哈尔滨工业大学

**<<数据结构与算法>>**

**实验报告**

**(2023年秋季学期)**

|  |  |
| --- | --- |
| **姓名：** | **阎发祥** |
| **学号：** | **2023140004** |
| **学院：** | **计算机学院** |
| **教师：** | **赵妍妍** |

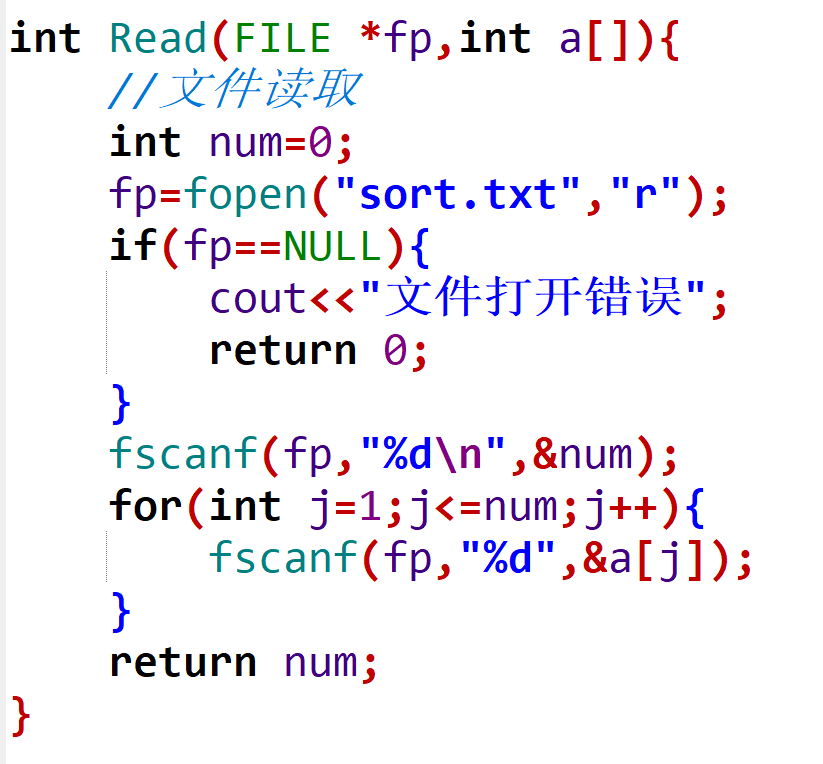
实验四 排序算法及其应用

## 实验内容

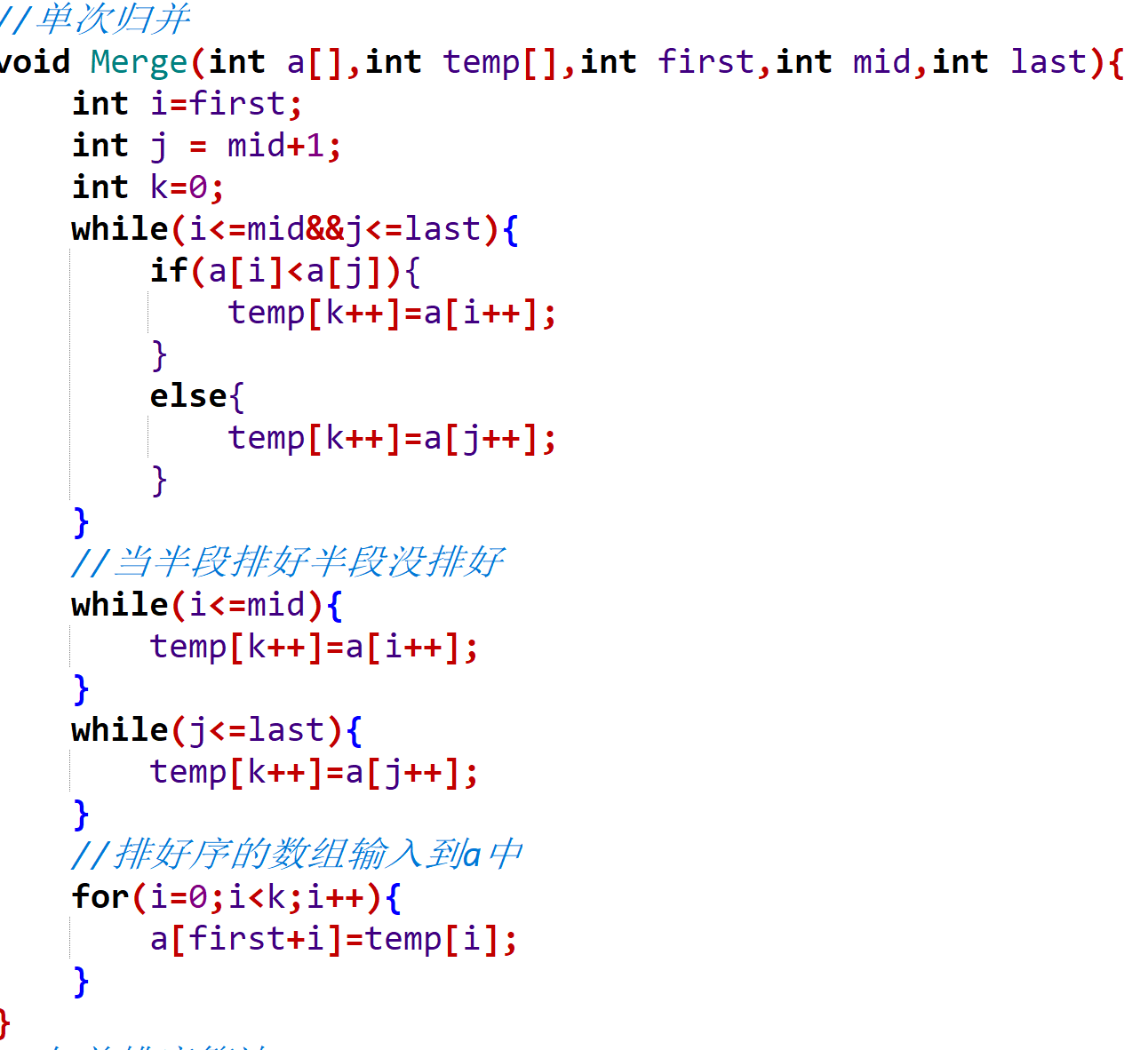
排序是计算机科学中的常见任务，它将一组无序的数据元素按照某种规则重新排列，以使得数据呈现有序的状态，便于后续的查找、统计和分析等操作。当数据量较小时，将数据全部读入内存并进行排序的算法称为内存排序算法，常见的内存排序算法有：插入排序、冒泡排序、归并排序、快速排序、堆排序、基数排序等。本实验要求设计并实现上述内存排序算法并比较其运行速度。

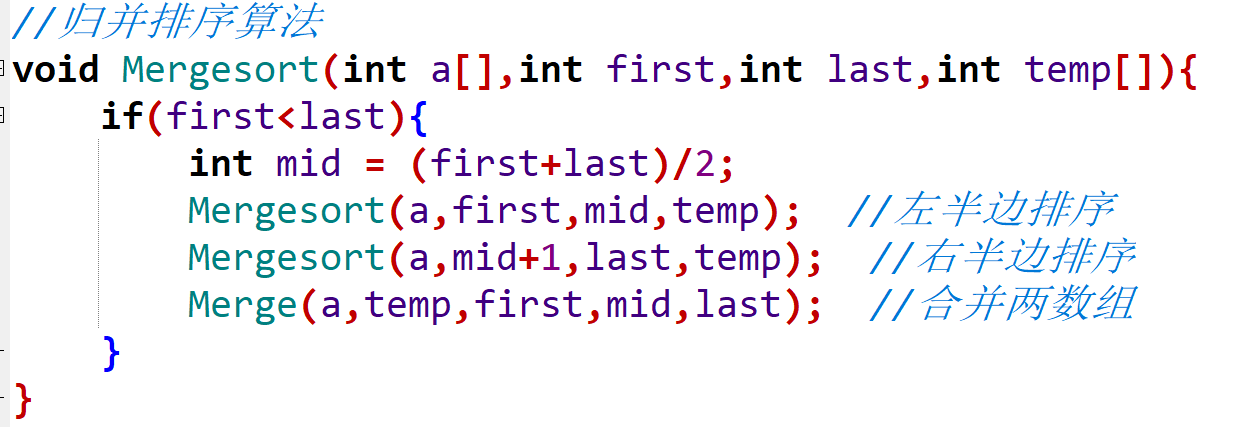
1. 从文本文件中将两行数据读入内存，其中第一行有一个整数 n(n≤100000)，表示待排序序列的长度，第二行有 n 个整数，用空格隔开，表示待排序序列。
2. 实现归并排序、快速排序算法，输出排序好的序列，并记录算法运行时间。
3. 实现选择排序算法或插入排序算法，并将其运行时间与归并排序、快速排序算法比较，随机生成多个适当规模的数据进行实验并绘制折线图，反映不同算法运行时间随着输入规模的变化趋势，并与理论分析结果进行比较。
4. 实验过程及结果

1.按照相应要求，使用fopen函数打开文件，然后使用fscanf函数将文件里的数字写入相应变量，存储到数组a中作为待排序序列。

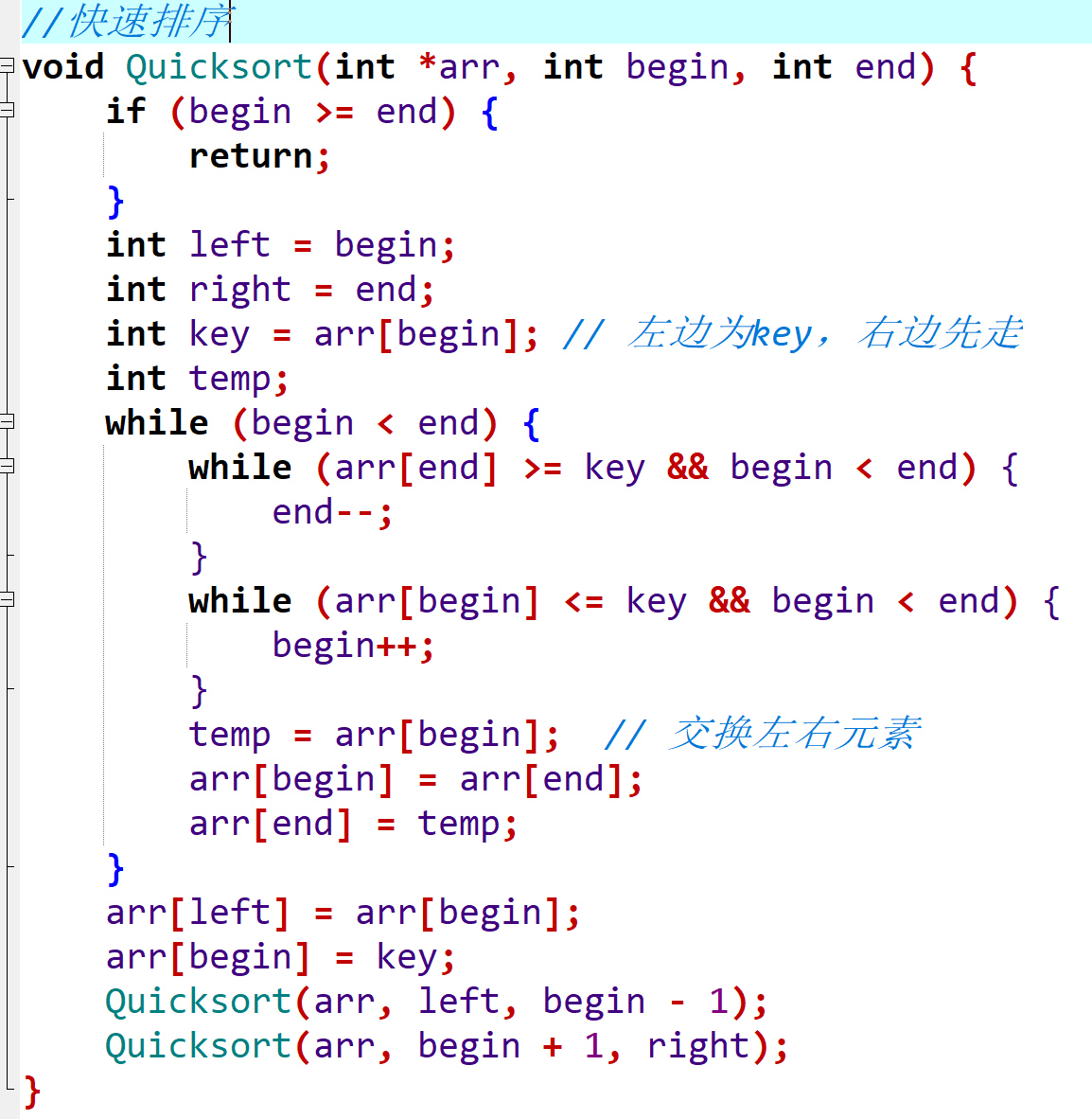


2.首先是归并排序，算法有两个基本的操作，一个是分，也就是把原数组划分成两个子数组的过程。另一个是治，它将两个有序数组合并成一个更大的有序数组。

将待排序的线性表不断地切分成若干个子表，直到每个子表只包含一个元素，这时，可以认为只包含一个元素的子表是有序表。将子表两两合并，每合并一次，就会产生一个新的且更长的有序表，重复这一步骤，直到最后只剩下一个子表，这个子表就是排好序的线性表。



然后是快速排序，思路：选出一个key，选择最左边的元素；定义一个begin和一个end，begin从左向右走，end从右向左走；在走的过程中，若end遇到小于key的数，则停下，begin开始走，直到begin遇到一个大于key的数时，将begin和right的内容交换，end再次开始走，如此进行下去，直到begin和end最终相遇，此时将相遇点的内容与key交换即可；此时key的左边都是小于key的数，key的右边都是大于key的数；将key的左序列和右序列再次进行这种单趟排序，如此反复操作下去，直到左右序列只有一个数据，或是左右序列不存在时，便停止操作，此时此部分已有序



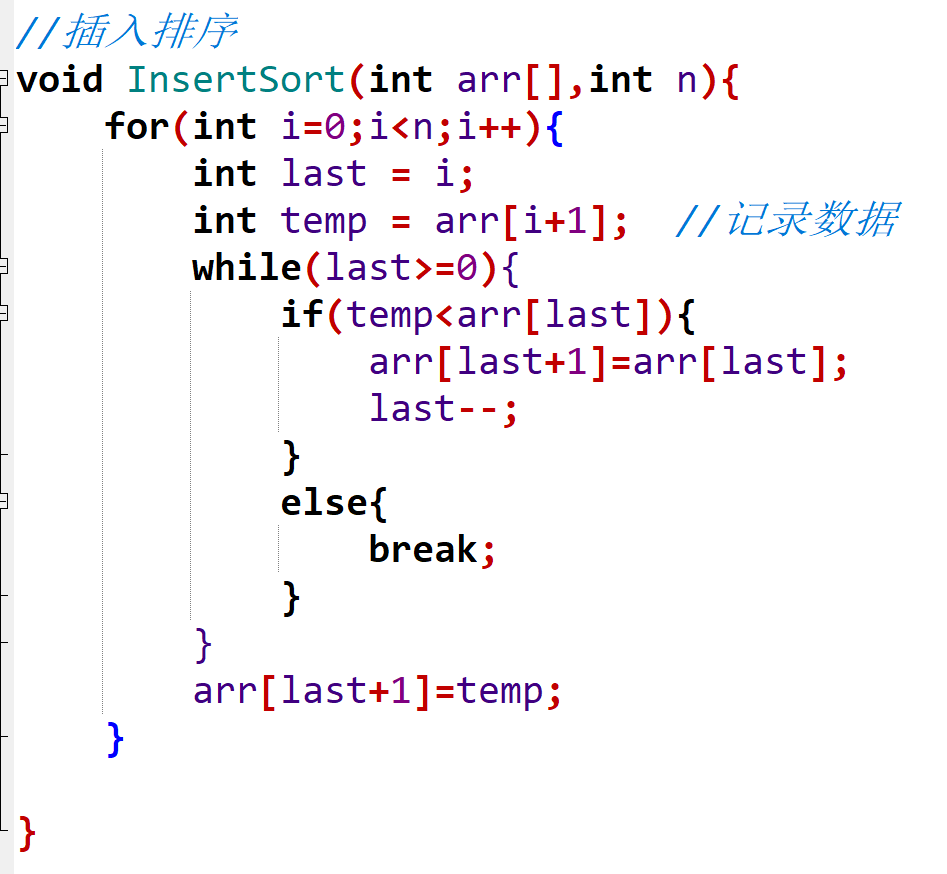
3.我选择的是插入排序，主要思路如下：

从第一个元素开始，该元素可以认为已经被排序

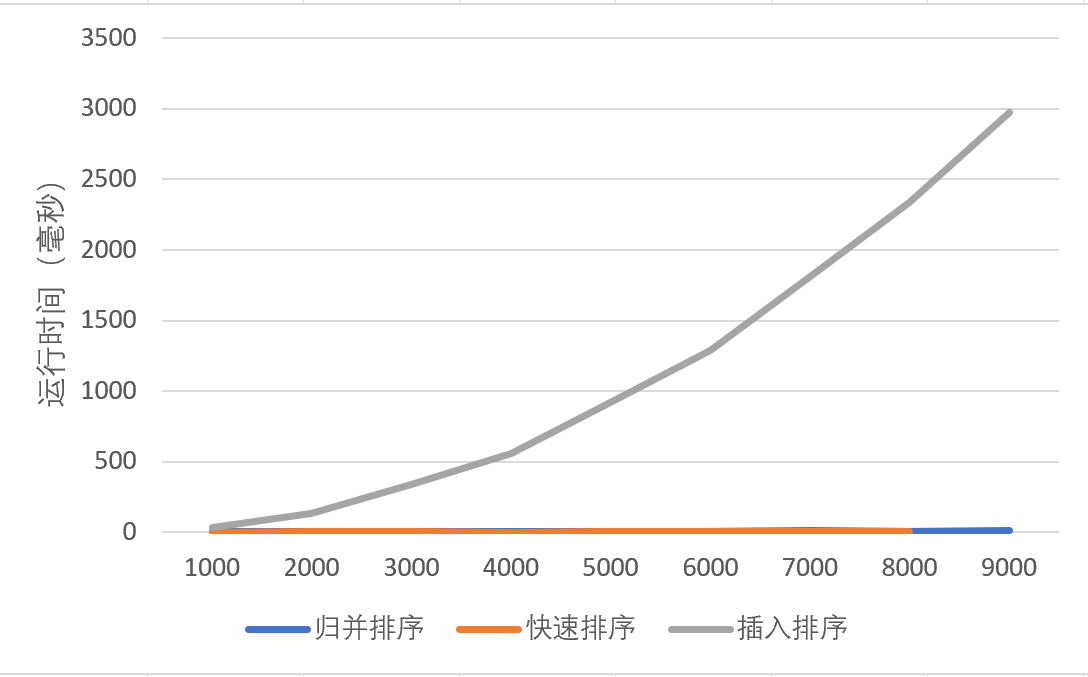
取下一个元素tem，从已排序的元素序列从后往前扫描

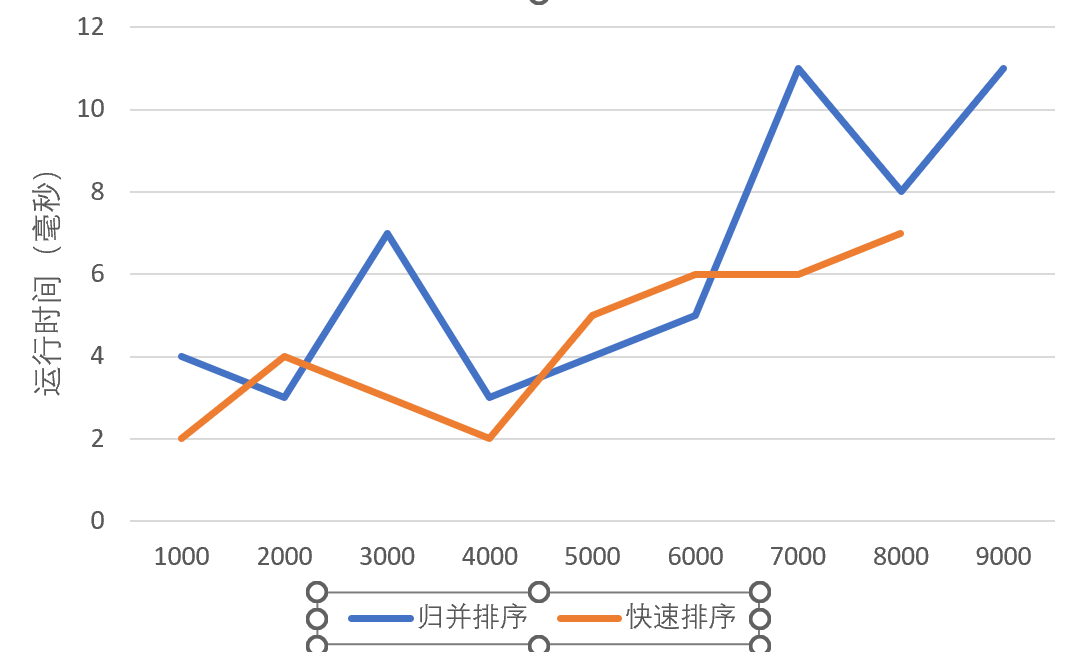
如果该元素大于tem，则将该元素移到下一位

重复步骤3，直到找到已排序元素中小于等于tem的元素

tem插入到该元素的后面，如果已排序所有元素都大于tem，则将tem插入到下标为0的位置，然后重复上述步骤

当数字个数较少时，排序函数的时间过小，导致无法准确测量。为此我们从10000个数字开始计算排序时间，每次增加10000个数字，记录相应的折线图。得到归并排序和快速排序平均时间复杂度是O（nlogn），插入排序时间复杂度是O（n^2）,但是由于数据量较少，排序数字较少精度不足，设备自身的影响，折线图并没有特别好的反映出时间的变化，只可用作建议参考。





三、实验心得

本次实现重点是学会了几个经典的排序操作，并练习了文件的读取操作，主要是加强了基础知识和算法的理解。通过绘制折线图更好地得到运行时间与输入规模的变化，让我也意识到了，时间复杂度是衡量算法好坏的重要评价指标。

四．测试结果

这个是待排序文件

输出结果如下

