CSE 101: Computer Science Principles (Fall 2019)

Lab #7

Assignment Due: Saturday, November 9, 2019, by 11:59 PM

Assignment Objectives

By the end of this assignment you should be able to develop original Python functions to solve simple programming problems involving strings, dictionaries, and files.

Getting Started

This assignment requires you to write Python code to solve several computational problems. To help you get started, we will give you a basic starter file for each problem. These files will contain *function stubs* and a few tests you can try out to see if your code seems to be correct (**note that the test cases we give you in the starter files are just examples; we will use different test inputs to grade your work!**). You need to complete (fill in the bodies of) these functions for the assignments. **Do not, under any circumstances, change the names of the functions or their parameter lists.** The automated grading system will be looking for functions with those exact names and parameter lists; if you change any function headers, the grading program will reject your solution and mark it as incorrect.

Directions

Solve each of the following problems to the best of your ability. We have provided a code skeleton for each of the programming problems. The automated grading system will execute your solution to each programming problem several times, using different input values each time. Each test that produces the correct/expected output will earn 1 or more points. This assignment contains 3 problems, and is worth a total of 30 points. **Note that not every problem may be worth the same number of points!**

- ⚠ Each starter file has a comment block at the top with various important information. Be sure to add your information (name, ID number, and NetID) to the first three comment lines, so that we can easily identify your work. **If you leave this information out, you may not receive credit for your work!**
- ⚠ Submit your work as a set of individual files (one per problem). **DO NOT** zip or otherwise compress your files, and **DO NOT** place all of your work in a single file. If you do so, we may not be able to grade your work and you will receive a failing grade for this assignment!
- ⚠ Every function **MUST** use the names and parameter lists indicated in the starter code file. Submissions that have the wrong function names (or whose functions contain the wrong number of parameters) can't be graded by our automated grading system, and may receive a grading penalty (or may not be graded at all).
- ⚠ Every function must explicitly ***return*** its final answer; the grading program will ignore anything that your code prints out. Along those lines, do **NOT** use `input()` anywhere within your functions (or anywhere before the `if __name__ == "__main__":` statement); your functions should get all of their input from their parameters. Programs that crash will likely receive a failing grade, so test your code thoroughly **with Python 3.7.4 or later** before submitting it.
- ⚠ Blackboard will provide information on how to submit this assignment. You **MUST** submit your completed work as directed by the indicated due date and time. We will not accept (or grade) any work that is submitted after the due date and time, or that is submitted before you have signed the course's Academic Honesty agreement.
- ⚠ **ALL** of the material that you submit (for each problem) **MUST** be your own work! You may not receive assistance from or share any materials with anyone else, except for the instructor and the (current) course TAs.

Part I: Counting Characters (10 points)

(Place your answer to this problem in the "counter.py" file)

Function Summary

Parameters: The first parameter is a string representing a file name. The second parameter is a string representing a single character to count.

Return Type: The function returns an integer: the total number of times the second parameter appears in the specified file (ignoring lines that begin with the # character).

Complete the `counter()` function, which takes two arguments: a filename and a single character to count. This function opens the file named by the first parameter and examines it line by line. It counts and returns the total number of times the specified character appears in the file, but it **skips over** lines that begin with a `"#"` character (lines that contain that character but do not begin with it should be included in the counting process). **NOTE:** This function is **case-sensitive;** it should distinguish between the uppercase and lowercase versions of a letter.

Examples:

Function Call	Return Value
`counter("no-comments.txt", "e")`	`76`
`counter("some-comments.txt", "s")`	`18`
`counter("no-letter-found.txt", "X")`	`0`

Part II: Tracking Long Words (10 points)

(Place your answer to this problem in the "long_words.py" file)

Function Summary

Parameters: The first parameter is an all-lowercase string of words, each separated by a single space character. The second parameter is an integer representing the minimum length of an "acceptable" word.

Return Type: The function returns a dictionary. Each key is a string corresponding to a word from the first parameter; its matching value is an integer representing the number of times that word appears in the first parameter.

Complete the `long_words()` function, which takes two arguments: an all-lowercase string and a positive (non-zero) integer, in that order. The function creates and returns a dictionary where each key is a word from the string whose length is greater than or equal to the integer parameter (e.g., if the second parameter was 4, every dictionary key must have 4 or more letters). A key's value is equal to the number of times that word appears in the original string. For example, `"science":2` indicates that the "science" appears twice in the original string.

Hint: As you process the string, the `setdefault()` operation may be helpful.

Examples:

Function Call	Return Value
`long_words("to be or not to be", 2)`	`{"to": 2, "be": 2, "or": 1, "not": 1}`
`long_words("to be or not to be", 5)`	`{}`
`long_words("i am the very model of a modern major general i've information vegetable animal and mineral i know the kings of england and i quote the fights historical from marathon to waterloo in order categorical i'm very good at integral and differential calculus i know the scientific names of beings animalculous in short in matters vegetable animal and mineral i am the very model of a modern major general", 5)`	`{"model": 2, "modern": 2, "major": 2, "general": 2, "information": 1, "vegetable": 2, "animal": 2, "mineral": 2, "kings": 1, "england": 1, "quote": 1, "fights": 1, "historical": 1, "marathon": 1, "waterloo": 1, "order": 1, "categorical": 1, "integral": 1, "differential": 1, "calculus": 1, "scientific": 1, "names": 1, "beings": 1, "animalculous": 1, "short": 1, "matters": 1}`
`long_words("Man: well what've you got waitress: well there's egg and bacon egg sausage and bacon egg and spam egg bacon and spam egg bacon sausage and spam spam bacon sausage and spam spam egg spam spam bacon and spam spam sausage spam spam bacon spam tomato and spam vikings: spam spam spam spam waitress: spam spam spam egg and spam spam spam spam spam spam spam baked beans spam spam spam vikings: spam lovely spam lovely spam waitress: or lobster thermidor au crevette with a mornay sauce served in a provencale manner with shallots and aubergines garnished with truffle pate brandy and with a fried egg on top and spam", 4)`	`{"Man:": 1, "well": 2, "what've": 1, "waitress:": 3, "there's": 1, "bacon": 7, "sausage": 4, "spam": 35, "tomato": 1, "vikings:": 2, "baked": 1, "beans": 1, "lovely": 2, "lobster": 1, "thermidor": 1, "crevette": 1, "with": 4, "mornay": 1, "sauce": 1, "served": 1, "provencale": 1, "manner": 1, "shallots": 1, "aubergines": 1, "garnished": 1, "truffle": 1, "pate": 1, "brandy": 1, "fried": 1}`

Part III: Multiplexing Madness (10 points)

(Place your answer to this problem in the "multiplexer.py" file)

Function Summary

Parameters: This function takes one parameter: a string that represents the name of a plain text data file

Return Type: This function returns a dictionary. Each key is a positive (non-zero) integer, and each value is a string.

Complete the `multiplexer()` function, which takes a single argument: a string representing the name of a file. Each line of this file begins with a positive (non-zero) integer and then contains some text; multiple lines may begin with the same integer. This function processes the file line-by-line. For each line, it uses the integer value as the key for a dictionary; every line that begins with the same integer has its contents appended to that dictionary entry. For example, given a file containing the lines

```
...
79 The rain in Spain
...
79 falls mainly on the plain.
...
```

the dictionary would contain the element `79: "The rain in Spain falls mainly on the plain. "`. Note that a single space has been inserted between the two strings; it's okay if the recombined string has extra space at the end.

Hint: After calling `strip()` on each line of the file, use `split()` and `join()`, or `find()` and slicing, to separate the key from the remaining text.

Examples:

`multiplexer("test-file-1.txt")` should produce the following dictionary:

```
{67:  "The first design for a program-controlled computer was Charles
Babbage's Analytical Engine in the 1830s.  A century later, in 1936,
mathematician Alan Turing published his description of what became
known as a Turing machine, a theoretical concept intended to explore
the limits of mechanical computation.  ", 84:  'The rules of modern
tennis have changed little since the 1890s.  Two exceptions are that
from 1908 to 1961 the server had to keep one foot on the ground at all
times, and the adoption of the tiebreak in the 1970s.  ', 28:  'Frank
Lloyd Wright (June 8, 1867 - April 9, 1959) was an American architect,
interior designer, writer, and educator, who designed more than 1,000
structures, 532 of which were completed.  Wright believed in designing
structures that were in harmony with humanity and its environment, a
philosophy he called organic architecture.  ', 39:  'Pepsi was first
introduced as "Bradś Drink" in New Bern, North Carolina, United States,
in 1893 by Caleb Bradham, who made it at his drugstore where the drink
was sold.  It was renamed Pepsi-Cola in 1898 after the Greek word for
"digestion", which the drink was purported to aid, and "cola" after the
kola nut.  '}
```

`multiplexer("test-file-2.txt")` should produce the following dictionary:

```
{25:  'Bruce McLaren Motor Racing was founded in 1963 by New Zealander
Bruce McLaren.  Bruce was a works driver for the British Formula One
team Cooper with whom he had won three Grands Prix and come second in
the 1960 World Championship.  ', 70:  'Four score and seven years ago
our fathers brought forth on this continent a new nation, conceived
in Liberty, and dedicated to the proposition that all men are created
equal.  Now we are engaged in a great civil war, testing whether
that nation or any nation so conceived and so dedicated, can long
endure.  ', 53:  'In mathematics, a Golomb ruler is a set of marks
at integer positions along an imaginary ruler such that no two pairs
of marks are the same distance apart.  The number of marks on the
ruler is its order, and the largest distance between two of its marks
is its length.  ', 35:  'Frank Lloyd Wright (June 8, 1867 - April
9, 1959) was an American architect, interior designer, writer, and
educator, who designed more than 1,000 structures, 532 of which were
completed.  Wright believed in designing structures that were in
harmony with humanity and its environment, a philosophy he called
organic architecture.  ', 66:  'Pepsi was first introduced as "Bradś
Drink" in New Bern, North Carolina, United States, in 1893 by Caleb
Bradham, who made it at his drugstore where the drink was sold.  It
was renamed Pepsi-Cola in 1898 after the Greek word for "digestion",
which the drink was purported to aid, and "cola" after the kola nut.
', 11:  "The first design for a program-controlled computer was Charles
Babbage's Analytical Engine in the 1830s.  A century later, in 1936,
mathematician Alan Turing published his description of what became
known as a Turing machine, a theoretical concept intended to explore
the limits of mechanical computation.  ", 40:  'The rules of modern
tennis have changed little since the 1890s.  Two exceptions are that
from 1908 to 1961 the server had to keep one foot on the ground at
all times, and the adoption of the tiebreak in the 1970s.  ', 69:
'The wyvern in its various forms is important to heraldry, frequently
appearing as a mascot of schools and athletic teams (chiefly in the
United States, United Kingdom, and Canada).  It is a popular creature
in European and British literature, video games, and modern fantasy.
'}
```

`multiplexer("test-file-3.txt")` should produce the following dictionary:

```
{43:  'Frank Lloyd Wright (June 8, 1867 - April 9, 1959) was an
American architect, interior designer, writer, and educator, who
designed more than 1,000 structures, 532 of which were completed.
Wright believed in designing structures that were in harmony with
humanity and its environment, a philosophy he called organic
architecture.  ', 24:  'The rules of modern tennis have changed little
since the 1890s.  Two exceptions are that from 1908 to 1961 the server
had to keep one foot on the ground at all times, and the adoption of
the tiebreak in the 1970s.  ', 40:  'Four score and seven years ago
our fathers brought forth on this continent a new nation, conceived
in Liberty, and dedicated to the proposition that all men are created
equal.  Now we are engaged in a great civil war, testing whether that
nation or any nation so conceived and so dedicated, can long endure.
', 72:  'The wyvern in its various forms is important to heraldry,
frequently appearing as a mascot of schools and athletic teams (chiefly
in the United States, United Kingdom, and Canada).  It is a popular
creature in European and British literature, video games, and modern
fantasy.  ', 87:  'In mathematics, a Golomb ruler is a set of marks at
integer positions along an imaginary ruler such that no two pairs of
marks are the same distance apart.  The number of marks on the ruler
is its order, and the largest distance between two of its marks is its
length.  ', 17:  'Bruce McLaren Motor Racing was founded in 1963 by
New Zealander Bruce McLaren.  Bruce was a works driver for the British
Formula One team Cooper with whom he had won three Grands Prix and come
second in the 1960 World Championship.  ', 5:  "The first design for a
program-controlled computer was Charles Babbage's Analytical Engine
in the 1830s.  A century later, in 1936, mathematician Alan Turing
published his description of what became known as a Turing machine,
a theoretical concept intended to explore the limits of mechanical
computation.  ", 45:  'Pepsi was first introduced as "Bradś Drink"
in New Bern, North Carolina, United States, in 1893 by Caleb Bradham,
who made it at his drugstore where the drink was sold.  It was renamed
Pepsi-Cola in 1898 after the Greek word for "digestion", which the
drink was purported to aid, and "cola" after the kola nut.  '}
```