【问题描述】

编写程序统计一个英文文本文件中每个单词的出现次数(词频统计),并将统计结果按单词字典序输出到屏幕上。

要求:程序应用二叉排序树(BST)来存储和统计读入的单词。

注:在此单词为仅由字母组成的字符序列。包含大写字母的单词应将大写字母转换为小写字母后统计。在生成二叉排序树不做平衡处理。

【输入形式】

打开当前目录下文件 article.txt, 从中读取英文单词进行词频统计。

【输出形式】

程序应首先输出二叉排序树中根节点、根节点的右节点及根节点的右节点的右节点上的单词(即 root、root->right、root->right 节点上的单词),单词中间有一个空格分隔,最后一个单词后没有空格,直接为回车(若单词个数不足三个,则按实际数目输出)。

程序将单词统计结果按单词字典序输出到屏幕上,每行输出一个单词及其出现次数,单词和其出现次数间由一个空格分隔,出现次数后无空格,直接为回车。

【样例输入】

当前目录下文件 article.txt 内容如下:

"Do not take to heart every thing you hear."

"Do not spend all that you have."

"Do not sleep as long as you want;"

【样例输出】

do not take

all 1

as 2

do 3

every 1

have 1

hear 1

heart 1

long 1

not 3

sleep 1

spend 1

take 1

that 1

thing 1

to 1

want 1

you 3

【样例说明】

程序首先在屏幕上输出程序中二叉排序树上根节点、根节点的右子节点及根节点的右子节点的右子节点上的单词,分别为 do not take, 然后按单词字典序依次输出单词及其出现次数。

【评分标准】

通过全部测试点得满分