Problem 1 (10 points)

Examine the schedule given below. There are three transactions, T1, T2, and T3. Initially, the salary = 1 and the tax = 2. The assignments happen within the local memory space of the transactions and the effects of these assignments are not reflected in the database until the WRITE operation.

```
     T1                      T2                  T3
-------------------------------------------------------------------
0                                                 start
1                                                 READ tax
2                                                 tax := tax + 1
3     start
4     READ salary
5     salary := salary + 1
6                                                 WRITE tax
7                                                 commit
8                             start
9                             READ tax
10                            READ salary
11                            tax := tax + salary
12                            WRITE tax
13                            commit
14   ------------------- checkpoint start -------------------------
15   READ tax
16   tax := tax + salary
17   WRITE salary
18   ------------------- checkpoint end -------------------------
19   commit
```

a) Show the undo/redo log file entries that would be generated by this execution. Use the same notation used in class. For each log entry, indicate what line above generates it.

b) Assume that the undo/redo recovery algorithm with checkpoints is being used. The database crashes immediately after statement 7. (Assume that all the log records up to this point are on disk.)

b.1) Which transactions would have to be undone?

b.2) Which transactions would have to be redone?

c) Again assume that the undo/redo recovery algorithm with checkpoints is being used, but now the databases crashes just after statement 18. (Assume that all the log records up to this point are on disk.)

c.1) Which transactions would have to be undone?

c.2) Which transactions would have to be redone?

Problem 2 (10 points)

Examine the schedule given below. There are four transactions, T1, T2, T3, and T4.

```
          T1              T2           T3              T4
-------------------------------------------------------------------
1                                                     READ tax
2     READ salary
3                                                     WRITE tax
4                     READ tax
5                     WRITE tax
6     READ tax
7     WRITE salary
8                                    READ salary
9     WRITE tax
10                                   WRITE salary
11                                                    READ salary
12                                                    WRITE salary
```

a) Draw the precedence graph for this schedule.

b) What is the equivalent serialization order for this schedule? If no order is possible, then state 'none'.

c) Assume that transaction T4 did not run at all. What is the precedence graph in this case?

d) What is the equivalent serialization order for this second schedule? If no order is possible, then state 'none'.

Problem 3 (15 points)

Consider a four-dimensional data cube S(W, X, Y, Z), where the following sum aggregates have been materialized:

Materialized: S(W, X, Y, Z)
S(*, X, Y, Z)
S(W, X, *, *)
S(W, *, Y, *)

For each of the following queries, give the materialized aggregates (or the cube itself) from which the query can most *efficiently* be computed. If the result can be computed from two aggregates S_1 and S_2, but the computation from S_1 is clearly less expensive than the computation from S_2, then only give S_1 in your answer. If with the information given here, you cannot infer that S_1 is better than S_2 or vice versa, then give both S_1 and S_2 in your answer.

Note that lower case variables represent particular values of the attribute (e.g., x is a value for X).

a) S(*, x, y, z)

b) S(w, *, *, *)

c) S(*, *, *, *)

d) S(w, *, y, z)