The Australian National University, School of Computing
COMP2400/6240 (Relational Databases)
Semester 2, 2022

Lab 6, Week 8

Normalisation (Solutions)

The purpose of this lab is to help you understand the normal forms 3NF and BCNF. In particular, you need to understand:

- What is BCNF? What is 3NF?
- What are the differences between 3NF and BCNF?

1 Normalisation - Inspection Example

Consider the following relation Inspection held at the MyHome real estate agency, in which {PropertyNo, Date} is the primary key:

PropertyNo	Address	Date	Time	StaffNo	StaffName	CameraID
PR4	6 Masson St	18-Oct-11	10:00	S137	Mike Jenk	C211
PR16	8 Berry St	22-Apr-12	09:00	S114	Sue Wang	C323
PR4	6 Masson St	01-Oct-13	12:00	S114	Sue Wang	C323
PR16	8 Berry St	21-Apr-12	13:00	S114	Sue Wang	C323

A set Σ of FDs for representing the business rules of Inspection is as follows:

- $\{\text{PropertyNo}\} \rightarrow \{\text{Address}\}$;
- $\{\text{StaffNo}\} \rightarrow \{\text{StaffName}\}$;
- $\{\text{PropertyNo, Date}\} \rightarrow \{\text{StaffNo, Time}\}$;
- $\{\text{StaffNo, Date}\} \rightarrow \{\text{CameraID}\}$;

- {StaffNo, Date, Time} → {PropertyNo};
- {Date, Time, CameraID} → {PropertyNo}.

(1) *Find all the keys and prime attributes w.r.t. Σ.*

Solution: The keys are:

- {PropertyNo, Date};
- {Date, Time, CameraID};
- {StaffNo, Date, Time}.

This is because the closure of {Date, PropertyNo}, {Date, Time, CameraID} or {StaffNo, Date, Time} is the set of all attributes of INSPECTION with respect to Σ and they are minimal.

The prime attributes are: Date, Time, PropertyNo, StaffNo and CameraID. The non-prime attributes are: StaffName and Address.

(2) *Is the given set of FDs minimal? If not, give a minimal cover.*

Solution: It's not minimal. One possible solution is as follows:

- {StaffNo} → {StaffName};
- {PropertyNo} → {Address};
- {PropertyNo, Date} → {StaffNo};
- {PropertyNo, Date} → {Time};
- {StaffNo, Date} → {CameraID};
- {Date, Time, CameraID} → {PropertyNo}.

The steps are as follows:

Starting with the given set Σ of FDs,

- by **Step 2** of the algorithm, we replace {PropertyNo, Date} → {StaffNo, Time} with {PropertyNo, Date} → {StaffNo} and {PropertyNo, Date} → {Time}. Then we have: {{StaffNo} → {StaffName}, {PropertyNo} → {Address}, {PropertyNo, Date} → {StaffNo}, {PropertyNo, Date} → {Time}, {StaffNo, Date} → {CameraID}, {StaffNo, Date, Time} → {PropertyNo}, {Date, Time, CameraID} → {PropertyNo}};

- by **Step 3** of the algorithm, we still have the same set of FDs as in the previous step;

- by **Step 4** of the algorithm, we calculate the closure of the determinant of a FD in terms of other FDs, if the closure contains the dependent of the FD, then the FD is redundant and can be removed. In doing so, we can only remove {StaffNo, Date, Time} $\rightarrow$ {PropertyNo} because the closure of {StaffNo, Date, Time} in terms of the other FDs contains PropertyNo.

After the above three steps, we can obtain the minimal cover.

(3) *Is* INSPECTION *in 3NF w.r.t. Σ? If not, determine a lossless and dependency preserving 3NF decomposition. Are the relation schemas you have obtained in the decomposition in BCNF? Justify your answers.*

Solution:

INSPECTION is not in 3NF w.r.t. Σ. This can be verified by testing each FD: X $\rightarrow$ A defined on INSPECTION: either X is a superkey or A is a prime attribute. In accordance with the results in Exercise (2), it is clear that {PropertyNo} $\rightarrow$ {Address} and {StaffNo} $\rightarrow$ {StaffName} are problematic. Using the minimal cover in Exercise (3) and the corresponding 3NF decomposition algorithm, we may decompose INSPECTION into the following relation schemas.

- STAFF={StaffNo, StaffName} with the FD: {StaffNo} $\rightarrow$ {StaffName};

- PROPERTY={PropertyNo, Address} with the FD: {PropertyNo} $\rightarrow$ {Address};

- INSPECTION1={PropertyNo, Date, staffNo, Time} with the FD: {PropertyNo, Date} $\rightarrow$ {StaffNo, Time};

- INSPECTION2={StaffNo, Date, CameraID} with the FD: {StaffNo, Date} $\rightarrow$ {CameraID};

- INSPECTION3={PropertyNo, Date, Time, CameraID} with the FDs: {PropertyNo, Date} $\rightarrow$ {Time}, and {Date, Time, CameraID} $\rightarrow$ {PropertyNo}.

The above decomposition into 3NF is lossless and dependency preserving.

Now let's discuss why the above 3NF decomposition preserves all FDs in the original set Σ. Assume the above minimal cover derived from Σ is Σ_{min}. Note that Σ and Σ_{min} must be equivalent according to the definition of the minimal cover. It is obvious that the 3NF decomposition preserves all FDs in Σ_{min} and therefore can imply any FD in the Σ.

The relation schemas Staff, Property, Inspection1 and Inspection2 are in BCNF. Inspection3 seems not in BCNF due to {PropertyNo, Date} → {Time} and {PropertyNo, Date} seems not a PK of Inspection3 based on two surviving FDs in Inspection3. However, we should consider all surviving FDs in this decomposition and thus {PropertyNo, Date} is actually a PK of Inspection3 because {PropertyNo, Date} → {StaffNo, Time} and {StaffNo, Date} → {CameraID} together imply {PropertyNo, Date} → {CameraID, Time} and thus Inspection3 is also in BCNF.

2 Normalisation - Meeting Example

Consider the following relation:

Meeting={CRN, Name, Date, Time, Officer, Cabin}

with the following set Σ of FDs:

fd1: {CRN, Date, Time} → {Officer};

fd2: {Date, Time, Cabin} → {CRN};

fd3: {Officer, Date, Time} → {CRN};

fd4: {Date, Officer} → {Cabin};

fd5: {CRN} → {Name}.

(4) *List all the keys of* Meeting *w.r.t.* Σ.

Solution: Let us find out all the keys for Meeting. Note the {Date, Time} must be part of all keys as they never appear in the dependent of any FD,

1. {CRN, Date, Time} is a key (minimal superkey)
2. {Cabin, Date, Time} is a key (minimal superkey)
3. {Officer, Date, Time} is a key (minimal superkey)

(5) *Find all the prime attributes of* Meeting *w.r.t.* Σ.

Solution:

{CRN, Date, Time, Officer, Cabin} is the set of all prime attributes of MEETING with respect to Σ.

(6) *Does* MEETING *satisfy 3NF w.r.t. Σ? If not, determine a minimal cover of Σ, and a lossless and dependency preserving 3NF decomposition. Justify your answers.*

Solution:

- MEETING doesn't satisfy 3NF because, in fd5: {CRN} $\rightarrow$ {Name}, neither CRN is a superkey nor Name is a prime attribute.
- Σ' = {{CRN, Date, Time} $\rightarrow$ {Officer}, {Date, Time, Cabin} $\rightarrow$ {CRN}, {Date, Officer} $\rightarrow$ {Cabin}, {CRN} $\rightarrow$ {Name}} is a minimal cover.
- By applying the corresponding algorithm, we can achieve a lossless and dependency preserving 3NF decomposition for MEETING as follows:
- From Σ', we add R_1={CRN, Date, Time, Officer}; R_2={Date, Time, Cabin, CRN}; R_3={Officer, Date, Cabin} and R_4={CRN, Name} to S.
- Since R_1 is a superkey, we don't need to add a key. Thus S:= {R_1, R_2, R_3, R_4}
- Therefore MEETING is decomposed into the following relations in 3NF:
  - R_1 ={CRN, Date, Time, Office} with Σ_1 = {CRN, Date, Time} $\rightarrow$ {Officer};
  - R_2 ={Date, Time, Cabin, CRN} with Σ_2 = {Date, Time, Cabin} $\rightarrow$ {CRN};
  - R_3 ={Officer, Date, Cabin} with Σ_3 = {Date, Officer} $\rightarrow$ {Cabin};
  - R_4 ={CRN, Name} with Σ_4 = {CRN} $\rightarrow$ {Name}.

Why the above 3NF decomposition preserves all fds in Σ?
The surviving fds ($\Sigma_1 \cup \Sigma_2 \cup \Sigma_3 \cup \Sigma_4$) is equivalent to Σ.

(7) *Does* MEETING *satisfy BCNF w.r.t. Σ? If not, determine a lossless decomposition for* MEETING *into BCNF. Does your decomposition preserve all dependencies of* MEETING*?*

Solution:

- Since the determinants of the FDs: {Date, Officer} $\rightarrow$ {Cabin} and {CRN} $\rightarrow$ {Name} are not superkeys, MEETING doesn't satisfy BCNF.
- By applying the corresponding algorithm,

- Let S := {Meeting}.
- Since Meeting is not in BCNF, we pick the FD: {Date, Officer} $\rightarrow$ {Cabin} that violates BCNF, and replace Meeting in S by two relation schemas R_1={CRN, Date, Time, Officer, Name} with Σ_1 ={ {CRN, Date, Time} $\rightarrow$ {Officer}, {Officer, Date, Time} $\rightarrow$ {CRN}, {CRN} $\rightarrow$ {Name}} and R_2={Date, Officer, Cabin} with Σ_2 = {{Date, Officer} $\rightarrow$ {Cabin}}. So we have S:={R_1, R_2}.
- Now we easily see that R_1 is still not in BCNF because the determinant of the FD {CRN} $\rightarrow$ {Name} and {CRN} is not a superkey with respect to Σ_1. We pick this problematic fd and further decompose R_1 into two relations R_{11}={CRN, Date, Time, Officer} with Σ_{11} = {{CRN, Date, Time} $\rightarrow$ {Officer}; {Officer, Date, Time} $\rightarrow$ {CRN}} and R_{12}={CRN, Name} with Σ_{12} = {{CRN} $\rightarrow$ {Name}}.
- Now we have S:= {R_{11}, R_{12}, R_2}.

• This decomposition is lossless, which is ensured by the algorithm. However the FD: Date, Time, Cabin $\rightarrow$ CRN isn't preserved in the decomposition of Meeting into R_{11}, R_{12} and R_2 because it cannot be inferred from Σ_{11}, Σ_{12} or Σ_2.