Learning of Semantic Relations between Ontology Concepts using Statistical Techniques

Athanasios Tegos[1,2], Vangelis Karkaletsis[1], and Alexandros Potamianos[2]

[1] Institute of informatics and Telecommunications, NCSR "Demokritos", Greece
{tegos,vangelis}@iit.demokritos.gr
[2] Department of Electronics and Computer Engineering, Technical University of Crete, Greece
potam@telecom.tuc.gr

Abstract. Acquiring domain knowledge for constructing ontologies is a resource demanding and time consuming task. Thus, the automatic or semi-automatic construction, enrichment and adaptation of ontologies, the so-called ontology learning task is highly desired. Although an emerging field, a significant amount of research has been performed in ontology learning, leading to a large number of proposed approaches and practical systems. This paper presents our approach on automated learning of ontologies from texts which are semantically annotated with instances of ontologies' concepts. Statistical techniques are applied to metadata extracted from the annotated texts, to discover semantic relations among the annotated concepts as well as to find cardinality restrictions for these concepts and their relations. The proposed method was applied to corpora from two different domains, athletics and biomedical, and was evaluated against the existing manually created ontologies for these domains.

1 Introduction

Ontologies are the backbone of the semantic web as well as of a growing number of knowledge-based systems. Research on automated development of ontologies from text has become increasingly important because manual construction of ontologies is labor intensive and costly and at the same time, a large amount of texts for specific domains is already available in electronic form. Domain ontologies consist of concepts, semantic relations among these concepts, and a set of inference rules. Thus, the process of ontology learning from text includes three core subtasks: learning of the concepts that will constitute the ontology, learning of the semantic relations among these concepts and finally, learning of a set of inference rules.

This paper presents our approach for the discovery of semantic relations between ontology concepts. In this work we distinguish between high-level and low-level concepts. For example, in a text document the name or the age of a person is an instance of a low-level concept because instances of these concepts are associated with relevant text portions. On the other hand, the concept person is not a low-level concept, as it is a "compound" concept in such a way that instances of this concept are related to instances of name, age, gender or maybe compound concepts. Compound concepts are referred to as high-level concepts, and instances of such concepts are not directly identiable in a text document. In comparison with other relevant approaches, we focus on the discovery of relations between high-level concepts, but we also show the applicability of the proposed approach to low-level concepts.

The discovery process is not based on the assumption that verbs typically indicate semantic relations between concepts and does not exploit lexico-syntactic patterns or clustering methods or any external knowledge sources like WorldNet. Our approach is based on the assumption that concepts which are semantically related, tend to be "near" in a plain text. This assumption arises from the principle of coherence on linguistics [4]. Based on this assumption, statistical methods are applied to metadata extracted from the annotated texts, to discover semantic directed relations between concepts. Moreover, our approach is able to find cardinality restrictions on concepts' relations.

The rest of this paper is organized as follows. Section 2 presents a short survey of the related work. Section 3 presents the proposed method for the discovery of relations between high-level concepts but also between low-level concepts. The experimental results and the evaluation of the proposed method for two different domain ontologies are presented in section 4. Finally, section 5 summarizes the paper and outlines directions for future work.

2 Related Work

Various techniques are presented in the literature for the learning of semantic relations among concepts. These techniques are divided in two categories, in those that learn taxonomic relations and in those that learn non-taxonomic relations between concepts.

Existing techniques for finding taxonomic (hierarchical) relations can be classified into pattern-based, clustering-based and combination of both. In pattern-based approaches, the user defines a set of lexico-syntactic patterns [5], which are applied to the texts to obtain instances of taxonomic relations. In clustering-based approaches [6], hierarchical clustering algorithms are used for finding the taxonomic relations between the concepts. In combined approaches [7], lexico-syntactic patterns are first applied in the text and then clustering techniques are used to filter the extracted taxonomic relations.

Few methods exist for the extraction of non-taxonomic relations from domain texts. In [1], relations extraction is considered as learning of selectional restrictions for verbs. According to this method, all terms co-occurring with a verb are clustered and each of the clusters is manually labeled. Methods presented in [8], [2], [9] exploit the syntactic structure and dependencies between words for relations extraction. Both [8] and [9] extract concept pairs exploiting dependency relations and use the chi-square test to verify the statistical significance of concept co-occurrence. Schutz et al. [9] technique builds the relation triples by extracting relevant pairs. In [8], the chi-square test is employed to learn patterns such as $SUBJ \rightarrow bind \rightarrow DIROBJ$ and then learned patterns are used to extract semantic relations. Kavalec et al. [2] approach, extracts triples (C_1, V, C_2) such that concepts C_1 and C_2 occur within a predefined distance from verb V in the domain text. Punuru et al. [3], extend Kavalec's [2] approach by directing the semantic relation between the concepts C_1 and C_2 ($C_1 \rightarrow C_2$) using the principle $SubjectConcept \rightarrow ObjectConcept$. Another approach is the one presented in [11]. This is a supervised machine learning approach which extractes binary relations between named entities (i.e. low-level concepts) already identified in texts using a named entity recognizer. Operating at the sentence level, a context-free grammar which captures the patterns connecting relating entities, is inferred from positive examples.

Our approach enables the discovery of both taxonomic and non-taxonomic relations among the concepts that have been annotated in the text. Furthermore, it does not require any training with a data set or the use any syntactic analysis in order to extract the relations.

3 The Proposed Method

The proposed method for ontology learning involves 2 major steps:

1. Finding the semantic relations of concepts that have been annotated in the corpus.
2. Finding the cardinality restrictions for the extracted relations.

The 34-year-old, World marathon record holder and two-time Olympic and four-time World 10,000m champion Haile Gebreselassie of Ethiopia today announced that he intends to compete in this 2008 FKB-Games - IAAF World Athletics Tour - in Hengelo, the Netherlands on 24 May in his bid to make Ethiopia's team for the Beijing Olympics in China.

Athlete (name:*Haile Gebreselassie*, age:*34*, nationality: *Ethiopia*, gender:*NotFound*)
SportsCompetition (sport-name:*10,000m*, city:*Hengelo*, stadium-name: *NotFound*, date:*24 May*)

Fig. 1. Text annotated with instances of two high-level concepts.

As noted before, the method requires the annotation of the corpus with instances of ontology's concepts. In the case of high-level concepts as instances we consider the fillers of the concept's attributes that have been found in a document. Fig. 1 shows an example of a text annotated with instances of the high-level concepts

Athlete, *SportsCompetition* of a domain ontology on athletics developed in the context of the project BOEMIE[3].

It is not necessary to find fillers for all the attributes of a high-level concept in the text, in order to annotate its instance. An instance of a high-level concept is annotated when specific attribute fillers are found that contain "enough" semantic information. As shown at Fig. 1, the athlete's instance does not contain a filler for the attribute gender. However, all instances of the athlete concept contain a filler for the athlete's name.

The application of the proposed method on high-level concepts is presented in sub-sections 3.1, 3.2, whereas the application on low-level concepts is examined in sub-section 3.3.

3.1 Discovery of Semantic Relations between High-Level Concepts

As noted previously, our approach is based on the assumption that concepts which are semantically related tend to co-occur near each other in a plain text, i.e., spatial proximity in text implies semantic similarity. Based on this assumption, we treat each document of the corpus as a sequence of symbols. We consider as symbols all the characters, including spaces and the punctuation marks that exist in the document. In this manner, each document is represented in a one-dimensional Euclidean space, depending on the place in which each symbol is found in the text. For example, the phrase "*The 34-year-old, World marathon record holder*" is represented with the set $[0, 44]$ because the text is a sequence of 45 symbols. In the same example, the offset of the phrase "*34-year-old, World marathon*" is the set $[4, 30]$, since the phrase starts from the 4^{th} symbol and ends at the 30^{th}.

Based on the aforementioned transformation of the documents, we first find for each document the offsets of the annotated instances. As mentioned in the previous section, each instance is formed of the fillers of the concept's attributes found in the text. Consequently, the offset of an instance is defined as the range from the first to the last symbol of the instance's fillers. For the example document shown at Fig. 1, whose offset is the set $[0 - 342]$, the offset for the athlete's instance is the set $[4 - 134]$, since it is the minimum part of text which encloses all its fillers.

For each document, we search for the different pairs of concepts that have overlapping instances. Specifically:

For the document doc_z, of the corpus:
$C_{doc_z} = \{C_1, C_2, \ldots, C_n\} \quad where \quad C_i = \{I_1, I_2, \ldots, I_m\}$
$where \quad I_k = [l, r] \bigcap \mathbb{N} \quad and \quad l < r,$
we compare the instances' offsets:
$\forall (I_x, I_y) \quad where \quad I_x \in C_i, \quad I_y \in C_j$
$and \quad C_i \in C_{doc_z} \quad and \quad C_j \in C_{doc_z} - \{C_i\}$

$$If\Big(I_x \bigcap I_y \neq \emptyset\Big) \quad then\ create\ a\ pair\ \Big(C_i, C_j\Big)\ for\ doc_z \tag{1}$$

[3] http://www.boemie.org

Where:
C_{doc_z}*: the set with the different concepts that have been annotated at least once in the document* doc_z
C_i*: the set with the instances of concept* C_i*, which have been annotated in the document* doc_z
I_k*: the offset of instance* I_k

For each document, a list of concept pairs is created according to (1). We assume that spatial overlap between concept instances in the text implies a semantic relationship. Thus after applying (1) to the corpus, we create a list per document with the different pairs of related concepts. Note that for each document we are interested only in finding the different pairs of related concepts and not the number of occurrences (or overlapping co-occurrence) for each of these pairs.

Then, in order to find the semantic relations between concepts, we propose the semantic-correlation metric $S(C_i \to C_j)$ between two concepts C_i and C_j. This metric measures the tendency of concept C_i to be semantically related, either taxonomically or non-taxonomically, with concept C_j, but not the inverse. The semantic-correlation metric (2), is defined as the product of the conditional probability $P(C_j|C_i)$ with the sum of the mutual information measure $I(C_i, C_j)$ plus 1. This definition is based on our initial assumption that concepts which are semantically related, tend to co-occur "near" each other in a plain text. Therefore, concepts whose instance offsets overlap frequently tend to be semantically related. For the above reason we use in our metric the conditional probability $P(C_j|C_i)$, in order to find for the concept C_i the most probable concept C_j with which is semantically related. Furthermore, the mutual information measure [14] is used in order to enhance our metric with the association between the concepts C_i and C_j. If there is a strong association between C_i and C_j, then the conditional probability $P(C_j|C_i) \gg P(C_i) \cdot P(C_j)$, and consequently $I(C_i, C_j) \gg 0$. If there is no interesting association between C_i and C_j, then $P(C_j|C_i) \approx P(C_i) \cdot P(C_j)$, and thus, $I(C_i, C_j) \approx 0$. If C_i and C_j are not associated, then $P(C_j|C_i) \ll P(C_i) \cdot P(C_j)$, forcing $I(C_i, C_j) \ll 0$. We estimate the probabilities by treating each of the different concepts, which have been annotated in the corpus, as a different event and the extracted pairs of related concepts per document (1) is the set of our observations for the different events. We use maximum likelihood estimation to estimate the probabilities of events, by counting event frequencies in the *set of documents*.

$$S(C_i \to C_j) = P(C_j|C_i) \cdot \Big(1 + I(C_i, C_j)\Big) = \\ = P(C_j|C_i) \cdot \left(1 + log\left(\frac{P(C_j|C_i)}{P(C_i) \cdot P(C_j)}\right)\right) \quad (2)$$

In order to find for a concept C_i the concept with which is semantically related (either taxonomically or non-taxonomically), we compute using our proposed metric the semantic-correlation scores between C_i and each of the rest of the concepts. The concept that maximizes this score (3) is the concept with which the concept C_i is related to.

Find how concepts are related:
$C_{corpus} = \{C_1, C_2, \ldots, C_n\}\,, \quad \forall C_i \in C_{corpus},$

$$RELATE \quad C_i \rightarrow C_j, \quad \arg\max_{C_j} S\Big(C_i \rightarrow C_j\Big), \tag{3}$$

where $C_j \in C_{corpus} - \{C_i\}$

Applying the aforementioned methodology to the annotated corpus, we manage to find the directed semantic relations between the annotated concepts. The proposed method does not use any lexicon-syntactic patterns and clustering methods, or any external knowledge like WorldNet. We simply apply statistical methods to document metadata that is, to the location of concept instances in text.

3.2 Finding of the Cardinality Restrictions for the Discovered Relations

Apart from the discovery of the semantic relations between ontology concepts, the proposed method is also able to find cardinality restrictions between the instances of the related concepts.

The types of connectivity among two related concepts, that the proposed methodology is able to specify, are $1 : N$ (one-to-many), $N : 1$ (many-to-one) and $M : N$ (many-to-many). We find the type of connectivity between two concepts, based on the initial assumption that concepts, whose instance offsets overlap, tend to be related. Hence, we specify as type of connectivity between the instances of two related concepts the type which occurs more often in the corpus. The proposed methodology consists of the following steps:

1. For each document in the corpus that contains instances of the concepts $C_A = \{I_{A_i}, \ldots\}$ and $C_B = \big\{I_{B_j}, \ldots\big\}$, we create a list with the overlapping instances, of the concepts C_A and C_B.
2. For each list, we find the type of connectivity, for each document, between the instances of concepts C_A and C_B as follows:
$$\left.\begin{array}{l} I_{A_i}, I_{B_j} \\ I_{A_i}, I_{B_m} \\ \ldots \end{array}\right\} \Rightarrow \big(1:N\big) \; or \; \left.\begin{array}{l} I_{A_i}, I_{B_j} \\ I_{A_k}, I_{B_j} \\ \ldots \end{array}\right\} \Rightarrow \big(N:1\big) \; or \; \left.\begin{array}{l} I_{A_i}, I_{B_j} \\ I_{A_j}, I_{B_k} \\ \ldots \end{array}\right\} \Rightarrow \big(M:N\big)$$
3. We specify as cardinality restriction, for the related instances of concepts C_A and C_B, the type of connectivity that occurs more often in the corpus.

3.3 Discovery of Semantic Relations between Low-Level Concepts

In order to find semantic relations among low-level concepts, we apply the proposed methodology with a variation on the denition of the instance offset of each low-level concept. Specically, we extend the offset of each instance by X symbols to the left and to the right. For example, in Fig. 1, concerning the *nationality* low-level concept, the offset of its instance (*Ethiopia*), with a window size X is $[(127-X),(134+X)]$. So, if the window size is 10 symbols, then this instance offset will be $[117, 144]$. The usage of a window size, is motivated by the fact that instances of low-level concepts contain very few words and thus semantically related concepts might be near each other in the text but not overlapping.

Using a named entity recogniser for the annotation of low-level concepts instances we will be able to further automate the application of the proposed variation to discover semantic relations between low-level concepts. Consequently this would enable us to apply the complete methodology, as described in subsections 3.1-3.2. In section 4.1, we present initial experimental results of this variation.

4 Experimental Results

The proposed method was applied on two corpora of different domains and the extracted ontologies were evaluated with respect to the corresponding manually created ontologies. The first corpus is from the athletics domain and consists of 2087 web pages, with content, collected mainly from the IAAF[4] web site. This corpus contains annotations of instances of 20 different high-level concepts. The second corpus is from the biomedical domain and consists of 286 abstracts of Pubmed[5] on allergens. The second corpus contains instance annotations of 6 different high-level concepts.

The corpus from the athletics domain was obtained from the EC-funded project BOEMIE. It contains 36,240 instances annotations for 20 high-level concepts and has already been used in [10], [11]. The corpus documents contain athletic articles for 10 different sports competitions. A part of the manually created ontology developed in the context of the same project, is presented in Fig. 2(a). The 20 high-level concepts with their attributes that have been annotated in the corpus are:

> ***Athlete, MaleAthlete, FemaleAthlete*** *(name, age, gender, nationality)*
> ***SportsRound*** *(round-name, date)*
> ***SportsEvent*** *(event- name, city, country, date)*
> ***SportsTrial*** *(performance, ranking)*
> ***SportCompetition, JumpingCompetition, ThrowingCompetition, RunningCompetition, TripleJumpCompetition, PoleVaultCompetition, HighJumpCompetition, LongJumpCompetition, HammerThrowCompetition, JavelingThrownCompetition , HurdlingCompetition, Running100mCompetition, MarathonCompetition, RaceWalkingCompetition*** *(sport name, city, date, stadium-name)*

[4] http://www.iaaf.org
[5] http://www.ncbi.nlm.nih.gov/pubmed/

Applying our method in this corpus, the ontology of Fig. 2(b) is constructed. Comparing to the manually created ontology, it can be noticed that only two out of the nineteen extracted relations are different. The method relates the concept *Athlete* with the concept *SportsCompetition*, instead of relating it with the concept *SportsRound*, which nevertheless is not semantically incorrect. The other missed relation is among the concepts *RunningCompetition* and *SportsEvents*, instead of relating it, with the concept *SportsCompetitions*. This was due to the fact that the Marathons' names in most times, in the documents, are mentioned with the city in which they took place (e.g. London-Marathon, Berlin-Marathon,. . .). That has as effect, the instances' offsets of MarathonCompetition to be overlapped with the instances' offsets of *SportsEvent*, because the *SportsEvent* concept has the attribute city. We have evaluated our method without the MarathonCompetition's instances in the corpus and the *RunningCompetition* concept was related correctly, with the *SportsCompetition* concept.

Figure 2(b) depicts also the type of connectivity found between the instances of the related concepts. For example, the method specified the relation between the concepts *Athlete* and *SportsCompetition* as of type $(N : 1)$, which is reasonable, because many athletes participate to one sport. Also the relation between *SportsRound* and *SportsCompetition* as of type $(N : 1)$, which is also reasonable, since one sport has many rounds.

The second corpus, from the biomedical domain, contains 1887 instances' annotations for 6 different high-level concepts and has also been used in [12]. The manually created ontology is depicted in Fig. 3(a). The 6 high-level concepts, instances of which have been annotated in the corpus are:

Allergens *(AllergenName common, Allergen Name scientific, Isoelectric Point, Molecular Weight, majorORminor)*

Protein *(protein Family, protein Name)*

Allergie *(Allergen Group)*

Allergen Sources *(source common name, source scientific name)*

Named Allergens *(AllergenName common, Allergen Name scientific, Isoelectric Point, Molecular Weight, majorORminor)*

Descriptive Allergens *(AllergenName common, Isoelectric Point, Molecular Weight, majorORminor)*

Applying our proposed approach in the corpus of allergens, the ontology depicted in Fig. 3(b) is constructed. Comparing to the manually created allergen ontology, it can be noticed that all semantic relations have been discovered. Figure 3(b) depicts also the types of connectivity, which have been specified automatically, between the related concepts.

Finally, one should bear in mind that the evaluation of ontologies when these ontologies are produced by an automated learning procedure is an open field of research. A standard methodology for automating ontology evaluation is still to be established [13].

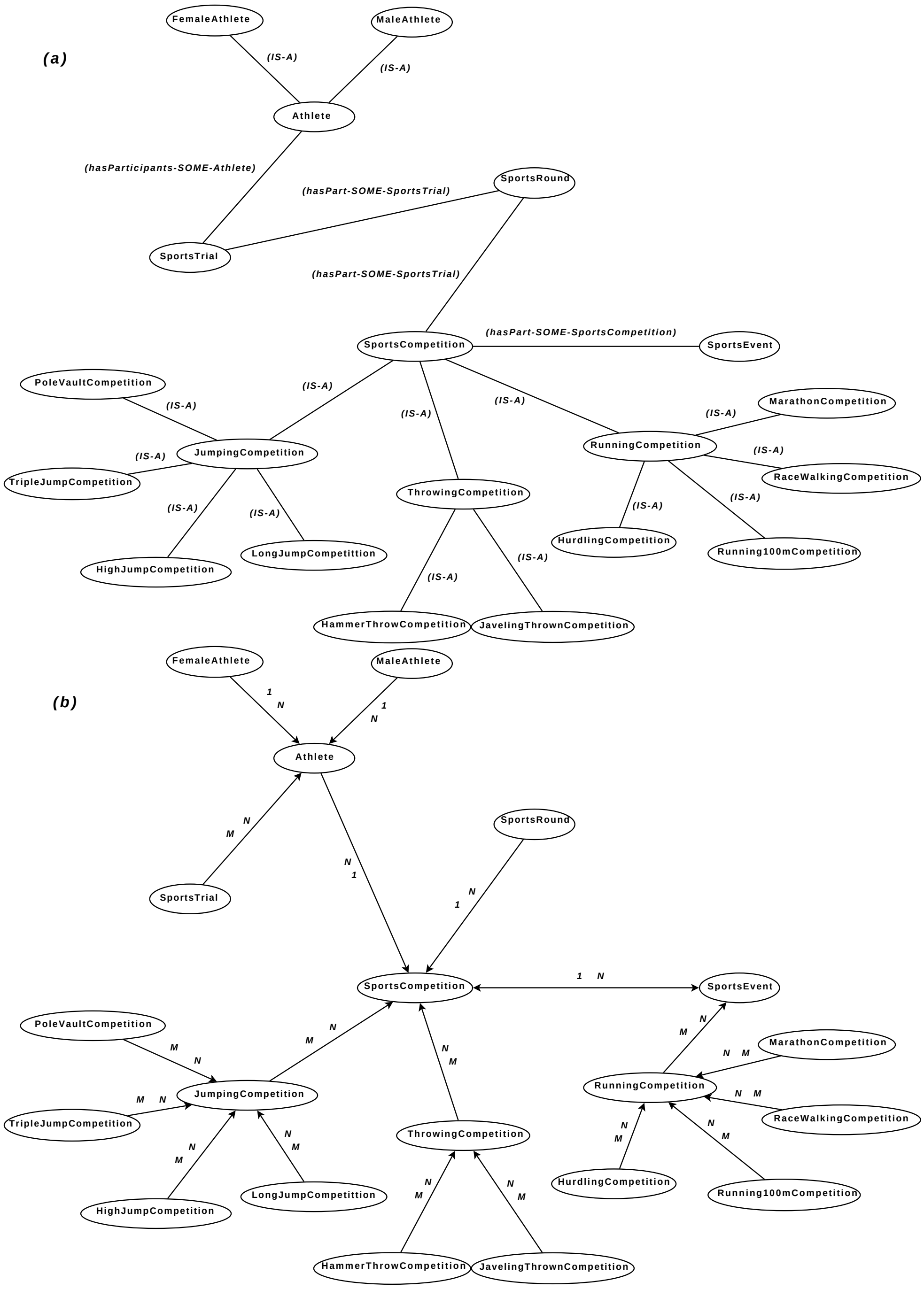

Fig. 2. (a) The manually created ontology for the domain of athletics. (b) The automatically extracted ontology.

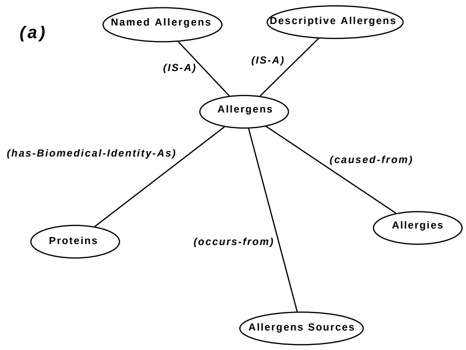

Fig. 3. (a) The manually created ontology for the domain of allergens. (b) The automatically extracted ontology.

4.1 Experimental Results for Low-Level Concepts

We applied the proposed methodology, as presented in section 3.3, on the aforementioned corpus from the athletics domain. Here the low-level concepts are the thirteen different attributes used in the 20 high-level concepts. The corpus contains 56494 concept instance annotations for the thirteen low-level concepts. The 13 low-level concepts are: *Name, Age, Nationality, Gender, Round-name, Date, Event-name, City, Country, Performance, Ranking, Sport-name, Stadium-name.*

Applying the proposed method on this corpus, for a window size X of 50-symbols, discovers the semantic relations among the low-level concepts and also their type of connectivity, which are depicted in Fig. 4.

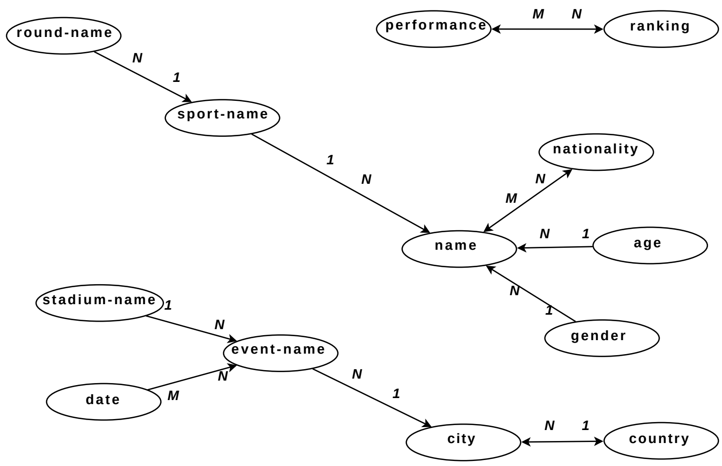

Fig. 4. The extracted semantic relations among low-level concepts, using a window size of 50-symbols.

Observing the discovered relations of the low-level concepts and also their type of connectivity, it can be noticed that they seem reasonable. The system relates many *round-names* with one *sport-name* and one *sport-name* with many *names*. Moreover, it relates the concepts *nationality, age* and *gender* with the concept *name*. One *stadium-name* is related with many *event-names*, which are also related with one *city* and many *cities* are related with one *country*. It is also remarkable that the method also clusters the low-level concepts. As depicted in Fig. 4, our proposed algorithm has discovered three clusters of related low-level concepts. Each of these clusters can be considered as a high-level concept which consists of low-level concepts.

The window size (*WS*) for this experiment was 50-symbols. The same results are also discovered for window size 100-symbols. For window size larger than 100-symbols, we observed that all the low-level concepts tend to be related with the concept *name*. This is expected since the concept *name* is the more frequently occurring one. In general, the usage of a large *WS* leads to over-generation of semantic relationships as an increasing number of concept instances are now overlapping. From experimentation with the WS for different corpora and different low-level concepts, we conclude that the best *WS* is related with the density of the annotated concept instances in the text. The rule of thumb is: the higher the density the lower the *WS* should be and vice versa.

5 Conclusions and Future Work

In this paper, we presented a novel method for discovering semantic relations between high-level and low-level concepts. We propose a statistical method which is able to extract directed semantic relations among the annotated concepts and also to find cardinality restrictions for these concepts and their relations. Our approach is based on the assumption that concepts which are semantically related, tend to co-occur near each other in a plain text. The proposed method was applied on two corpora of different domains and the extracted ontologies were evaluated with respect to the corresponding manually created ontologies. The results proved to be very promising in both domains.

Our next step is to use existing techniques for the automatic annotation of concepts' instances. In the case of low-level concepts, named entity recognition techniques will be employed. This is an approach already adopted in [11] for the discovery of binary relations between named entities on the BOEMIE corpus. Another approach for the automatic annotation of the low-level concepts is the one presented in [15]. In the case of high-level concepts, the work for the discovery of high-level concepts, performed in the context of the BOEMIE project Castano et al. [16], will be examined. We also plan to examine the proposed approach in combination with other works on relation discovery and specifically the work presented in [11]. Another aspect for future work is to apply already existing methods in order to label the directed extracted relations.

6 Acknowledgments

We would like to thank BOEMIE project for providing us access to the semantically annotated athletic corpus. Many thanks also to colleagues from NCSR SKEL laboratory for the fruitful discussions we had on ontology learning.

References

1. Faure, D., Nedellec, C.: A corpus-based conceptual clustering method for verb frames and ontology acquisition. In: LREC, Spain, (1998)
2. Kavalec, M., Maedche, A., Svatek, V.: Discovery of Lexical Entries for Non-taxonomic Relations in Ontology Learning. In: SOFSEM (2004)
3. Punuru, J., Chen, J.: Extraction of Non-hierarchical Relation from Domain Texts. In: IEEE Symposium on Computational Intelligence and Data Mining (CIDM 2007)
4. Robert-Alain de Bengrade, Dressler, W, : Introduction to text Linguistics. Longman (August 24, 1981)
5. Hearst, M.: Automatic Acquisition of Hyponynms from Large Text Corpora. In: 14^{th} International Conference on Computational Linguistics. (1992)
6. Caraballo, S.: Automatic Construction of a Hypernym-labeled Noun Hierarchy from Text. In: Assosiation of Computational Linguistics. (1999)
7. Cederberg, S., Widdows, D.: Using LSA and Noun Coordination Information to Improve the Precision and Recall of Automatic Hyponym Extraction. In: Conference on Natural Language Learning. (2003)
8. Ciaramita, M., Gangemi, A., Ratsch, E., Jasmin, S., Isabel, R.: Unsupervised Learning of Semantic Relations between Concepts of Molecular Biology Ontology. In: International Joint Conference on Artificial Intelligence. (2005)
9. Schutz, A., Buitelaar, P., RelExt: A Tool for Relation Extraction from Text in Ontology Extension. In: 4^{th} International Semantic Web Conference. (2005)
10. Espinosa, I.S., Kaya, A., Melzer, S., Moeller, R.: On Ontology Based Abduction For Text Interpretation. In: Proceedings of CICLing 2008, pp. 194205 (2008).
11. Petasis, G., Karkaletsis, V., Paliouras, G., Spyropoulos, C.: Learning context-free grammars to extract relations from text. In: Proceedings of ECAI-2008, pp. 303-307, (2008)
12. Valarakos, A., Karkaletsis, V., Alexopoulou, D., Papadimitriou, Spyropoulos, C., Vouros, G.: Building an Allergens Ontology and Maintaining it using Machine Learning Techniques. In: Computers in Biology and Medicine Journal (CBM), 36 (10): 1155-1184, (2006)
13. Zavitsanos, E., Paliouras, G., Vouros, G.: Ontology Learning and Evaluation: A survey. Technical Report, NCSR Demokritos DEMO-2006-3
14. Church, K., and Hanks, P.: Word Association Norms, Mutual Information and Lexicography. In: Computational Linguistics, Vol 16:1, pp. 22-29, (1991)
15. Iosif, E., Tegos, A., Pangos, A., Fosler-Lussier, E., Potamianos, A.: Unsupervised Combination Of Metrics For Semantic Class Induction. In: IEEE Spoken Language Technology Workshop, (2006)
16. Castano, S., Espinosa, S., Ferrara, A., Karkaletsis, V., Kaya, A., Moller, R., Montanelli, S., Petasis, G., Wessel, M.,: Multimedia Interpretation for Dynamic Ontology Evolution. In: Journal of Logic and Computation (to appear)