Bridgewater challenge: Concordance

My name is `Andre Pena`. This is my Bridgewater challenge resolution for the `Concordance` problem. The objective is to count occurrences of words on a text file.

The project is also available on GitHub.

Building and running

I've written the program in C# and Visual Studio 2015. I would use 2013 if I had it, but I don't. However, I intentionally didn't use any C# 6.0 feature, so it will be easy to port it to a Visual Studio 2013 solution only by copying the source files if openening the solution doesn't work.

In order to run the program, open the WordCounter.sln solution and run the `WordCounter` project. It is a simple command line application.

`WordCounter` accepts a single argument which is the full name of the file from which the words should be counted.

If you are running from Visual Studio, right click the `WordCounter` project on the solution, click `Debug` and add the fullname of the file you want to analyze on the `Command line arguments` input. If you are running from the command line, just pass the same fullname as the first argument.

About the program

`WordCounter` is focused on performance, so…

- If does not completely load the target file in memory. Therefore, it should work with files that are greater than the available memory in the system.
- It iterate through each character of the file only once.
- It outputs the total number of milliseconds it took to run.

`Wordcounter` has basic support for numbers. The characters `.` and `,` are not considered sentence or word breakers if they have adjoining digits. There's no validation for numbers

whatsoever.

Unit tests

I've made basic unit tests, they can be found here. I've tested basic edge cases. It's not complete, but works well enough for what I intended to do.

Time needed to complete

`WordCounter` took me 5 hours to complete, not considering this documentation.

Contact me

You can reach me out at andrerpena@gmail.com