

# طراحی پایگاه داده رابطه ای

قمرناز تدین

## بستار مجموعه خصیصه

► برای مجموعه خصیصه های  $\alpha$ ، بستار  $\alpha$  تحت  $F$  (که با  $\alpha^+$  نشان داده می شود) مجموعه خصیصه هایی است که از نظر تابعی با  $\alpha$  و تحت  $F$  تعیین می شوند.

►  $R = (A, B, C, G, H, I)$

►  $F = \{A \rightarrow B$

$A \rightarrow C$

$CG \rightarrow H$

$CG \rightarrow I$

$B \rightarrow H\}$

►  $(AG)^+$

1.  $result = AG$

2.  $result = ABCG (A \rightarrow C \text{ and } A \rightarrow B)$

3.  $result = ABCGH \quad (CG \rightarrow H \text{ and } CG \subseteq AGBC)$

4.  $result = ABCGHI \quad (CG \rightarrow I \text{ and } CG \subseteq AGBCH)$

► آیا  $AG$  کلید کاندیدا است؟

1. آیا  $AG$  سوپرکلید است؟

1. Does  $AG \rightarrow R? == \text{Is } (AG)^+ \supseteq R$

2. آیا زیرمجموعه ای از  $AG$  سوپرکلید است؟

1. Does  $A \rightarrow R? == \text{Is } (A)^+ \supseteq R$

2. Does  $G \rightarrow R? == \text{Is } (G)^+ \supseteq R$

# کاربرد بستار خصیصه

## ► بررسی سوپرکلید:

- برای بررسی اینکه آیا  $\alpha$  سوپرکلید است،  $\alpha^+$  را محاسبه میکنیم و بررسی میکنیم آیا شامل همه خصیصه های  $R$  هست..

## ► بررسی وابستگی تابعی

- برای بررسی برقراری وابستگی تابعی  $\alpha \rightarrow \beta$  باید  $\alpha^+ \subseteq \beta^+$  بررسی شود.

## ► محاسبه بستار $F$

- برای هر  $\gamma \subseteq R$  بستار  $\gamma^+$  را محاسبه میکنیم و برای هر  $S \subseteq \gamma^+$  وابستگی تابعی  $\gamma \rightarrow S$  را ایجاد میکنیم.

# تجزیه بدون از دست دادن اطلاعات

► برای حالتی که  $R = (R_1, R_2)$ ، باید برای همه رابطه های ممکن  $r$  با شمای  $R$  داشته باشیم:

$$r = \Pi_{R_1}(r) \bowtie \Pi_{R_2}(r)$$

► تجزیه  $R$  به  $R_1$  و  $R_2$  بدون از دست دادن اطلاعات است اگر حداقل یکی از وابستگیهای زیر در  $F^+$  وجود داشته باشد:

- $R_1 \cap R_2 \rightarrow R_1$
- $R_1 \cap R_2 \rightarrow R_2$

# Example

▶  $R = (A, B, C)$

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

◦ میتواند به دو شکل تجزیه شود:

▶  $R_1 = (A, B), \quad R_2 = (B, C)$

◦ تجزیه بدون از دست دادن اطلاعات

$$R_1 \cap R_2 = \{B\} \text{ and } B \rightarrow BC$$

◦ حفظ وابستگیها

▶  $R_1 = (A, B), \quad R_2 = (A, C)$

◦ تجزیه بدون از دست دادن اطلاعات

$$R_1 \cap R_2 = \{A\} \text{ and } A \rightarrow AB$$

◦ وابستگیها حفظ نمی شوند.

بدون محاسبه  $R_1 \bowtie R_2$  نمی توان  $B \rightarrow C$  را بررسی کرد.

## حفظ وابستگیها

► اگر  $F_i$  مجموعه ای در وابستگیهای  $F^+$  باشد که فقط شامل خصیصه های  $R_i$  است، تجزیه «وابستگیها را حفظ میکند» اگر:

$$(F_1 \cup F_2 \cup \dots \cup F_n)^+ = F^+$$

# Example

▶  $R = (A, B, C)$

$F = \{A \rightarrow B$

$B \rightarrow C\}$

Key =  $\{A\}$

▶  $R$  is not in BCNF

▶ Decomposition  $R_1 = (A, B), R_2 = (B, C)$

- $R_1$  and  $R_2$  in BCNF
- Lossless-join decomposition
- Dependency preserving



## مثال تجزیه BCNF

▶  $R = (A, B, C)$

$F = \{A \rightarrow B$

$B \rightarrow C\}$

Key =  $\{A\}$

▶  $R$  در BCNF نیست. ( $B \rightarrow C$  اما  $B$  سوپرکلید نیست)

▶ تجزیه

◦  $R_1 = (B, C)$

◦  $R_2 = (A, B)$

# مثال تجزیه BCNF

- ▶ *class (course\_id, title, dept\_name, credits, sec\_id, semester, year, building, room\_number, capacity, time\_slot\_id)*

▶ وابستگیهای تابعی

- *course\_id* → *title, dept\_name, credits*
- *building, room\_number* → *capacity*
- *course\_id, sec\_id, semester, year* → *building, room\_number, time\_slot\_id*

▶ یک کلید کاندیدا:

*{course\_id, sec\_id, semester, year}*.

▶ تجزیه BCNF:

- *course\_id* → *title, dept\_name, credits*
  - اما *course\_id* سوپرکلید نیست.
  - Class به صورت زیر جایگزین می شود:
- *course(course\_id, title, dept\_name, credits)*
- *class-1 (course\_id, sec\_id, semester, year, building, room\_number, capacity, time\_slot\_id)*

## مثال تجزیه BCNF (ادامه)

- ▶ *course* is in BCNF

▶ *class-1* در *building, room\_number* → *capacity* برقرار است.

◦ اما  $\{building, room\_number\}$  سوپرکلید در *class-1* نیست.

◦ *class-1* به صورت زیر جایگزین میشود:

- *classroom* (*building, room\_number, capacity*)
- *section* (*course\_id, sec\_id, semester, year, building, room\_number, time\_slot\_id*)

- ▶ *classroom* and *section* are in BCNF.

# 3NF Example

## ▶ Relation *dept\_advisor*.

- *dept\_advisor*(*s\_ID*, *i\_ID*, *dept\_name*)  
 $F = \{s\_ID, dept\_name \rightarrow i\_ID, i\_ID \rightarrow dept\_name\}$
- Two candidate keys: *s\_ID*, *dept\_name*, and *i\_ID*, *s\_ID*
- *R* is in 3NF
  - $s\_ID, dept\_name \rightarrow i\_ID \quad s\_ID$ 
    - *dept\_name* is a superkey
  - $i\_ID \rightarrow dept\_name$ 
    - *dept\_name* is contained in a candidate key

# 3NF Decomposition: An Example

- ▶ Relation schema:

*cust\_banker\_branch = (customer\_id, employee\_id, branch\_name, type)*

- ▶ The functional dependencies for this relation schema are:

*1. customer\_id, employee\_id → branch\_name, type*

*2. employee\_id → branch\_name*

*3. customer\_id, branch\_name → employee\_id*

*(customer\_id, employee\_id, type)*

*(customer\_id, branch\_name, employee\_id)*

# وابستگی چندمقداری Multivalued Dependencies

► نام فرزندان و شماره تلفنهای استاد:

- *inst\_child*(*ID*, *child\_name*)
- *inst\_phone*(*ID*, *phone\_number*)

► از ترکیب دو رابطه فوق:

- *inst\_info*(*ID*, *child\_name*, *phone\_number*)
- Example data:
  - (99999, David, 512-555-1234)
  - (99999, David, 512-555-4321)
  - (99999, William, 512-555-1234)
  - (99999, William, 512-555-4321)

► رابطه در BCNF است.

# Multivalued Dependencies (MVDs)

► اگر  $R$  شمای رابطه باشد و  $\alpha \subseteq R$  و  $\beta \subseteq R$  ، وابستگی چندمقداری زیر در  $R$  برقرار است:

$$\alpha \twoheadrightarrow \beta$$

اگر در هر رابطه منطقی  $r(R)$  برای همه جفت رکوردهای  $t_1$  و  $t_2$  در  $r$  که  $t_1[\alpha] = t_2[\alpha]$  ، رکوردهای  $t_3$  و  $t_4$  در  $r$  باشند به طوری که:

$$t_1[\alpha] = t_2[\alpha] = t_3[\alpha] = t_4[\alpha]$$

$$t_3[\beta] = t_1[\beta]$$

$$t_3[R - \beta] = t_2[R - \beta]$$

$$t_4[\beta] = t_2[\beta]$$

$$t_4[R - \beta] = t_1[R - \beta]$$

# MVD (Cont.)

►  $\alpha \twoheadrightarrow \beta$

	$\alpha$	$\beta$	$R - \alpha - \beta$
$t_1$	$a_1 \dots a_i$	$a_{i+1} \dots a_j$	$a_{j+1} \dots a_n$
$t_2$	$a_1 \dots a_i$	$b_{i+1} \dots b_j$	$b_{j+1} \dots b_n$
$t_3$	$a_1 \dots a_i$	$a_{i+1} \dots a_j$	$b_{j+1} \dots b_n$
$t_4$	$a_1 \dots a_i$	$b_{i+1} \dots b_j$	$a_{j+1} \dots a_n$



# Example

► فرض کنید  $R$  یک شمای رابطه ای باشد که خصیصه های آن به سه مجموعه غیرتهی تقسیم شده اند:  
 $Y, Z, W$

► میگوییم  $Y \twoheadrightarrow Z$  (  $Y$  به طور چندمقداری  $Z$  را تعیین میکند) اگر و تنها اگر برای همه روابط ممکن  $r(R)$ :

$$\langle y_1, z_1, w_1 \rangle \in r \text{ and } \langle y_1, z_2, w_2 \rangle \in r$$

then

$$\langle y_1, z_1, w_2 \rangle \in r \text{ and } \langle y_1, z_2, w_1 \rangle \in r$$

►  $Y \twoheadrightarrow Z$  if  $Y \twoheadrightarrow W$

# Example (Cont.)

▶ در مثال:

$ID \twoheadrightarrow child\_name$

$ID \twoheadrightarrow phone\_number$

# شکل نرمال چهارم Fourth Normal Form

► شمای رابطه ای  $R$  با مجموعه وابستگیهای تابعی و چند مقداری  $D$  در **4NF** است اگر برای همه وابستگیهای چندمقداری در  $D^+$  به شکل  $\beta \twoheadrightarrow \alpha$ ، که  $\alpha \subseteq R$  و  $\beta \subseteq R$ ، حداقل یکی از موارد زیر برقرار باشد:

- $\beta \subseteq \alpha$  or  $\alpha \cup \beta = R$  بدیهی باشد
- $\alpha$  برای شمای  $R$  سوپرکلید باشد.
- اگر رابطه در 4NF باشد در BCNF هم هست.

►  $R = (A, B, C, G, H, I)$

$F = \{ A \twoheadrightarrow B$

$B \twoheadrightarrow HI$

$CG \twoheadrightarrow H \}$

►  $R$  در 4NF نیست چون  $A \twoheadrightarrow B$  و  $A$  سوپرکلید نیست.

► تجزیه:

a)  $R_1 = (A, B)$  ( $R_1$  is in 4NF)

b)  $R_2 = (A, C, G, H, I)$  ( $R_2$  is not in 4NF, decompose into  $R_3$  and  $R_4$ )

c)  $R_3 = (C, G, H)$  ( $R_3$  is in 4NF)

d)  $R_4 = (A, C, G, I)$  ( $R_4$  is not in 4NF, decompose into  $R_5$  and  $R_6$ )

$A \twoheadrightarrow I$  و بنابراین (تعدی MVD) ،  $B \twoheadrightarrow HI \Rightarrow A \twoheadrightarrow HI$  و  $A \twoheadrightarrow B$  (در  $R_4$ )

e)  $R_5 = (A, I)$  ( $R_5$  is in 4NF)

f)  $R_6 = (A, C, G)$  ( $R_6$  is in 4NF)