**2020春《数据库系统》作业5**

主讲教师: 邹兆年(znzou@hit.edu.cn)

姓名: \_\_\_\_\_陶飞宇\_\_\_\_ 学号: \_\_\_\_117030720\_\_\_\_\_\_\_\_ 班号: \_\_\_\_1737301\_\_\_\_\_

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 题目 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 总分 |
| 得分 |  |  |  |  |  |  |  |  |  |  |  |

一个页大小应为8kb。

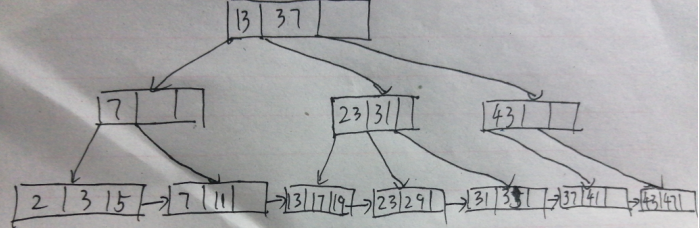
原因：一个t占据的空间：8B+4B+1KB（头部 int char（1024））

可以看到七个t后进行了换页，证明七个t占满了一个页（剩余的空间放置头部和指向每个t元组的指针）。

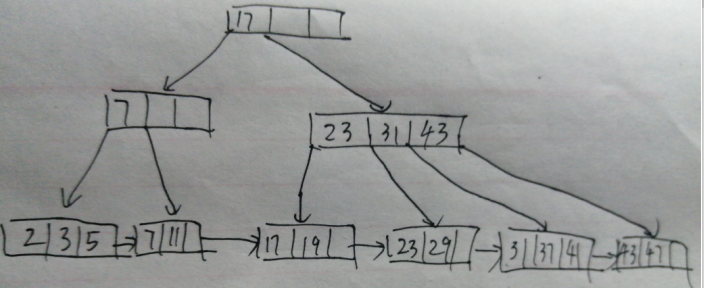
故而一个页大小为8KB

2.

a)



b)



3.

a)根据题目说明，最佳算法为哈希连接，因为在所有缓冲池页数满足的算法中，哈希算法与归并算法I/O开销都为最低。但归并排序中需要对数据进行排序，故而选择哈希算法。

b)开销为3B（R）+3B（S）=45000

c)不需要物化到文件中，产生的结果数量很小，而且很快就要进行下一次连接操作，故而应该进行流水线操作，将这49块直接放入下一次操作的输入缓冲区中

d)应使用一趟连接算法，因为可以将49块全部装入内存中，此时得到代价最低

e)代价为10000+49=10049

f)因为此时无法保证能将全部的结果装入内存中，故而仍应采取哈希算法，在保证算法能够执行的前提下执行

g)代价为b中计算的45000