Homework 9

Introductory Programming **Allen B. Downey**
Fall 2004

The reading for this assignment is Chapters 10–11 of *How to think...*.

9.1 Spellcheck!

The goal of this homework is to implement a spellchecker. The simplest kind of spellchecker uses a dictionary of "legal" words and flags anything that is not in the dictionary.

1. Start with either your solution from the previous homework or my `ip_hw08_soln.py`, which is available on the class web page. Either way, you should have a program that reads a file and builds a dictionary that contains the unique words in the file.

2. In my solution, `process_file` has the side-effect of printing the 10 most common words in the file. This is a weird thing for a fruitful function to do becase you wouldn't necessarily want to print something every time you process a file.

   Modify `process_file` so that it prints nothing (has no side effects) and returns the dictionary. Then change the place where `process_file` is invoked so that it stores the return value from `process_file` in a variable named `words`, instead of printing it. Finally, add a line of code at the end of the program to print the top ten words in the dictionary.

3. Now we have a general-purpose function that we could use for more than one purpose. For example, we can use it to read a list of "legal" words and store it in a dictionary.

   Add a line of code to read the file `/usr/share/dict/words` and put the contents into the dictionary. How many words are there in this file? How many of them are unique?

4. At this point you should have two dictionaries, one containing the words from *The Great Gatsby* (or at least the first 100 lines) and the other containing the "legal" words.

   Write a loop that prints all the words from *Gatsby* that are not "legal". Hint: break this step into several smaller steps!

5. Encapsulate the code you just wrote in a function named `spellcheck` that takes two dictionaries as parameters and that prints all the words from the first dictionary that are not in the second dictionary. Hint: make sure your function gets defined before you try to invoke it.

6. Modify `spellcheck` to make it fruitful; that is, instead of printing the words, it should create and return a list of words.

7. Challenge (mild): what are the 10 most common "illegal" words in *Gatsby*?