

# Project 1

**Due: Mar 19 11:59 PM**

You will be doing a number of tasks related to parsing English sentences.

## Background:

Rules for legal English sentences can be represented as grammar rules. For the purposes of this project, we will only have the following rules:

1. `<sentence> ::= <noun><verb><noun>`
2. `<sentence> ::= <noun><verb><adjective><noun>`
3. `<sentence> ::= <pronoun><verb><noun>`
4. `<sentence> ::= <pronoun><verb><adjective><noun>`
5. `<sentence> ::= <pronoun><verb><pronoun>`

For example, the first rule says that a noun followed by a verb followed by a noun (and nothing else) is one way to make a legal sentence.

Of course the actual English language contains many more such rules. Also note that we are concerned purely with syntax, and can thus construct sentences that are semantic nonsense (e.g., "I perform depression") or do not take into account other grammatical rules such as those about subject-verb agreement (e.g., "I are large bird").

## Input:

You have available to you 4 input files<sup>1</sup>, named nouns, verbs, adjectives, and pronouns. These files have been filtered to exclude words that can be multiple parts of speech (e.g., dance). Each file contains one word per line, all lowercase.

The input files are supplied in tar form. To extract them, type "tar -xvf words.tar" without the quotes - you will learn about tar later in 136.

## Assignment

### Part A:

Prompt the user for a sentence. Depending on whether or not the sentence is a legal English sentence according to the above rules, output one of the following 6 messages (exactly as stated) to the screen, and terminate:

Your sentence is a legal sentence by rule {1,2,3,4,5}.

Your sentence is not a legal sentence.

You may assume that the input sentence contains only letters and spaces, though letters may be either upper or lower case. Words in the sentence are separated by at least one space.

### Part B:

Elbonian is similar to English, but has 2 differences:

1. Verbs always occur at the beginning of the sentence (rumor has it that Elbonians like to start doing things before understanding the problem). Other words are still in the same order as in English. For example, rule 1 above would be as follows in Elbonian:  
<sentence> ::= <verb><noun><noun>
2. There are no adjective words, and the modified noun is changed instead. More specifically, the modified noun is changed to have the first letter of the adjective as a prefix and the remaining letters as a suffix.

For example, the English sentence "I avoid large thing" would be "avoid I lthingarge" in Elbonian.

Given a file of sentences named "engsents", produce a second file named "elbsents" containing the equivalent Elbonian sentences. Both files contain one sentence per line. If an input is not a legal sentence, you should output "Illegal sentence found" on its own line in the output file and proceed to the next sentence. Your translation should preserve cases, though the output should have exactly one space between words.

## Program and Submission Guidelines

For this program you may only use those concepts covered through strings in 135 lectures. This includes file I/O, functions, and all control structures, but EXCLUDES any libraries (e.g., string library functions not covered in class) or data structures (e.g., arrays, vectors, structs, pointers) coming later this semester. This means that you will need to read the input files many times.

As always, you should have the following header on all programs:

```
/*
  Author: <name>
  Course: {135,136}
  Instructor: <name>
  Assignment: <title, e.g., "Lab 1">

  This program does ...
*/
```

As with 136, a program that does not run on the computers in Lab G is considered incorrect. You are graded on the same guidelines as 136. For the testing component of the grade, write in comments what test cases you used for each task.

You should submit one program per task, named <firstname>.<surname>.prog1{A,B}.cpp. You should also submit your input test file from part B. Although this only counts towards your 135 grade, it will be submitted to your 136 instructor for grading (there will be a submission link on your 136 bb page).

## Footnotes

1 The input files are downloaded from talkenglish.com, and have been filtered to exclude words that may be multiple forms of speech. Errors in the source have not been checked (there are a few), though it shouldn't affect you for this program.