# CZ2007 Tutorial 4: BCNF + 3NF

Week 6

# Question 1

- A medical clinic database schema contains the following:
- APPOINTMENT (patient-id, patient-name, doctor-id, doctor-name, appointment-date, appointment-time, clinic-room-no)
- Identify the functional dependencies in the schema, stating any assumptions made.
  - There could be different sets of FDs depending on how you interpret them.
- Using these functional dependencies, normalise the schema to Third Normal Form.
  - We may get different sets of normalized relations.

# Question 1

- Let's map the attribute names to simpler letters:
  - patient-id to A
  - patient-name to B
  - doctor-id to C
  - doctor-name to D
  - appointment-date to E
  - appointment-time to F
  - clinic-room-no to G
- Since A is a key, we have default functional dependency:

  $A \rightarrow BCDEFG$, assuming no other FDs.
- Does this FD form a minimal basis (MB)?

# Question 1

- Since A is a key, we have default functional dependency: $A \rightarrow BCDEFG$, assuming no other FDs.
- Does this FD form a minimal basis (MB)?
- Condition 1 of MB: RHS is single attribute. So we break the FD into: $A \rightarrow B$, $A \rightarrow C$, $A \rightarrow D$, $A \rightarrow E$, $A \rightarrow F$, $A \rightarrow G$
- Conditions 2 and 3 are satisfied since there are no other FDs to reason.
- So MB = $\{A \rightarrow B, A \rightarrow C, A \rightarrow D, A \rightarrow E, A \rightarrow F, A \rightarrow G\}$ and we can form relations AB, AC, AD, AE, AF, AG from it.
- Do these relations make sense? Any other way to do this?

# Question 1

- Suppose other than A→BCDEFG, we derive other FDs using common sense:
  - doctor-id → doctor-name; i.e., C→D
  - appointment-date, appointment-time, clinic-room-no → patient-id, doctor-id; i.e., EFG→AC
- So altogether we have: A→BCDEFG, C→D, EFG→AC
- Let's check MB conditions: Condition 1 says RHS must be single attribute, so we have: A→B, A→C, A→D, A→E, A→F, A→G, C→D, **EFG→A, EFG→C**
- Condition 2 says no redundant FDs. **EFG→C** and A→D are redundant and removed. Condition 3 says no redundant LHS attributes. None of LHS attribute of EFG→A is redundant.
- So MB={A→B, A→C, A→E, A→F, A→G, C→D, EFG→A} and we form relations AB, AC, AE, AF, AG, CD, AEFG.
- Do these relations make sense? Any other way to do this?

# Question 1

- Suppose we have the following FDs using common sense:
  - patient-id → patient-name; i.e., $A \to B$
  - doctor-id → doctor-name; i.e., $C \to D$
  - appointment-date, appointment-time, clinic-room-no → patient-id, doctor-id; i.e., $EFG \to AC$
- So altogether we have: $A \to B$, $C \to D$, $EFG \to AC$
- Let's check MB conditions: Condition 1 says RHS must be single attribute, so we have: $A \to B$, $C \to D$, $EFG \to A$, $EFG \to C$
- Condition 2 says no redundant FDs. There are none. Condition 3 says no redundant LHS attributes. There are none.
- So MB=$\{A \to B, C \to D, EFG \to A, EFG \to C\}$ and we form relations AB, CD, ACEFG.
- Do these relations make sense? Yes. This last approach seems the best.

# Question 1

- Using common sense, we derive functional dependencies:
  - FD1: patient-id $\rightarrow$ patient-name
  - FD2: doctor-id $\rightarrow$ doctor-name
  - FD3: appointment-date, appointment-time, clinic-room-no $\rightarrow$ patient-id, doctor-id
- Using 3NF normalization, we have decomposed relations:
  - PATIENT( <u>patient-id</u>, patient-name )
  - DOCTOR( <u>doctor-id</u>, doctor-name )
  - APPOINTMENT( <u>appointment-date, appointment-time, clinic-room-no</u> , patient-id, doctor-id )

# Question 2

Consider the relation Courses(C, T, H, R, S, G) whose attributes may be thought informally as course, teacher, hour, room, student, and grade. Let the set of FD's of Courses be:

$C \rightarrow T$, $HR \rightarrow C$, $HT \rightarrow R$, $HS \rightarrow R$, and $CS \rightarrow G$.

(a) What are all the keys for Courses?

(b) Verify that the given FDs are their own minimal basis.

(c) Use the 3NF decomposition algorithm to find a lossless-join, dependency-preserving decomposition.

# Question 2(a)

- The usual procedure to find keys is to take the closure of all 63 nonempty subsets.
- However, we notice that none of the right sides of the FDs contains attributes H and S; we may conclude that **H** and **S** must be part of any key.
- Given FDs $C \rightarrow T$, $HR \rightarrow C$, $HT \rightarrow R$, $HS \rightarrow R$, $CS \rightarrow G$, let's start with HS.
  - $HS \rightarrow R \Rightarrow HS \rightarrow HR$ and $HR \rightarrow C \Rightarrow HS \rightarrow C$; $HS^+=\{CHRS\}$
  - $CS \rightarrow G$ and $C \rightarrow T \Rightarrow$ **$HS^+=\{CGHRST\}$**
- Using the closure method, we eventually find out that **HS** is the only key in the Courses relation.

# Question 2(b)

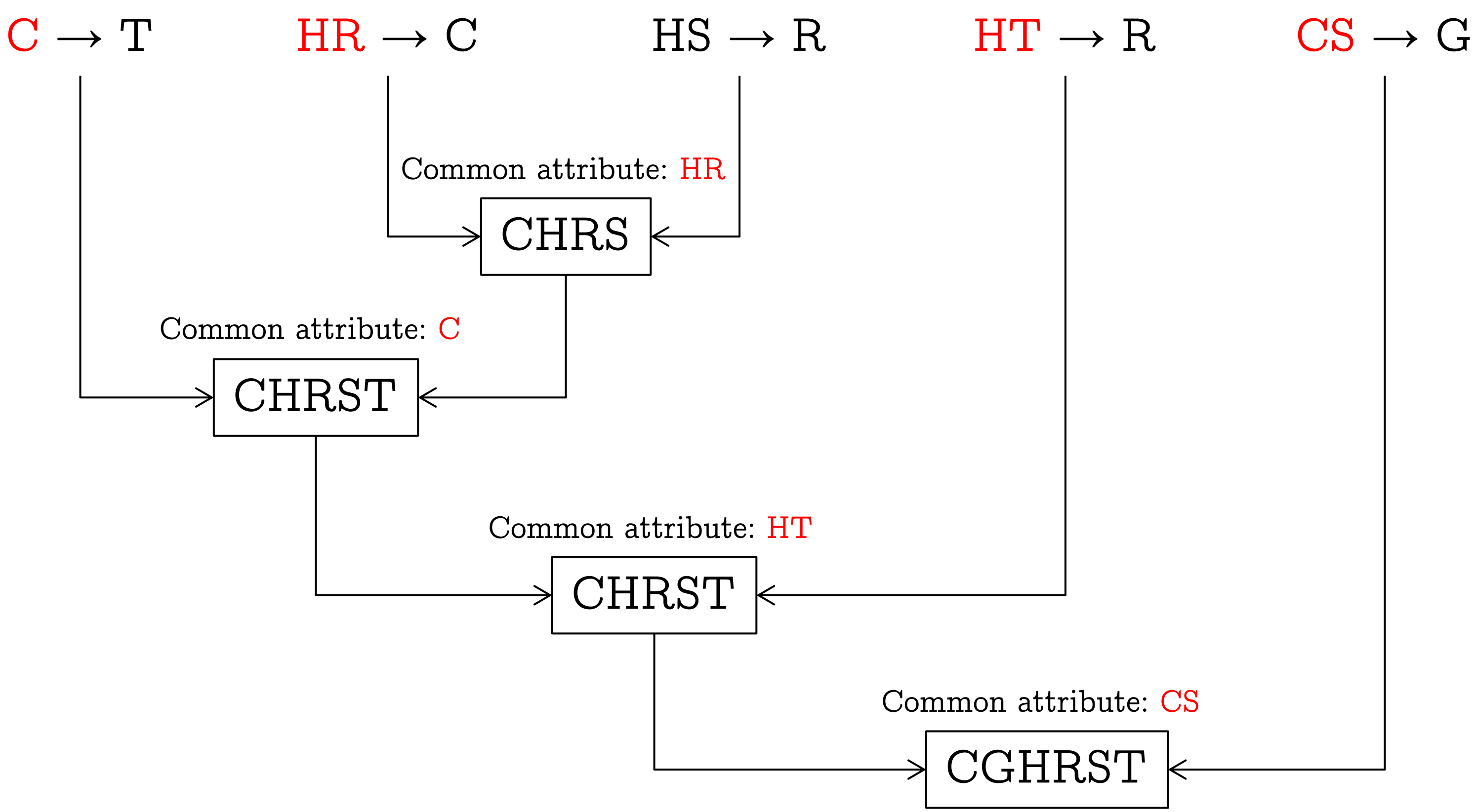
- Given FDs: C→T, HR→C, HT→R, HS→R, CS→G
- Check if any of the FDs is redundant.
  - None. If we remove any one of the five FDs, the remaining four FDs do not imply the removed FD.
- Check if any of the LHS attribute of an FD can be removed without losing the dependencies.
  - None. The attributes on the left side of the four FDs are not redundant.
- Thus, the given set of FDs is a minimal basis.

# Question 2(c)

- Since the only key is HS, the given set of FDs has some dependencies that violate 3NF.
  - Violating FDs: $C \rightarrow T$, $HR \rightarrow C$, $HT \rightarrow R$, $CS \rightarrow G$
- We also know that the given set of FDs is a minimal basis. Thus, the decomposed relations are (CT), (HRC), (HTR), (HSR) and (CSG).
- Since the relation HSR contains a key, we do not need to add an additional relation. The final set of decomposed relations is (CT), (HRC), (HTR), (HSR) and (CSG).
- Since each decomposed relation came from a FD, the decomposition is FD preserving.

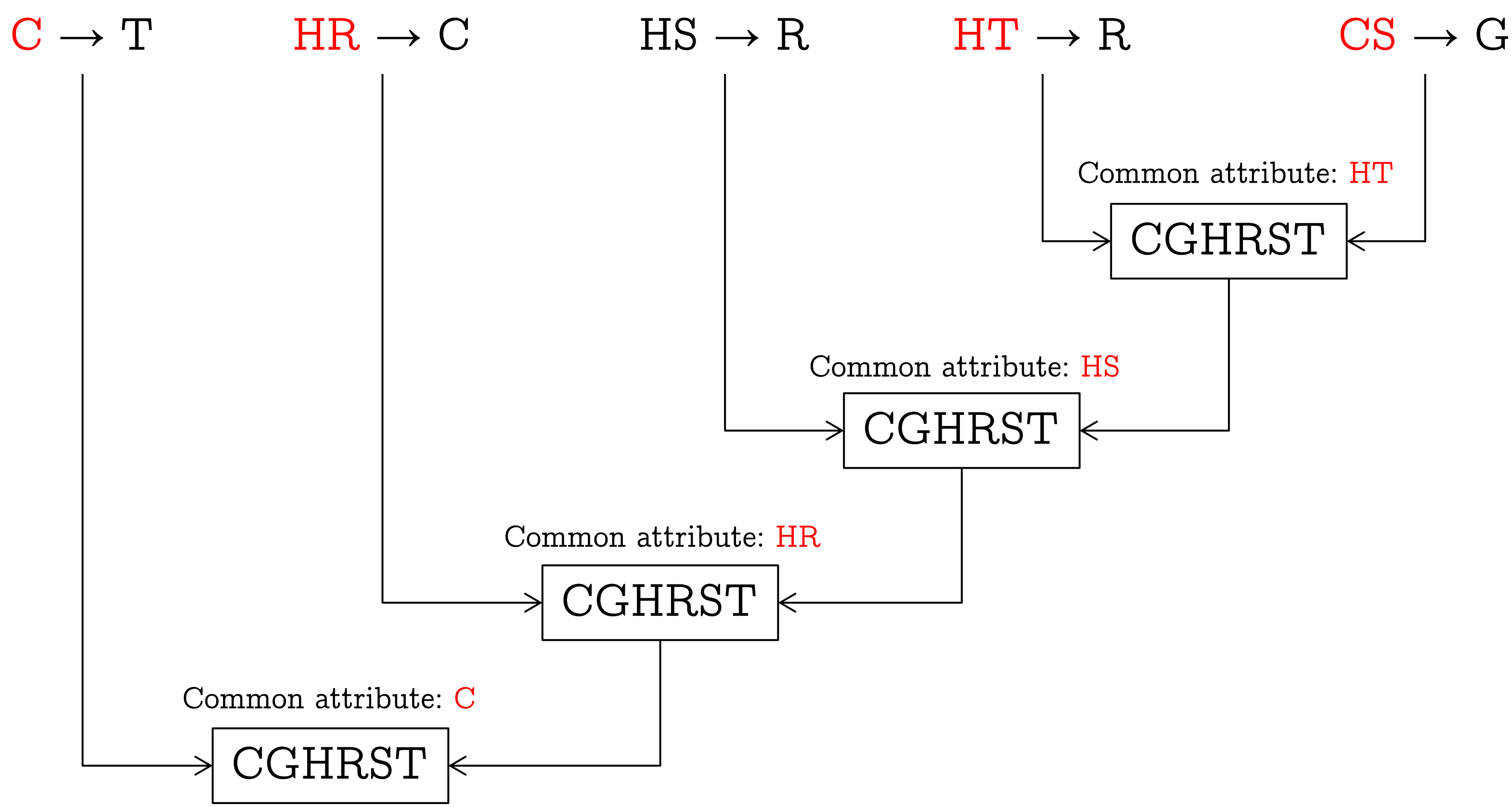
# Question 2(c)

The following sequence of joins shows that the decomposition is lossless:



# Question 2(c)

Another sequence of joins to show that the decomposition is lossless:



# Question 3

Consider a relation R(W, X, Y, Z) which satisfies the following set of FDs G = {Z→W, Y→X, Y→Z, XW→Y}, where G is a minimal basis.

(a) Decompose R into a set of relations in 3NF.

(b) Is the decomposition also in BCNF? Explain your answer.

# Question 3(a)

- Given G ={Z→W, Y→X, Y→Z, XW→Y}, where G is a minimal basis
  - Decomposed relations: R1(Z, W), R2(X, Y, Z), R3(X, Y, W)
- To determine whether BCNF satisfied, find keys using reasoning:
  - Y→X, Y→Z, Z→W yield Y→WXYZ; so Y is a key
  - XW→Y and Y is a key means XW is a key.
  - XW is a key and Z→W yields XZ→XW; so XZ is a key.
  - Keys: Y, WX, XZ

# Question 3(b)

- Keys: Y, WX, XZ; check FDs in R1(Z,W), R2(X,Y,Z), R3(X,Y,W):
  - FDs in R1: 2-attribute relation is in BCNF
  - FDs in R2: $Y \to X$, $Y \to Z$; LHS are keys
  - FDs in R3: $Y \to X$, $XW \to Y$; LHS are keys
- Since all LHS of FDs are keys; relations are in BCNF

# Question 4

Consider a relation R(A,B,C,D,E) and FD's $A \rightarrow BC$, $CD \rightarrow E$, $E \rightarrow A$, and $B \rightarrow D$.

(a) Is the decomposition R1(A,B,C) and R2(A,D,E) of R lossless or lossy? Justify your answer. Is this decomposition dependency preserving? If your answer is NO, then what is not preserved?

(b) Is the decomposition R3(A,B,C,D) and R4(C,D,E) of R lossless or lossy? Justify your answer. Is this decomposition dependency preserving? If your answer is NO, then what is not preserved?

# Question 4

- FDs: $A \rightarrow BC$, $E \rightarrow A$, $CD \rightarrow E$, $B \rightarrow D$
- (a) Decomposition: R1(A,B,C) and R2(A,D,E)
- (b) Decomposition: R3(A,B,C,D) and R4(C,D,E)
- Is decomposition lossless (can be joined back)?
- Is decomposition dependency preserving (no FDs lost)?
- Keys: A, E, CD, BC

# Question 4(a)

- Decomposition: R1(A,B,C) and R2(A,D,E); FDs: $A \rightarrow BC$, $E \rightarrow A$, $CD \rightarrow E$, $B \rightarrow D$; Keys: A, E, CD, BC
- Decomposition R1(A,B,C) and R2(A,D,E) is lossless because
  - R1 and R2 have a common attribute A, and
  - A is a superkey for R1(A,B,C)
- FDs that hold on R1(A,B,C)
  - $A \rightarrow BC$, $BC \rightarrow A$ since R1 contains A, B, and C
- FDs that hold on R2(A,D,E)
  - $E \rightarrow A$, $A \rightarrow E$ since R2 contains A and E
- Two other FDs need to be checked: $CD \rightarrow E$, $B \rightarrow D$
  - From $A \rightarrow BC$, $BC \rightarrow A$, $E \rightarrow A$, $A \rightarrow E$, we have:
  - $\{B\}+ = \{B\}$, so $B \rightarrow D$ is not preserved
  - $\{CD\}+ = \{CD\}$, so $CD \rightarrow E$ is not preserved
- Decomposition is NOT dependency-preserving

# Question 4(b)

- Decomposition: R3(A,B,C,D) and R4(C,D,E); FDs: $A \to BC$, $E \to A$, $CD \to E$, $B \to D$; Keys: A, E, CD, BC
- Decomposition R3(A,B,C,D) and R4(C,D,E) is lossless because
  - R3 and R4 have common attributes CD, and
  - CD is a superkey for R4(C,D,E)
- FDs that hold on R3(A,B,C,D)
  - $A \to BC$, $BC \to A$, $\cancel{CD \to E}$, $CD \to A$, and $B \to D$ since R1 contains A, B, C, and D
- FDs that hold on R4(C,D,E)
  - $CD \to E$, $E \to CD$
- One other FD needs to be checked: $E \to A$
  - From $A \to BC$, $BC \to A$, $CD \to A$, $B \to D$, $CD \to E$, $E \to CD$ and, we have:
  - $\{E\}+ = \{A, B, C, D, E\}$, $E \to CD$ holds in R4, $CD \to A$ holds in R3, $E \to A$ is preserved
- Decomposition is dependency-preserving


Do the best
you can until you know better.
Then when you know better, do better.
~ Maya Angelou
ALL THE BEST