

Fondamenti di Informatica

Prof. Evelina Lamma

29 July 2014

Exercise (20 points) (1 hour and 45 min)

A binary file, `parole1.bin`, contains some Italian words. The file is unsorted, and each word can be repeated. The words contain at most 20 characters, including the string terminator.

Create a C program, organized in at least three functions, that should perform the following tasks:

- create in memory a list `L`, that contains each word of the file exactly once, together with an integer field representing the number of times in which the word occurs in the file. In the list, the words must be sorted in alphabetical order.

Note: since each word should be inserted only once in the list, the usual function for inserting a value in the list should be modified accordingly: if a word is already present in the list, it should not be inserted again, but its number of occurrences should be incremented.

This function, called **functionA**, should take as parameters a file pointer, plus other parameters (if necessary), and should return the pointer to the first element of the list.

- show on the screen the content of the list `L`;
- by accessing `L`, compute the word that occurs most often in the file (i.e., the word for which the integer field takes the maximum value in the list).

This function, called **functionB**, takes as parameters the pointer to the first element of the list `L`, plus other parameters (if necessary), and returns a structure containing the word with maximum number of occurrences and its number of occurrences.

After invoking **functionB**, the `main` function should print on an output file, named `uscita.txt`, the returned word and its number of occurrences.

For example, if the file `parole1.bin` contains

```
il mare a genova e' piu' bello del mare a savona ed anche del mare adriatico
```

then the list should contain:

a	2
adriatico	1
anche	1
bello	1
del	2
e'	1
ed	1
genova	1
il	1
mare	3
piu'	1
savona	1

and the output file should contain:

mare 3

The output file **uscita.txt** should be submitted together with the source code.

For students that take the full A+B exam (Further 20 points, on a total of 60 for the full A+B exam; time +45 min; tot 2.30 h):

In the `parole1.bin` file there are at least 10 distinct words. Insert the first 10 distinct words of this file in an array `V` of size 10. Show the content of the array on the screen.

Sort the array `V` by using the function `qsort`, then print on a text file, named **outputAB.txt**, all the words in the array `V`. This task should be performed in a function called **functionC** that has, as parameters, the array `V` and the file pointer to file `parole1.bin`(plus other parameters, if necessary), and returns `void`.

The file **outputAB.txt** should be submitted together with the source code.

Note: Submit the source files, the executable file, and the output file that have been generated by your program. In case the application consists of more files, all user libraries should be submitted as well.