

CS 448 Spring 2024 Homework 4

Sean Rettig

March 26, 2024

1

Consider the universal relation $R = \{A, B, C, D, E, F, G, H, I, J\}$ and the set of functional dependencies F .

$F =$

$\{\{A, B\} \rightarrow \{C, D\},$
 $\{D\} \rightarrow \{E, F, G\},$
 $\{F, G\} \rightarrow \{H\},$
 $\{A\} \rightarrow \{I\},$
 $\{A, B\} \rightarrow \{E, G\},$
 $\{A, I\} \rightarrow \{J\}\}$

(a) **What is the closure of the set $\{A, D\}$?**

$\{A, D\}^+ = \{A, D, E, F, G, H, I, J\}$

(b) **What is the key for R ?**

Key = $\{A, B, D\}$

(c) **What is the minimum cover for the set of functional dependencies F ?**

Convert to canonical form:

$\{\{A, B\} \rightarrow \{C\},$
 $\{A, B\} \rightarrow \{D\},$
 $\{D\} \rightarrow \{E\},$
 $\{D\} \rightarrow \{F\},$
 $\{D\} \rightarrow \{G\},$
 $\{F, G\} \rightarrow \{H\},$
 $\{A\} \rightarrow \{I\},$
 $\{A, B\} \rightarrow \{E\},$
 $\{A, B\} \rightarrow \{G\},$
 $\{A, I\} \rightarrow \{J\}$
 $\{A, I\} \rightarrow \{J\}\}$

Eliminate extraneous stuff:

$\{\{A, B\} \rightarrow \{C\},$
 $\{A, B\} \rightarrow \{D\},$
 $\{D\} \rightarrow \{E\},$
 $\{D\} \rightarrow \{F\},$

$\{D\} \rightarrow \{G\},$
 $\{F, G\} \rightarrow \{H\},$
 $\{A\} \rightarrow \{I\},$
 $\{A, B\} \rightarrow \{E\},$
 $\{A, B\} \rightarrow \{G\},$
 $\{A\} \rightarrow \{J\}$

Eliminate redundancies:

$\{\{A, B\} \rightarrow \{C\},$
 $\{A, B\} \rightarrow \{D\},$
 $\{D\} \rightarrow \{E\},$
 $\{D\} \rightarrow \{F\},$
 $\{D\} \rightarrow \{G\},$
 $\{F, G\} \rightarrow \{H\},$
 $\{A\} \rightarrow \{I\},$
 $\{A\} \rightarrow \{J\}\}$

(d) **Decompose R into 2NF.**

$R = \{A, B, C, D, E, F, G, H, I, J\}$

Split up:

$R_1 = \{\underline{A}, \underline{B}, C, D, I, J\}; R_2 = \{\underline{D}, E, F, G, H\}$

Split up again:

$R_1 = \{\underline{A}, \underline{B}, C, D\}; R_2 = \{\underline{A}, I, J\}; R_3 = \{\underline{D}, E, F, G, H\}$

(e) Decompose R into 3NF

$R_1 = \{\underline{A}, \underline{B}, C, D\}; R_2 = \{\underline{A}, I, J\}; R_3 = \{\underline{D}, E, F, G\} R_4 = \{\underline{F}, \underline{G}, H\}$

2

- (a) **What are the possible sources of the information that defines the functional dependencies that hold among the attributes of a relation schema? Discuss two possible sources. Explain your answer.**

Functional dependencies must be defined explicitly by someone who knows the semantics of the attributes in the relation. So what are some possible sources for these semantics?

A possible source could be the entity that wants to use the database. Perhaps they have a bunch of attributes that they want to be able to keep track of. They can put together the attributes and then make some sort of ID for the attributes and then that defines all these functional dependencies for the relation.

Another possible source could be some sort of relationship that is already defined in the nature of whatever you are already trying to model. For example, the relationship between a child and their mother. There is a functional dependency from the child to the mother because the child can only have one biological mother while the mother could have multiple children.

- (b) **What is meant by the closure of a set of functional dependencies? Illustrate with an example.**

The closure of a set of FD really just means that it includes all inferences that can be made from a sparser set of functional dependencies. It's the most explicit you can be about a group of relations.

For example, if there is a functional dependency between my SSN and me, and also a functional dependency between me and my mother, the closure of the set of those two would also include the functional dependency that can be inferred between my ssn and my mother.

- (c) **What is the lossless (or nonadditive) join property of a decomposition? Why is it important?**

This property ensures that spurious tuples, or tuples that were not in the original relation before decomposition, do not appear after rejoining the relations.

This is important so we can complete the normalization of the data and relations without losing information. We need the benefits that come with normalizing relations, but we cannot sacrifice data to achieve it.