Tutorial 4: Logical Database Design

Jin, Ziyang   Kim, Joon Hyung
\# 34893140   \# 35183128
f4a0b   l1m8

February 2018

```
StudentInfo(S, N, M, A, C, T, I, L, G)
```

with the following FDs:

1. S -> N

2. C -> T, I

3. I -> L

4. S, C, M -> G

5. S, M -> A

6. A -> M

(a) Give an instance of `StudentInfo` (i.e., a relation) that illustrates these three anomalies: insertion, deletion, and update.

- **Insertion Anomaly:** We know an instructor I and his/her InstructorLocation L, but we cannot store this piece of information into `StudentInfo` without storing S, N, M, A, C, T, G together with I, L.

- **Deletion Anomaly:** We would like to delete an instructor I and his/her InstructorLocation L since he/she no longer works in the university. However, if we delete any tuple that has I, L, the corresponding student information like student number S and name N will be lost as well.

- **Update Anomaly:** We would like to update an instructor's InstructorLocation L since he/she just changed an office. However, if we only update one tuple that has I and its corresponding L, an inconsistency is created because for the same I there are more than one values of L, unless all tuples with the same I are updated with the new L.

(b) Consider the decomposition of the relation `StudentInfo` into: `SI1(S, N, M, A, C)` and `SI2(C, T, I, L, G)`. Is this a lossy or lossless-join? Justify your answer.

Answer:
It is a lossy join.
The shared attribute between `SI1` and `SI2` is `C`.
According to FD (2) and FD (3), C^+ = (`C T I L`)
`SI2` has (`C T I L G`), so `C` cannot uniquely determine a tuple in `SI2`, i.e. `C` is not a key for `SI2`. So it is possible for `SI2` to have 2 tuples that share the same `C` value but have different `G` values.
Therefore, when we join `SI1` and `SI2`, a tuple in `SI1` may have more than 1 tuples in `SI2` to choose from. In this case, we don't know which one is the original tuple and which one is newly created by the join.
Therefore, the join is lossy.

(c) Repeat (b) for the decomposition: `SI3(S, M, A, C)` and `SI4(S, N, M, C, T, I, L, G)`.

Answer:
It is a loss-less join.
The share attributes between `SI3` and `SI4` are `S, M, C`.
According to FDs, $(SMC)^+$ = (`S N M A C T I L G`).
So (`S M C`) can uniquely determine tuples in `SI3` and `SI4`. Therefore, when we join `SI3` and `SI4`, there is only one tuple in `SI3` and one tuple in `SI4` can join.
Therefore, the join is loss-less.