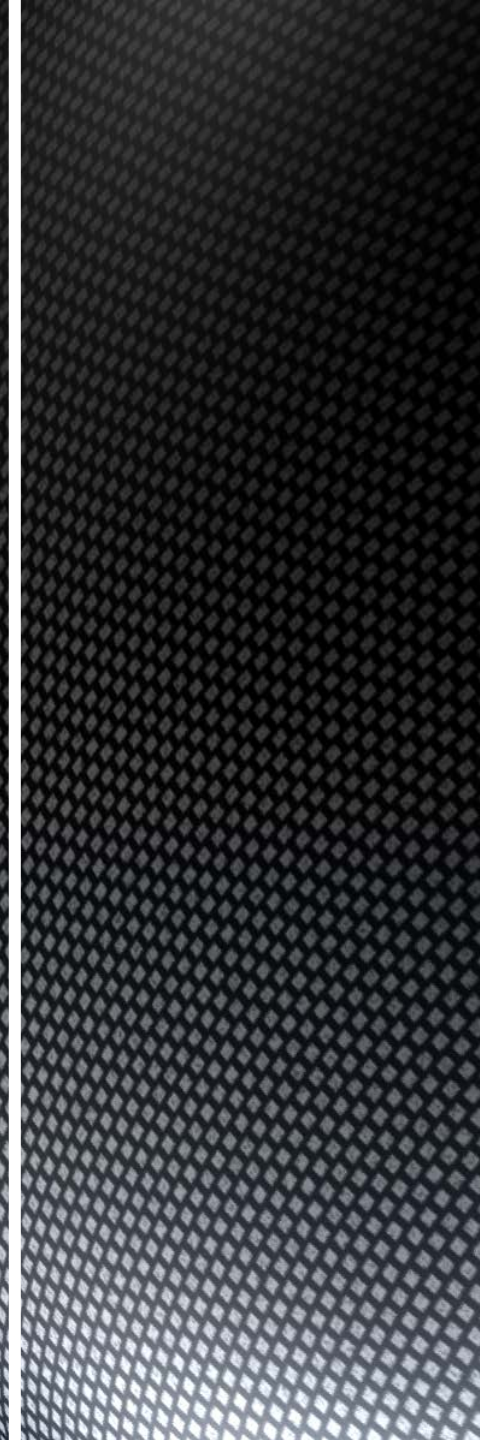# COMPSCI 107
Computer Science Fundamentals

Lecture 06 – Reading and Writing Text Files

Learning outcomes

- At the end of this lecture, students should be able to:
  - Read text files
  - Write text files

- Example code:
  - Word count of a text file
  - Sum the numbers in a text file

Working with Random numbers

- Random numbers are pseudo-random
  - Use a seed value

- import random

- random.seed(x)

- random.choice(<sequence>)

- Get a reference to the file on disk
  - open(filename, mode)

- Modes
  - 'r' = read
  - 'w' = write
  - 'r+' = read and write
  - Default is text mode, append 'b' for binary

```
f = open(filename)
f = open(filename, 'r')
f = open(filename, 'wb')
```

Reading a text file

- Different ways to read the contents of the file
  - When the end of file is reached, empty string is returned

f.read() # returns the entire file as a string

f.readlines() # returns a list of lines (strings) from the file

f.readline() # returns a single line (string) from the file

for line in f:
 print(line)   #do something with the line

Avoiding memory issues for large files

- With
- automatically closes the file
- handles some of the exceptions

```
with open('filename') as f:
    data = f.readlines()
```

```
with open('filename', 'w') as f:
    f.write('Hello World')
```

```
with open('filename') as f:
    for line in f:
        print(line)
```

Reading data from a file

```python
def read_data(filename):
    '''Reads data from a file

    Given the name of a file, reads all the data from a file and
    return the data as a list where each element in the list is a
    single piece of datum from the file.

    test2.txt contains the text:
    This is a small test
    file with a few words.

    >>> read_data('test2.txt')
    ['This', 'is', 'a', 'small', 'test', 'file', 'with', 'a', 'few', 'words.']
    '''
    data = []
    with open(filename) as f:
        for line in f:
            data += line.split()
    return data
```

Exercise

- Write a function that returns the word count for a text file.

- Write a function that returns the number of unique words in a text file (where words are defined as a sequence of characters separated by white space).

- Read data as text
  - Convert text to number

- Functions that change types
  - int()
  - float()
  - str()

```
for line in f:
    number = int(line)    #one value per line
```

Exercise

- Write a function that reads in a file of numbers and returns the sum of those numbers.

Exercise

- Write a function called unique_words that accepts the name of a text file as an argument and produces a list of all the words in that file, without repeating any of the words.

Exercise

- Write a function called **difference** that accepts two file names as arguments. The function should print out any line in the first file that is not identical to the same line in the second file

- **Similar to reading data**
  - Write all the data
  - Close the file after all the data is written.

```
f.write( string_to_write )
f.close()
```

Exercise

- Write a function called word_list that accepts the name of a text file as an argument and produces a list of all the words in that file, sorted in order with one word per line.

Summary

- Files need to be opened before you can read or write to them
  - f = open(*filename[, mode]*)

- Files should be closed when you finish with them.
  - f.close()

- Reading a file
  - f.read()
  - f.readline()

- Writing a file
  - f.write(*string*)