# Programming Exercises

*For Software Engineer Candidates*

## The Bigram Parsing Problem:

**Task**
Create an application that can take as input **any** text file and output a histogram of the bigrams in the text.

### Description:

A bigram is any two adjacent words in the text disregarding case and punctuation. A histogram is the count of how many times that particular bigram occurred in the text.

A well-formed submission will be runnable from command line and have accompanying unit tests for the bigram parsing and counting code. You may do this in any language you wish and use any framework or data structures you wish to handle reading the files, building up the output, and running the unit tests. However the bigram parsing and counting code must be implemented by yourself.

**Please note that it is important to meet all of the requirements, including unit tests, for the submission to be considered complete.**

### Example:

Given the text: “The quick brown fox and the quick blue hare.” The bigrams with their counts would be.

- “the quick” 2
- “quick brown” 1
- “brown fox” 1
- “fox and” 1
- “and the” 1
- “quick blue” 1
- “blue hare” 1