

## Homework 2

### Due Tuesday, 02/12/08

Please turn in a print-out of your homework solutions at the beginning of class. If a program is required for a problem then you should provide sample input and output. (If the input is data from an on-line source then you may just indicate where the file can be found.) Programming solutions to problems should be placed in separate files whose names indicate the problem number (e.g. prob1.py). These separate files should be put into a folder whose name includes the assignment number and your name (e.g., Hmwk1-yourname). This folder should be dragged into the class dropoff folder at

`/common/cs/cs181/dropbox`

While I expect that I will be able to do most of the grading based on the papers that you hand in at the beginning of class, I would like to have access to the programs so that I can test them if necessary. I should be able to load your file into python and have it compile and run without error. All programming solutions should be fully documented and use good variable names.

Note that the files on our system can be accessed using ssh via the Mac server xserv.cs.pomona.edu or the linux server linux.cs.pomona.edu. Files and/or folders can be deposited into the dropbox remotely by using a program utilizing the sftp protocol.

*Warning: There are several different versions of Bird et al floating around. From now on all assignments will be relative to the pdf edition of the entire book dated January 24, 2008, and available via the course web page.*

1. Please do Problem 5 in section 3.2.4 on page 78 of Bird.
2. Please do Problem 11 in section 3.2.4 on page 78 of Bird.
3. Please do Problem 1b (money only) in section 3.4.4 on page 87 of Bird.
4. Please do Problem 3 in section 3.4.4 on page 88 of Bird. Use the text in “austin-persuasion” in the Gutenberg corpus of nltk. Compare the entries for the second 200 words using each of the two stemmers and explain the differences.
5. Please do Problem 3.4 on page 39 of JM.
6. Please do Problem 3.9 on page 39 of JM.
7. Please do Problem 3.10 on page 39 of JM.
8. Please do Problem 3.12 on page 39 of JM. You may start with my program in file minEditDistance.py in `/common/cs/cs181/programs`.

Your output should list the table in a style similar to that shown in Lecture 4 (though you’ll have to figure out another way to produce something to take the place of arrows). Also present a list of exactly the changes to be made to get from the start to the finish and its cost. Do your best to make the output as readable as possible.