**Experiment 2**

* Open the text file and make a list of unique words contained in the file.
* 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.

* Write a program to read through the mbox-short.txt and figure out who has sent the greatest number of mail messages. The program looks for 'From ' lines and takes the second word of those lines as the person who sent the mail. The program creates a Python dictionary that maps the sender's mail address to a count of the number of times they appear in the file.

**1.**

file\_name = input("Enter file name: ")

fh = open(file\_name)

lst = list()

for line in fh:

line.rstrip()

word = line.split()

for j in word:

lst.append(j)

print(lst)

**OUTPUT:**

****

**2.**

file\_name = input("Enter file name: ")

fh = open(file\_name)

count = 0

lis = list()

for line in fh:

line.rstrip()

word = line.split()

for j in range(len(word)):

if word[j].lower() == 'from' and '@' in word[j + 1]:

lis.append(word[j + 1])

print(word[j + 1])

count = count + 1

print(count)

print("There were", count, "mail addresses")

print(lis)

**OUTPUT:**



**3.**

handle = open('mbox-short.txt')

count = 0

lst = list()

for line in handle:

if not line.startswith("From:"):

continue

line = line.split()

lst.append(line[1])

counts = dict()

for word in lst:

counts[word] = counts.get(word, 0) + 1

bigcount = None

bigword = None

for word, count in counts.items():

if bigcount is None or count > bigcount:

bigcount = count

bigword = word

print(bigword, bigcount)

**OUTPUT:**

