Introduction to Database design

Dennis Thinh Tan Nguyen

January 5, 2016

1 ER-Diagrams

1.1

```
CREATE TABLE User(
userID int PRIMARY KEY,
name varchar(20),
password varchar(30)
);
```

Figure 1: User DDL

1.2 Follows DDL

```
CREATE TABLE Follows(
userToFollowId int,
followerId int,
PRIMARY KEY (userToFollowId,followerId),
FOREIGN KEY (userToFollowId)  REFERENCES User(userID),
FOREIGN KEY (followerId)  REFERENCES User(userID)
);
```

Figure 2: Follows DDL

1.3 Tweet DDL

```
CREATE TABLE Tweet(
tweetID int,
text varchar(140),
userNumber int,
dateAdded DATE,
PRIMARY KEY (tweetID),
FOREIGN KEY (userNumber) REFERENCES User(userID),
);
```

Figure 3: Tweet DDL

1.4 Explaination of Compose relation

By the assumption that the given E/R diagram is correct, the relation *compose* is not to be created as a table, since the cardinality between User and Tweet is a one-to-many relationship. One can only create the *compose* table if the relation itself contained any attributes or if the cardinality between User and Tweet is a many-to-many relationship. Hence, no table is created for the table *compose*.

2 Relational Algebra

2.1

The answear is: 3 - It lists all tuples in the UserProfile table with a reputation greater than or equal to 200.

2.2 List all where reputation greater than 75 expressions

Below is a list of valid expressions that returns titles of questions asked by top users, (i.e., users with a reputation greater than 75.)

- 1
- 3

2.3 Evaluation of vote count greater than 25 expression

Answer for 1: - Since the result set is a joined table of the *Answer* table and *Question* table we can list the following attriibutes:

- UserID
- name
- reputation
- questionID
- title
- voteCount

Answer for 2 - The tuples returned are tubles where the attribute *vote-count* in the *Answer* table is greater than 25. 5 tuples are returned in this case.

userID	name	reputation	questionID	text	Vote_count
U3	Jesper	44	Q0	[...]	44
U2	Emma	131	Q1	[...]	131
U4	Louise	200	Q2	[...]	118
U5	Michael	75	Q3	[...]	45
U1	Jesper	44	Q3	[...]	87

Figure 4: Result set of expression. (Text has been omitted)

2.4 Evaluation of division expression

Answer for 1: - The attributes in the result set for the given expression are as following

- userID

Answer for 2 - Two tuples are returned that is returned is

userID	questionID
U5	Q2
U3	Q2

Figure 5: Resultset of the expression.

2.5 Query for super users

```
SELECT userID, MAX(vote_count)
FROM Answer,
Group BY (userID);
```

Figure 6: Query to retrieve super users

3 Functional Dependencies

3.1 Canoncial cover

The answer is 3

3.2 Deducing functional dependencies

- 1 is true
- 2 is true
- 3 is true
- 4 is false

3.3 Attribute closure of B+

The answear is 4 - The attribute closure for $B+$ is B,C,F,G

3.4 Attribute closure of A,D+

The answear is 4 - The attribute closure for $B+$ is A,B,C,D,E,F,G,H

4 Normalization

4.1 Candidate keys

Below the candidate key

- WX

  Since WX = $U, U => VW, VW => V, V => XY$

4.2 BCNF

No because it is not in 3NF and can there not be in BCNF.

4.3 3NF

No because there is a transitve dependency in the relation. Thus, it violates 3NF.

4.4 3NF decomposition

```
R1{Z,U}
R2{V,X,Y}
R3{W,Z}
R4{U,W,V)
```

Figure 7: 3NF Decomposition

4.5 BCNF decomposition

```
R1{Z,U}
R2{V,X,Y}
R3{W,Z}
R4{U,W,V)
```

Figure 8: BCNF decomposition

5 Serializability and 2PL

5.1 Question about serializability and 2PL

Answers

- true
- true
- true
- true
- false

5.2 Is given schedule serial

No a serial schedule does not exists, since T1 is dependent on T2 and vice versa. See Figur:9 for dependency graph.

5.3 Dependency graph

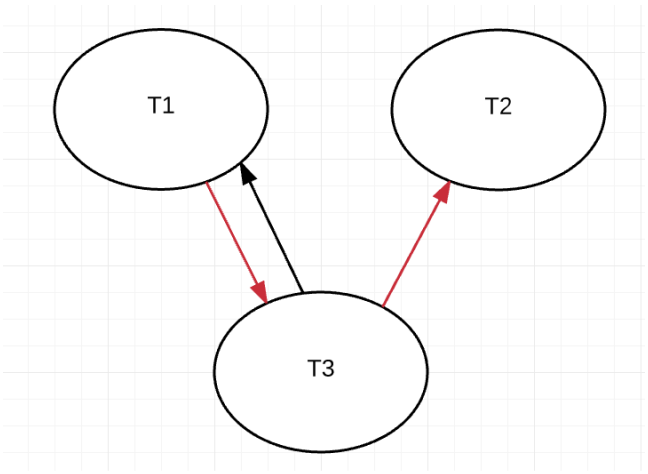

Figure 9: Dependency graph of given schedule

5.4 Conflict serializable

Since the dependency graph is cyclic the schedule is not conflict serializable.

5.5 Is the schedule produced by 2pl

No, because 2PL guarantees serializability, but this schedule is not serializable. Thus it can not be produced by 2Pl