Midterm 2 Practice

Due No due date **Points** 30 **Questions** 22 **Available** Mar 11 at 10am - Mar 18 at 3pm 7 days
Time Limit None **Allowed Attempts** 7

Take the Quiz Again

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	1,334 minutes	4.67 out of 30 *

* Some questions not yet graded

Submitted Mar 16 at 12:15pm

Question 1 — 0.33 / 1 pts

Select all of the **true** statements about the query processor/optimizer.

- Correct! ☑ A SQL query is converted from a string to a parse tree by the parser component.
- ☐ A logical query tree has relational algebra operators and has selected algorithms to use to implement those operations.
- Correct! ☑ The parser will generate a syntax error if the SQL query is not valid according to the SQL grammar.
- You Answered ☑ The preprocessor validates table and field names and runs before the parser builds the parse tree.
- Correct Answer ☐ A rule in a grammar may be applied more than once when parsing a query.

Question 2 — 1 / 1 pts

Select all of the **true** statements about heuristic optimization.

- ☐ Heuristic optimization rules always improve query plan performance.
- Correct! ☑ Joins are commutative and associative.
- Correct! ☑ Selection and cross-product operators can be combined into a join.
- Correct! ☑ Conjunctive selections can be separated into a sequence of selection operations.
- Correct! ☑ It is better to perform selections closer to the bottom (leaf nodes) of the query tree.

Question 3

0 / 1 pts

Select the **true** statement about optimization.

- A canonical logical query tree shows the join order in the query plan.
- **Correct Answer** Cost-based optimization uses data statistics to estimate operator costs.
- **You Answered** Heuristic and cost-based optimization are not used together.
- The selectivity is the fraction of tuples eliminated by a selection operator.
- The size of a join result is never larger than the size of the largest relation being joined.

Question 4

0 / 1 pts

Select the **true** statement about transactions.

- A transaction is an atomic program that always executes completely.
- **You Answered** A consistent database always accurately reflects real-world reality.
- A transaction can be in both aborted and committed states during its lifetime.
- A schedule may interleave (reorder) operations within a transaction.
- **Correct Answer** A schedule can be used to detect concurrency issues but does not prevent them.

Question 5

0 / 1 pts

Select the **true** statement about serializability and schedule properties.

- **Correct Answer** Two operations conflict if they use the same data item and one of them is a write.
- A schedule is conflict serializable if its precedence graph contains a cycle.
- A conflict serializable schedule is made equivalent to a serial schedule by swapping conflicting operations.

A recoverable schedule is always cascade-free.

You Answered

A strict schedule guarantees that no committed data item is overwritten.

Use the SQL statement below to answer the following questions.

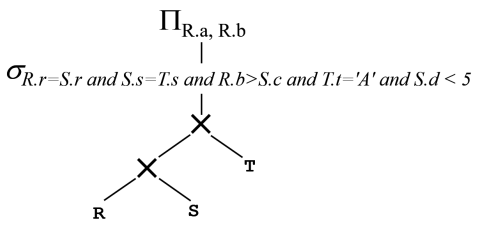
```
SELECT R.a, R.b
FROM R, S, T
WHERE R.r = S.r and S.s = T.s and R.b > S.c and T.t = 'A' and S.d < 5
```

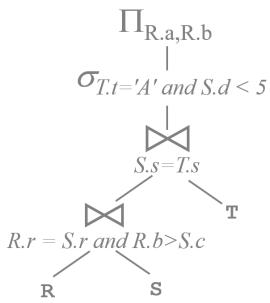
Question 6

0 / 2 pts

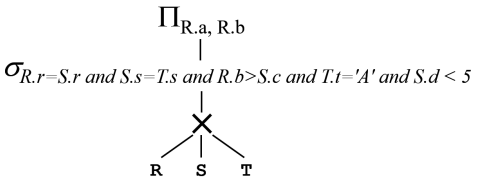
Select a canonical, logical query tree for the given SQL statement. Do NOT apply any optimization.

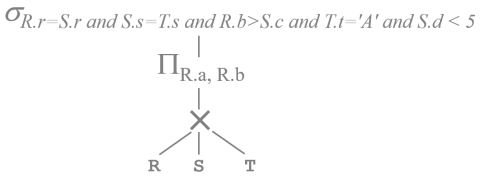
You Answered

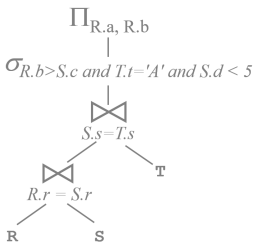


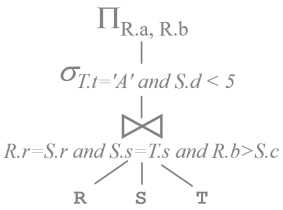


Correct Answer





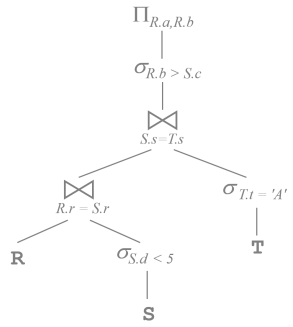


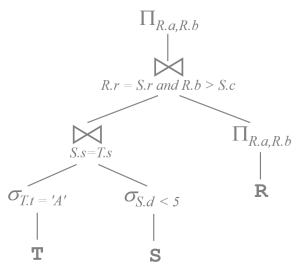


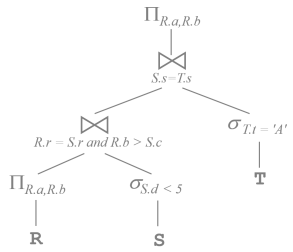
Question 7

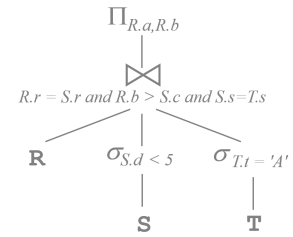
0 / 3 pts

Select the best, optimized logical query tree.

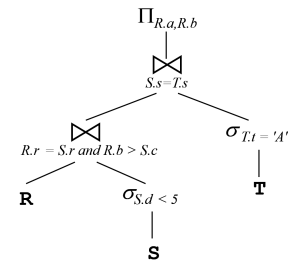




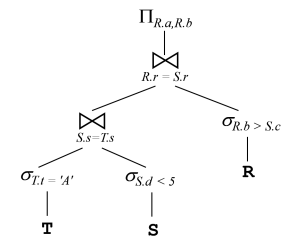




Correct Answer



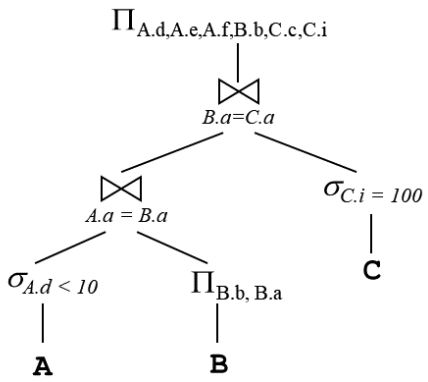
You Answered



Given the following statistics and logical query tree, perform the requested calculations and physical optimization. Assume no spanning is used. **Answer the questions by giving the number of tuples and the number of blocks.**

Statistics:

```
A(a,d,e,f,g) - T(A)=30000   V(A,a)=10000  V(A,d)=1000
B(b,a,h)     - T(B)=100000  V(B,a)=50000
C(c,a,i)     - T(C)=50000   V(C,a)=5000   V(C,i)=250
Block size is 1000 bytes with no spanning.
Attribute sizes: a=b=c=i=4 bytes ; h=1000 bytes ; d=6 bytes ; e=50 bytes ; f=g=20 bytes
There are indexes on A.a, A.d, B.a, C.a, and C.i.
Maximum memory size per operator is 100 blocks.
```



Unanswered

Question 8

0 / 1 pts

Calculate size of: $\sigma_{A.d < 10}$ **(A)**

In tuples:

In blocks:

Answer 1:

You Answered: (You left this blank)

Correct Answer: 10000

Answer 2:

You Answered: (You left this blank)

Correct Answer: 1000

Unanswered

Question 9

0 / 1 pts

Calculate size of: $\sigma_{C.i = 100}$ ***(C)***

In tuples:

In blocks:

Answer 1:

You Answered: (You left this blank)

Correct Answer 200

Answer 2:

You Answered (You left this blank)

Correct Answer 3

Unanswered

Question 10

0 / 1 pts

Calculate size of: $\Pi_{B.b,\ B.a}$ **(B)**

In tuples:

In blocks:

Answer 1:

You Answered (You left this blank)

Correct Answer 100000

Answer 2:

You Answered (You left this blank)

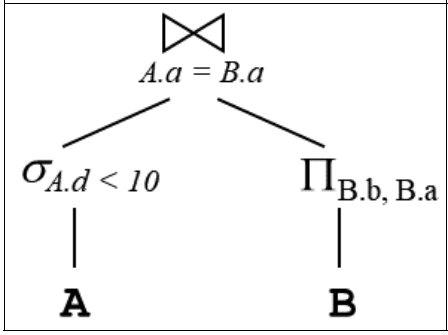
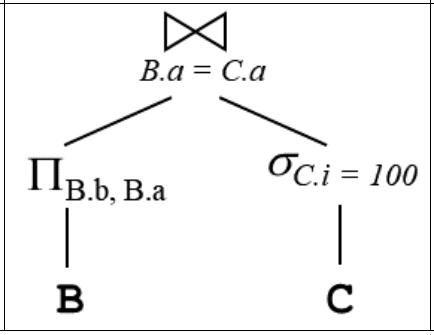
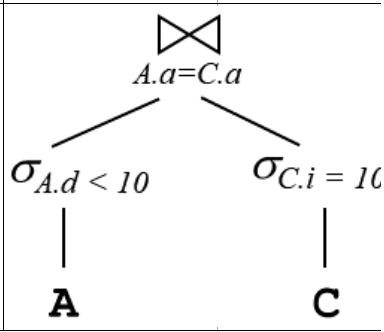
Correct Answer 800

Unanswered

Question 11

0 / 3 pts

Calculate the following three join sizes in tuples and blocks:

1	2	3
		

1) **Join of A and B**

In tuples:

In blocks:

2) **Join of B and C**

In tuples:

In blocks:

3) **Join of A and C**

In tuples:

In blocks:

Answer 1:

You Answered (You left this blank)

Correct Answer 20000

Answer 2:

You Answered (You left this blank)

Correct Answer 2222

Answer 3:

You Answered (You left this blank)

Correct Answer 400

Answer 4:

You Answered (You left this blank)

Correct Answer 8

Correct Answer 7

Answer 5:

You Answered (You left this blank)

Correct Answer 200

Answer 6:

You Answered (You left this blank)

Correct Answer 25

Correct Answer 23

Question 12

1 / 1 pts

Apply cost-based optimization to answer the next few questions.

What is the best join order?

- A ⋈ B ⋈ C
- Correct! A ⋈ (B ⋈ C)
- A ⋈ C ⋈ B
- B ⋈ (A ⋈ C)
- C ⋈ A ⋈ B

Question 13

0 / 1 pts

What join algorithms should be used?

- first join is hash join, second join is hash join
- You Answered first join is hash join, second join is one-pass join
- Correct Answer first join is one-pass join, second join is one-pass join
- first join is sort join, second join is sort join
- first join is one-pass join, second join is hash join

Question 14

0 / 1 pts

What is the best way to implement $\sigma_{C.i = 100}(C)$?

- table scan
- Correct Answer index scan

You Answered

- (●) one-pass join
- ( ) operator is eliminated in physical tree

Question 15

0 / 1 pts

Consider the optimized, physical query tree. Select all of the **true** statements.

You Answered

- [x] The operator ΠB.b, B.a (B) was eliminated after cost-based optimization as it is not improve cost.

You Answered

- [x] There is a selection operator that is above a join.
- [ ] The selection σA.d < 10 (A) should use an index scan.

Correct Answer

- [ ] The join order is different than the original query plan (given at start of question just above Q8).

Correct Answer

- [ ] A projection operator will not use an index scan.

Unanswered

Question 16

0 / 5 pts

Determine if the following schedule is conflict serializable by drawing the precedence graph.

```
r1(A);r2(A);w2(A);r3(B);w3(B);r4(C);w4(A);w2(B);r1(C);w1(D);w2(D);r4(A);w3(A)
```

Answer the questions about your precedence graph.

nodes:

edges with label A:

edges with label B:

edges with label C:

edges with label D:

Is the schedule serializable (y/n)?

Enter # of nodes in a cycle (put 0 if no cycle):

If a cycle exists, is T2 in the cycle (y/n)?

Is there an edge T1-> T4 with label A (y/n)?

Is there an edge T3-> T1 with label A (y/n)?

Answer 1:

You Answered (You left this blank)

Correct Answer 4

Answer 2:

You Answered (You left this blank)

Correct Answer 6

Answer 3:

You Answered (You left this blank)

Correct Answer 1

Answer 4:

You Answered (You left this blank)

Correct Answer 0

Answer 5:

You Answered (You left this blank)

Correct Answer 1

Answer 6:

You Answered (You left this blank)

Correct Answer n

Correct Answer no

Answer 7:

You Answered (You left this blank)

Correct Answer 2

Answer 8:

You Answered (You left this blank)

Correct Answer y

Correct Answer yes

Answer 9:

You Answered (You left this blank)

Correct Answer y

Correct Answer yes

Answer 10:

You Answered (You left this blank)

Correct Answer n

Correct Answer no

Answer the following questions on transaction schedules and schedule properties given these transactions:

```
T1: r1(A);r1(B);w1(A);c1;
T2: r2(B);r2(A);w2(A);w2(B);c2;
T3: r3(C);w3(B);c3;
```

Question 17

0 / 1 pts

TRUE or FALSE: This schedule is conflict-serializable and recoverable:

```
r3(C);w3(B);r2(B);r2(A);w2(A);c3;r1(A);w2(B);r1(B);w1(A);c1;c2;
```

You Answered ◉ True

Correct Answer ○ False

Question 18

1 / 1 pts

TRUE or FALSE: This schedule is cascade-free but not strict:

```
r2(B);r3(C);r2(A);w2(A);w2(B);w3(B);c2;r1(A);c3;r1(B);w1(A);c1;
```

Correct! ◉ True

False

(cascade free as always reads committed values, not strict as w3(B) overwrites w2(B))

Question 19

1 / 1 pts

TRUE or FALSE: It is possible to create a schedule that is view serializable but not conflict serializable from these transactions.

Correct!

True

False

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

(Not conflict serializable as cycle between T1 and T3 on B, view serializable to T1, T3, T2)

Question 20

0 / 1 pts

Select **ALL** of the properties of the following schedule:

```
r3(C);w3(B);r2(B);r2(A);w2(A);c3;r1(A);w2(B);r1(B);w1(A);c2;c1;
```

You Answered

serial

Correct!

conflict serializable

Correct!

recoverable

You Answered

cascade-free

strict

Question 21

0.33 / 1 pts

Select **ALL** of the properties of the following schedule:

```
r1(A);r1(B);r3(C);w3(B);r2(B);w1(B);c1;r2(A);c3;w2(A);w2(B);c2;
```

Correct! ☑ non-serial

You Answered ☑ conflict serializable

Correct! ☑ view serializable

Correct Answer ☐ recoverable

☐ cascade-free

Unanswered

Question 22

Not yet graded / 0 pts

NOTE: This question shows as 0 marks but is worth 3 bonus marks.

Using the GitHub repository: **https://classroom.github.com/a/CbuniShh (https://classroom.github.com/a/CbuniShh)**

Modify the file `src/textdb/question/PracticeQuestion.java`. Test with `src/textdb/junit/TestPracticeQueryTree.java`.

Write an optimized query plan for the SQL query:

```
SELECT n_name, c_custkey, c_name, o_orderkey, o_totalprice
FROM  Nation AS N, Customer AS C, Orders AS O
WHERE N.n_nationkey = C.c_nationkey AND O.o_custkey = C.c_custkey
      AND n_name = 'CANADA' AND o_totalprice > 410000
```

Assumptions:

- T(Nation) = 25, T(Customer) = 15000, T(Orders) = 150000
- There is index on c_nationkey. No index on o_totalprice.
- Assume a selection above a TextFileScan will use an index if possible.
- Size of SELECT * FROM Orders WHERE o_totalprice > 410000 is 70.
- Only using BlockNestedLoopJoin. Smallest table is left input (input #0).

You will get 1 bonus mark for passing the first test that produces a CORRECT plan. You will get 2 more bonus marks for passing the second test that checks for the optimal query plan.

When done, commit/push your repository. Submit the URL of the repository on Canvas.

Your Answer: