

# CS143: Database Systems

## Homework #8

1. Consider the relation `Employee(name, salary)` where `name` is the key. The following three transactions are being executed:

$T_1$ :

```
SELECT SUM(salary) FROM Employee;
COMMIT;
```

$T_2$ :

```
UPDATE Employee SET salary = salary + 200;
UPDATE Employee SET salary = salary + 1000 WHERE name = 'Tony';
COMMIT;
```

$T_3$ :

```
UPDATE Employee SET salary = salary + 100 WHERE name = 'James';
UPDATE Employee SET salary = salary + 200 WHERE name = 'Tony';
COMMIT;
```

The table `Employee` originally has two tuples, ('Tony', 1000) and ('James', 1000). Please assume that individual SQL statements are executed atomically.

- (a) Assume that all three transactions run under the isolation level `SERIALIZABLE`. List all possible values that can be returned by  $T_1$ . Briefly explain your answer.

**ANSWER:**

2000, 2300, 3400, 3700.

When all transactions run under `SERIALIZABLE`, any possible schedule is conflict equivalent to a serial schedule. Possible schedules for  $T_1, T_2$  and  $T_3$  are:  $T_1T_2T_3$ ,  $T_1T_3T_2$ ,  $T_2T_1T_3$ ,  $T_2T_3T_1$ ,  $T_3T_1T_2$ , and  $T_3T_2T_1$ . The outputs from  $T_1$  are 2000, 2000, 3400, 3700, 2300, 3700, respectively.

- (b) Assume that  $T_1$  runs under the isolation level `READ UNCOMMITTED` and  $T_2$  under `REPEATABLE READ` and  $T_3$  under `SERIALIZABLE`. List all possible values that can be returned by  $T_1$ . Briefly explain your answer.

**ANSWER:**

2000, 2100, 2300, 2400, 2700, 3400, 3500, 3700.

Only  $T_2$  and  $T_3$  are updating values, so let us focus on these two transactions first. Under `REPEATABLE READ`, the only exception to ACID is phantom, but because  $T_2$  and  $T_3$  do not insert any tuple, we do not need to worry about ACID exceptions for the two. So the possible schedules for  $T_2$  and  $T_3$  are equivalent to  $T_2T_3$  or  $T_3T_2$ . Regarding  $T_1$ , since it is `READ UNCOMMITTED`, its `SELECT` statement may do a dirty read.

Now let us consider the schedule  $T_2T_3$ . Under this schedule, the total salary value changes from  $2000 \rightarrow 2400 \rightarrow 3400 \rightarrow 3500 \rightarrow 3700$ .  $T_1$  may read any of these salary sum values.

For the schedule  $T_3T_2$ , the total salary value changes from  $2000 \rightarrow 2100 \rightarrow 2300 \rightarrow 2700 \rightarrow 3700$ . Again,  $T_1$  may read any of these salary sum values.

Therefore, possible outputs from  $T_1$  are 2000, 2100, 2300, 2400, 2700, 3400, 3500, 3700.