Project Description

Chloe Chen (group partner: Katie Roberts)

9/9/2016

1. Give it a title

“Spatial Role Labeling in Text – annotation algorithm with Natural Language Processing and Machine Learning Techniques”

1. Describe the problem for the audience who is not familiar with your field and terminology

The problem we will be focusing on is a machine-learning algorithm that could take a sentence or phrase and returns properly labeled roles regarding the spatial relationship between the objects included. In a sentence, the object that this spatial relationship refers to is called the trajector, the token that indicates the specific relationship is labeled as the spatial indicator, and the object that serve to complement the relationship is labeled the landmark. In a given example “Give me the grey book on the table”, “the grey book” will be seen as the trajector, “on” is the spatial indicator, and “the table” is the landmark.

More complications arise in the example “Bring me the book on the table in the afternoon.” Here we have the first preposition “on” as a spatial indicator, but the other preposition “in”, though could also be taken as a spatial indicator in another phrase, doesn’t function to reference any spatial relationships here. For the machine to be able to tell the difference here is a challenging task.

1. Why this problem is important or interesting

This problem is interesting to me because as a Linguistics major, I’ve been very interested in syntax and semantics, and what roles words play in daily communications and how they do it. It is a perfect example of how words of the same category function differently and we, as humans, utilize them without noticing what changed. It becomes very interesting when trying to teach it to a machine, and I think labeling of such function is essential for any kind of text processing.

1. What is your input space

The input space will be annotated texts from database.

1. What is your output space

The output space will be texts properly labeled with spatial roles, i.e. the trajector, the spatial indicator, and the landmark.

1. Do you know any data available for your learning task

The data used for this project will come from three datasets from Dr. Kordjamshidi’s previous research on this topic. Originally from KU Leuven University, the data was specifically created for Spatial Role Labeling and used in Semantic Evaluation Campaigns in 2012 and 2013.[[1]](#footnote-1)

1. What learning paradigm you think you will use (supervised, unsupervised, ..)

I think we will be using supervised learning paradigm because we have desired output labels for each word to be marked with. Therefore, we’re training the machine to sort the input text based on given categories.

1. Parisa Kordjamshidi, Martijn van Otterlo, and Marie-Francine Moens. Spatial Role Labeling Annotation Scheme*.* http://www.cs.tulane.edu/~pkordjam/SpRL.htm [↑](#footnote-ref-1)