

L555: Programming for (Computational) Linguists

September 1, 2010

In-Class Exercise #1

- 1) Go to the directory `/Users` and find out how many users there are.
- 2) Go back to your home directory.
- 3) Create a directory called *useless*. Go to this directory and create two subdirectories called *waste_of_money* and *waste_of_time*.
- 4) Download the file *data.tgz* from oncourse and move it to *waste_of_time*.
- 5) Unpack the file.
- 6) Move the file *numbers.txt* to *waste_of_money*.
- 7) Sort the file. Sort it again so that the numbers are in ascending order.
- 8) Go back to *waste_of_time* and search the file *vm.pos* for adjectives (JJ) using `grep`. Save the adjectives in a separate file called *adj.txt*.
- 9) Then search for personal pronouns (PP) and save them in a file called *pron.txt*.
- 10) Find out how many adjectives and pronouns there are in these files. How many different ones?
- 11) Sort the words from *vm.pos* alphabetically (don't worry about the POS tags) and kick out duplicates. Save the result in *single.txt*.
- 12) Go back to your home directory and delete the whole directory *useless*.