

1. Why the 4-way relationship *Births* in Figure 2.15 of the textbook is a bad idea.

Answers:

The 4-way relationship will be converted into a relation *Births*(*baby*, *mother*, *doctor*, *nurse*) with key $\{baby, doctor, nurse\}$. We note that there are several nontrivial FDs and MVDs that hold for this relation, e.g.

$$\begin{array}{lll} baby & \rightarrow & mother \\ mother & \twoheadrightarrow & baby \\ baby & \twoheadrightarrow & doctor \\ baby & \twoheadrightarrow & nurse \end{array}$$

In other words, *Births* is not in 4NF, BCNF, or 3NF, and suffers from various types of redundancy, mostly caused by representing multiple many-to-many relationships in one relation. Therefore, the 4-way relationship is not a good design.

2. Exercise 3.6.1 (c) and (e).

Answers:

(c)

We start by finding all the keys, which happen to be AB , BC , CD , and AD . At this point we can already say that R is in 3NF, because the RHS of any FD will surely be part of a key. However, we don't know yet whether there is any BCNF violation caused by nontrivial FDs that can be *derived* from the given ones. To check for those, theoretically we need to compute $\{A\}^+$, $\{B\}^+$, ..., $\{A, B, C, D\}^+$, or in other words, all possible FDs. Fortunately, we only need to check $\{A\}^+$, $\{B\}^+$, $\{C\}^+$, $\{D\}^+$, $\{A, C\}^+$, and $\{B, D\}^+$, because the rest of the combinations are all superkeys thereby cannot lead to BCNF violations. And since none of the closures above reveal any nontrivial FD, we can say that there is no BCNF violation and R is already in BCNF.

(e)

We note from the given FDs that nothing determines $\{A, B, E\}$, so they must be part of any key. By computing $\{A, B, E\}^+$ we find out that $\{A, B, E\}$ is indeed the one and the only key, so all the given FDs are BCNF violations, and 3NF violations for that matter since C and D are not part of any key.

We take $AB \rightarrow C$ (or any other BCNF-violating FD) for decomposition. Since $\{A, B\}^+ = \{A, B, C, D\}$, we decompose R into $R_1 = \{A, B, C, D\}$ and $R_2 = \{A, B, E\}$. Note that

the key of R_1 is $\{A,B\}$ and $B \rightarrow D$ is a BCNF violation for R_1 , so R_1 can be further decomposed into $R_{11} = \{A, B, C\}$ and $R_{12} = \{B, D\}$.

Finally, note that throughout the whole process, all violations are both BCNF and 3NF violations, and the relations after the decomposition are in BCNF (and thereby 3NF).

3. Exercise 3.7.3 (b) and (d).

Answers:

(b)

Since there's no nontrivial FD that holds for R , the key consists of all attributes, therefore both MVDs are 4NF violations. We take $A \twoheadrightarrow B$ and decompose R into $R_1 = \{A, B\}$ and $R_2 = \{A, C, D\}$, and that's it.

Note that if you take $B \twoheadrightarrow CD$ to decompose you'll have $R_1 = \{A, B\}$ and $R_2 = \{B, C, D\}$, which are also correct.

(d)

First find the key, which is $\{A,B,C\}$. So it's clear that all given MVDs and FDs (which are also MVDs by definition) are 4NF violations, and furthermore, the FDs are BCNF violations as well. Now the question is whether we do BCNF decomposition first or 4NF decomposition first. I prefer BCNF first since FDs are easier to deal with.

First choose $AB \rightarrow E$ and decompose R into $R_1 = \{A, B, E\}$ and $R_2 = \{A, B, C, D\}$. Since $\{A,B\}$ is the key for R_1 and $A \twoheadrightarrow B$ holds for R_1 , we decompose R_1 further into $R_{11} = \{A, B\}$ and $R_{12} = \{A, E\}$. Since $\{A,B,C\}$ is the key for R_2 and $A \twoheadrightarrow D$ holds for R_2 , we decompose R_2 into $R_{21} = \{A, B, C\}$ and $R_{22} = \{A, D\}$. Finally, $\{A,B,C\}$ is the key for R_{21} and $A \twoheadrightarrow B$ holds for R_{21} , so we continue to decompose R_{21} into $R_{211} = \{A, B\}$ and $R_{212} = \{A, C\}$.

In the end, the relations after the decomposition are $\{A,B\}$, $\{A,C\}$, $\{A,D\}$, $\{A,E\}$.

And for extra credit, prove that given R with the MVDs and FDs, $A \twoheadrightarrow C$.

4. Proof.

Answers:

(a) If $AB \twoheadrightarrow BC$, then $AB \twoheadrightarrow B$ and $AB \twoheadrightarrow C$.

The first part comes directly from the definition of trivial MVD, and the second part can be easily proven using the *proof by chase* method we discussed in class.

(b) If $A \twoheadrightarrow BC$, then $A \twoheadrightarrow B$ and $A \twoheadrightarrow C$.

You can find a counter-example in the textbook.

5. Exercise 5.2.4 (e), (f), (g), (h)

Answers:

Note that in all the queries below I assume the *Set* semantics. For *Bags*, simply add duplicate elimination at proper places.

(e)

$$\begin{aligned} S(ship, class, launched) &:= Ships \\ R &:= Outcomes \bowtie S \bowtie Classes \\ Ans &:= \pi_{ship, displacement, numGuns}(R) \end{aligned}$$

(f)

$$\pi_{ship}(Outcomes) \cup \pi_{name \rightarrow ship}(Ships)$$

(g) A self join of *Ships* would give us the classes that have more than one member, then the difference between these classes and all classes would be the answer we want.

$$\begin{aligned} Ships1(name1, class, launched1) &:= Ships \\ R &:= \pi_{class} \sigma_{name <> name1}(Ships \bowtie Ships1) \\ Ans &:= \pi_{class}(Classes) - R \end{aligned}$$

(h) A self join of *Classes* would do.

$$\begin{aligned} Countries1 &:= \pi_{country, type \rightarrow type1}(Classes) \\ Countries2 &:= \pi_{country, type \rightarrow type2}(Classes) \\ Ans &:= \pi_{country} \sigma_{type1 <> type2} Countries1 \bowtie Countries2 \end{aligned}$$

6. Exercise 5.5.1 (c), (e)

Answers:

(c)

$$\sigma_{type=pc}(Product) \cap \sigma_{type=laptop}(Product) = \emptyset$$

(e)

$$\sigma_{Laptop.ram > PC.ram \text{ AND } Laptop.price < PC.price}(Laptop \times PC) = \emptyset$$