PWr Week 8
Data Structures and Algorithms
Laboratory – **List 8**

1. Find ten most often words in a huge text file (a word is a **consecutive sequence of letters** from set ‘a’..’z’ or ‘A..’Z’, before count words, change all letters to small ones). For counting words use:
   a. Binary search tree (without balancing)
   b. Balanced binary search tree (implemented on your own or from library)
   c. Hash table (implemented on your own or from library)

   Present the solution on the screen. Write the time of computation for used structures.

   Examples of huge text files:
   http://norvig.com/big.txt
   http://www.gutenberg.org/ (choose a big book in English in TXT format)

   Example what are words in this task.

   For a line:

   It’s a ver5y good #” idea of a line. You know it?

   The words are:
   it
   s
   a
   ver
   y
   good
   idea
   of
   line
   you
   know
   Words “it” and “a” are twice in the example text.

For **10 points** present solutions for this list till **Week 10**.
For 8 **points** present solutions for this list till **Week 11**.
For 5 **points** present solutions for this list till **Week 12**.
After Week 12 the list is closed.