Automatic Extraction of Part-whole Relations

Debela Tesfaye[1] and Michael Zock[2]

[1] IT doctoral program, Addis Ababa University, Addis Ababa, Ethiopia

[2] Laboratoire d'Informatique Fondamentale (LIF), CNRS, UMR 6166, Case 901 - 163 Avenue de Luminy, F-13288, Marseille Cedex 9, France

Abstract. The Web has become a universal repository of knowledge allowing to share information in a scale never seen before. Yet, with the growing size of the resource have grown the difficulties to access information. In this paper we present a system that automatically extracts Part-Whole relations among nouns. The approach is unsupervised, quasi language independent and does not require a huge resource like WordNet. The results show that the patterns used to extract Part-Whole relations can be learned from N-grams.

1 Introduction

The Web has become a universal repository of knowledge allowing to share information in a scale never seen before. Yet, with the expanding size of the resource have grown the difficulties to access information. Finding the targeted information is often a tedious and difficult task. Computers still lack too much of the needed knowledge to understand well the meaning of messages. Hence, information extraction remains a real challenge.

To overcome this problem researchers have started to develop methods to mine information from the web. Text mining is knowledge discovery in a corpus via NLP- and machine-learning techniques. To achieve this goal, i.e. to access the information contained “inside” the documents, one has to be able to go beyond the information given explicitly. Put differently, to find a given document one must understand its content which implies that one is able to identify and understand the concepts and relations evoked in the document.

There are many kinds of relations, for example: Cause-Effect, Instrument-Agency, Product-producer, Origine-entity, Theme-tool, Part-whole, Content-container, etc. We will focus here only on one of them, Part-Whole relations, and their automatic extraction from corpora. Relations can be used for various tasks: information extraction, question answering, automatic construction of ontologies, index building to enhance navigation in electronic dictionaries [17], etc.

The basic premise underlying this work is the idea that the meaning of a word depends to a large extent on neighborhood be it direct, or indirect: words occurring in similar contexts tend to have similar meanings [9]. This idea is by no means new. It is known as the "distributional hypothesis"[1] and is generally related to scholars like

[1]http://en.wikipedia.org/wiki/Distributional_hypothesis

Tesfaye D. and Zock M..
Automatic Extraction of Part-whole Relations.
DOI: 10.5220/0004113801300139
In *Proceedings of the 9th International Workshop on Natural Language Processing and Cognitive Science* (NLPCS-2012), pages 130-139
ISBN: 978-989-8565-16-7

Harris [8], Firth [3] or Wittgenstein [16]. To understand the meaning of a word requires that one takes its use, i.e. its context, into account. This latter can be expressed in terms of (more or less direct) neighbors, i.e. words co-occuring within a defined window (phrase, sentence, paragraph). One way of doing this is to create a vector space composed of the target word and its neighbors [11]. We use this kind of approach to measure the similarity between a pair of words. The vector space model (VSM) has been developed by Salton and his colleagues [14] for information retrieval. The idea is to represent each document in a collection as a point in a space, i.e. a vector in a vector space. Semantic similarity is relative to the distance of two points: closely related points signal similarity, while distant points signal remotely related words. We are concerned here with word similarity rather than document similarity. The meaning of a word is represented as a vector based on the n-gram value of all co-occuring words. The use of the VSM to extract Part-Whole relations has two advantages: it requires little human effort and few corpora, at least far less than Girju's work [4] which relies on annotated corpora and WordNet.

2 Related Work

Several scholars have proposed a taxonomy of Part-Whole relations [4, 14]. We will follow Winston's [15] classical proposal:

(1) *component – integral object* *handle – cup*
(2) *member –collection* *tree – forest*
(3) *portion –mass* *grain – salt*
(4) *stuff – object* *steel – bike*
(5) *feature – activity* *paying – shopping*
(6) *place – area* *oasis – desert*

Integral objects have a structure; their components can be torn apart and their elements have a functional relation with respect to the whole. For example, 'kitchen–apartment' and 'aria–opera' are typical component–integral relations.

'Chairman-committee', 'soldier-army', 'professor-faculty' or 'tree-forest' are typical representatives of Member-Collection relations.

Portion-Mass captures the relations between portions, masses, objects and physical dimensions. For example: 'meter-kilometer'.

The Stuff-Object category encodes the relations between an object and the stuff of which it is made of, be it partly or entirely. For example: 'steel-car', or grape-wine'.

Place-Area captures the relation between an area and a sub-area, or a place within the area. For example, 'Addis Ababa-Ethiopia'.

Relations can also be categorized according to their frequency. Present part-whole relations are relations that are always true, while generally absent relations are those that have happened only at some point in time (episodic presence). Such fine-grained distinctions are necessary for knowledge representation (acquisition of facts) as well as for other tasks like question answering, ontologies, etc. Part-whole relations play a key role in many domains. For example, they are a structuring principle in artifact design (ships, cars), in chemistry (structure of a substance) and in medicine (anatomy).

Most researchers use supervised learning techniques, relying on numerous features and large datasets. For example Girju et al. [4] formed a large corpus of 27,963 negative and 29,134 positive examples by using WordNet, the LA Times (TREC9) and the SemCor 1.7 text to develop three clusters of classification patterns. Matthew and Charniak [13] applied statistical methods to a very large corpus. This kind of approach requires manual tagging of the semantic relation of the concepts occuring in the training set that is expensive in terms of time and use of human resources.

Beamer et al. [1] used the ISA relations, i.e. WordNet's noun hierarchy and the semantic boundaries built from this hierarchy in order to classify new, i.e. not yet encountered noun-noun pairs. Unfortunately, the reliance on non-incremental lexical resources (like WordNet) makes this kind of approach unfeasible for real-world applications. In addition, it cannot be used for languages devoid of such a resource.

Other approaches are domain dependent. For example, Hage et al. [7] focused on food ingredients, while Grad's [6] work dealt with bioscience. These approaches are well suited for a specific domain, but they cannot be used as a general tool.

In addition, the above-mentioned approaches rely on grammar rules (phrase structure, sentence structure rules, part of speech), which again makes them very domain- or language dependent, but not well suited as a general tool.

3 Methodology

Our approach is unsupervised, little dependent on language specific and domain knowledge, and it allows the automatic extraction of meronymic relations (component–integral object; part-whole). All the system needs is a 'Part of Speech Tagger' (POS) or a 'part-of-speech tagged corpus' (our case). This is quite different from other peoples' work in that it does not require a resource like WordNet, which makes the system useful even for under-resourced language, i.e. languages possibly lacking a resource like WordNet. The method works for English and it can be generalized. Hence, it can be adapted to other languages than the one for which it was initially designed. To achieve this we relied on the 'The Corpus of Historical American English' (COHA) that contains 400 million words of the period of 1810-2009 [12]. It is an N-gram corpus tagged for parts of speech [12]. For languages lacking this kind of tagged corpus, plain text can be used, as the system is able to identify the concepts' N-gram value in the corpus. This feature is very convenient for under-resourced languages, making their organization (preparation) easier than annotating corpora manually. The system contains various components (see 3.1-3.6).

3.1 A POS Tagger

This component identifies the part of speech of the sentence elements. Since *part-whole relations* connect only nouns, the system requires only a tagger able to identify nouns. As mentioned already, we used the 'Corpus of Historical American English' (COHA) [12]. This is an N-gram corpus whose elements are tagged in terms of part of speech. The development of a POS tagger being beyond the scope of this paper, we will not address this issue here.

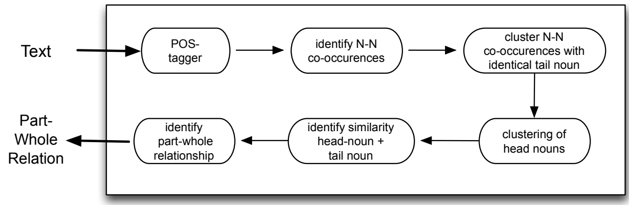

Fig. 1. System information flow.

3.2 Identifier of Noun-noun Co-occurrences

This component takes the tagged corpus as input to identify then noun-noun co-occurrences. There are two types of co-occurrences: nouns occurring directly together, that is, in adjacent position (NN NN) and nouns whose co-occurrence is mediated via another type of word occurring in between them (possibly a preposition, adjective, verbs). Both types need to be identified.

For example, being co-occurrences, the following words are extracted from the corpus: ‘corolla car’, ‘door of car’, ‘car engine’, ‘engine of car’, ‘car design’, ‘network design’,’ airplane engine’, ‘search engine’ etc. As long as the words linking two nouns are syntactically different from the linked nouns (verbs, adjectives, adverbs, preposition), nouns can be easily extracted regardless the type and the number of words in between them (provided these words are not nouns), even if the nouns co-occur in a relatively distant location.

At this level we call the noun at the left hand side the *head* and the one at the right hand side the tail. 'Car' and 'engine' are respectively the head and the tail in the 'car-engine' co-occurrence, while they are the reverse in the 'engine-car' co-occurrence. Hence, cases where the part appears both before and after the whole object will be retrieved. Since the conclusion that an element is the head or the tail may be linguistically incorrect we have decided to delay the decision until the end. In other words, this decision will only be made by the last module (3.6).

This is our first attempt to identify the concepts. The extracted co-occurrences in this module are the potential concepts. The size of the window of words depends on the distance between the noun occurrences (the number of words other than noun occurring between the nouns). Starting from the current noun, we expand the window size till the next noun occurs. Only concepts inside of the window are considered.

Sentences containing nouns expressing a part whole relationship but whose head and tail are separated by another noun cannot be retrieved by the system. Example, ‘some *tables* from the beginning of the 19th century have three *legs*’. ‘Table’ and ‘legs’ express a part-whole relationship, but this relationship cannot be extracted by this module as other nouns ‘century/beginning’ occur in between. However, this relationship can be expressed by a different sentence, contained in the corpus: ‘Some sophisticated tables have three legs.’ In this case the relationship can be extracted, as there is no noun in between. This example shows why we should have a well-balanced corpus, i.e. a corpus rich and representative enough to compensate for certain cases.

Obviously, the type and number of words other than a noun, yet linking the nouns, differ from language to language, but this does not have any impact on this module.

3.3 Clustering of Noun Co-occurrences with Identical Tail Noun

This module clusters noun co-occurrences on the basis of their *tail* noun. Noun co-occurrences sharing the *tail* noun belong to the same cluster. Hence, 'corolla car' and 'door of car' belong to one cluster, both of them having the same *tail* noun: 'car', while 'car design' and 'network design' belong to another cluster. The same holds true for 'airplane engine,' 'search engine' and 'car engine'.

car [corolla, door],
design [car, network],
engine [airplane, search, car]

3.4 Clustering of the Head Nouns

This module clusters the *head* nouns of the noun pairs according to the *tail* noun. This clustering is based on the vector similarity value of the *head* nouns belonging to the same cluster. The classification is made within the clusters created above. It is based on the *head* noun and it is a cluster within a cluster created above. The vector similarity value is derived from the bi-gram value of a word with all other words in the corpus (language ?). For this work we used the co-occurrence value (bi-gram) of a word with all the words occurring in the COCA corpus [12].

All words are represented as a vector of their bi-gram value. Hence, each word has an N-gram value, represented as a vector. In order to calculate the similarity between the *head* nouns we used the cosine value of the vectors of the *head* noun. *Head* nouns whose cosine values are above a certain threshold are clustered together.

Identifying the vector similarity value of words for languages devoid of an N-gram corpus is very expensive in terms of processing (execution time). Thus, the N-gram portion of the system will be executed only once, and the result will be stored in a file to be used later on. This allows using the file at any moment without taxing the system at the wrong moment.

According to the example given here above 'airplane' and 'car' belong to the same cluster, while 'search' belongs to another cluster in the 'airplane-engine', 'search-engine' and 'car-engine' cluster. 'Corolla' and 'door' belong to different clusters in the 'corolla-car' and 'door of car' cluster. Likewise, 'car' and 'network' belong to a different cluster in 'car-design' and 'network-design'. The clusters are shown below.

car { [corolla] [door] }
design { [car] [network] }
engine {[airplane, car] [search]}

3.5 Identification of the Similarity Value between the Head-noun and the Tail-noun

This module is the most important one. It identifies the similarity value between the *tail* noun and the clusters (*head* noun). The basic idea is that the *tail* nouns of the noun pairs presenting the 'Component-Integral object' or a 'Part-Whole relation' have a strong similarity value in their clusters.

In this module two types of similarity values are calculated. We call them S_1 and S_2. The two vectors created for S_1 and S_2 in this module are different from the one used in the module above. The vector for S_1 is built by considering the words co-occurring with the *tail* noun only.

If a word co-occurs both with the *tail* and the *head* noun, the N-gram value is recorded in both vectors, otherwise their respective vector values will be 1 (*tail* noun) and zero (*head* noun). Words co-occuring only with the *head* noun will not be included in the vector. Hence, the size of the vector is equal to the size of the number of words co-occurring with the *tail* noun. However, in order to create a vector for S_2, we will also consider words co-occurring with the *head* noun. The similarity value (C, for cosinus) between the *head* nouns in the clusters is calculated on the basis of the bi-gram value of their co-occurences. Hence, words like 'airplane' and 'car' have a strong similarity value with respect 'engine', while 'search' has only a small value: airplane-engine, 'search-engine', 'car-engine'.

3.6 Module for Identifying Part-whole Relations

This module identifies whether two nouns are linked via an *integral component Part-Whole relation* or not. To this end, the system relies on information provided by the above-mentioned modules. From the clusters and the similarity values identified above in the training corpus, the system extracts automatically a production rule (if <condition> then <action>) identifying whether two words are linked via an integral component Part-Whole relation. This rule is very simple, as it requires only a few steps.

The rules: Given a pair of nouns as described in section 3.4

If the similarity value $S_2 > 0.4$ && if the similarity value $S_1 > 0.8$

If the noun pairs occurred at least once as compound noun

Then the *head* noun refers to the whole and the *tail* to the part

Else

If the average similarity value (C) between the noun and the co-occuring nouns in the cluster > 0.4

If one of the nouns in the cluster has $S_2 > 0.4$ and $S_1 > 0.8$

If the noun pairs occurred at least once as compound noun

Then the *head* noun refers to the whole and the *tail* to the part

Else

The relationship between the nouns is other than a whole-part relation

The rule stipulating that ‘noun pairs occurring at least once as compound noun’, does not imply that the noun referring to the 'part' is always the second noun, and the 'whole' the first. Indeed, the two may be separated by other types of words, for example, a preposition, in which case the position of the arguments will swap, the 'part' preceding the 'whole'. Both cases will be handled as discussed in section 3.2. After extracting the nouns for both cases, we can find the pairs as a compound noun at least once in a well-balanced corpus. For example, ‘engine of car’ can be extracted as explained in section 3.2 and the system interpretes the pair as 'part-whole' if it exists as ‘car engine’, which is always the case in a well-balanced English corpus. For other languages this should change according to the rules of the langue. Concerning the syntax of part-whole relations it should be noted that, while the two arguments (the 'whole' or the 'part') can in principle be expressed in either order, some languages allowing even for both of them, every culture, i.e. language, tends to have its own preferences. For example, in Amharic the order is like in English, in Afan Oromo, the opposite is the case: the *part* is mentioned before the *whole*.

Example:

Cluster 1: Vehicles [car-engine, train-engine, airplane-engine]

Comment: There is an integral component Part-Whole relation, 'engine' being part of the set of holistic entities: vehicles, car, train, and airplane.

Cluster 2: Oil [benzene-engine, gasoline-engine]

Comment: 'Engine' is not part of 'oil' (benzine or gasoline).

The above two clusters are created within a cluster having engine as supporting ('root' ?) noun. The clusters are identified based on the similarity value among the *head* nouns. Accordingly 'car', 'train', and 'airplane' have a strong similarity value. 'Benzine' and 'gasoline' belong to the same cluster as they have a strong similarity value. The vector similarity of 'engine' and 'oil' cluster is below the threshold value, while the one of 'engine' and 'vehicle' is above it. If the similarity valueswere below it, we would conclude that there is no Part-Whole relationship.

Relational knowledge harvesting is generally based on patterns or clustering. Yet, it can be improved by vector-based methods. Cederberg and Widdowson [2] showed how Latent Semantic Analysis (LSA)[10] can help to improve hyponymy extraction from free text. While LSA is certainly a very powerful method for semantic analysis, we used a different method. Both approaches rely on word distribution for calculating semantic similarity, but our approach is different in a number of ways:

1. LSA uses a 'term-by-document-space' to create a vector for a word. This requires thousands of documents, whereas our approach requires only bi-gram information of a single well-balanced corpus in order to create a vector of words.

2. LSA uses SVD to reduce the size of the vectors. We do not need SVD at all, as the vectors are built only on the basis of important N-gram values.

3. LSA considers every word co-occurrence in all documents. In our model the vector space built for S1 considers only the N-gram value of the tail.

4. LSA uses second order co-occurrence information while we rely only on simple co-occurrences.

In sum, according to our knowledge, there is no other work trying to extract part-whole relation via a vector-based approach as we do (including LSA).

4 Evaluation

In order to test our system for the extraction of part–whole relations we used the text collections of SemEval [4]. The test corpus is POS-tagged and annotated in terms of WordNet senses. The corpus has positive and negative semantic relations. The part–whole relations extracted by the system were validated by comparing them with the valid relations labeled in the test set answer key. The format of the test set is described in the sample here below:

"Some sophisticated <e2>tables</e2> have three <e1>legs</e1>."
WordNet(e1) = "n3", WordNet(e2)="n2"; Part-Whole(e1, e2) = "true"

This format has been defined by Girju et al [5]. Since this does not correspond to a real text format, we have changed the corpus accordingly, to obtain the following text: "Some sophisticated tables have three legs". To evaluate the performance of our system we defined precision, recall, and F-measure performance metrics in the following way:

Recall	$\frac{\text{Number of correctly retrieved relations}}{\text{Number of correct relations}}$
Precision	$\frac{\text{Number of correctly retrieved relations}}{\text{Number of relations retrieved}}$
F-measure	$\frac{2}{\frac{1}{precision}+\frac{1}{recall}}$

Our system identified ALL of the *present Component-Integral object part-whole* relation pairs in the SemEval test set. *Component-Integral object part-whole relations* in the Semeval training and test set are both present and non-present. This being so, we considered the present relations to evaluate the performance of our approach. The performance of the system is evaluated towards identifying *present Component-Integral part whole relation* in our test set. As explained in section 3.2 there are sentences with nouns expressing part whole relations that cannot be detected due to the fact that there is one or more nouns between the head and the tail. We did not encounter such sentences in the SemEval test set. However, the nouns can be extracted from other sentences linking the nouns with words belonging to another syntactic category. We believe that the nouns can be extracted from a well-balanced corpus containing other sentences expressing the very same relation by using patterns.

It should be noted that sentence-based evaluation (as there are sentences not supported by this approach) and corpus-based evaluation yield different results. Of course, the system should cover both types. To allow for this one could pair all nouns in a given sentence to form a matrix. The remaining modules will check all the pairs and those satisfying the constraint specified by rule will be selected.

As the number of *present Component-Integral object part-whole relation* in the SemEval test set is small, we added the ones from the SemEval training set and some of our own in order. The resulting number of relation pairs accounts now for 20% of our test set. 80 % of this set contains negative examples coming either from the SemEval test set (all of them) or from our own. As result we achieved a precision of 95, 2%; a recall and F-measure of 95%. All the encountered errors are hyponyms ('car' and 'vehicle'), which does not imply by any means that all the hyponyms in the test are incorrectly retrieved as part-whole relation. Actually, only 12% of the hyponyms in the test set are incorrectly retrieved as part-whole relation.

It should also be noted that the majority (80%) of our test set relations are not *present Component-Integral object part-whole relations.* Therefore, the probability of randomly selecting *present Component-Integral object part-whole relation* is 20/80 (0.25), which shows the effectiveness of this approach for discriminating this kind of relations.

4.1 Comparison with Related Works

As mentionned earlier, other researchers have tried to extract part-whole relation, and their work [4] generally requires a lot of human resources as it relies on the manual annotation of the training corpus to extract all the types of the part-whole relations mentioned in section 2. In addition, it requires the use of WordNet features to extract the pattern. Also, Girju did not distinguish between 'present part-whole relations' and 'absent' relations as defined by us.

Yet, this kind of distinction is very important for knowledge representation (acquisition of facts) and question answering. Part-whole relations play a key role in many applications and in many domains. For example, they are the central structuring principle in artifact design (ships, cars), in chemistry (structure of a substance) and in medicine (anatomy). Such relations are omni-present in these domains. Therefore, we need systems being able to distinguish present part-whole relations from non-present ones.

Also, most other approaches are highly language dependent. The features used to build the rules are extracted from a corpus pertaining to a specific language. For example in Girju et al [4] and Matthew and Charniak [13] most of the rules are based on prepositions like 'of', 'in', the 's' of the genitive (the NOUN's) and the auxiliary verb 'to have'. This makes them fairly useless for other languages. Nevertheless, one must admit that they have covered a wide range of part-whole relation types and that they did obtain encouraging results.

By contrast, our approach depends little on language: the classification features are based on word distributions. The features are mainly dependent on the N-gram value of the concepts, the latter depending entirely on the corpus. The N-gram value measures the co-occurrence value of the concepts with all other content-bearing words in the training corpus. This value is represented as a vector allowing to measure the similarity between a given pair of concepts in a given context. The rules are learned from word distributions in a corpus only, sentence-structure or phrase-structure is not considered at all. However, our approach needs to be tested for other languages in order to attest the validity and the scope of our claims.

5 Conclusions and Future Work

We proposed a new approach to perform semantic analysis. It depends on the distribution of words for calculating their similarity. The results obtained are quite promising despite the fact that few resources are used compared to other work. . We believe that this approach can also be used to extract other types of semantic relations. The results also showed that the patterns used to extract semantic relations could be learned from N-gram information.

This latter can be extracted in different ways; we have shown two of them (S_1 and S_2) for extracting *present Component-Integral object Part-Whole relation*. We have also noted that some hyponyms and *present Component-Integral part-whole relation* tend to exhibit to a large extent the same kind of N-gram pattern.

While our current goal has been the identification of part-whole relations, we have addressed so far only a small part of the problem. Our next step will consist in extracting the remaining part-whole relations (see, section 2). We also need to check the applicability of our approach for other languages. This should allow us to evaluate the relative efficiency of the system in language dependency. There are cases where nouns cannot be extracted by this approach. However, we believe that our approach can be used to extract *present component-integral part-whole relation* from a well balanced-corpus to help building resources like WordNet or an ontology in a given language. For applications requiring sentence-based semantics, the approach needs to be modified in order to cover sentences containing nouns with a part-whole relation, but whose head and tail are separated by one or several nouns.

References

1. Beamer, B., Rozovskaya, A. and Girju, A., (2008). Automatic Semantic Relation Extraction with Multiple Boundary Generation. Association for the Advancement of Artificial Intelligence.
2. Cederberg, S. and D. Widdows (2003). Using LSA and Noun Coordination Information to Improve the Precision and Recall of Automatic Hyponymy Extraction. In Conference on Natural Language Learning (CoNLL-2003), Edmonton, Canada, pp. 111-118.
3. Firth, J. R., (1957). A synopsis of linguistic theory 1930-1955. In Studies in Linguistic Analysis, pp. 1-32. Oxford: Philological Society.
4. Girju R., Moldovan D., Tatu, M. and Antohe, D., (2005). Automatic Discovery of Part–Whole Relations. ACM 32(1)
5. Girju, R., Hearst, M., Nakov, P., Nastase, V., Szpakowicz; S., Turney, P. and Yuret D., (2007). Classification of Semantic Relations between Nominals: Dataset for Task 4 in SemEval, 4th International Workshop on Semantic Evaluations, Prague, Czech Republic.
6. Grad, B., (1995). Extraction of semantic relations from bioscience text. University of Trieste, Italy.
7. Hage, W., Kolb, H. and Schreiber, G., (2006). A Method for Learning Part-Whole Relations. TNO Science & Industry Delft, Vrije Universiteit Amsterdam.
8. Harris, Z., (1954). Distributional structure. Word 10 (23), 46–162.
9. Harshman, R., (1970). Foundations of the parafac procedure: Models and conditions for an “explanatory” multi-modal factor analysis. UCLA Working Papers in Phonetics, 16.
10. Landauer, T., Foltz P., & Laham, D., (1998). Introduction to Latent Semantic Analysis. Discourse processes, 25, 259-284.
11. Lund, K. and Burgess, C., (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. Behavior Research Methods, Instruments, and Computers, 28(2), 203–208.
12. Mark, D., (2011). N-grams and word frequency data from the Corpus of Historical American English (COHA). http://www.ngrams.info.
13. Matthew, B. and Charniak, E., (1999). Finding parts in very large corpora. In Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL1999), pages 57–64, University of Maryland.
14. Salton, G., Wong, A. and Yang, C.-S., (1975). A vector space model for automatic indexing. Communications of the ACM, 18(11), 613–620.
15. Winston, M., Chaffin, R. and Hermann, D., (1987). Taxonomy of part-whole relations. Cognitive Science, 11(4), 417–444.
16. Wittgenstein, L., (1922). Tractatus Logico-Philosophicus. London: Routledge & Kegan Paul.
17. Zock, M.: Needles in a haystack and how to find them? The case of lexical access. In E. Miyares Bermudez and L. Ruiz Miyares (Eds.) "Linguistics in the Twenty First Century". Cambridge Scholars Press, 2006, United Kingdom, pp. 155-162.