CSC 370

Activity Worksheet:
Third Normal Form

Dr. Sean Chester

Fall 2022

Notes

This worksheet should provide extra practice questions for decomposing a relation into 3NF using the Synthesis Algorithm. In each question you are given a relation and a set of functional dependencies; you should show your work to decompose the relation into a set of relations that are all in 3NF. The first question is answered as a model solution

Questions

1. R(A, B, C, D)

   AB → CD
   C → D
   D → B

   Solution:

   There is some lack of clarity as to whether we should first check if a relation is in 3NF before decomposing it. This step is not in the algorithm in the textbook. The argument against doing so is that it involves computing a combinatorial number of closures to find out which attributes are prime. However, on an example where there are only four attributes, the power set really isn't that large and it is worth checking whether the decomposition is actually necessary. It doesn't matter whether we do this before or after calculating a minimal basis, because the set of closures will be equivalent anyway.

   First, we check if it is already in 3NF by evaluating each FD independently to see if any of them are 3NF violations. For this, it will help to identify any keys. We can do this by exhaustively searching the power set of the attribute set. This will yield the following keys:

   $\{AB\}$
   $\{AC\}$
   $\{AD\}$

   So, then we check each FD:

   AB → CD

   $\{AB\} \subseteq \{AB\}$ // i.e., the key $\{AB\}$ is a subset of the left-hand side of this FD. Thus the FD is in BCNF and 3NF

   D → B

   $\{AB\} \supseteq \{B\}$ // i.e., the key $\{AB\}$ is a superset of the right-hand side of this FD. Thus the FD is in 3NF.

   C → D

   $\{AD\} \supseteq \{D\}$ // i.e., the key $\{AD\}$ is a superset of the right-hand side of this FD. Thus the FD is in 3NF.

   ∴ R is already in 3NF and a decomposition should not be done.

   (But for practice, let's do it anyway):

   First, we need a minimal basis. To determine that, we first apply the splitting rule to every FD to which it applies:

   AB → C

AB → D
C → D
D → B

Next, we check if any FD's follow from the others. We can do this by removing each FD and checking if the right-hand side is still part of the closure of the left-hand-side.

After removing AB → C
$\{A, B\}^+ \supseteq \{A, B\}$ // trivially
$\{A, B\}^+ \supseteq \{A, B, D\}$ // from AB → D
$\{A, B\}^+ = \{A, B, D\}$ // because we cannot productively apply any more FD's.
∴ AB → C is not redundant, because AB does not determine C from the remaining FD's

After removing AB → D
$\{A, B\}^+ \supseteq \{A, B\}$ // trivially
$\{A, B\}^+ \supseteq \{A, B, C\}$ // from AB → C
$\{A, B\}^+ \supseteq \{A, B, C, D\}$ // from C → D
$\{A, B\}^+ = \{A, B, C, D\}$ // because we have already established that $\{AB\}$ is a key.
∴ AB → D is redundant, because AB determines D from the remaining FD's

After removing C → D
$\{C\}^+ \supseteq \{C\}$ // trivially
$\{C\}^+ = \{C\}$ // because there are no FD's with a subset of $\{C\}$ on the left-hand side
∴ C → D is not redundant, because C cannot determine D from the remaining FD's

After removing D → B
$\{D\}^+ \supseteq \{D\}$ // trivially
$\{D\}^+ = \{D\}$ // because there are no FD's with a subset of $\{D\}$ on the left-hand side
∴ D → B is not redundant, because D cannot determine B from the remaining FD's

At this point, our minimal basis is:
AB → C
C → D
D → B

As a final condition, we check whether it is possible to remove any attributes from the left-hand side of the first FD:

Changing AB → C to A → C
Observe that now $\{A\}^+ = \{A, B, C, D\}$, but it used to be $\{A\}$. Thus, this is not a basis.

Changing AB → C to B → C
Observe that now $\{B\}^+ = \{B, C, D\}$, but it used to be $\{B\}$. Thus, this is not a basis.

In summary, a minimal basis for R is:
AB → C
C → D
D → B

As a second step, we create one projection for each FD:

S(A ,B, C)
T(C, D)
U(B, D)

(Observe that none of these are subsets of each other so none are redundant and we should keep all of them.)

Finally, we confirm whether any of the projections involve a superkey for R. Checking each relation in turn:
$\{A, B, C\}^+ \supseteq \{A, B, C\}$ // trivially
$\{A, B, C\}^+ \supseteq \{A, B, C, D\}$ // from C $\rightarrow$ D
$\{A, B, C\}^+ = \{A, B, C, D\}$ // because we have already established that $\{A, B, C\}$ is a superkey.

$\therefore$ we already have a superkey among our projections and the decomposition is complete.

Summary:
S(A, B, C)
T(C, D)
U(B, D)

2. MovieScreening(title, genre, theatre, city)

   title $\rightarrow$ genre
   title city $\rightarrow$ theatre genre
   theatre $\rightarrow$ city

   Solution:

3. R(A, B, C, D, E)

   $AB \rightarrow CD$
   $CD \rightarrow AB$

   Solution:

4. Provide a set of functional dependencies for a relation R(A, B, C, D, E, F) such that either:

   R is in BCNF but not 3NF;
   or R is in 3NF but not BCNF

   Solution:

Solutions

Question 1

There is some lack of clarity as to whether we should first check if a relation is in 3NF before decomposing it. This step is not in the algorithm in the textbook. The argument against doing so is that it involves computing a combinatorial number of closures to find out which attributes are prime. However, on an example where there are only four attributes, the power set really isn't that large and it is worth checking whether the decomposition is actually necessary. It doesn't matter whether we do this before or after calculating a minimal basis, because the set of closures will be equivalent anyway.

First, we check if it is already in 3NF by evaluating each FD independently to see if any of them are 3NF violations. For this, it will help to identify any keys. We can do this by exhaustively searching the power set of the attribute set. This will yield the following keys:

$\{AB\}$
$\{AC\}$
$\{AD\}$

So, then we check each FD:

AB $\rightarrow$ CD

$\{AB\} \subseteq \{AB\}$ // i.e., the key $\{AB\}$ is a subset of the left-hand side of this FD. Thus the FD is in BCNF and 3NF

D $\rightarrow$ B

$\{AB\} \supseteq \{B\}$ // i.e., the key $\{AB\}$ is a superset of the right-hand side of this FD. Thus the FD is in 3NF.

C $\rightarrow$ D

$\{AD\} \supseteq \{D\}$ // i.e., the key $\{AD\}$ is a superset of the right-hand side of this FD. Thus the FD is in 3NF.

$\therefore$ R is already in 3NF and a decomposition should not be done.

(But for practice, let's do it anyway):

First, we need a minimal basis. To determine that, we first apply the splitting rule to every FD to which it applies:

AB $\rightarrow$ C
AB $\rightarrow$ D
C $\rightarrow$ D
D $\rightarrow$ B

Next, we check if any FD's follow from the others. We can do this by removing each FD and checking if the right-hand side is still part of the closure of the left-hand-side.

After removing AB $\rightarrow$ C

$\{A, B\}^+ \supseteq \{A, B\}$ // trivially
$\{A, B\}^+ \supseteq \{A, B, D\}$ // from AB → D
$\{A, B\}^+ = \{A, B, D\}$ // because we cannot productively apply any more FD's.
∴ AB → C is not redundant, because AB does not determine C from the remaining FD's

After removing AB → D
$\{A, B\}^+ \supseteq \{A, B\}$ // trivially
$\{A, B\}^+ \supseteq \{A, B, C\}$ // from AB → C
$\{A, B\}^+ \supseteq \{A, B, C, D\}$ // from C → D
$\{A, B\}^+ = \{A, B, C, D\}$ // because we have already established that $\{AB\}$ is a key.
∴ AB → D is redundant, because AB determines D from the remaining FD's

After removing C → D
$\{C\}^+ \supseteq \{C\}$ // trivially
$\{C\}^+ = \{C\}$ // because there are no FD's with a subset of $\{C\}$ on the left-hand side
∴ C → D is not redundant, because C cannot determine D from the remaining FD's

After removing D → B
$\{D\}^+ \supseteq \{D\}$ // trivially
$\{D\}^+ = \{D\}$ // because there are no FD's with a subset of $\{D\}$ on the left-hand side
∴ D → B is not redundant, because D cannot determine B from the remaining FD's

At this point, our minimal basis is:
AB → C
C → D
D → B

As a final condition, we check whether it is possible to remove any attributes from the left-hand side of the first FD:

Changing AB → C to A → C
Observe that now $\{A\}^+ = \{A, B, C, D\}$, but it used to be $\{A\}$. Thus, this is not a basis.

Changing AB → C to B → C
Observe that now $\{B\}^+ = \{B, C, D\}$, but it used to be $\{B\}$. Thus, this is not a basis.

In summary, a minimal basis for R is:
AB → C
C → D
D → B

As a second step, we create one projection for each FD:

S(A ,B, C)
T(C, D)
U(B, D)

(Observe that none of these are subsets of each other so none are redundant and we should keep all of them.)

Finally, we confirm whether any of the projections involve a superkey for R. Checking each relation in turn:
$\{A, B, C\}^+ \supseteq \{A, B, C\}$ // trivially
$\{A, B, C\}^+ \supseteq \{A, B, C, D\}$ // from C $\rightarrow$ D
$\{A, B, C\}^+ = \{A, B, C, D\}$ // because we have already established that $\{A, B, C\}$ is a superkey.

$\therefore$ we already have a superkey among our projections and the decomposition is complete.

Summary:
S(A, B, C)
T(C, D)
U(B, D)

Question 2

First, let us calculate all the keys so that we can ascertain which attributes are prime and establish whether this relation is already in 3NF. An exhaustive search will reveal the following keys:

$\{title, city\}$
$\{title, theatre\}$

We can thus see that we will have a 3NF violation on the first FD, since neither is $\{title\}$ a superkey nor is genre prime. We thus proceed with the decomposition.

To start, we need to find a minimal basis. We use the splitting rule to arrive at: title → genre
title city → theatre
title city → genre
theatre → city

Immediately, we can see that title city → genre is redundant, because we already have title → genre. Thus we end up with the minimal basis:
title → genre
title city → theatre
theatre → city

Next, we confirm that there are no 3NF violations in the minimal basis.

As a second step, we create one projection for each FD:
S(title, genre)
T(title, city, theatre)
U(theatre, city)

Observe that the attributes of U are a subset of those of T; thus, U is redundant and we can strike it out.

Finally, we check if we have a superkey with respect to MovieScreening among our projections:
$\{title, genre\}^+ \supseteq \{title, genre\}$ // trivially
$\{title, genre\}^+ = \{title, genre\}$ // because the only FD with a left-hand side that is a subset of $\{title, genre\}$ does not expand the closure

Since T is not a projection onto a superkey of MovieScreening, we try U:
$\{title, city, theatre\}^+ \supseteq \{title, city, theatre\}$ // trivially
$\{title, city, theatre\}^+ \supseteq \{title, city, theatre, genre\}$ // because title → genre
$\{title, city, theatre\}^+ = \{title, city, theatre, genre\}$ // because we have confirmed that we have found a superkey

Thus, U is a projection of MovieScreening onto a superkey and we are finished.

Summary (semantically renaming the projections):

Movie(title, genre)
Screening(title, city, theatre)

Question 3

First, let us determine all the keys by exhaustive search so we can establish whether this relation is already in 3NF. This reveals the following keys:

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

We see that neither $\{A, B\}$ nor $\{C, D\}$, the left-hand sides of the FD's, are superkeys so this is not in BCNF; however, A, B, C, and D are all prime. Therefore neither FD is a 3NF violation and this relation is already in 3NF and we need not decompose it. Nonetheless, we will anyway just for extra practice with the decomposition algorithm and to see an example that includes non-determinism.

To start, we construct a minimal basis by splitting every FD:
AB → C
AB → D
CD → A
CD → B

Next, we form our projections from each FD:
S(A, B, C)
T(A, B, D)
U(A, C, D)
V(B, C, D)

Finally, we confirm if any of the projections are onto a superkey of R.
$\{A, B, C\}^+ \supseteq \{A, B, C\}$ // trivially
$\{A, B, C\}^+ \supseteq \{A, B, C, D\}$ // because AB → D
$\{A, B, C\}^+ = \{A, B, C, D\}$ // because there are no more FD's that we can apply to expand the closure

$\{A, B, D\}^+ \supseteq \{A, B, D\}$ // trivially
$\{A, B, D\}^+ \supseteq \{A, B, C, D\}$ // because AB → C
$\{A, B, D\}^+ = \{A, B, C, D\}$ // because there are no more FD's that we can apply to expand the closure

$\{A, C, D\}^+ \supseteq \{A, C, D\}$ // trivially
$\{A, C, D\}^+ \supseteq \{A, B, C, D\}$ // because CD → B
$\{A, C, D\}^+ = \{A, B, C, D\}$ // because there are no more FD's that we can apply to expand the closure

$\{B, C, D\}^+ \supseteq \{B, C, D\}$ // trivially
$\{B, C, D\}^+ \supseteq \{A, B, C, D\}$ // because CD → A
$\{B, C, D\}^+ = \{A, B, C, D\}$ // because there are no more FD's that we can apply to expand the closure

Thus, we do not have a superkey and need to add another relation that is a key.

To find a key for R, we can search exhaustively through the powerset of its attribute set:
$\{A\}^+ = \{A\}$ // not a key
$\{B\}^+ = \{B\}$ // not a key
$\{C\}^+ = \{C\}$ // not a key
$\{D\}^+ = \{D\}$ // not a key

$\{E\}^+ = \{E\}$ // not a key
$\{A, B\}^+ = \{A, B, C, D\}$ // not a key
$\{A, C\}^+ = \{A, C\}$ // not a key
$\{A, D\}^+ = \{A, D\}$ // not a key
$\{A, E\}^+ = \{A, E\}$ // not a key
$\{B, C\}^+ = \{B, C\}$ // not a key
$\{B, D\}^+ = \{B, D\}$ // not a key
$\{B, E\}^+ = \{B, E\}$ // not a key
$\{C, D\}^+ = \{A, B, C, D\}$ // not a key
$\{C, E\}^+ = \{C, E\}$ // not a key
$\{D, E\}^+ = \{D, E\}$ // not a key
$\{A, B, C\}^+ = \{A, B, C, D\}$ // not a key
$\{A, B, D\}^+ = \{A, B, D\}$ // not a key
$\{A, B, E\}^+ = \{A, B, C, D, E\}$ // is a key

($\{C, D, E\}$ is also a key.)

Thus we add a second projection onto one of the keys: W(A, B, E)

Summary:

S(A, B, C)
T(A, B, D)
U(A, C, D)
V(B, C, D)
W(A, B, E)

Question 4

Note that only the second option is possible, because the 3NF condition is strictly more permissive than the BCNF condition.

Thus, we need to construct an example in which there is an FD with a left-hand side that is not a superkey but the right-hand side consists only of prime attributes.

Let us begin by creating a key so that we have some prime attributes:
AB → CDEF

Next, let us create an FD in which the left-hand side is not a superkey, but we use one (or more) of the prime attributes on the right-hand side:
C → A

Thus, our final solution is:
AB → CDEF
C → A

To verify the solution, we can confirm that:
$\{C\}^+ = \{A, C\}$ // BCNF violation
$\{A, B\}^+ = \{A, B, C, D, E, F\}$ // a key; therefore, A and B are both prime.
Moreover, AB → CDEF is in BCNF.