

WordCount 需求说明

1 WordCount 需求说明

WordCount 的需求可以概括为：对程序设计语言源文件统计字符数、单词数、行数，统计结果以指定格式输出到默认文件中，以及其他扩展功能，并能够快速地处理多个文件。
可执行程序命名为：wc.exe，该程序处理用户需求的模式为：

```
wc.exe [parameter] [input_file_name]
```

存储统计结果的文件默认为 result.txt，放在与 wc.exe 相同的目录下。

1.1 基本功能

```
wc.exe -c file.c      //返回文件 file.c 的字符数
wc.exe -w file.c      //返回文件 file.c 的单词总数
wc.exe -l file.c      //返回文件 file.c 的总行数
wc.exe -o outputFile.txt //将结果输出到指定文件 outputFile.txt
```

注意：
空格也算字符。
由空格或逗号分割开的都视为单词，且不做单词的有效性校验，例如：thi#,that 视为用逗号隔开的 2 个单词。
-c, -w, -l 参数可以共用同一个输入文件，形如：wc.exe -w -c file.c 。
-o 必须与文件名同时使用，且输出文件必须紧跟在-o 参数后面，不允许单独使用-o 参数。

1.2 扩展功能

```
wc.exe -s              //递归处理目录下符合条件的文件
wc.exe -a file.c       //返回更复杂的数据（代码行 / 空行 / 注释行）
wc.exe -e stopList.txt // 停用词表，统计文件单词总数时，不统计该表中的单词
```

[file_name]: 文件或目录名，可以处理一般通配符。

其中，
代码行：本行包括多于一个字符的代码。
空 行：本行全部是空格或格式控制字符，如果包括代码，则只有不超过一个可显示的字符，例如 “{” 。
注释行：本行不是代码行，并且本行包括注释。一个有趣的例子是有些程序员会在单字符后面加注释：
 } //注释
在这种情况下，这一行属于注释行。
-e 必须与停用词文件名同时使用，且停用词文件必须紧跟在-e 参数后面，不允许单独使用-e 参数。
stopList.txt 中停用词可以多于 1 个，每个单词占一行，不区分大小写，形如：

```
while
if
switch
```

则 while, if, switch 作为 3 个停用词，在单词统计的时候不予考虑。停用词表仅对单词统计产生影响，不影响字符和行数的统计。

1.3 高级功能

```
wc.exe -x              //该参数单独使用，如果命令行有该参数，则程序会显示图形界面，用户可以通过界面选取单个文件，程序就会显示文件的字符数、单词数、行数等全部统计信息。
```

1.4 举例

```
wc.exe -s -a *.c
```

返回当前目录及子目录中所有.c 文件的代码行数、空行数、注释行数。

```
wc.exe -s -a -c -w *.c -e stop.txt -o -output.txt
```

返回当前目录及子目录中所有.c 文件的字符数、单词总数、代码行数、空行数、注释行数，并将结果保存在 output.txt 中，且统计单词时忽略 stop.txt 中的单词。

1.5 输出结果的格式

输出结果格式形如：

```
[文件名], [统计标志]: [总数]
```

举例如下：
wc.exe -c file.c, 则输出结果保存在 result.txt 中，内容格式如下（注意大小写）：

file.c, 字符数: 50

wc.exe -w file.c, 则输出结果保存在 result.txt 中, 内容格式如下 (注意大小写) :

file.c, 单词数: 30

wc.exe -l file.c, 则输出结果保存在 result.txt 中, 内容格式如下 (注意大小写) :

file.c, 行数: 10

如果同时涉及多项统计, 如同时需要统计字符、单词和行数, 则按照**字符→单词→行数→代码行数/空行数/注释行的顺序, 依次分行显示**。显示顺序与输入参数的次序无关。例如:
wc.exe -l -c file.c, 则统计 file.c 中的字符数和行数, 输出结果保存在 result.txt 中, 内容格式如下:

file.c, 行数: 10

file.c, 字符数: 50

例如:
wc.exe -s -a -c -w *.c-e stop.txt -o -output.txt
输出结果保存在 output.txt 中, 内容格式为:

file1.c, 字符数: 50

file1.c, 单词数: 30

file1.c, 代码行/空行/注释行: 5/2/3

file2.c, 字符数: 40

file2.c, 单词数: 26

file2.c, 代码行/空行/注释行: 6/2/1