1.设教学管理数据库有如下3个关系模式：

S(S#, SNAME, AGE, SEX)

C(C#, CNAME, TEACHER)

SC(S#, C#, GRADE)

其中S为学生信息表、SC为选课表、C为课程信息表；S#、C#分别为S、C表的主码，(S#, C#)是SC表的主码，也分别是参照S、C表的外码

用户有一查询语句：

Select SNAME

From S, SC, C

Where SC.S#=S.S# and SC.C#=C.C# and CNAME=“数据库”

检索选学“数据库”课程的学生的姓名。

(1)写出以上SQL语句所对应的关系代数表达式。

(2)画出上述关系代数表达式所对应的查询计划树。使用启发式查询优化算法，对以上查询计划树进行优化，并画出优化后的查询计划树。

(3)设SC表有10000条元组，C表有50条元组，S表中有1000条元组，SC中满足选修数据库课程的元组数为150，计算优化前与优化后的查询计划中每一步所产生的中间结果大小

2.设*ri(X)*与*wi(X)* 表示事务*Ti*读和写数据库元素*X*，则一个并发调度可以抽象为读、写串。基于上述表示，请判断下面两个并发调度是否是可串行化的，为什么？（5分）

调度1：

r1(A); r2(A); w2(A); r1(B); w1(B); r2(B); w1(A); w2(B);

调度2：

r1(B); r2(A); w1(B); r2(B); w2(A); w2(B); r3(A); w3(A);

3.已知关系R(w,x),S(x,y),T(y,z)的块数分别为5000, 10000, 10000。我们准备执行关系代数查询(R ⋈ S) ⋈ T。假设缓冲池中有M = 101个页可用，R,S,T上均无索引且未按连接属性排序。请回答下列问题。

(a) 使用什么算法执行R ⋈ S最适合? 说明理由。

(b) 使用(a)中选择的算法执行R ⋈ S的I/O代价是多少?

# (c) 如果R⋈S的结果不超过49块，那么在使用(a)中选择的算法执行R ⋈ S时，R ⋈ S的结果是否需要物化(materialize)到文件中? 说明理由。

(d) 如果R ⋈ S的结果不超过49块，那么使用什么算法将R ⋈ S的结果与U进行自然连接最合适? 说明理由。

(e) 使用(d)中选择的算法计算连接结果的I/O代价是多少?

(f) 如果R ⋈ S的结果大于49块，那么使用什么算法将R ⋈ S的结果与U进行自然连接最合适? 说明理由。

(g) 使用(f)中选择的算法计算连接结果的I/O代价是多少?

4.设T1、T2、T3是如下三个事务：

T1：A:=A+4

T2：A:=A\*3

T3：A:=A2

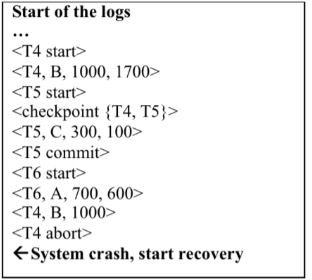
初始A=2

设三个事务都遵守两段锁协议，按 T2-T3-T1的顺序执行，请给出一个不产生死锁的可串行化调度（在下表中填写相应代码即可），并给出最终 A 的结果

|  |  |  |  |
| --- | --- | --- | --- |
| 时间 | T1 | T2 | T3 |
| 1 |  |  |  |
| 2 |  |  |  |
| 3 |  |  |  |
| 4 |  |  |  |
| 5 |  |  |  |

A=

5. 一个带检查点的日志内容如下，结束处发生了故障，请简述恢复算法的过程，并给出Undo-List和Redo-List，以及数据库系统恢复后的A,B,C的值。



6. 有一个带检查点的日志内容如下，请完善日志。

<T0 start> //T0 事务开始

<T0,A,-30,10>

<T0, commit> //T0 事务提交

<T1, start> //T1 事务开始

<T1,B,-10,10>

<T2, start> //T2 事务开始

<T2,A,10,20>

<T2,A,\_\_,-20>

<checkpoint{\_\_,\_\_}> //检查点标记

<T2, commit> //T2 事务提交

<T3, start> //T3 事务开始

<T3,C,10,-20>

<T3,D,-20,30>

结束处<T3,D,-20,30>发生了故障，请问恢复后数据库中 A,B,C,D 的值各为多少？ (注：更新日志记录的格式为< Ti, X, Vold, Vnew>)