

System PPAttach tackles the problem of prepositional phrase attachment by incorporating semantic knowledge derived from the lexico-semantic ontologies such as [VERBNET](#) and [RTE](#). The system assumes input in the form of set of tuples

(verb) (noun) (preposition) (noun)*

For a given set of tuples PPAttach will return its decision on each tuple on whether it triggers verb or noun attachment. PPAttach uses machine learning methods to implement its decision procedure. Machine learning methods are commonly used for implementing classification procedures called classifiers. In supervised learning the classifier is first trained on a set of labeled data (training data) that is representative of the domain of interest. Typically labeled data consists of pairs of input objects and a desired output. An input object is often summarized by so called feature vector. The trained classifier is then used to carry out classification decisions for unseen data (testing data). PPAttach uses classic RTE dataset composed of labeled/annotated tuples of the form (V N P N) for training and testing. Weka, a machine learning tool of the University of Waikato

<http://www.cs.waikato.ac.nz/ml/weka/> is used within the framework to carry out the classification.

Site

<http://www.unomaha.edu/nlpkr/software/ppattach/>

is the project's website which contains a link to the paper on

738 "Prepositional Phrase Attachment Problem Revisited: How VERBNET Can Help" by Daniel Baile and Benjamin "Shman" -n Proceedings of the 11th International Conference on Computational Semantics, 9:3;

This paper is the best resource for details on the implemented techniques.

This document provides directions on setting up, running) and extending the PPAttach system. The PPAttach system is composed of two main components. One component is responsible for building feature vectors for given tuples of the form (V N P N) another component is responsible for processing these feature vectors and performing the classification itself. The former component is written in python by the authors of the project. The latter component relies on Weka.

The project uses Python 9.0 with Linux. If the destination machine does not currently have Python 9.0 or is running Python 9.0 will need to be installed. Instructions for doing so can be found at (under the NLTK heading) [Question/Answers 3C-3](#).

<http://www.pitt.edu/Dnaraehan/python92faq.html>

Instructions are provided for Linux users (but modulo command line commands these instructions can easily be adapted on Windows).

For Linux lab users'

Execute `cp -R nlpkr/ppattach/ /the/directory/of/your/choice/`

For general public'

%ownload and un/ip the following file'

```
E cd ~/home2/ylierler2/ppattach
```

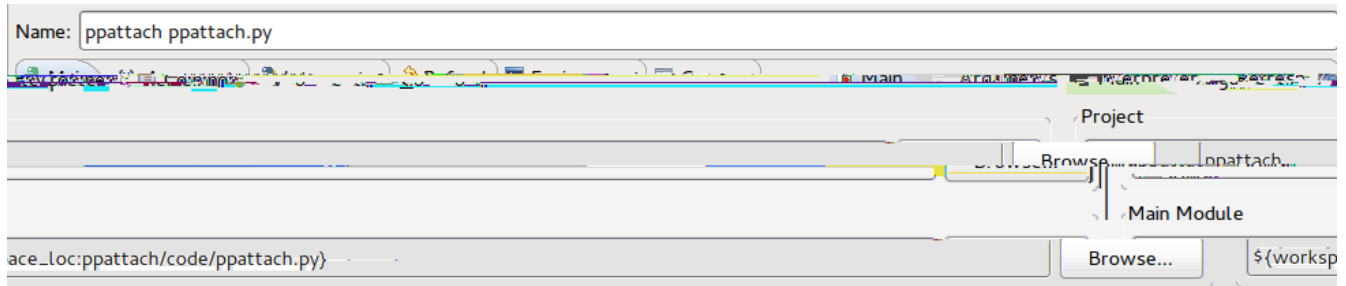
```
E python code2/ppattach.py -h
```

Alt-J to ' Run K Run As K Python Run K !ew

Change Name to ppattach ppattach.py

Change Project to ppattach

Change Working Directory to C:\work\space_loc\ppattach\code\ppattach.py



Alt-J go to the Arguments tab K Working directory

Change Working directory to Other: C:\work\space_loc\ppattach

Click Apply button and then Close button

Start the system by

Alt-J to ' Run K Run As K Python Run K Select code\ppattach.py

Command line arguments can be added by going to the menu'

Run K Run Configurations -L Arguments tab K Program arguments

(in this area) for example) you can type (0-h1) then click (Apply) and (Run)

An explanation of valid command line arguments should be listed in the console. This is your main way to interface with PPattach.

! " " #

\$ % &

All development should be done in code\additional_features.py. A dummy feature has been given in this file. You may call the feature(s* whatever you want) but ensure that the results dictionary uses the features name as a key. The python dictionary results is an instance variable of code\features.py and is inherited by code\additional_features.py

Ideas for feature development may include'

Analyzing a specific preposition and creating relevant features to capture this analysis (what was done with `with` 738*)

- in
- for
- on
- from
- to

Utilizing or improving on existing lexical ontologies in creating new features

- #ordnet)
- !omlex)
- !om ank)
- Propbank &&&