Spring Semester 2018
Iowa State University

ENGL 516 - Formal Methods of Linguistic Analysis

Assignment 7
Submission Deadline: 14 APR 2018

Instructions: This assignment consists of two questions. Each question carries 5 marks. Both the questions involve writing python code and using an external module. Save your code as two .py files (one for each question), and submit them on Canvas. If any file does not run and throws errors, you cannot get a grade unless you defend it in the office hours.

Question 1

Beautiful Soup is a python library that gives us easy methods to parse HTML files and extract information from them. You have to now write a python code using beautiful soup, which takes a url from stackoverflow.com as input, and returns the following information as output:

1. title of the post

2. Number of votes the question received

3. Number of stars the question received

4. Number of answers that question received

5. Who is the original poster of the question?

For example, if the user enters a URL: `http://stackoverflow.com/questions/3207219/how-to-list-all-files-of-a-directory-in-python`, your program should output:
The title of this thread is: "How do I list all files of a directory?"
Number of votes the question received: 2372 (as on 18th Jan 2018)
Number of stars the question received: 580
Number of answers the question received: 27
The original poster of this question is: duhhunjonn
If the user enters any other URL of some other website, or if the stackoverflow url the user entered does not have forum discussions (e.g., careers page), then just print a message saying: "Please give the url to a discussion post on stackoverflow.com"

Question 2

LanguageTool.org is a open-source proof reading program which does spelling and grammar check, in many languages. Figure out how to use languagetool from Python (there are python libraries), read the documentation, and write a program that asks user to enter a sentence, and prints out: a) number of errors identified by LanguageTool for this sentence the user gave b) error categories for each of the errors.

A note of caution: LanguageTool makes errors. So, don't take its error tags as ground truth. This is just an exercise to practice using external libraries in your Python code.