Advances in Computer Science and its Applications (ISSN 2166-2924)
Vol. 1, No. 2, June 2012

Lightweight ontologies mapping and the Semantic Similarity based on WordNet

[1]Jialing Li

[*1] Faculty of Computer and Information Science, Southwest University, Chongqing, China

Email: ljl333@swu.edu.cn

Abstract – Ontology mapping is an important way to solve two mainly heterogeneity problem metadata and instance heterogeneity between ontologies [1], the former focuses on the intended meaning of described information, while the latter is more concerned about different representations of instances. In this paper, we mapped lightweight ontology exacted from WordNet tree-like structure with ontology based on semantic lexicon build by W3C to attain a XML schema file which preserving hierarchy information meanwhile each node contains URI linked to more complete information. Also we implement a method to calculate semantic similarity based on the cosine distance of 3 vectors exacted from WordNet's semantic information. Experiment shows that it's a feasible way for similarity calculation.

Keywords – Lightweight ontologies; Semantic similarity; W3C ontology, WordNet

1. Introduction

Ontology technique is aiming at problems such as reusing of data [2], knowledge automatic reasoning [3], precise definition of concepts [4]. As the information explosion and different applications require different data structures to meet their corresponding needs, mapping on ontology become a primary way to solve the problem. Ontology mapping [5] appeared to solve the interoperability among representation with diverse data structures and languages expressing. There're two kinds of ontology heterogeneity: metadata heterogeneity and instance heterogeneity. Metadata heterogeneity mainly be solved by semantic calculation on described information. Marc [6] present an approach determining similarity through rules manually formulated by ontology experts. Most works focus on supervised learning [7] to construct neural network [8], Markov [9], Beysian model [1] for semantic mapping, they usually have good performance. Other semantic mapping is based on knowledge taxonomy WordNet [10], focusing on building framework which combines multiple matchers in a flexible way [11]. Generally we concentrate on large ontology, Hu [12] proposed a divide-and-conquer approach to deal with large ontology in ontology fragment. But classification structures such as taxonomies, business catalogs, web directories in rich tree-like structure named lightweight ontology [13] contain rich semantic information. In this paper, we propose full use of lightweight instance information [14] to map with onto logy-based semantic lexicon WordNet. The mapping result could be used for retrieving , disambiguating, recommending , integrating information, classifications, etc.

WordNet not only stores a word's stand-alone information, but also saves a variety of contacts with other words, this allows us to observe the word in multi-angle, multi-direction. Many works devote to unearth the semantic meaning wrapped by WordNet content of least common subsumer(LCS) and path length relation structure. Exploiting wealth of word's information in WordNet , such as synonyms, sense explanation, relation tree, we implement a method calculating two noun words' similarity based on the 3 feature vectors , experiments confirmed that it could be a way to measure two noun words' similarity.

The rest of the paper is structured as follows: Section 2 introduces background knowledge about WordNet and W3C ontology. Section 3 recommend our idea about gaining WordNet information through interface function, mapping lightweight ontology with W3C WordNet ontology and semantic similarity calculation formula. Section 4 present the implementation of our method to compute semantic similarity , the result evaluating our approach. Finally , a discussion and a future work for improvement provided in section 5.

2. Background knowledge

2.1 WordNet

WordNet [15] is a large lexical database of English based on Cognitive Linguistics which was created by psychologists, linguists and computer engineers of Princeton University. It was generated not only to produce a combination of dictionary and thesaurus that is more intuitively usable, but also to support automatic text analysis and artificial intelligence applications.

Words in WordNet are organized into sets of cognitive synonyms called synset. Each synset contains rich information, such as explanation and a set of words who share the gloss, etc. WordNet consists of synsets and semantic relations connecting the synsets together. The relations are varying on the different types of Parts of Speech (POS), there're mainly 4 POS in WordNet ,namely, Noun, Verb, Adjective, and Adverb.

Nouns and Verbs are both organized into hierarchies with the hypernym-hyponym relation or IS-A relation. The hypernym-hyponym relation is the backbone of the network established by the relationship between noun synset, and accounting for nearly 80% of the relationship. A is hypernym of B means that B is a kind of A, or we can say B is hyponym of A.

Entering “sport” in search box of the WordNet 2.0 Browser, the interface shown in fig. 1 comes up. As shown, there’re 4 noun senses , 2 verb senses of the word. The item highlighted in red rectangle represents a synset. The synset is formed in the pattern as we shown in fig. 2.

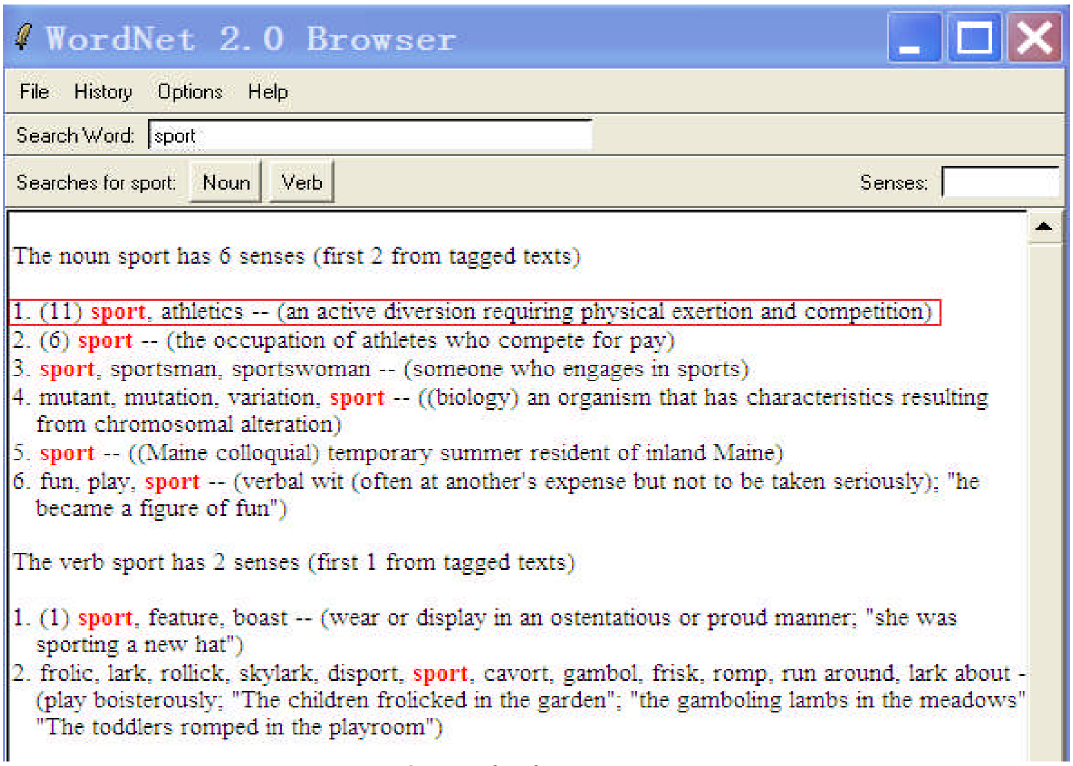

Figure 1. WordNet interface for querying noun “Sport”

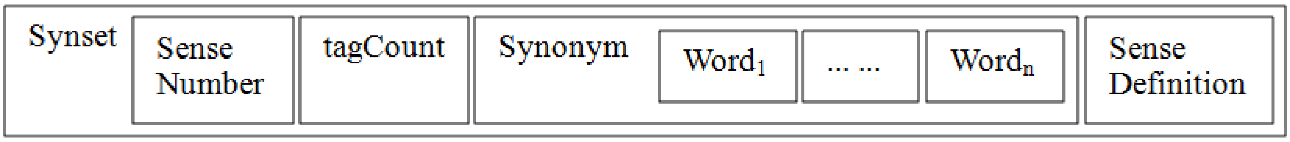

Figure 2. Structure of Synset in WordNet

- Synset
  - AdjectiveSynset
    - AdjectiveSatelliteSynset
  - AdverbSynset
  - NounSynset
  - VerbSynset
- WordSense
  - AdjectiveWordSense
    - AdjectiveSatelliteWordSense
  - AdverbWordSense
  - NounWordSense
  - VerbWordSense
- Word
  - Collocation

Figure 3. Class structure of W3C WordNet Ontology

Where “Sense Number” refers to the ordinal in all the senses the word have. Generally the smaller the number the more used the synset. “tagCount” records the count of the sense is tagged in the corresponding semantic concordance. The “Synonym” consist of a series of words have a same explanation. The “Sense Definition” is the gloss of the synset. Fig. 2 provides only partial information of a synset, We can get more through JWNL in 3.1.

2.2 W3C WordNet Ontology

WordNet can eliminate characteristics of polysemy of words and be widely used in the fields of

Natural Language Processing(NLP) and Information Retrieval(IR). WordNet also works as knowledge-base database of lexical items organized from the human cognitive point of view. Without formatting the data into the Resource Description Framework (RDF) or the Web Ontology Language (OWL), the knowledge provided by WordNet can't be derived from reasoning rules. In order to extend the usage of WordNet, many efforts were devoted to convert WordNet's database to RDF/OWL.

We focus on the standard conversion presented by The World Wide Web Consortium (W3C) based on the WordNet 2.0 version. The WordNet schema of W3C [16] has three main classes: Synset, WordSense and Word. The first two classes have subclasses for the lexical groups presented in WordNet as shown in fig. 3. Each class has its own pattern of URI.

2.3 Our Works

Lightweight ontologies represented by subsumption relations like hypernym-hyponym in WordNet can be used to retrieve in IR, recommend related conceptual node and disambiguate word sense in NLP, etc. With a classification structure stored in XML schema [17], we can easily extend the use of knowledge in WordNet to many other applications.

In this paper, our works based on WordNet involves three aspects: firstly, we exacted the hierarchical hyponym relation of a noun word, recording the information about each node from the tree-like structure, such as the node's synsetID, synonym, word lexical form, gloss etc. Secondly, we mapped each node with the corresponding standard unique URI in W3C WordNet ontology, through which can obtained more completed information representation about a synset. Lastly, we proposed a way to compute similarity between two noun words based on the cosine feature vectors exacted from each word's synsets information, experiment described in the next part shows that it is a feasible way to calculate the word similarity.

3. Experiment

3.1 JWNL

WordNet is heavily-used lexical resource, we intend to extend its use to other application over a lexical dictionary. Sourceforge [18] has developed an open-source Java Framework library for WordNet. With the Java WordNet Library (JWNL) working as an API, we accessed WordNet-style relational dictionaries. Besides data accessing, it also provide functionality to relationship discovery and morphological processing. The JWNL in this paper is based on the version 2.0 of Princeton WordNet.

WordNet was designed to establish the connections in four types of POS-noun, verb, adjective and adverb. In our experiment, we focus on the nouns and typically on the relation hypernym and hyponym of nouns.

The relations likes IS-A or hypernym-hyponym are presented hierarchical mapping to the human cognitive view of classification. Hypernym-hyponym relationships among the noun synsets can be interpreted as specialization relations between conceptual categories. The lexical taxonomy structure provided by hypernym-hyponym could afford potential useful information for user in an application.

Our application receives a lexical form of word (for example "sport") checking it's pos to affirm it's a noun. Using the JWNL class- PointerUtil to generate a hierarchical hyponym tree for the first synset of noun "sport". Traveling the tree under depth first rule, we record the information of each synset, such as synset ID, gloss, and synonyms contained in the synset and also its depth in the tree.

```
[PointerTargetTreeNode: [Synset: [Offset: 311477] [POS: noun] Wc
[PointerTargetTreeNode: [Synset: [Offset: 410363] [POS: noun] Wc
  [PointerTargetTreeNode: [Synset: [Offset: 422482] [POS: noun]
    [PointerTargetTreeNode: [Synset: [Offset: 422936] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 423036] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 423118] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 423251] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 423424] [POS: nour
  [PointerTargetTreeNode: [Synset: [Offset: 424160] [POS: noun]
    [PointerTargetTreeNode: [Synset: [Offset: 424481] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 424559] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 424665] [POS: nour
  [PointerTargetTreeNode: [Synset: [Offset: 439578] [POS: noun]
  [PointerTargetTreeNode: [Synset: [Offset: 444092] [POS: noun]
    [PointerTargetTreeNode: [Synset: [Offset: 445245] [POS: nour
      [PointerTargetTreeNode: [Synset: [Offset: 446257] [POS: nc
    [PointerTargetTreeNode: [Synset: [Offset: 446129] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 446541] [POS: nour
    [PointerTargetTreeNode: [Synset: [Offset: 453585] [POS: nour
[PointerTargetTreeNode: [Synset: [Offset: 410566] [POS: noun] Wc
  [PointerTargetTreeNode: [Synset: [Offset: 428522] [POS: noun]
```

Figure 4 Code example of using PointerUtil and corresponding result

3.2 Lightweight ontology mapping

The lightweight ontology exacted in the last part can be used for querying translation, semantic annotation and data integration, etc. In order to identify each item with a standard URI with completed information linked, we mapped each node in the hierarchical taxonomy to the corresponding presentation of RDF/OWL provided by W3C. The mapping progress is shown in fig. 5 and the result XML file recording the hyponym as subclasses of "sport" is shown in fig. 6.

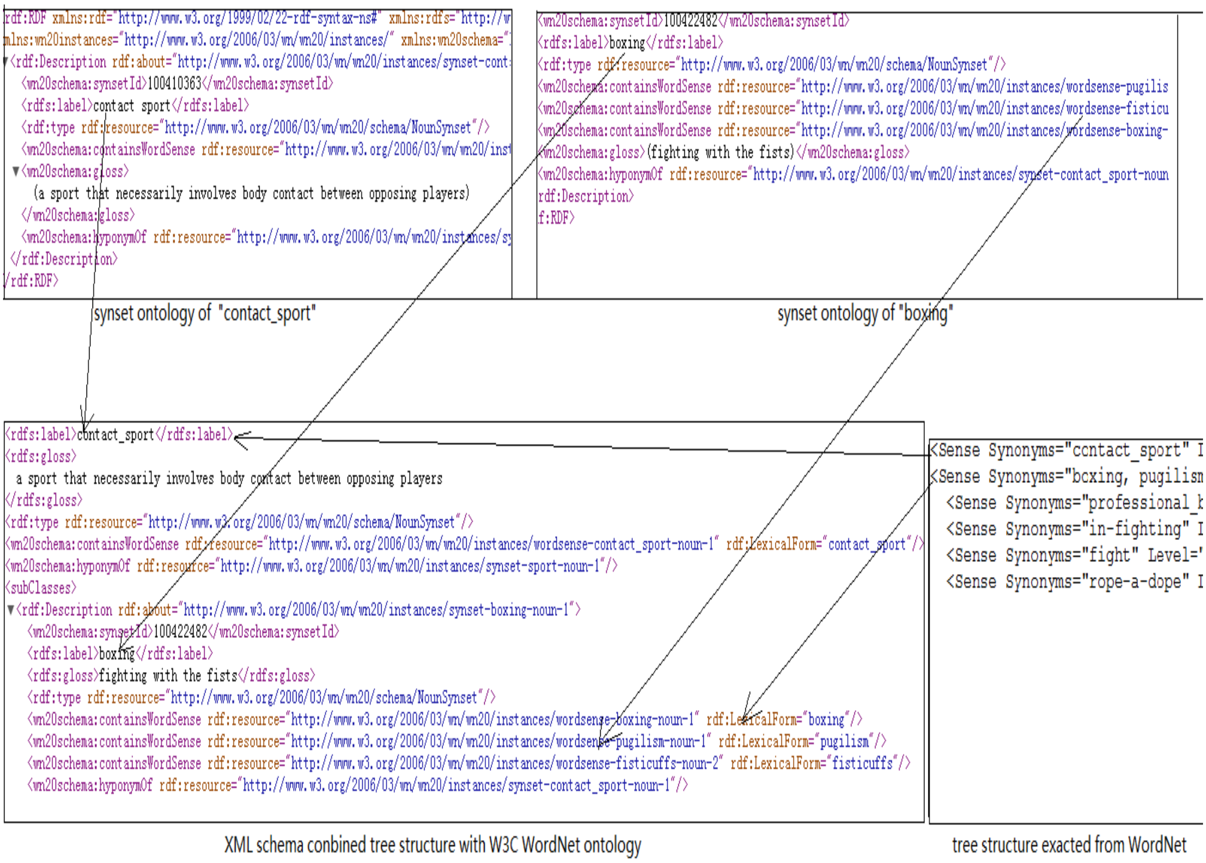

Figure 5 Mapping progress

Figure 6. "sport" taxonomy

Where each fold contains information about corresponding sense of noun "sport". There're 6 senses above, the URI in the red big rectangle is formed in the pattern that "synset +'word' +noun+ ordinal". Where word is specified as the first word in the synonyms of the corresponding synset. The first sense's information structure is expanded in fig. 7.

Information in the red rectangle shows the subclasses in depth 1 of the first synset. Folds the in blue rectangle shows subclasses of the first synset's second subclass. Unfold the third one in fig. 8.

The synset shown in figure 11 is the leaf synset without "subclasses" property.

Contrast to W3C WordNet ontology, we reserve the hierarchy of noun word recording it's hyponym as "subClasses" property in a more intuitive way. It fit more well to the human cognitive view of classification. Contrast to the lightweight ontology exacted from WordNet hierarchical relation we complement the information through the standard unique URI provided by W3C.

3.3 Semantic Similarity

WordNet provides sufficient information of a word, such as its synsets, synonyms, sense definition, hypernym-hyponym etc, so we can mining the similar information between two words.

In our experiment, we calculate semantic similarity between two noun words by computing the cosine distance between two vectors of words exacted from each word's synonym, hypernym-hyponym relation and sense explanation.

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#" xmlns:wn20instances="http://www.w3.org/2006/03/wn/wn20/instances/"
xmlns:wn20schema="http://www.w3.org/2006/03/wn/wn20/schema/">
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-sport-noun-1">
  <wn20schema:synsetId>100407449</wn20schema:synsetId>
  <rdfs:label>sport</rdfs:label>
  <rdfs:gloss>...</rdfs:gloss>
  <rdf:type rdf:resource="http://www.w3.org/2006/03/wn/wn20/schema/NounSynset"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-sport-noun-1" rdf:LexicalForm="sport"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-athletics-noun-1" rdf:LexicalForm="athletics"/>
  <wn20schema:hyponymOf rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/synset-diversion-noun-1"/>
  <subClasses>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-rock_climbing-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-contact_sport-noun-1">
    <wn20schema:synsetId>100410363</wn20schema:synsetId>
    <rdfs:label>contact_sport</rdfs:label>
    <rdfs:gloss>...</rdfs:gloss>
    <rdf:type rdf:resource="http://www.w3.org/2006/03/wn/wn20/schema/NounSynset"/>
    <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-contact_sport-noun-1" rdf:LexicalForm="contact_sport"/>
    <wn20schema:hyponymOf rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/synset-sport-noun-1"/>
    <subClasses>
     <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-boxing-noun-1">...</rdf:Description>
     <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-wrestling-noun-2">...</rdf:Description>
     <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-ice_hockey-noun-1">...</rdf:Description>
     <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-football-noun-1">...</rdf:Description>
    </subClasses>
   </rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-outdoor_sport-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-gymnastics-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-track_and_field-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-skiing-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-water_sport-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-rowing-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-archery-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-sledding-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-skating-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-racing-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-riding-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-cycling-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-blood_sport-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-athletic_game-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-judo-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-spectator_sport-noun-1">...</rdf:Description>
   <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-team_sport-noun-1">...</rdf:Description>
  </subClasses>
 </rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-sport-noun-2">...</rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-sport-noun-3">...</rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-mutant-noun-1">...</rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-sport-noun-5">...</rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-fun-noun-2">...</rdf:Description>
</rdf:RDF>
```


Figure 7. Expansion of "sport" taxonomy


```
<subClasses>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-boxing-noun-1">...</rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-wrestling-noun-2">...</rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-ice_hockey-noun-1">
  <wn20schema:synsetId>100439578</wn20schema:synsetId>
  <rdfs:label>ice_hockey</rdfs:label>
  <rdfs:gloss>
   a game played on an ice rink by two opposing teams of 6 skaters each who try to knock a flat round puck into the opponents' goal with hockey sticks
  </rdfs:gloss>
  <rdf:type rdf:resource="http://www.w3.org/2006/03/wn/wn20/schema/NounSynset"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-ice_hockey-noun-1" rdf:LexicalForm="ice_hockey"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-hockey-noun-2" rdf:LexicalForm="hockey"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-hockey_game-noun-1" rdf:LexicalForm="hockey"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-ice_hockey-noun-1" rdf:LexicalForm="hockey"/>
  <wn20schema:containsWordSense rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/wordsense-hockey_game-noun-1" rdf:LexicalForm="hockey_game"/>
  <wn20schema:hyponymOf rdf:resource="http://www.w3.org/2006/03/wn/wn20/instances/synset-contact_sport-noun-1"/>
 </rdf:Description>
 <rdf:Description rdf:about="http://www.w3.org/2006/03/wn/wn20/instances/synset-football-noun-1">...</rdf:Description>
</subClasses>
</rdf:Description>
```


Figure 8. Fragment of expansion for "sport" taxonomy

-Feature exacting

Firstly, we get the word's synset, hypernym-hyponym hierarchical tree and gloss through JWNL, then exacting candidate words from the 3 collection.

Feature(SW) = {{Ws},{Wc}{We}}

{Ws} contains all the synonym words in synset W of the noun word

{Wc} contains words from synset W's hypernym and hyponym trees

{We} contains reserved content words after using stoplist to dispose synset W's gloss

-Sense similarity

In order to calculate the information similarity between two words, firstly, we should compute the sense similarity between synset i of the fist word and synset j of the other word.

$$Similarity(SW_i, SW_j) = \frac{\sum_{w_i \in \{Ws_i\} \cap \{Ws_j\}} Ks \times IDF(w_i)^2 + \sum_{w_i \in \{Wc_i\} \cap \{Wc_j\}} Kc \times IDF(w_i)^2 + \sum_{w_i \in \{We_i\} \cap \{We_j\}} Ke \times IDF(w_i)^2}{No(SW_i) \times No(SW_j) \times \sqrt{\sum_{i \in Q_U, K \in \{Ks, Kc, Ke\}} K \times IDF(w_i)^2} \times \sqrt{\sum_{j \in Q_V, K \in \{Ks, Kc, Ke\}} K \times IDF(w_j)^2}}$$

(1)

No(SW) : the ordinal of this synset stands in all the synsets of the word

$IDF(w_i)$: the reverse of the w_i's use frequency(tagCount)

Ks = 0.45: the weight assigned to the synonym feature

Kc = 0.35 : the weight assigned to the hypernym-hyponym feature

Ke = 0.2: the weight assigned to the synset explanation feature

Qu : the collections containing w_i

Qv : the collections containing w_j

Since the word's hypernym contains more common feature information than the hyponym one, in the experiment, we use two parameters to coordinate hypernym and hyponym's contribute to the similarity measure:

Kchyper = 0.9: the weight multiplied to Kc if the w_i is belong to the hypernym tree of the word we calculate

Kchypon = 0.1: the weight multiplied to Kc if the w_i is belong to the hyponym tree of the word we calculate

-Similarity between two words

After computing sense similarity between each synset contained by word1 (W1) and every synset contained by word2 (W2), in order to remove the inequity caused by order, we calculate the words' similarity:

$$Similarity(W_1,W_2)=\frac{\sum_{\substack{i\in\{1,\dots,|SW1|\},\\ j\in\{1,\dots,|SW2|\}}}\max(Similarity(SW1_i,SW2_j))+\sum_{\substack{i\in\{1,\dots,|SW2|\},\\ j\in\{1,\dots,|SW1|\}}}\max(Similarity(SW2_i,SW1_j))}{|SW1|+|SW2|}$$

(2)

|SW1| : the count of synsets contained by W1
|SW2| : the count of synsets contained by W2

4. Evaluation

4.1 Implementation

Our implementation was based on the IDE of java MyEclipse 8.0 in Windows platform and WordNet version 2.0, the noun portion of WordNet 2.0 was selected as the taxonomy to compute the similarity between two words. It contains about 115,424 synsets, among it there're 79,685 noun synsets.

Since the similarity measure lays more on the structure the words share, we handle the antonym like "holiness" and "unholiness" which has the meaning in the contrary but sharing most words in structure by multiplying a weight as 0.01 to the final result in the proposed method. Before we calculate two words in the proposed way we do a preprocessing to find if the two words are synonyms.

Because the deepest noun hierarchical tree can have 16 nodes, we limit the number of words getting from hyponym tree to 20, hypernym tree to 5 to guarantee the efficiency of program, and it influence the result little.

4.2 Result

Table 1 lists the result for word pair. These word pairs and corresponding scores of two control groups Jiang & Conarth and Lin are selected from the Test Collection of Finkelstein [19]. The semantic similarity scores of our Method, Jiang & Conarth and Lin are corresponding to the last 3 columns in table 1.

Table 1 *Comparison of Semantic Similarity for Our Method, Jiang&Conarth. And Lin*

	Word Pair	Our Method	Jiang &Con	Lin
Tiger	Tiger	1.000	1.00	1.00
Midday	Noon	1.000	1.00	1.00
Dollar	Buck	1.000	1.00	1.00
Car	Automobile	1.000	1.00	1.00
King	Queen	1.000	0.27	0.89
Calculation	Computation	1.000	1.00	1.00
Wood	Forest	1.000	1.00	1.00
Murder	Manslaughter	0.929	0.17	0.76
Football	Soccer	0.611	0.27	0.88
Computer	Keyboard	0.583	0.08	0.43
Physics	Chemistry	0.572	0.23	0.81
Vodka	Brandy	0.545	0.14	0.73
Television	Radio	0.505	0.14	0.73
Psychology	Science	0.415	0.24	0.81
Professor	Doctor	0.385	0.20	0.77
Vodka	Gin	0.368	0.12	0.70
Harvard	Yale	0.359	0.17	0.79

Century	Year	0.318	0.13	0.52
Planet	Sun	0.314	0.28	0.84
Doctor	Nurse	0.290	0.25	0.83
Announcement	News	0.283	0.10	0.47
Closet	Clothes	0.277	0.08	0.31
Phone	Equipment	0.285	0.24	0.80
Bishop	Rabbi	0.258	0.16	0.74
Hundred	Percent	0.247	0.07	0.21
Seafood	Lobster	0.229	0.24	0.84
Weather	Forecast	0.223	0.05	0.00

Since only 1 out of 8 of the word “Wood” ’s synsets focus on the tree attribute expressed by the word, so it could be so different with the score of the other two control group.

5. Future Work

Ontology mapping is an important way to solve linguistic and structural comparability between ontologies. In this paper, we focus on mapping lightweight ontology in semantic lexicon with W3C WordNet ontology, the result XML schema can be used to eliminate polysemy of words, recommend related words to users, or extend searching progress, etc. The proposed method proved to be a viable way for similarity calculation [20-21].

References

[1] T, J., Li, J., Bangyong Liang, et al. Using Bayesian decision for ontology mapping[J].*Web Semantic: Science, Services and Agents on the World Wide Web*,2006,4(4):243-262.

[2] Simperl,E.:Reusing ontologies on the Semantic Web: A feasibility study[J].*Data & Knowledge Engineering*, 2009,68(10):905-925.

[3] Stuckenschmidt, H., Klein, M.: Reasoning and change management in modular ontologies[J].*Data & Knowledge Engineering*, 2007,63(2):200-223.

[4] Formica, A.: Ontology-based concept similarity in Formal Concept Analysis[J]. *Information Sciences*, 2006,176(18):2624-2641.

[5] Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art[J]. *The Knowledge Engineering Review*, 2003, 18(1):1-31.

[6] Doan, A., Domingos, P., Halevy, A.: Learning to Match the Schemas of Data Sources: A Multistrategy Approach[J]. *AI Magazine*, 2005,26(1):83-94.

[7] Huang, J., Dang, J., Vidal, J.M., et al. Ontology Matching Using an Artificial Neural Network to Learn Weights[C]. *In Proceeding of IJCAI Workshop on Semantic Web for Collaborative Knowledge Acquisition (SWeCKa-07)*. Hyderabad, India, 2007.

[8] Zhong, C., Zhong, M., Bai, C.:A high-order discrete scheme of Lattice Boltzmann method for cavitation simulation. Advances in Computer Science and its Applications, 2012:1(1).

[9] Kaza, S., Chen, H.: Evaluating ontology mapping techniques: An experiment in public safety information sharing[J]. *Decision Support Systems*, 2008,45(4):714-728.

[10] H.Do, E. Rahm. Coma c a system for flexible combination of schema matching approaches[C]. *In Proceedings of the Twenty-eighth VLDB Conference*, Hong Kong, China, 2002.

[11] Bassil, Y.: Distributed, Cross-Platform, and Regression Testing Architecture for Service-Oriented Architecture. Advances in Computer Science and its Applications, 2012, 1(1).

[12] Giunchiglia, F., Autayeu, A., Pane, J.: S-Match: an open source framework for matching lightweight ontologies[J]. Semantic Web Journal, 2010.

[13] Su, X., Gulla, J.A.: An information retrieval approach to ontology mapping[J]. *Data & Knowledge Engineering*, 2006, 58(1): 47-69.

[14] WordNet[EB/OL]. http://en.wikipedia.org/wiki/Wordnet.

[15] WordNet W3C Ontology[EB/OL].http://www.w3.org/TR/wordnet-rdf/.

[16] Download form[EB/OL]: http://www.w3.org/2006/03/wn/wn20/download/

[17] Deshmane, V.: XML driven SCPI interpreter. Advances in Computer Science and its Applications, 2012,1(1).

[18] JWNL[EB/OL].http://sourceforge.net/projects/jwordnet/.

[19] Finkelstein, L., Gabrilovich, E., Matias, Y., et al: Placing Search in Context: The Concept Revisited[J]. *ACM Transactions on Information* Systems, 2002,20(1):116-131.

[20] G. S. Sarma, A.S.N. Chakravarthy, Ch. MoniSucharitha, Khan Zabiullahkhan, A Novel Method of SAR Image Regularization With Fast Approximate, Advances in Digital Multimedia, 2012, 1(2): 64-79

[21] Yudong Zhang, Lenan Wu, Rigid Image Registration by PSOSQP Algorithm, Advances in Digital Multimedia, 2012, 1(1): 4-8