

# 软件工程基础 - 结对项目

请从以下3个项目中任选一个完成

2位同学结对共同完成，分别在各自的博客中对完成过程进行描述

代码仍然需要使用GitHub进行管理

- [软件工程基础 – 结对项目](#)
  - [项目1: WordCount \(单词计数\)](#)
    - [1.1 项目要求](#)
    - [1.2 标准测试集, 正确性和速度评比](#)
  - [项目2: 四则运算题目生成](#)
    - [2.1 项目要求](#)
      - [2.1.1 第1阶段](#)
      - [2.1.2 第2阶段](#)
      - [2.1.3 第3阶段](#)
    - [2.2 测试](#)
  - [项目3: 地铁](#)
    - [3.1 项目要求](#)
    - [3.2 扩展需求](#)
    - [3.3 测试](#)
  - [博客要求](#)
  - [附录](#)

## 项目1：WordCount（单词计数）

### 1.1 项目要求

wc.exe是一个常见的工具，它能统计文本文件的字符数、单词数和行数。这个项目要求写一个命令行程序，模仿已有wc.exe 的功能，并加以扩充，给出某程序设计语言源文件的字符数、单词数和行数。

实现一个统计程序，它能正确统计程序文件中的字符数、单词数、行数，以及还具备其他扩展功能，并能够快速处理多个文件。

- 具体功能要求：

程序处理用户需求的模式为：

```
wc.exe [parameter] [file_name]
```

- 基本功能列表：

```
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 的词数目
wc.exe -l file.c      //返回文件 file.c 的行数
```

- 扩展功能：

```
-s  递归处理目录下符合条件的文件。
-a  返回更复杂的数据（代码行 / 空行 / 注释行）。
```

**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“{”。

**代码行：**本行包括多于一个字符的代码。

**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。

- 高级功能：

**-x 参数。**这个参数单独使用。如果命令行有这个参数，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、行数等全部统计信息。

需求举例：

```
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。

## 1.2 标准测试集，正确性和速度评比

以子目录的形式把目前所有同学的程序都集中在一个总的测试目录下，作为测试集合。然后大家看看各自的程序要花多少时间才能正确并且较快地完成任务。在这里，同学们要记下满足了标准测试集之后，每人实际花费的时间[实际值]，并且按照附录中PSP的表格统计自己在软件开发的各个阶段所花费的时间。

## 项目2: 四则运算题目生成

## 2.1 项目要求

### 2.1.1 第1阶段

写一个能自动生成小学四则运算题目的命令行“软件”， 分别满足下面的各种需求。

下面这些需求都可以用命令行参数的形式来指定：

a) 一次可以出一千道题目， 并且没有重复的， 把题目写入一个文件中。

任何两道题目不能通过有限次交换+和×左右的算术表达式变换为同一道题目。例如， $23 + 45 =$  和  $45 + 23 =$  是重复的题目， $6 \times 8 =$  和  $8 \times 6 =$  也是重复的题目。 $3+(2+1)$ 和 $1+2+3$ 这两个题目是重复的，由于+是左结合的， $1+2+3$ 等价于 $(1+2)+3$ ，也就是 $3+(1+2)$ ，也就是 $3+(2+1)$ 。但是 $1+2+3$ 和 $3+2+1$ 是不重复的两道题，因为 $1+2+3$ 等价于 $(1+2)+3$ ，而 $3+2+1$ 等价于 $(3+2)+1$ ，它们之间不能通过有限次交换变成同一个题目。

b) 当有多于一个运算符的时候，如何对一个表达式求值？逐步扩展功能和可以支持的表达式类型，最后希望能支持下面类型的题目（最多 10 个运算符，括号的数量不限制）：

```
25 - 3 * 4 - 2 / 2 + 89 = ?  
1/2 + 1/3 - 1/4 = ?  
(5 - 4 ) * (3 +28) =?
```

c) 除了整数以外，还要支持真分数的四则运算。（例如： $1/6 + 1/8 = 7/24$ ）

d) 让程序能接受用户输入答案，并判定对错。最后给出总共 对/错 的数量。

### 2.1.2 第2阶段

增加一个运算符，要支持乘方(power)运算。乘方运算的优先级高于乘除法。如何表示乘方，有两种表示方法：

```
4 ^ 2 = 16,    4 的二次方等于 16。  这里， ^ 表示乘方  
4 ** 2 = 16,   4 的二次方等于16。   这里， ** 表示乘方 (** 之间不能有空格， 否则是错误的算式)
```

两种表示方法都要支持，可以通过设置来选择。

### 2.1.3 第3阶段

结对的同学商量一下，从以下几个方向中选择一个，对程序进行扩展。

- 把程序变成一个 Windows/Mac/Linux 电脑图形界面的程序（取决于你目前使用的电脑），同时增加“倒计时”功能， 每个题目必须在 20 秒钟完成，如果完不成，则得0 分并进入下一题。增加“历史纪录”功能， 把用户做题的成绩记录下来并可以展现历史记录。

- 把程序变成一个智能手机程序（你正在用什么手机，就写那个手机的程序），增加倒计时，和历史纪录功能（见上）。
- 把程序变成一个网页程序，用户通过设定参数，就可以得到各种题目。
- 选一个你从来没有学过的编程语言，试一试实现基本功能。  
估计做好这个软件需要的时间，并且写出大概的设计步骤和实现算法。
- 把这个程序的思路变成一个可以一步一步演示的动画。写一个带有图形界面的程序：

输入：一个正常的四则运算句子

输出：程序用动画表示分词的过程，后序转换的过程，处理不同运算符优先级的过程，根据调度场工作的原理，逐步算出得出结果的过程。

## 2.2 测试

请自行准备相应的测试数据，并在博客中描述为什么选择这些测试数据？

## 项目3：地铁

请上网搜到最新的[北京地铁线路图](#)。



## 3.1 项目要求

- 把这个图的各个线路，各个站点，换乘信息等用文本文件（假设名字叫 `beijing-subway.txt`）的形式保存起来，应该保存的信息有 {线路号，线路的各个站名，车站的换乘信息}，应用程序可以读取这个文件，就能掌握关于北京地铁线路的所有信息，应该用什么样的格式呢？
- 写一个命令程序（不妨叫 `subway.exe`），这个程序启动的时候，会读取 `beijing-subway.txt` 的信息，然后这个程序就等待用户的输入，用户可以输入地铁的线路编号，然后程序就输出此地铁线路的所有站名（按某一方向顺序输出即可）。输出站名后，程序又进入等待状态。
- 找到两点之间的最有效线路？请实现下面这个需求：

```
subway.exe /b 知春路 中关村
```

返回经历的站名的个数，和路径，如果有换乘，请列出换乘的线路，例如：

```
4

知春路

知春里

海淀黄庄 换乘10号线

中关村
```

- 如果乘客有钱又有闲，那么，怎样才能尽可能快地遍历地铁的所有车站呢（只用经过一次，不用下车，就算经过车站）。例如，注意到13号线和10号线相交的**知春路地铁站**，我们选它作为一个起始站，从这个站出发，要经历多少站（换乘不出地铁系统，即不能从一个地铁口走到路面，然后从另一个站进去），才能把所有地铁站都遍历呢？

扩展命令程序，让它接受一个地铁站名。例如这个格式：`subway.exe /a 知春路`

程序输出总共经历多少站，以及经历的站名，举一个特例，假如地铁系统只有知春路，西土城两个站，那么这个程序应该输出：

```
3

知春路

西土城

知春路
```

- 在地铁系统中换乘是比较麻烦的一件事情，花费额外的时间和精力（下车，步行到下一个线路，等车，挤上车，再找座位）。做一个改进，每次换乘，相当于额外经历了 3 个车站，

还是要求遍历所有车站的最优线路，程序应该怎么修改呢？

总的车站数量 = 实际经过的车站数量 + 换乘等价的车站数量

- 能否做成图形界面呢？首先我们要给每个站点一个坐标信息。
  - 修改 beijing-subway.txt 文件，加入适当的信息，为实现图形界面做准备。如何设计这些坐标呢？
  - 改进 subway.exe，加一个 `/g` 的命令行参数，让它根据你提供的坐标画出各个地铁站，线路，以及换乘站。
  - 存储有下面4个方案：
    - 用文本文件，用自定义的格式存放
    - 用XML格式存放
    - 用数据库的方式来存放，例如 [sqlite](#)
    - 还可以用自己定义的二进制的格式来存放。

请比较几种方式的优缺点，特别是，这些方式如何应对变化的内部，外部因素。

- 在GUI（图形界面）中实现遍历的解法。  
扩展 subway.exe，处理下面参数的时候，subway.exe /g 知春路

程序在图形界面中显示地铁地图（各个站点的相对位置和官方地图类似即可），然后用一个小亮点表示乘客，乘客正在经过的车站就会闪亮，乘客走过的路用不同的颜色标识，同时在适当的地方有数字表明乘客已经经过车站的数目。

## 3.2 扩展需求

请结对的同学商量一下是否要进行扩展，如果有时间，可从以下2个方向选择一个进行。

- 让程序能处理上海的地铁地图，或者其它城市的地图。把程序由“固定处理一个地图”升级为“能处理多个地图”，程序的什么模块需要变化？
- 把程序移植到网页/手机，用户指定起点和终点，程序就报告这两个点的最优路径。把程序从PC平台搬到Web或Mobile，原来写的所有代码都要扔掉么？还是有很多部分可以重用？怎样重用？

## 3.3 测试

```
subway.exe /z filename
```

filename 指向一个文本文件，里面放了 /a 参数的答案（就是题目 3 的输出，一个数字加上站名）

这个 /z 参数要求在命令行输出：

- true：如果filename 文件中的数据的确覆盖了整个地铁的所有站点至少一次，并且车站的数量是对的，车站的遍历次序是合理的。
- false: 车站的遍历次序仍然合理，但是有遗漏的站点，或者车站的数量错误。如果有遗漏的



站点，这个程序要至少输出一个遗漏的车站名。

- error: 如果车站的遍历次序不合理（例如直接从“知春路”站到了“中关村”站），打印出出错的两个站名。

## 博客要求

发表在你的个人博客上。具体要求如下：

- 1) 在开始实现程序之前，使用下述PSP表格记录下你们**估计**将在程序的各个模块的开发上**耗费的时间**。
- 2) 在你们实现完程序之后，使用下述PSP表格记录下各自在程序的各个模块上**实际花费的时间**。
- 3) 记录你在改进程序性能上花费了多少时间，描述你改进的思路，并展示一张性能分析的图。如果可能，展示你程序中消耗最大的函数。
- 4) 说明你在结对项目中学到了什么。

## 附录

附：PSP 2.1表格

PSP2.1	Personal Software Process Stages	预估耗时 (分钟)	实际耗时 (分钟)
Planning	计划		
· Estimate	· 估计这个任务需要多少时间		
Development	开发		
· Analysis	· 需求分析 (包括学习新技术)		
· Design Spec	· 生成设计文档		
· Design Review	· 设计复审 (和同事审核设计文档)		
· Coding Standard	· 代码规范 (为目前的开发制定合适的规范)		
· Design	· 具体设计		
· Coding	· 具体编码		
· Code Review	· 代码复审		
· Test	· 测试 (自我测试, 修改代码, 提交修改)		

Reporting	报告		
· Test Report	· 测试报告		
· Size Measurement	· 计算工作量		
· Postmortem & Process Improvement Plan	· 事后总结, 并提出过程改进计划		
	合计		