

数据库连接、分组、排序算法

赛题方向

基于 C/C++语言运用合理的算法和数据结构知识设计并实现关系型数据库的连接、分组、排序功能。

赛题简介

关系型数据库常见的操作有排序、分组、连接等，例如“SELECT * FROM t1 JOIN t2 group by t1.id1 order by t2.id1”这样的语句。

现有两张数据表 t1(id1 int primary key, id2 int, id3 int)、t2(id1 int primary key, id2 int, id3 int)，其中表 t1 的数据存储在文件 input1.csv 中，表 t2 的数据存储在文件 input2.csv 中。每个文件中有三个整形字段，字段之间用逗号分开，记录条目之间换行。要求针对以上给出的 t1、t2 两个表数据，采用 C/C++开发能模拟数据库连接(JOIN)、分组(GROUP BY)、排序(ORDER BY)功能的算法程序，正确计算出下面语句的结果。

```
select max(t1.id1),min(t2.id1) from t1 join t2 on t1.id3 = t2.id3  
group by t1.id2,t2.id2 order by max(t1.id1),t2.id2,t1.id2;
```

赛题说明

1. 输入文件路径为：/home/web/ztedatabase/input1.csv 和 /home/web/ztedatabase/input2.csv
2. 输出请输出到“标准系统输出”中，程序不要输出结果集以外的任何内容, 输出结果形式同输入一致；

例如：

输入/home/web/ztedatabase/input1.csv

1, 2, 3

2, 2, 3

3, 1, 2

/home/web/ztedatabase/input2.csv

1, 1, 2

2, 2, 2

3, 1, 2

输出结果为

3, 1

3, 2

验收准则：

- (1) 输出结果正确
- (2) 时间越快越优
- (3) 内存控制合理

赛题赛制想法

本题重点考察做题者对关系型数据库连接、分组、排序的理解程度以及编码能力，需要一定的数据库专业知识以及扎实 C/C++开发和运用算法能力。