|  |
| --- |
| 哈尔滨工业大学(深圳) |
| **《数据库》实验报告** |
|  |
| **实验五**  **查询处理算法的模拟实现**  学 院: 计算机科学与技术   |  |  | | --- | --- | | 姓 名: | 孙铎 | | 学 号: | 200110503 | | 专 业: | 计算机科学与技术 | | 日 期: | 2023-01-01 | |

# 实验目的

*阐述本次实验的目的。*

* 理解索引、散列的作用；
* 掌握关系选择、投影、连接、集合的交、并、差等操作的实现算法；
* 加深对算法I/O复杂性的理解
* 理解两阶段多路归并排序算法的思想、实现与应用
* 理解、实现并运用简单的索引

# 实验环境

*阐述本次实验的环境。*

* Windows 10 操作系统
* gcc version 8.1.0 (x86\_64-win32-seh-rev0, Built by MinGW-W64 project)
* CodeBlocks

# 实验内容

*阐述本次实验的具体内容。*

* 基于ExtMem程序库，模拟实现数据库的一些查询处理算法
* 实现关系选择、连接操作算法
* 实现简单的索引，并实现基于索引的关系选择算法
* 实现集合并、交、差操作算法

# 实验过程

*对实验中的5个题目分别进行分析，并对核心代码和算法流程进行讲解，用自然语言描述解决问题的方案 。并给出程序正确运行的结果截图。*

1. **修改extmem.c**

* 修改freeBlockInBuffer函数：freeBlockInBuffer函数原本只将内存块标志位设置为可用，并没有真正清空内存块数据，这里将其修改为，用memset将内存块全部置0并设置标志位为可用。
* 修改writeBlockToDisk函数：将writeBlockToDisk函数中将内存块标志位设置为可用的代码修改为对freeBlockInBuffer函数的调用，这样实现的效果是一致的。

1. **实现基于线性搜索的关系选择算法**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

1. **实现两阶段多路归并排序算法（TPMMS）**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

1. **实现基于索引的关系选择算法**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

1. **实现基于排序的连接操作算法（Sort-Merge-Join）**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

1. **实现基于排序的两趟扫描算法，实现交集合操作算法**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

# 附加题

*对剩余的两种集合操作进行问题分析，并给出程序正确运行的结果截图。*

1. **实现基于排序的两趟扫描算法，实现并集合操作算法**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

1. **实现基于排序的两趟扫描算法，实现差集合操作算法**

问题分析：

* xxxxx
* 核心代码与算法流程：
  1. xxx

实验结果：

# 总结

*总结本次实验的遇到并解决的问题、收获及反思。*

* 之后可以不用Codeblocks吗，Bug太多了
* 4个字节的存储方式，请说清存的是字符而不是二进制数
* 写个OJ可以不
* 明确要求，明确连续和离散，明确内存块利用率