline Hagège

Xerrox Research Centre Europe, 6 chemin de Maupertuis 38240 Meylan, France
{Caroline.Brun, Caroline.Hagege}@xrce.xerox.com

Abstract. In this paper, we describe a method that automatically generates lexico-syntactic patterns which are then used to extract semantic relations between named entities. The method uses a small set of seeds, i.e. named entities that are a priori known to be in relation. This information can easily be extracted from encyclopedias or existing databases. From very large corpora we extract sentences that contain combinations of these attested entities. These sentences are then used in order to automatically generate, using a syntactic parser, lexico-syntactic patterns that links these entities. These patterns are then re-applied on texts in order to extract relations between new entities of the same type. Furthermore, the patterns that are extracted not only provide a way to spot new entities relations but also build a valuable paraphrase resource. An evaluation on the relation holding between an event, the place of the event occurrence and the date of the event occurrence has been carried out on French corpus and shows good results. We believe that this kind of methodology can be applied for other kinds of relation between named entities.

1 Introduction

In this paper we describe a system that extracts accurately semantic relations between named entities from raw text. Taking as input a small set of already known relations that can be extracted from encyclopedias or from databases, our system first learn from a large corpus a wide range of lexico-syntactic patterns conveying the desired semantic relation. These learned patterns are then further applied on texts, and as a result, new occurrences of the given semantic relations linking new entities are detected. As all the patterns extracted represent a comparable semantic situation, they can be considered as paraphrase patterns. These patterns can then be used both in generation and for information extraction tasks.

2 Related Work

Many research works on extraction of relation between entities have already been performed since this kind of information is useful for a wide range of applications of information extraction. For instance [8] describe an algorithm to extract relations between named entities and the resulting improvement of a question answering

© A. Gelbukh (Ed.)
Advances in Computational Linguistics.
Research in Computing Science 41, 2009, pp. 35-46

Received 07/11/08
Accepted 09/12/08
Final version 04/02/09

system. Semantic relation detection between named entities has also been investigated in the context of the semantic web (try to obtain a rich and accurate metadata annotation from web content) as for instance in [6]. In the biomedical domain, [7] and [14] present methods to automatically extract interaction relations between genes and/or protein using machine learning techniques.

Some of these approaches rely on pattern matching exploiting simple syntactic relations as the Subject-Verb-Object relation. Sometimes, additional ontological knowledge is also exploited. These approaches take advantage of the fact that a certain syntactic configuration can be mapped onto a semantic relation. This is particularly well described in [12] where a shallow parser and a deeper parser are used to extract SVO relations between triples. In [7] and [14], a previous dependency analysis is performed to derive necessary information for learning algorithms.

Other approaches based on syntactic analysis rely on the fact that the type of syntactic relations is predefined: it has to be stated beforehand that a SVO syntactic relation or an appositive relation is meaningful for the kind of semantic relation the system extracts.

Other approaches, which do not presupposes the type of syntactic relations (see [7] and [14]) holding between entities, take into account both positive and negative examples of possible relations (relying on the closed-world assumption, stating that if no link is found between two entities, then these entities are never related).

Our approach neither needs the a priori knowledge of relevant syntactic links conveying the semantic relational information of interest, nor assumes the presence of negative examples: in some cases, like the one we present for illustration, the notion of negative example is not pertinent since one of the entities in focus convey temporal information, which is not compatible with close world hypothesis. This approach has the advantage to extract relations between entities that are less predictable in advance.

In the following steps we describe our method by exemplifying it on a concrete case: the extraction of relations between events, the place where the event occurs and the date of the event occurrence. The same methodology can be applied for other kind of relations and in other contexts. It must be stressed that our approach needs quite limited resources, namely a small list of attested and trustable relations between entities, that we can find in freely available encyclopedia as Wikipedia and a linguistic engine, which is able to detect named entities and to syntactically relate linguistic units appearing in texts.

3 A System for NE Relation Extraction

In this section, we describe the general methodology we used to build our system; we also describe the robust parser we use as core component of the system, and then present the resulting prototype.

3.1 Description of the Method

The first step of our method is to extract attested related entities from a trustable external resource. In other words, we need a resource which stores n-uples of entities which are linked by some kind of semantic relation. For instance, from the French Wikipedia in the article about Olympic Games, we can obtain the following list of triples which links a date (date of the event), a place (place of the event at that date) and the event name (in this case, the Olympic Games).

1896 Grèce Athènes Jeux Olympiques.
1900 France Paris Jeux Olympiques.
1904 États-Unis Saint-Louis Jeux Olympiques.
1908 Royaume-Uni Londres Jeux Olympiques.
1912 Suède Stockholm Jeux Olympiques.
1916 Berlin Allemagne Jeux Olympiques.
1920 Belgique Anvers Jeux Olympiques.
1924 France Paris Jeux Olympiques.
1928 Pays-Bas Amsterdam Jeux Olympiques.
1932 États-Unis Los Angeles Jeux Olympiques.
1936 Allemagne Berlin Jeux Olympiques.
1940 Helsinki Finlande Jeux Olympiques.
1944 Londres Grande-Bretagne Jeux Olympiques.
1948 Royaume-Uni Londres Grande-Bretagne Jeux Olympiques.
1952 Finlande Helsinki Jeux Olympiques.
1960 Italie Rome Jeux Olympiques.
1964 Japon Tôkyô Jeux Olympiques.
1968 Mexique Mexico Jeux Olympiques.
1972 Allemagne Munich Jeux Olympiques.
1976 Canada Montréal Jeux Olympiques.
1980 Moscou URSS Jeux Olympiques.
1984 États-Unis Los Angeles Jeux Olympiques.
1988 Séoul Corée du Sud Jeux Olympiques.
1992 Espagne Barcelone Jeux Olympiques.
1996 États-Unis Atlanta Jeux Olympiques.
2000 Australie Sydney Jeux Olympiques.
2004 Grèce Athènes Jeux Olympiques.
2008 Chine Pékin Jeux Olympiques.
2012 Royaume-Uni Londres Jeux Olympiques.

The second step of the method consists in extracting from very large corpora all sentences that contain all or a part of one of the triples appearing in the list. In other words, for our example, we extract all the sentences containing the name of the event together with the date and/or the place(s). For example, for “1992 Espagne Barcelone Jeux Olympiques.”, we extract all sentences containing Espagne/1992/Jeux Olympiques, Barcelone/1992/Jeux Olympiques, Espagne/Jeux Olympiques, Barcelone/Jeux Olympiques, and 1992/Jeux Olympiques.

Sentences like the following ones are extracted from our initial list:

1. Le CIO a fixé des objectifs de lutte contre le dopage durant les jeux Olympiques de Sydney. (Event + Place)
(IOC has given objectives for fighting doping during the Olympic Games of Sydney)

2. Toutefois, les premières mesures contre le dopage n'ont été prises qu'après les jeux olympiques d'Helsinki de 1952. (Event + Place + Date)
(The first measures against doping, however, have been only taken after the Olympic Games of Helsinki in 1952)
3. En 2008, la Chine accueillera les Jeux olympiques, et le pays est un membre permanent du Conseil de sécurité des Nations unies. (Event + Place + Date)
(In 2008, China will receive the Olympic Games, and the country is a permanent member of the United Nations Security Council.
4. Ce n'est qu'en 1928 que décision a été prise d'autoriser la participation des femmes aux Jeux olympiques. (Event + Date).
Etc.
(It is only in 1928 that the decision has been taken of authorizing the participation of women to the Olympic Games)

This extraction is a simple processing step, because only pattern matching is required (possibly with some normalization of case).

The third step consists in applying a robust syntactic dependency parser on these sentences in order to extract the syntactic relationships linking the attested elements. It is important to stress that there is no a priori about the possible syntactic relations that can hold between the entities. These links can also be direct links or indirect links, transitivity being taken into account.

In addition to the syntactic parsing, the linguistic engine also performs named entity recognition, in order to be able to generalize the patterns extracted.

The following example shows the relations that the linguistic engine extracts from sentence 3. Binary relations correspond to grammatical links between lemmatized lexical units of the sentence while unary relations correspond to the identification of Named Entities.

En 2008, la Chine accueillera les Jeux olympiques, et le pays est un membre permanent du Conseil de sécurité des Nations unies.
(In 2008, China will receive the Olympic Games, and the country is a permanent member of the United Nations Security Council.

SUBJ(accueillir,Chine)
OBJ(accueillir,Jeux olympiques)
VMOD(accueillir,2008)
EVENT(Jeux Olympiques)
DATE(2008)
PLACE(Chine)

The fourth step consists in generalizing the set of relations captured by the parser in order to obtain a generic lexico-syntactic rules pattern. This generalization is performed by abstracting the Named Entities by their type (in our example, event, date and place) and keeping the lemmas of the lexical elements which are present in the grammatical binary relations.

From our previous example, we obtain automatically from the set of extracted relationships the following lexico-syntactic pattern (where & represent a conjunction).

SUBJ(accueillir, PLACE(X)) & OBJ(accueillir,EVENT(Y)) & VMOD(accueillir,DATE(Z))
==>DATE-and-PLACE-of-EVENT(Z,X,Y)

Once learned, as our syntactic parser is rule-based, these learned rules can be integrated as such on top of the syntactic parser set of rules, and applied as any other kind of rules on any corpora.

As a result, thanks to the generalization of entity types, the application of these rules, will enable to extract new n-uples of related named entities that were not present in our initially extracted list (i.e. other kind of events that are not necessarily Olympic games or even sport event associated with their date of occurrence and their place of occurrence may be discovered).

The whole process is summarized on the following figure:

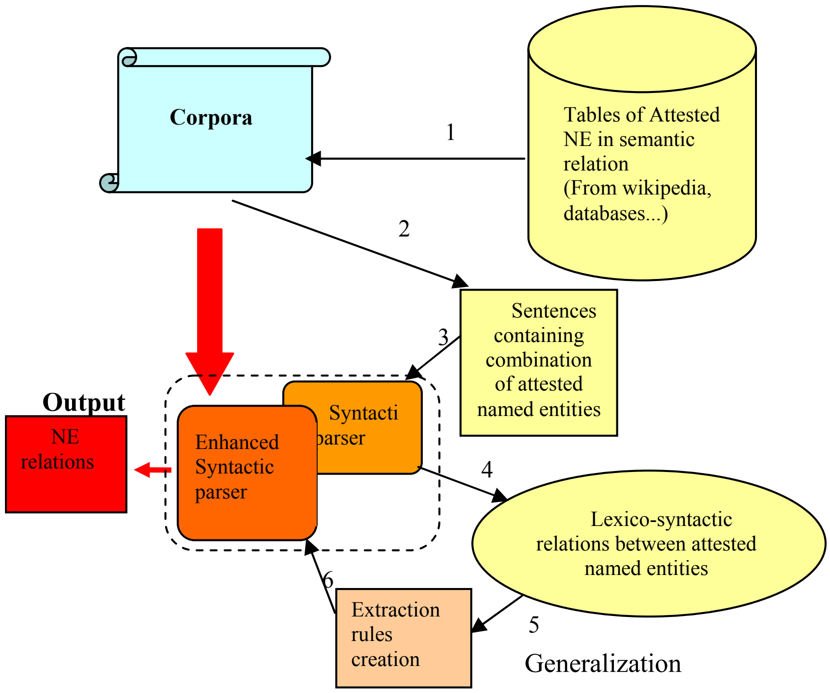

Fig. 1.Summary of the process

3.2 Robust and Deep Parsing using XIP

As a fundamental component of the system we designed for named entity relation extraction, we use the Xerox Incremental Parser (XIP, see [3]) in order to perform robust and deep syntactic analysis. Deep syntactic analysis consists here in the construction of a set of syntactic relations[1] from an input text. These relations link lexical units of the input text and/or more complex syntactic domains that are constructed during the processing (mainly chunks, see [1]). These relations are labeled with deep syntactic functions. More precisely, a predicate (verbal or nominal) is linked with what we call its deep subject (SUBJ-N), its deep object (OBJ-N), and modifiers.

In addition, the parser calculates more sophisticated and complex relations using derivational morphologic properties, deep syntactic properties (subject and object of infinitives in the context of control verbs), and some limited lexical semantic coding (Levin's verb class alternations, see [9], and some elements of the Framenet[2] classification [11]). These deep syntactic relations correspond roughly to the agent-experiencer roles that is subsumed by the SUBJ-N relation and to the patient-theme role subsumed by the OBJ-N relation (see [5] and [4]). Not only verbs bear these relations but also deverbal nouns with their corresponding arguments.

The use of such sophisticated relations in the pattern extraction process enables us to extract "normalized patterns" that have a wide coverage. For example, a single pattern extracted with the normalization grammar will match different surface realization such as a passive form, active form or nominalization of a given predicate.

This parser includes also a module for Named Entity recognition, i.e. detection of numerical expressions, dates, person, organization, location names, and events. This module is built within the XIP parser presented above, on top of a part-of-speech tagger. This system is purely rule-based, and consists in a set of ordered local rules that use lexical information combined with contextual information about part-of-speech, lemma forms and lexical features.

These rules detect the sequence of words involved in the entity and assign a feature (loc, org, date, event, etc.) to the top node of the sequence, which is a noun in most of the cases. This system has been evaluated internally and show a performance of .90 in F-measure on the whole set of types of entities.

Here is an example of an output (Named entities, chunk tree and deep syntactic relations) of the most sophisticated version of the grammar:

"Lebanon still wanted to see the implementation of a UN resolution."

TOP{SC{NP{Lebanon} FV{still wanted}} IV{to see} NP{the implementation} PP{of NP{a UN resolution}} .}

PLACE_COUNTRY(Lebanon)
ORGANISATION(UN)
MOD_PRE(wanted,still)
MOD_PRE(resolution,UN)
MOD_POST(implementation,resolution)
EXPERIENCER_PRE(wanted,Lebanon)
EXPERIENCER(see,Lebanon)
CONTENT(see,implementation)
EMBED_INFINIT(see,wanted)
OBJ-N(implement,resolution)

[1] Inspired from dependency grammars, see [10] and [13].
[2] http://framenet.icsi.berkeley.edu/

3.3 Description of the System

In order to validate our method, we implemented a prototype which uses as relation seeds the relation holding between the Olympic Games and the corresponding date and place of occurrence. As shown in the first subsection, we first extracted from Wikipedia a first list of triples.

To set up the prototype, we used a French corpus of about 1.3 million sentences, provided by the European community about "acquis communautaires" (i.e. "community acquis", the rights and obligations that EU countries share). We divide this corpus in two parts, and, then, on the first half, we extract sentences that contain the triplets "Olympic game/date/place" given by the Wikipedia attested list, and if the triplet is not present, then the couples "Olympic game/date" and "Olympic game/place". From the initial corpus, we extracted 150 sentences containing either a triplet or couple of attested entities.

We parse these sentences with XIP, which provided us with a list of dependencies involving the attested entities.

This list of dependencies is automatically transformed[3] into a set of XIP rules that can be applied on top of the parser previously used.

For example, when parsing:

« Londres organisera les Jeux Olympiques en 2012 »
(London will organize the Olympic Games in 2012)
XIP outputs the following dependencies:

SUBJ(organiser,Londres)
OBJ(organiser,jeux olympiques)
VMOD_POSIT1(organiser,2008)
PREPOBJ(2008,en)
DETERM_DEF_SPORT(jeux olympiques,le)
DATE(2008)
PLACE_CITY(Londres)
EVENT_SPORT(jeux olympiques)

We then select from this output all dependencies that:

- Involve one of the named entity in focus (here entities of type event, date & place)
- Involve only non-functional words (noun, verbs, adjectives and not preposition or determiner for example)

In this example, the selected dependencies are then:

SUBJ(organiser,Londres)
OBJ(organiser,jeux olympiques)
VMOD_POSIT1(organiser,2008)
PLACE_CITY(Londres)
DATE(2008)
EVENT_SPORT(jeux olympiques)

---

[3] By a python script developed for that purpose.

We then abstract on the named entity types, and consequently deduce the following XIP rule:

If(SUBJ(#1[lemma:="organiser"],#2) & PLACE(#2) & OBJ(#1,#3) & EVENT(#3) & VMOD(#1,#4) & DATE(#4))
==> DATE-and-PLACE-of-EVENT(#4,#2,#3)

This rule can then be applied as such incrementally on top of the parser, to discover new entities in semantic relation.

When applying the enhanced parser integrating the new learned rules, we get for example the following result, on a different event:

« Le Championnat d'Europe sera organisé en 2004 par le Portugal. »
(European Championship will be organized in 2004 by Portugal).

SUBJ_PASSIVE(organiser,championnat d'Europe)
SUBJ(organiser,Portugal)
OBJ(organiser,championnat de Europe)
VMOD_POSIT1(organiser,2004)
NMOD_POSIT1(2004,Portugal)
AUXIL_PASSIVE(organiser,être)
DATE(2004)
PLACE_COUNTRY(Portugal)
EVENT_SPORT(championnat d'Europe)
DATE-and-PLACE-of-EVENT(2004, Portugal, championnat d'Europe)

Additionally, this example shows the interest of using deep syntax ("syntactic normalization", cf. [6]), which enables to map active and passive cases.

On the 150 sentences about the Olympic Games that we extracted in the original corpus, we automatically build about 60 XIP rules that are incrementally applied on top of the parser, in a new layer of rules.

While many of the research focuses on extracting subject-verb-object patterns, our method does not make a priori hypothesis for the type of syntactic dependencies that links the entities in semantic relation (typically, dates have verbal or nominal modifier functions). It can therefore account for examples like:

« Les jeux olympiques de 1992 se déroulaient à Albertville. »
(The Olympic Games of 1992 were happening in Albertville).

SUBJ(dérouler,jeux olympiques)
VMOD_POSIT1(dérouler,Albertville)
NMOD_POSIT1(jeux olympiques,1992)
PREPOBJ(Albertville,à)
PREPOBJ(1992,de)
DETERM_DEF(jeux olympiques,le)
REFLEX(dérouler,se)
DATE(1992)
PLACE_CITY(Albertville)
EVENT_SPORT(jeux olympiques)
DATE-and-PLACE-of-EVENT(1992, Albertville, jeux olympiques)

Here, the entities of date and location have modifier syntactic functions and do not follow the pattern subject-verb-object of the above-mentioned two previous examples.

Applying the system on different kind of corpora, such as newspapers, shows that the system extract relations concerning other types of event, such as cultural events, since there are recognized by the named entity module and since the relation patterns hold also for them:

« La Biennale d'art contemporain aura lieu à Lyon du 16 septembre 2009 au 3 janvier 2010 »

(The Biennale of contemporary art will take place from the 16 September 2009 until the 3 January 2010)

SUBJ(aura,Biennale d'art contemporain)
OBJ2(aura,lieu)
VMOD_POSIT1(aura, du 16 septembre 2009 au 3 janvier 2010)
VMOD_POSIT1(aura, Lyon)
DATE_INTERVAL(du 16 septembre 2009 au 3 janvier 2010)
PLACE_CITY(LYON)
EVENT_CULTURAL(Biennale d'art contemporain)
DATE-and-PLACE-of-EVENT(du 16 septembre 2009 au 3 janvier 2010, Lyon, Biennale d'art contemporain)

3.4 Extraction of Paraphrase Patterns

One of the interesting points of the system we developed is the construction of a valuable resource of paraphrase templates that have been automatically constructed for extracting new relations between EVENTS, DATES and PLACES. Our templates correspond of XIP grammar rules as shown in section 3.3. and are very precise descriptions of the grammar links holding between the different NE.

As all our patterns semantically denote a situation where an event occurs in a certain place and possibly at a certain date, we can consider two things:

First, the text segments that enable to extract the templates are paraphrases of one another;

Second, the templates themselves can be considered are basis for the generation of paraphrases.

Let's take some example to illustrate:

if ((EVENT(#1) & SUBJ(#2[lemme:"avoir"],#1) & OBJ2(#2,#3[lemme:"lieu"]) & VMOD(#2,#4) & PLACE(#4) & PREPD(#4,?[lemma: "à"]))
==> PLACE-of-EVENT(#1,#4)

This template expresses the situation of an event taking place in a certain place (NEW-PLACE-EVENT situation).

It expresses that the EVENT is the subject of support verb "avoir lieu" and that this verb has a modifier of type PLACE which is introduced by preposition "à"

Following the canonical sentence order for French, this corresponds for instance to expressions like:

(1) <EVENT> a eu lieu à <PLACE>

(<EVENT> took place in <PLACE>)

Let's now take another example of pattern denoting the same situation:
if(SUBJ(#1[lemme:"accueillir"],#2) & OBJ(#1,#3) & EVENT(#3) & COREF[rel](#4,#2) & PLACE(#4))
==> PLACE-of-EVENT(#3,#4)

This time the template expresses that the verb "accueillir" has to have a subject of type PLACE and at the same time a direct object as type EVENT.
Following the canonical order for French, this corresponds for instance to expressions like:
(2) <PLACE> a accueilli <EVENT>
(<PLACE> received <EVENT>)

(1) and (2) are paraphrase patterns and linguistic realizations of those patterns convey the same information.
These two examples can be considered as exact paraphrases candidate, however we also extract approximate paraphrase patterns like:

<PLACE> organize <EVENT>
or
<PLACE> prepare <EVENT>,

that do not denote exactly the same situation (the occurrence of the event) but a situation which is presupposed by the second one. A manual verification or an automatic access to WordNet data could enable however to take into account this difference (organize and prepare are in a same WordNet synset). For the moment we did not experiment the automatic verification, but this is a possible future development.
Following the experiment described above, we extracted 35 paraphrase patterns expressing the situation of an event occurrence in a certain place and at a certain date. They include both exact and approximate paraphrases as explained before. We believe that this methodology can be applied to other and can be a valuable resource for different kind of information extraction application (like QA) or even for textual entailment.

4 Evaluation

In order to test our system, we applied it on the second half of the initial corpora. On this part of the corpora, we extract a subset that is potentially in focus for our prototype, i.e. sentences that contain at least a named entity event extracted by the robust parser XIP. This subset consists in about 1500 sentences. We annotate them manually in terms of relations (DATE-AND-PLACE-of-EVENT, DATE-of-EVENT, PLACE-of-EVENT). The confrontation of this corpus with our system gives the following results in term of precision and recall (see Table 1).
It shows that our system gets a very high precision, with an acceptable recall. The recall is a bit low for the triplet relation, because in many cases, our system didn't

Table 1. Evaluation results

	Date-and-place-of-event	Date-of-event	Place-of-Event	All
Precision	90.3	90.9	92.9	92.3
Recall	49.1	80.0	83.7	77
F-measure	63.6	85.1	88.1	83.9

catch the full semantic relation between the 3 elements, but on partial relation (date or place of event). This evaluation shows however that our method is promising. We can indeed relate in a certain way our results with the results obtained[4] by [8] where the authors show that they obtain an f-measure of 72.8 in the detection of most common binary relations between named entities using machine learning methods.

Related work for English (relation detection task between named entities obtained in last ACE competition) shows an overall F-score of 21.6. However it has to be stated that the kind of text provided by ACE included broadcast transcripts and web logs which make the task much more complicated. Furthermore the relations to be recognized are of different kind, see [2].

5 Conclusion

In this paper we present a method illustrated by an effective experiment to extract relations between named entities. Although this field has been studied a lot and that there exists many research work on that matter, our proposal has the particularity to be able to detect semantic relations that are not necessarily conveyed by prototypical syntactic relations (like Subject-Verb-Object relations). This is particularly suitable when one of the related elements is a date or a place as they are very often in very variable syntactic configurations (noun modifier, verb modifier). Our method relies on a first reliable small set of related NE, a parser and a NER system. The first set of related NE can be obtained from various sources. Online encyclopedias are one of them, but this kind of data is also available from specialized databases in specific domains which relate terms, or domain-dependence NE. The prototype we developed showed interesting results in terms of precision for the discovering of new entities relations and we believe that the methodology we adopted can be applied for other kind of relations between named entities. A side effect of the method is that it produces paraphrase or pseudo-paraphrases patterns.

From this first experiment, we expect future work to be done in different directions:

- Apply the method for other kind of relations;
- Try to extend the generalization of the lexico-syntactic patterns by replacing the specific lemmas that anchored the syntactic relations by word-senses;
- Study the effectiveness of paraphrase patterns for paraphrase detection and generation.

[4] Authors are working on the Korean language and they only are interested in binary relations.

References

1. Abney S.: Parsing by Chunks. In Robert Berwick, Steven Abney and Carol Teny (eds.). Principle-based Parsing, Kluwer Academics Publishers.(1991)
2. ACE 2007. NIST 2007 Automatic Content Extraction Evaluation Official Results. http://www.nist.gov/speech/tests/ace/ace07/doc/ace07_eval_official_results_20070402.htm
3. Aït-Mokhtar S., Chanod, J.P., Roux, C.: Robustness beyond Shallowness: Incremental Dependency Parsing. Special issue of NLE journal (2002)
4. Brun, C., Hagège C.: Normalization and Paraphrasing Using Symbolic Methods, Proceeding of the Second International Workshop on Paraphrasing. ACL 2003, Sapporo, Japan (2003)
5. Hagège C., Roux C.: Entre syntaxe et sémantique: Normalisation de l'analyse syntaxique en vue de l'amélioration de l'extraction d'information. Proceedings TALN 2003, Batz-sur-Mer, France. (2003)
6. Iria J., Ciravegna Fabio.: Relation Extraction for Mining the Semantic Web. In proceedings Machine Learning for the Semantic Web, Dagstuhl, DE (2005)
7. Kim S., Yoon J., Yang J.: Kernel approachees for genic interaction extraction. Bioinformatics, Vol. 24, no. 1. pp. 118—126 (2008)
8. Lee Changki, Yi-Gyu Hwang, Myung-Gil Jang.: Fine-Grained Named Entity Recognition and Relation Extraction for Question Answering. Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. Amsterdam, The Netherlands (2007)
9. Levin, B.: English Verb Classes and Alternations – A preliminary Investigation. The University of Chicago Press. (1993)
10. Mel'čuk I.: Dependency Syntax. State University of New York, Albany, N.Y.: The SUNY Press. (1988)
11. Ruppenhofer, J., Michael Ellsworth, Miriam R. L. Petruck, Christopher R Johnson and Jan Scheffczyk. Framenet II: Extended Theory and Practice (2006)
12. Specia Lucia, Motta Enrico.: A hybrid approach for extracting semantic relations from texts. Proceedings of the 2nd Workshop on Ontology Learning and Population: Bridging the Gap between Text and Knowledge, July 2006, Sydney, Australia (2006)
13. Tesnière L,: Eléments de Syntaxe Structurale. Klincksiek Eds. (Corrected edition Paris 1969). (1959)
14. Van Landeghem S., Saeys Y., Van de Peer Y.: Extracting protein-protein interactions from text using rich feature vectors and feature selection. Proceedings of Third International Symposium on Semantic Mining in Biomedicine (MBM 08). pp. 54--84 (2008)