School of Engineering and Computer Science

SWEN304 **Database System Engineering**

Assignment 3

Due date: 23:59, **Monday 23 May**

The objective of this assignment is to test your understanding of functional dependencies, normal forms, database normalization, The assignment is worth **10%** of your final grade. It will be marked out of 100.

Submission instructions:

- Submit your assignment in **pdf** via the submission system

Note: Assignments not in **pdf** will incur a deduction of 3 marks.

Question 1. Functional Dependencies and Normal Forms [4 marks]

A	*B*	*C*
a	*b*	*3*
d	*b*	*3*

Consider a relation schema *N (R, F)* where $R = \{A, B, C\}$. Suppose we find the following two tuples in an instance of this relation schema.

Which of the following functional dependencies does definitely **not** hold over the relation schema *N*? Justify your answer.

1) $A \rightarrow B$

2) $B \rightarrow A$

3) $B \rightarrow C$

4) $C \rightarrow A$

1. **A can determine B. Because A is uniquely for B.**
2. **B can not determine A. Because two duplicate B have different A is not a valid functional dependency.**
3. **B can determine C. B can identify the C accurately.**
4. **C can not determine A. Because two duplicate C have different A is not a valid functional dependency.**

Question 2. Normal Forms [16 marks]

Consider a relation schema N (R, F) where R = {A, B, C, D}. For each of the following sets F of functional dependencies, determine which normal form (1NF, 2NF, 3NF, BCNF) the relation schema N is in. Justify your answer.

Hint: Note that in all four cases AB is the only key for N.

1) $F = \{AB \rightarrow C, C \rightarrow D\}$

2NF. Because C is fully depended on the key AB. And D is fully depended on the C. Therefore it is 2NF.

2) $F = \{AB \rightarrow D, B \rightarrow C\}$

1NF. Because AB is the key for N. For the $B \rightarrow C$, C is partially depended on the key.

3) $F = \{AB \rightarrow C, AB \rightarrow D\}$

BCNF. Because all attributes on the left side are super keys of N.

4) $F = \{AB \rightarrow CD, C \rightarrow B\}$

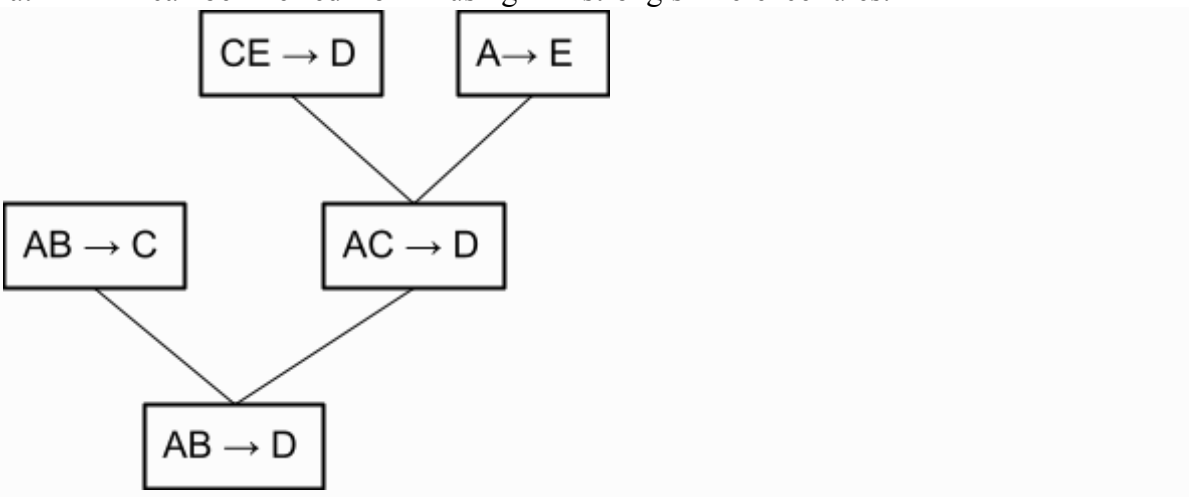
3NF. Because AB on the left side is the key for N and CD on the right side is not prime attribute. C on the left side is not the key for N and B on the right side is the prime attribute. Therefore it is 3NF.

Question 3. Functional Dependency [7 marks]

Consider a relation schema $N(R, F)$ where $R=\{A, B, C, D, E\}$ with the set of functional dependencies

$$F = \{AB \rightarrow C, CE \rightarrow D, A \rightarrow E\}$$

Show that $AB \rightarrow D$ can be inferred from F using Armstrong's inference rules.



- **If $CE \rightarrow D$ and $A \rightarrow E$, then $AC \rightarrow D$.**
- **If $AB \rightarrow C$ and $AC \rightarrow D$, then $AB \rightarrow D$.**

Question 4. Minimal Cover of a set of Functional Dependencies [18 marks]

Consider the set of functional dependencies $F = \{A \rightarrow B, B \rightarrow CD, D \rightarrow A, AC \rightarrow D\}$. Compute a minimal cover of F. Justify your answer.

1. ***First step*** *− Convert RHS attribute into singleton attribute.*

   $F = \{A \rightarrow B, B \rightarrow C, B \rightarrow D, D \rightarrow A, AC \rightarrow D\}$.

2. ***Second step*** *− Remove the extra LHS attribute. Find the closure of A, B, C, D.*

   A+ = {A, B, C, D}

   B+ = {B, C, D, A}

   C+ = {C}

   D+ = {A, B, C, D}

   So, AC -> D can be converted into A -> D

   $F = \{A \rightarrow B, B \rightarrow C, B \rightarrow D, D \rightarrow A, A \rightarrow D\}$.

3. **Third step** − Remove the redundant FDs.

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

Question 5. 3NF Normalization [25 marks]

Consider a relation schema $N(R, F)$ where $R = \{A, B, C, D\}$ and $F = \{A \rightarrow B, C \rightarrow D\}$. Perform the following tasks. Justify your answers.

1) [5 marks] Identify all keys for N. Show your process.

The only minimal key is AC. Because A and C are both on the left side to make the right side.

2) [5 marks] Identify the highest normal form (1NF, 2NF, 3NF, BCNF) that N satisfies.

Because AC is the key for N and there is not any AC on the left side therefore it is not BCNF. Also there are not any prime attributes on the right side therefore it is not 3NF. Both B and D are partially depended on the key. Therefore it is 1NF.

3) [10 marks] If N is not in 3NF, compute a lossless transformation into a set of 3NF relation schemas using the Synthesis algorithm.

1. (A →B) (C → D) S = {({A, B }, {A }), ({C, D }, {C })}
2. A + = AB

   C + = CD

   K (U, F) = AC
3. S = {({A, B }, {A }), ({C, D }, {C }), ({A, C }, {AC })}

4) [5 marks] Verify explicitly that your result has the lossless property, satisfies 3NF, and that all functional dependencies are preserved.

U = AB U CD U AC = {A, B, C, D} which contains all attributes.

F = $\{A \rightarrow B\}$ U $\{C \rightarrow D\}$ U $\{A \rightarrow C\}$ = $\{A \rightarrow B, C \rightarrow D, A \rightarrow C\}$ all FDs are preserved.

Because A is the key for $A \rightarrow B$ and C is the key for $C \rightarrow D$ and C is the prime attribute for $A \rightarrow C$. Therefore it is in 3NF.

Question 6. BCNF Normalization [30 marks]

Consider a relation schema $N(R, F)$, where $R = \{A, B, C, D\}$ and $F = \{A \rightarrow C,\ D \rightarrow B,\ BC \rightarrow A,\ BC \rightarrow D\}$. Perform the following tasks. Justify your answers.

1) [5 marks] Identify all keys for N. Show process.

A+ = AC

B+ = B

C+ = C

D+ = D

AB+ = ABCD

AC+ = AC

AD+ = ADCB

BC+ = BCAD

BD+ = BD

CD+ = CDBA

ALL MINIMAL KEYS ARE: AB, AD, BC, CD

2) [4 marks] Identify the highest normal form (1NF, 2NF, 3NF, BCNF) that N satisfies.

Because for $A \rightarrow C,\ D \rightarrow B,$ both A and D are not the key for N therefore it is not BCNF. Because $BC \rightarrow A$, $BC \rightarrow D$ the key for N is BC so B and C are prime attributes. Therefore it is 3NF.

3) [16 marks] If N is not in BCNF, compute a lossless decomposition into a set of BCNF relation schemas using the BCNF decomposition algorithm.

$R = \{A, B, C, D\}$

$F = \{A \rightarrow C,\ D \rightarrow B,\ BC \rightarrow A,\ BC \rightarrow D\}$

R is not in BCNF because there are FDS, $A \rightarrow C$ and $D \rightarrow B,$ of which the LHS is not a super key.

Decompose R using A → C into R1= BCD with F1= {D → B}, R2=AC with F2= { A→ C}. R2 is in BCNF.

S1 = (R1 ({B,C,D},{$D \rightarrow B$}), R2 ({A,C},{$A \rightarrow C$}))

Decompose R1 along D → B, R1 is decomposed into R3 = DC with F3 ={}, and R4 = BD with F4 ={$D \rightarrow B$} . Both R3 and R4 is in BCNF

S2 = (R1 ({B,C,D},{$D \rightarrow B$}), R2 ({A,C},{$A \rightarrow C$}), R3 ({D,C},{}), R4 ({B,D},{$D \rightarrow B$}))

4) [5 marks] Verify explicitly whether your result satisfies BCNF, and all functional dependencies are preserved.

U = DB U AC U DC = {A, B, C, D}

F = {$A \rightarrow C$} U {} U {$D \rightarrow B$} = {$A \rightarrow C,\ D \rightarrow B$}

Compare to the original one there are $BC \rightarrow A$, $BC \rightarrow D$ are lost.

According to 3) R2, R3 and R4 are in BCNF. Therefore it is in BCNF.

****