Week 4 : Programs

**8.4** Open the file **romeo.txt** and read it line by line. For each line, split the line into a list of words using the **split()** method. The program should build a list of words. For each word on each line check to see if the word is already in the list and if not append it to the list. When the program completes, sort and print the resulting words in alphabetical order.

Explanation:

**fname = input("Enter file name: ")**

**fh = open(fname)**

**lst = list() # create an empty list and store in variable "lst"**

**for line in fh: # we check every line in the file**

**line=line.split() # we split every word after ' ' is found in the line**

**for x in line : # we take a variable to check each word in the list of each line**

**if x in lst : # if the word is present in the list, go to the next word in the list**

**continue # go again to the for loop to check the next word in the list**

**else :**

**lst.append(x) # if the word is not found, append it into the list**

**lst.sort()**

**print(lst)**

**8.5** Open the file **mbox-short.txt** and read it line by line. When you find a line that starts with 'From ' like the following line:

From stephen.marquard@uct.ac.za Sat Jan 5 09:14:16 2008

You will parse the From line using split() and print out the second word in the line (i.e. the entire address of the person who sent the message). Then print out a count at the end.

**Hint:** make sure not to include the lines that start with 'From:'.

Explanation:

**fname = input("Enter file name: ")**

**fh = open(fname) count = 0**

**for line in fh :**

**if line.startswith('From '): # consider the lines which start from the word "From "**

**y=line.split() # we split the line into words and store it in a list**

**print(y[1]) # print the word present at index 1**

**count=count+1 # increment the count variable**

**print("There were", count, "lines in the file with From as the first word")**