CS143–Winter 2020—FINAL EXAM

Your Name Your Student ID PageNo:

(Write your ID and the page number on the top of every page of the final.)

On the first page you must write & sign the following Academic Honesty statement:
I thereby certify with my signature that I completed this exam entirely on my own, without reference to any prohibited sources or materials, nor communications with anyone.

Signed: Junhong Wang

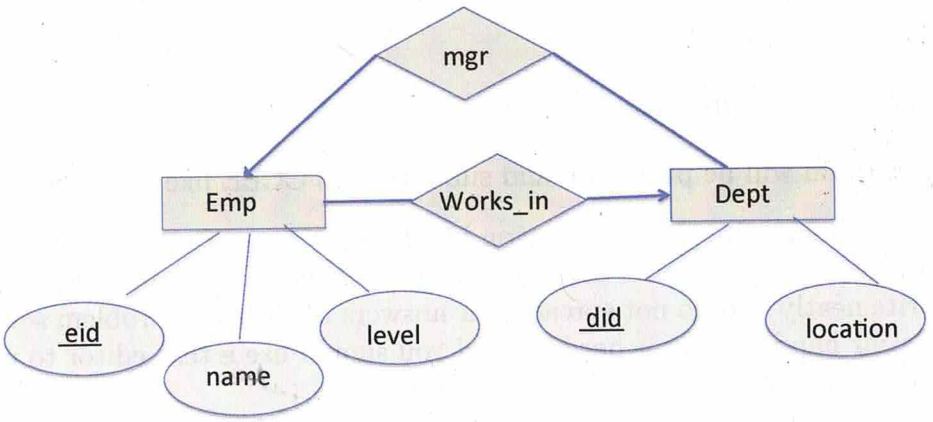
Total number of pages in your submitted final: 4

General Instructions

- The exam will be posted on and submitted to CCLE like previous assignments.
- You should write your answers on blank paper pages (letter-size recommended) and scan/take a picture of them with your device (e.g. your cell-phone). Please write neatly and do not spread your answers to the same problem several pages. If your hand-writing is hard to read you should use a text editor to prepare the PDF document that you'll submit to CCLE.
- For submissions, there will be a grace period of 30 Minutes (beyond the regular three hours).
- Make sure that you submission contains cover page with your Name, ID, total number of pages, and the signed statement certifying your academic honesty. Also write your ID and the page number on top of every page.
- Your exam will be proctored via Zoom. You will need to keep plenty of social distancing around you and only the exam pages and the cheat-sheet on your desk (one double-sided cheat sheet allowed).
- You must have the camera and the video enabled. Make sure that you have the bandwidth required to support this. Your mike should on mute, but you can unmute and use it to ask special question of questions of general interest
- For questions, Piazza will be on, but Chat is much better. Also, you can use chat to ask for the cell-phone number of a TA: if everything else fails, my office number is 1-310-454-7120.

Problem A. 20 points, 5 points each question. Given the ER-diagram

Problem A: Question A1

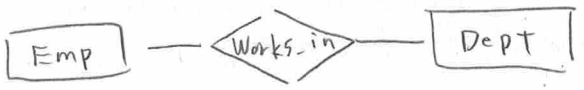


A1. Generate a BCNF schema with minimal relation count and show the relations with all their keys underlined (as in `Mytable(A1, A2, A3, ...)`.

manager's eid

Emp (eid , did , name , level) Dept (did , eid , location)

A2. Describe in words how the E-R changes if an employee can work in several departments.



A3. Generate a BCNF schema with minimal relation count and keys underlined for the ER diagram modified as described in Q2.

Emp (eid , name , level)

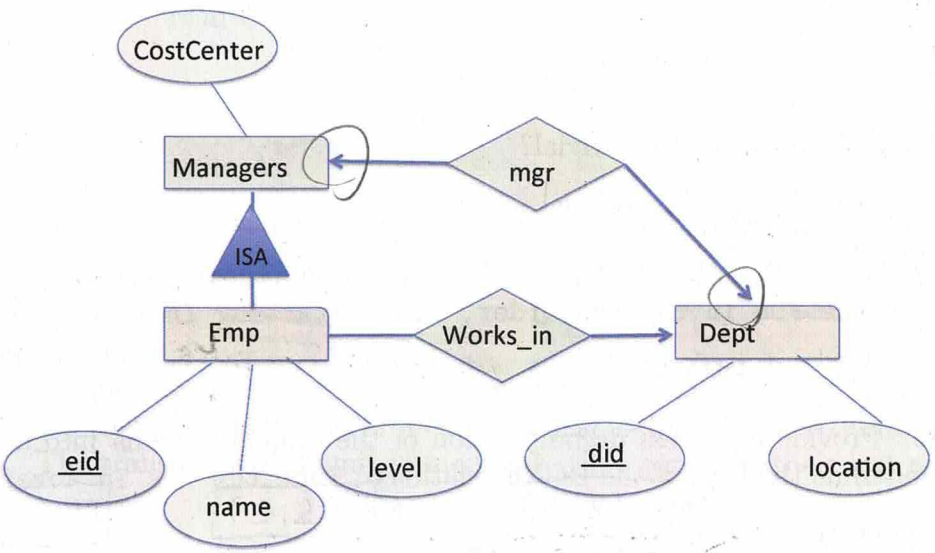
work (did , eid)

Dept (did , eid location)

↑ manager's eid

A4. Consider now the ER-diagram below where Managers have an additional attribute (say `CostCenter`) which regular employees do not have (however managers are still considered to work in the department they manage).

Problem A: Question A4



From this new ER diagram, generate a BCNF schema with minimal relation count and keys underlined

Emp (eid , did, name, level)

Dept (did , eid , location)

↑

manager's eid

Managers (eid , did, CostCenter)

↑

did it manages

Problem B. 20 Points: 5 per question.

Consider the relational scheme $R(A, B, C, D, E)$ with the following FDs:

1. $AB \rightarrow C$
2. $C \rightarrow B$
3. $C \rightarrow D$
4. $C \rightarrow E$
5. $ED \rightarrow E$
6. $B \rightarrow D$

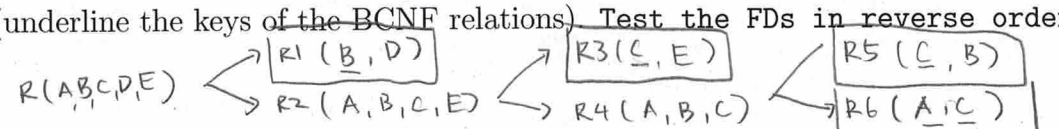
B1. Is any of these FDs trivial?

5 is trivial

B2. Is $R(A, B, C, D, E)$ BCNF? (You must explain the reasons for your answer.Please test the FDs in reverse order, i.e. starting from FD 6 and moving up .

$\{B\}^+ = \{BD\}$ B is not key, so R is not BCNF.

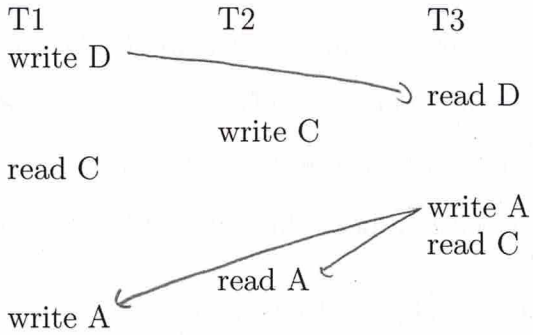
B3. Provide a lossless decomposition of the original schema into a BCNF schema (underline the keys of the BCNF relations) Test the FDs in reverse order!



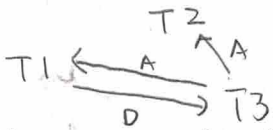
B4. Is the BCNF decomposition you just produced FD-preserving? (Explain the reasons for your answer.)

AB → C is lost so this is not FD-preserving

Problem C 20 Points: 5 poins per question. T1 is the oldest transaction, and T3 is the youngest.



C1. Is this first schedule conflict-serializable? (Justify your answer using the applicable graph.)



NO, because there is a cycle between T1 and T3 in the dependency graph

C2

In the next two questions, assume that the transaction manager uses a **strict 2PL protocol** where each shared/exclusive lock is set just before the read/write action requiring it. The transactions commit immediately they complete their last read or write.

C2. If we do not use any deadlock prevention strategy, will the resulting transactions (i) complete, or (ii) deadlock? If your answer is (i), show a completed schedule; if your answer is (ii), then show the schedule up to the deadlock.
Also state clearly and concisely what happen to each transaction: e.g., abort, restart, wait-for resource, kill other transactions, etc.

X(R): exclusive lock on R
S(R): shared lock on R
R(R): read R
W(R): write R

← (i) complete, see left

C3

C3. If we use a wound-wait deadlock prevention strategy will the resulting transactions (i) complete, or (ii) deadlock? If your answer is (i), show a completed schedule (there could be more than one–any correct one will do); if your answer is (ii), then show the schedule up to the deadlock.

(i) complete, see right →

C4. Now assume that we are using a timestamp-based concurrency control method, with Thomas' write rule. Which transactions will abort (and why) and which will complete?

(TS1 = 1) (TS2 = 2) (TS3 = 3)

T1: W(D)

T3: R(D)

T2: W(C)

Abort because TS-W(C) = 2

T3: W(A), R(C), T3 completes

Abort because TS-W(A) = 3

	A	C	D
R		3	3
W	3	2	1

T1, T2 abort.
T3 complete.
See schedule on the left

Problem D: 20 points: 2 per question. Answer the questions by a list such as D1:Yes|No, D2:Yes|No, ..., D10:Yes|No. An explanation is not required.

In general, what are the reasons for normalizing relational schemas?

D1 Removing update anomalies. YES

D2 facilitating maintenance of integrity constraints: YES

D3 Making query execution more efficient: NO

What does a conservative 2PL protocol assures?

D4 Absence of dirty reads? YES

D5 Absence of deadlocks? NO

In Transaction Recovery after a crash:

D6 The Deferred Update log protocol eliminates the need for any REDO. YES

D7 The execution of a CHECKPOINT forces all the running transactions to commit or abort. NO

Our Database Transactions use a Time-Stamp based protocol:

D8 Our transactions use a Time-Stamp based protocol. Can these transactions incur a deadlock situation? NO

D9 Do rolled-back transactions keep their original timestamp? YES

D10 Does this protocol guarantee cascadeless schedules? YES

Problem E: 20 points: 5 points per question.

Recovering from a crash using the immediate database modification protocol, the database system finds the following log (stored in three pages):

```
T0 start
T0 C, 900, 700
T1 start
Checkpoint L1
-------------log page boundary
T2 start
T1 C, 700, 955
T1 B, 400, 800
T1 commit
T3 start
T0 Commit
Checkpoint L2
T2 A, 160, 200
-------------log page boundary
T3 B, 880, 600
T4 start
T4 C, 955, 2400
T5 start
T4 commit
T5 D, 2100, 4000
```

E1. For each checkpoint the recovery system maintains a list of transactions that satisfy a certain property What property is that, and which are the the transactions in the lists L1 and L2 shown next to the two `Checkpoint` in the above schedule?

It checks for active transaction.

L1 = T0, T1; L2 = T2, T3

E2. How many pages of the log will the system read for recovery (explain)?

It checks last check point which is T2, T3. we need to undo T2, T3, T5 and redo T4. Thus 2 pages. because we need find start of these transactions.

E3. Which transactions will be undone and which will be redone?

undo : T2, T3, T5 redo : T4

E4. What are the values of `A`,`B` and `C` be at the end of a successful recovery (explain the reason for your statement)?

A : 160 (T2 is undone)

B : 800 (T1 is completed but T3 is undone)

C : 2400 (T4 is redone)