Spatial Role Labeling

By :

Satvik Gupta

Garvit Pahal

Mentor:

Amitabha Mukerjee

Introduction

Spatial role labeling is the task of automatic labeling of words and phrases in a sentence with a set of spatial roles such as trajector, landmark, spatial indicator, distance, direction etc

Example

John(Trajector) is sitting on(Spatial Indicator) the ground(Landmark)

The vase(Trajector1) is on(Spacial Indicator1) the ground(Lanmark1,Trajector2) on(Spatial Indicator2) your left(Landmark2)

Spatial indicators: Signals a spatial relation between objects (trajectors and landmarks) of a spatial scene.

Trajector : Entities like a person, object or event whose location is described.

Landmark: The reference entity in relation to which the location or trajectory of the trajectors motion is specified.

Dataset

The data we got from the organizers of SemEval’12 in XML format.

It consists of around 600 labeled sentences

Examples

About 20 kids in traditional clothing and hats waiting on stairs

<TRAJECTOR id="tw2"> Kids</TRAJECTOR>

<LANDMARK id="lw10"> Stairs</LANDMARK>

<SPATIAL_INDICATOR id="sw9">on </SPATIAL_INDICATOR>

Work Flow

Tokenize and POS

- Parsing the xml file and add part of speech tag for the sentence

Parse Tree

- Build a dependency parse tree using stanford dependency code

POS + Lemma

- For each word that depends on preposition and words on which preposition depends we assign the part of speech tag and Lemma tag
- We also assign a parameter of percentage spatial character to the preposition using the TPP Dataset

Naïve Byes

- Classify various preposition in a sentence as spatial Indicator using Naïve Byes Algorithm

Landmark Trajector

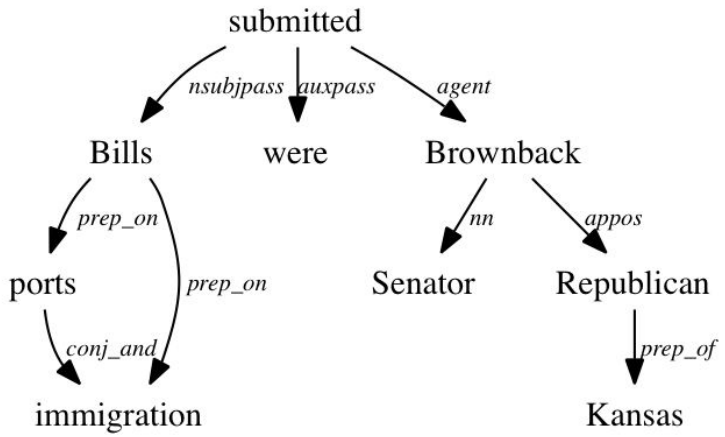
- Using the words that depends on preposition as Trajector and other as Landmark

Stanford Dependency Parsing

Input :

Bills|NNS on|IN ports|NNS and|CC immigration|NN were|VBD submitted|VBN by|IN Senator|NNP Brownback|NNP ,|, Republican|NNP of|IN Kansas|NNP

Output



nsubjpass(submitted, Bills)
auxpass(submitted, were)
agent(submitted, Brownback)
nn(Brownback, Senator)
appos(Brownback, Republican)
prep_of(Republican, Kansas)
prep_on(Bills, ports)
conj_and(ports, immigration)
prep_on(Bills, immigration)

Image Src : http://nlp.stanford.edu/software/stanford-dependencies.shtml

Naïve Bayes

Given $x_{1..n}$ as n words of a sentence and $y_{1..n}$ as n labels. Then in Naïve Bayes algorithm

$$P(x,y) = P(y) \prod_{i=1}^{n} P(x_i|y_i)$$

In Naïve Bayes algorithm we consider conditional independence among different words of a sentence

$$\left(P(x_i \,|y_i, x_j) = \;\; P(x_i|y_i)\right)$$

Naïve Bayes Training

Input Format

[[Trajector,Trajector.Lemma,Trajector.POS] , [Landmark, Landmark.Lemma,Landmark.POS],[Preposition,Preposition.Lemma,Preposition.POS], Spatial sense]] – TAG (1 means is a spatial indicator)

The spatial sense is calculated using the tpp dataset(which a large corpus of preposition along with it senses

Example

About 20 kids in traditional clothing and hats waiting on stairs

[[kids,kid,NNS], [clothing,clothing,NN], [in,IN,prep],0.13]] - 0

[[waiting,waiting,VBG], [stairs,stair,NNS], [on,IN,prep],0.25]] – 1

Results

Spatial Role	Accuracy
Spatial Indicator	83%
Trajector	76%
LandMark	79%

These Results are obtained on taking 480 sentences for training and 120 dataset for testing

Future Work

- Append Word to vector as an additional feature in the feature dataset
- Using learning approaches like Conditional Random Field which is basically maximization of conditional probability of labels given all the set of observations

Acknowledgement

- TPP Data set : (Semevall 2007 Task of word sense disambiguation for preposition)
- Semeval 2012 Dataset : http://www.cs.york.ac.uk/semeval-2012/task3/
- Code Modified: https://code.google.com/p/pln-pmt-pract/
- Paper Referred : 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.
- Python NLTK Library (nltk.pos_tag(), nltk.lemma())