Spatial Role Labeling

CS365 Course Project

Amit Kumar , akkumar@iitk.ac.in
Chandra Sekhar, gchandra@iitk.ac.in

Supervisor : Dr.Amitabha Mukerjee

ABSTRACT

In natural language processing one of the important functions is to convey spatial relationships between objects and their relative/absolute location in a space. Spatial Role labeling as defined in Sem Eval Task concerns with the extraction of spatial semantics from natural language. The sentence "Give me the gray book on the large table." expresses information about the spatial configuration of two objects (book, table) in some space. Understanding such spatial utterances is a problem in many areas, including robotics, navigation, traffic management, and query answering systems [Tappan 2004].Through this project we aim to extract the three important spatial aspects from a natural language sentence i.e. spatial indicators, landmarks and trajectors.We would be using machine learning techniques to determine the spatial semantics of a sentence.We are using a pipelining approach to identify the spatial indicators and the results we got are promising.

Introduction :

Task definition.

We define spatial role labeling (SpRL) as the automatic labeling of natural language with a set of spatial roles. The sentence-level spatial analysis of text deals with characterizing spatial descriptions, denoting the spatial properties of objects and their location (e.g. to answer "what/who/where" questions). A spatial term (typically a preposition) establishes the type of spatial relation and other constituents express the participants of the spatial relation (e.g. a location). The roles are drawn from a pre-specified list of possible spatial roles from the annotated dataset [1] and the role-bearing constituents in a spatial expression must be identified and their correct spatial role labels assigned.

For example, consider the sentence

Give me the [gray book]$_{tr}$ [on]$_{si}$ [the big table]$_{lm}$.

- The phrase headed by the token "book" is referring to a trajector object. The **trajector (TR)** is an entity whose location is described in the sentence.
- The phrase headed by the token "table" is referring to the role of a landmark . The **landmark**(LM) is a reference object for describing the location of a trajector.
- These two spatial entities are related by the spatial expression "on" denoted as **spatial indicator (SP).**

The spatial indicator (often a preposition in English, but sometimes a verb, noun, adjective, or adverb) indicates the existence of spatial information in the sentence and establishes the type of a spatial relation.

A major obstacle when dealing with unrestricted language is the scarcity of annotated data available for training machine learning models. We therefore start with the available resources. In our primitive approach , we learn prepositions' spatial senses by exploiting annotated data from the preposition project (TPP) employed in SemEval-2007 [Litkowski and Hargraves 2007] and then use the results of preposition disambiguation in a spatial role labeler that identifies trajector and landmark roles.The results for identifying spatial indicators is promising on the given dataset.

2. Although we have defined spatial indicators, trajectors and landmarks as arbitrary segments of a sentence, we focus on single words, each as one segment. However, a phrase in the sentence commonly plays a role, and we thus assume that the head word of the phrase is the role-holder.

A head word determines its phrase’s syntactic type; analogously, it is a stem that determines the semantic category of its component’s compound. The other elements of a phrase modify the head. For example, in ”the huge blue book”, ”book” is the head word, and ”huge” and ”blue” are modifiers. In our data, the labeling scheme reflects this fact and only assigns roles to head words and labels the remaining words (e.g., modifiers) as “none”. Hence, a sentence is hereafter assumed to be a sequence of words.

Approach

There are 2 stages in our project:

1. Identification of Spatial indicator:: Since most of the time prepositions are the spatial indicator in the sentence. So for the identification of spatial indicator first we have to identify the prepositions in the sentence.

For the identification of preposition we are using Stanford Dependency Parser. When we pass a sentence to dependency parser as a input it will give us the relation between each word to all other words which are dependent on that.

So for example if we pass a sentence “ a yellow building with white columns in the background “ to the Stanford Parser it gives a parsed tree and relations between all the words which are dependent on each other as output (shown in fig 1 .)

```
SENTENCE ::  a yellow building with white columns in the background .

PARSED TREE ::  (ROOT (NP [59.367] (NP [20.107] (DT [1.419] a) (JJ [9.132]
yellow) (NN [6.856] building)) (PP [21.709] (IN [3.398] with) (NP [17.909]
(JJ [6.762] white) (NNS [8.150] columns))) (PP [14.223] (IN [1.850] in) (NP
 [11.972] (DT [0.650] the) (NN [9.135] background))) (. [0.021] .)))


 PARSER OUTPUT

det(building-3, a-1)
amod(building-3, yellow-2)
prep(building-3, with-4)
amod(column-6, white-5)
pobj(with-4, column-6)
prep(building-3, in-7)
det(background-9, the-8)
pobj(in-7, background-9)
```

Fig. 1

Here word “a” depends on the word “building” and word “yellow” also depends on word “building” and so on.

On the basis of this dependency relationship we form a feature set by using only the words dependent on prepositions. We can identify those words by looking at the words corresponding to “prep” and “pobj” tagged line on the dependency relation.
Note :: Prepositions are the words which are tagged as “prep” in the output of Stanford Dependency Parser .

For making the feature set we are using a python code from “Universitat de Barcelona” which disambiguates the prepositions from other words of the sentence on the basis of the dependency relationship. Each feature set contains one preposition. So there will be as many as feature set for a sentence as number of prepositions in the sentence.
SO for the example sentence (above) we have 2 set one corresponding to each preposition(with and in). Each feature set contains the words (Head1 and Head2) which are dependent on the preposition (as shown in the fig 2).

```
amit@ubuntu: ~/SpatialRoleLabeling/SIdetection                    0:32   5:05 PM
GEtting spatial sense of ::  with
Features extracted:
{'head1': u'building',
 'head1_LEMMA': u'building',
 'head1_POS': 'NN',
 'head2': u'columns',
 'head2_LEMMA': u'column',
 'head2_POS': 'NNS',
 'preposition': u'with',
 'preposition_POS': 'IN',
 'preposition_spatial': 0.0}

GEtting spatial sense of ::  in
Features extracted:
{'head1': u'building',
 'head1_LEMMA': u'building',
 'head1_POS': 'NN',
 'head2': u'background',
 'head2_LEMMA': u'background',
 'head2_POS': 'NN',
 'preposition': u'in',
 'preposition_POS': 'IN',
 'preposition_spatial': 0.13333333333333333}
```

Fig. 2

Then we construct a list of these feature sets and assign a class to each feature set of the list to SI or NSI where SI tells that the preposition present in the feature set acts as a spatial indicator in the sentence and NSI says that preposition doesn't act as a spatial indicator in the sentence.(shown in Fig 3).

```
</DEPENDENCIES>
- <PREPS>
  - <PREPOSITION>
      <HEAD1>building</HEAD1>
      <HEAD1_LEMMA>building</HEAD1_LEMMA>
      <HEAD1_POS>NN</HEAD1_POS>
      <HEAD2>columns</HEAD2>
      <HEAD2_LEMMA>column</HEAD2_LEMMA>
      <HEAD2_POS>NNS</HEAD2_POS>
      <PREP>with</PREP>
      <PREP_POS>IN</PREP_POS>
      <PREP_SPATIAL>0.0</PREP_SPATIAL>
      <CLASS>NSI</CLASS>
    </PREPOSITION>
  - <PREPOSITION>
      <HEAD1>building</HEAD1>
      <HEAD1_LEMMA>building</HEAD1_LEMMA>
      <HEAD1_POS>NN</HEAD1_POS>
      <HEAD2>background</HEAD2>
      <HEAD2_LEMMA>background</HEAD2_LEMMA>
      <HEAD2_POS>NN</HEAD2_POS>
      <PREP>in</PREP>
      <PREP_POS>IN</PREP_POS>
      <PREP_SPATIAL>0.133333333333</PREP_SPATIAL>
      <CLASS>SI</CLASS>
    </PREPOSITION>
```


Fig. 3

For the computation of the accuracy we take one sentence at a time identify that our prepositions which are marked as SI are matching with the original spatial indicators. If it matched with the original we increment our score and finally divide that score from the total number of spatial indicators present in the training set.

2. Identification of Trajectors and Landmarks : In the second phase of the project we have to identify the landmarks and trajectors. For this we assumed that the Head1 and Head2 present in the feature set are the words which are acting as trajector and landmark in the sentence.
So for each sentence we have to compare these guessed trajectors and landmarks present in the sentence to the trajectors and landmarks given in the training set corresponding to the spatial indicator and compute the accuracy as computed for spatial indicators.

Limitations :

Our main limitation is our Training Data set. Since this is a new field in the natural language processing so we don't have a large training data set for our project.

We have some limitations in our approach.

- We cannot apply this approach on those sentences which doesn't have any prepositions because we our whole approach revolves around prepositions present in the sentence.

- We cannot apply this approach on those sentences in which spatial indicators are multi word.

Experiments :

Dataset : The main annotated corpus is a subset of IAPR TC-12 Benchmark. It consists of 600 sentences annotated with basic spatial roles of trajector, landmark and their corresponding spatial indicators. SemEval-2007 data of TPP [Litkowski and Hargraves 2007] was used for preposition disambiguation.

Tools Used : NLTK , Stanford Parser , Naïve Bayes Classifier.

- Used K- Fold Cross Validation on different amount of datasets and calculated the accuracy with Naive Bayes Classifier for spatial indicators on the rest of the dataset. Achieved a maximum of **86.2%** accuracy when we split the data into 7:3 for training and testing purposes.

Our Work :

We had arranged the source code for the preposition disambiguation from the Universitat de Barcelona, Spain. With the help of this code we are able to assign class (SI or NSI) to each feature set of the sentence. Then we take this updated feature set (with the classes assgined) as input and identify the Trajectors and Landmarks. Then we have computed the accuracy of our approach for the identification of trajector and landmark.

Results :

Accuracy	Identification of Spatial Indicators	Identification of Trajectors	Identification of Landmarks
Maximum	86 %	78.2 %	83 %
Minimum	77.2 %	69.2 %	63.3 %
Average	82.5	75.14 %	78.26 %

Future Scope :

Since we are taking few assumptions in our approach like we have only those sentences which contains 1 word preposition. So if a preposition “in front of” is acting as a spatial indicator in the sentence then we cannot identify this by our approach.

We are only taking prepositions for the identification of spatial indicators but other, non-prepositional, words can also act a spatial indicator in the sentence. So for the sentences, in which non prepositional words are acting as a spatial indicator, our approach fails.

So we can work on our approach on these fields to improve our approach.

Taking into account the probability of a preposition being a Spatial Indicator we could achieve favourable results on the test data set. Through this pipelining approach , we will try to extend to identification of landmarks and trajectors from the sentences taking into the account difficulty level in adapting the training data with features extracted to the HMM Network.

Acknowledgements :

We would like to thank our mentor and guide Prof. Amitabha Mukerjee for guiding us throughout the project and giving us helpful insights.Also we have used the libraries and some part of code developed by a team headed by Victor Ponce , Universitat de Barcelona, Spain , we acknowledge him for sharing the code.

References :

[1] P. Kordjamshidi, M van Otterlo, and M. F. Moens. Spatial role labeling: task definition and annotation scheme. In LREC, 2010.

[2] P. Kordjamshidi, M van Otterlo, and M. F. Moens. From language towards formal spatial calculi.In Workshop on Computational Models of Spatial Language Interpretation (CoSLI 2010, at SpatialCognition 2010), 2010.

[3] Parisa Kordjamshidi, Martijn van Otterlo, and Marie-Francie Moens. Spatial role labeling: Towards extraction of spatial relations from natural language. ACM Transactions on Speech and Language Processing, Nov. 2011.

[4]. Code used and modified : https://code.google.com/p/pln-pmt-pract/