

Fall Semester 2016  
Iowa State University

## ENGL 520 - Computational Analysis of English

### Assignment 2

**Submission Deadline: 27 SEP 2016, end of the day**

**Instructions:** This assignment consists of three questions. Each question carries 5 marks. First question does not involve programming. Second and third question involve programming and you can use any programming language you want. Create a zip file with three components (a pdf file for first question, 2 perl or python files). If any of the programs does not run and throws errors, you will not get credit for that. Late submissions are allowed, but will not be awarded full credit.

### Question 1

Write 3 rules to remove the prefixes and 3 rules to remove suffixes of English words. Example rules include: "if the word ends in 'ed', remove the 'ed'" (cooked - cook), "if the word begins with re, remove re" (e.g., relocate - locate). For each rule, give two examples of words for which these rules work, and two examples for which these rules will result in wrong stripping (e.g., applying the "re" rule to reel will result in el, but re is not a prefix to el).

### Question 2

Write a program that identifies date expressions in user input. Try to cover as many possible patterns as possible in your program. Example expressions: March 15, 22nd of November, 15-3-2016, 3-15-2016, 15/3/2016, 3/15/2016, 15th March, March 15 2016, 15th March 2016, 1st March, 2nd May 2016.(Not just for march!! Patterns should cover all months)

### Question 3

Write a program to convert a name (first name or last name) into soundex representation. Details about Soundex: <http://www.archives.gov/research/census/soundex.html>.