Spatial Text Labelling

Under Supervision Of
Prof. Amitabha Mukherjee

PRACHI GUPTA (12485)
SATYAM KUMAR SHIVAM (11917658)
Indian Institute of Technology, Kanpur

November 16, 2015

Abstract

The objective of this project is to identify and label spatial keywords in the given corpus. Instead of using any parser based method, our approach is based on word vector models. We use Google's Word2Vec for creating word vectors. We train a backpropagation neural network to extract the relationship between three important spatial aspects, that is, trajectors, landmarks and spatial indicators. Given a pair of a trajector and a landmark, we try to predict a spatial indicator best describing the relationship between the given trajector and landmark.

Acknowledgement

The project is a great opportunity to try hands on research work and get a taste of what a research is really like. We are grateful to have this opportunity and tried to justify to our ends.

We wish to thank Prof. Amitabha Mukherjee (Department of Computer Science & Engineering) for providing continuous mentor-ship throughout the duration of the project. Without his guidance, the project would not have been possible.

Prachi Gupta
gprachi@iitk.ac.in
Satyam Kumar Shivam
satyamks@iitk.ac.in

Contents

1

Introduction

One of the major tasks in natural language processing is to talk about spatial relationships between objects. The sentence "*Give me the book lying on the table*" expresses information about the spatial configuration of the 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).

Different types of spatial roles can be assigned to different words in a sentence. In this project, we mainly use three different types of spatial roles:

1. Trajector: denotes a central object of a spatial scene.
2. Landmark: denotes a secondary object of a spatial scene.
3. Spatial Indicator: signals a spatial relation between objects.

For example, consider the sentence,

"*A lake in the forest*"

Here, '*lake*' is the trajector, '*forest*' is the landmark and '*in*' is the spatial indicator connecting the two.

A lot of work has been done earlier in this field but most of those works use a grammar based parser to find relationships among different objects. So, it becomes difficult to extend them to other languages because the parsers for other languages are not as easily available as for English. Hence, we try to use Word2Vec and backpropagation neural network model to obtain these relationships.

2

Literature Review

Most of the related work in this field has been done using standard parsers available for specific languages like English.

One such work is in the field of geo-location route recognition. It is needed there to recognize the geographical entities and disambiguate various forms of the same route. The Stanford parser is used to determine the dependencies between prepositions and place names. The dependencies given by the parser help to find the source and the destination easily.

Other work done specifically for Spatial Role Labelling uses a dependency parser to find the prepositions in a sentence since prepositions are mostly the spatial indicators in a sentence. The trajector and landmark are predicted based on these dependencies as well.

3

Data Set

The data-set SpRL for the Spatial Role Labelling shared task at SemEval-2013 has been taken. This data is available on Github.
It contains data for static spatial relations in the form of an XML file where each child is a sentence description containing CONTENT, TRAJECTOR, LANDMARK, SPATIAL_INDICATOR and RELATION as tags.

Here is a screen-shot of the data:


```xml
<DOC>
<SENTENCE id='s1'>
<CONTENT>About 20 kids in traditional clothing and hats waiting on stairs .</CONTENT>
<TRAJECTOR id='tw2'> kids</TRAJECTOR>
<LANDMARK id='lw10'> stairs</LANDMARK>
<SPATIAL_INDICATOR id='sw9'>on </SPATIAL_INDICATOR>
<RELATION id='r0'       sp='sw9'        tr='tw2'        lm='lw10'       general_type='region'/>
</SENTENCE>

<SENTENCE id='s2'>
<CONTENT>a house and a green wall with gate in the background .</CONTENT>
<TRAJECTOR id='tw1'> house</TRAJECTOR>
<LANDMARK id='lw10'> background</LANDMARK>
<SPATIAL_INDICATOR id='sw8'>in </SPATIAL_INDICATOR>
<RELATION id='r0'       sp='sw8'        tr='tw1'        lm='lw10'       general_type='region'/>
<TRAJECTOR id='tw5'> wall</TRAJECTOR>
<RELATION id='r1'       sp='sw8'        tr='tw5'        lm='lw10'       general_type='region'/>
</SENTENCE>

<SENTENCE id='s3'>
<CONTENT>a sign saying that plants can't be picked up on the right .</CONTENT>
<TRAJECTOR id='tw1'> sign</TRAJECTOR>
<LANDMARK id='lw14'>undefined</LANDMARK>
<SPATIAL_INDICATOR id='sw10'>on the right </SPATIAL_INDICATOR>
<RELATION id='r0'       sp='sw10'       tr='tw1'        lm='lw14'       general_type='direction'/>
</SENTENCE>

<SENTENCE id='s4'>
<CONTENT>a rock hole in the ground surrounded by sand .</CONTENT>
<TRAJECTOR id='tw2'> hole</TRAJECTOR>
<LANDMARK id='lw5'> ground</LANDMARK>
<SPATIAL_INDICATOR id='sw3'>in </SPATIAL_INDICATOR>
<RELATION id='r0'       sp='sw3'        tr='tw2'        lm='lw5'        general_type='region'/>
<LANDMARK id='lw8'> sand</LANDMARK>
```

4

Approach

There are two main steps in our approach:

1. Create Word Embeddings: We first build word vocabulary by training Google's Word2Vec on sufficiently large dataset.

2. Train Backpropagation Neural Network: Once we create word embeddings, we train a backpropagation neural network with word embeddings for trajectors and landmarks as input and embedding for spatial indicators as output.

4.1 Implementation

1. Pre-processing:
   Firstly, we processed the data-set available to us and extracted the trajectors, landmarks and spatial indicators from each child in the XML file. In total, we had 482 different pairs of them, with 56 distinct spatial indicators. We merged prepositional phrases into a single word, for example, '*in front of*' was changed to '*in_front_of*'.

2. Creating Word Vectors:
   This data is then appended to the Wikipedia data and a Word2Vec model is trained using the new data. The word vector dimensions have been taken to be 50.

3. Training Backpropagation Neural Network:
   A backpropagation neural network with 100 neurons in the input layer, 100 neurons in the hidden layer and 50 neurons in the output layer is built. It is trained using a pair of a trajector and a landmark word vector as input and the corresponding spatial indicator word vector as the output, with the termination condition being full epoch cycles till convergence.

5

Results

Of the 480 sentences, 400 sentences have been taken in the training set and rest 80 sentences have been taken in the test set. Different pairs of trajectors and landmarks are picked from the test set and tested using the neural network. The 50 dimensional word vector obtained as output is matched with the word vectors earlier obtained from Word2Vec model and the most similar word is reported as the answer. Cosine distance between predicted spatial indicator and actual spatial indicator is also calculated to demonstrate accuracy of prediction by our trained backpropagation neural network.

Trajector, Landmark	**Correct Spatial Indicator (cs)**	**Predicted Spatial Indicator (ps)**	**Cosine distance between cs and ps**
sky, background	in	in	2.7186372520
entrance, background	in	in	2.5142250183
arch, stars	below	on	1.7129730747
pool, house	in	in_front_of	2.0555872556
briefcase, chair	on	on	1.8701356229

5.1 Cross-Validation

We have taken 400 sentences in training set and 80 sentences in validation set for performing cross-validation. So, there will be 6 rounds of cross-validation.

Cross-Validation Round	**Total number of sentences**	**Total number of sentences having correct output**	**Accuracy**
1	80	36	45.00%
2	80	34	42.50%
3	80	43	53.75%
4	80	50	62.50%
5	80	49	61.25%
6	80	31	38.75%
Average	80	40.50	50.625%

6

Conclusion

Our approach of spatial role labelling without using any parser based method gives comparable result, in comparison to results obtained by parser based methods. It is thus possible to use word vector models for this purpose.

The total accuracy for the task of predicting a spatial indicator given a landmark and a trajector (for our 80 sentences test set) by our approach is around 50% to 60%. Before this, accuracy of simple tasks of identifying a spatial indicator in a sentence using parsing has been around 80%.

An important advantage of using this approach in the task of spatial role labelling is that it can be easily extended to other languages like Hindi for which parsers are not readily available.

The approach fails in some cases where more than one spatial indicator seems to be valid. Of course, it would work better if the context is also taken into account rather than simply taking pairs of trajectors and landmarks. This would lead to resolve the ambiguity of more than one correct answer and will hopefully give the one as already used in the sentence.

Bibliography

[1] Kolomiyets, Oleksandr, et al. "Semeval-2013 task 3: Spatial role labeling." *Second joint conference on lexical and computational semantics (* SEM), Volume 2: Proceedings of the seventh international workshop on semantic evaluation (SemEval 2013).* 2013.

[2] Kordjamshidi, Parisa, and Marie-Francine Moens. "Global machine learning for spatial ontology population." *Web Semantics: Science, Services and Agents on the World Wide Web* 30 (2015): 3-21.

[3] Pustejovsky, James, Jessica Moszkowicz, and Marc Verhagen. "A linguistically grounded annotation language for spatial information." (2013).

[4] Clough, Paul. "Extracting metadata for spatially-aware information retrieval on the internet." *Proceedings of the 2005 workshop on Geographic information retrieval.* ACM, 2005.

[5] Ramisa, Arnau, et al. "Combining geometric, textual and visual features for predicting prepositions in image descriptions." *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP), Lisboa, Portugal, September. Association for Computational Linguistics.* 2015.

[6] Zheng, Yingjie, et al. 'Geo-location Route Recognition' *CS224N Final Project, Stanford University.* 2010.

[7] "Kolomiyets/sprl2013." GitHub. N.p., n.d. Web data repository.