

# طراحی پایگاه داده رابطه ای

قمرناز تدین

## ترکیب شماها

- ▶ فرض کنید دو رابطه *instructor* و *department* را در یک رابطه *inst\_dept* ترکیب کنیم.
- ▶ در نتیجه احتمال تکرار اطلاعات وجود دارد.

ID	name	salary	dept_name	building	budget
22222	Einstein	95000	Physics	Watson	70000
12121	Wu	90000	Finance	Painter	120000
32343	El Said	60000	History	Painter	50000
45565	Katz	75000	Comp. Sci.	Taylor	100000
98345	Kim	80000	Elec. Eng.	Taylor	85000
76766	Crick	72000	Biology	Watson	90000
10101	Srinivasan	65000	Comp. Sci.	Taylor	100000
58583	Califieri	62000	History	Painter	50000
83821	Brandt	92000	Comp. Sci.	Taylor	100000
15151	Mozart	40000	Music	Packard	80000
33456	Gold	87000	Physics	Watson	70000
76543	Singh	80000	Finance	Painter	120000

## تعیین شماهای کوچکتر

► فرض کنید رابطه *inst\_dept* را داریم. چگونه میتوانیم تشخیص دهیم که میتوان این رابطه را به دو رابطه *instructor* و *department* شکست؟

► یک قانون: اگر این شما را داشته باشیم (*dept\_name, building, budget*) آنگاه *dept\_name* می تواند کلید کاندیدا باشد.

► به عنوان وابستگی تابعی نشان داده می شود: *dept\_name* → *building, budget*

► در *inst\_dept* به دلیل این که *dept\_name* کلید کاندیدا نیست، بنابراین ساختمان و بودجه یک گروه ممکن است تکرار شود.  
◦ بنابراین لازم است که رابطه *inst\_dept* شکسته شود.

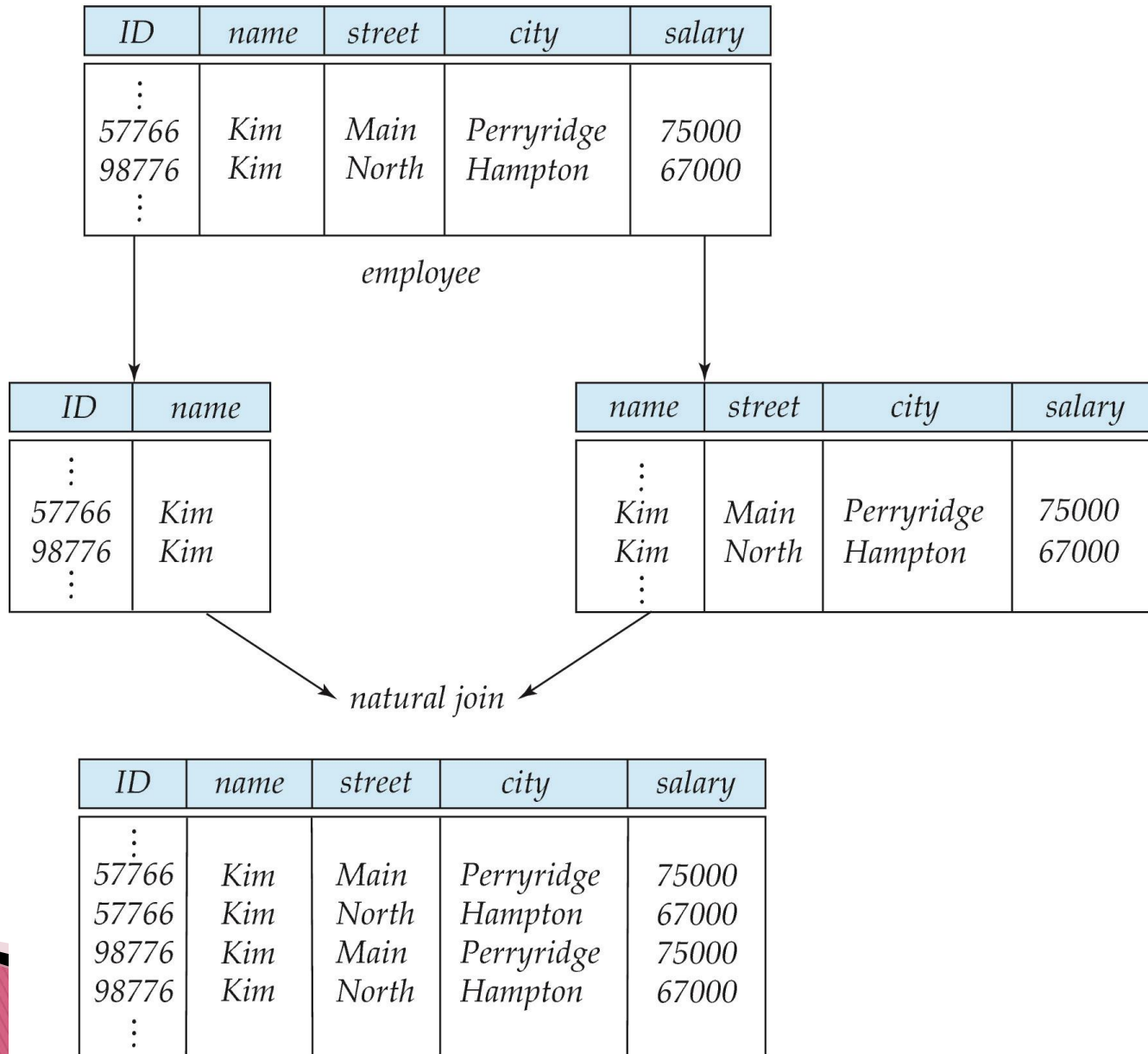
► همیشه شکستن رابطه مفید نیست. فرض کنید رابطه کارمند  
*employee(ID, name, street, city, salary)* را به دو  
رابطه بشکنیم:

*employee1 (ID, name)*

*employee2 (name, street, city, salary)*

► نشان می دهیم که چگونه اطلاعات را از دست می دهیم، به این معنی  
که نمیتوانیم مجدداً رابطه اصلی *employee* را به دست آوریم.

# تجزیه با از دست دادن اطلاعات A Lossy Decomposition



# مثالی از تجزیه بدون از دست دادن اطلاعات

## ► Lossless join decomposition

►  $R = (A, B, C)$

$R_1 = (A, B) \quad R_2 = (B, C)$

$A$	$B$	$C$
$\alpha$	1	A
$\beta$	2	B

$r$

$A$	$B$
$\alpha$	1
$\beta$	2

$\Pi_{A,B}(r)$

$B$	$C$
1	A
2	B

$\Pi_{B,C}(r)$

$\Pi_A(r) \bowtie \Pi_B(r)$

$A$	$B$	$C$
$\alpha$	1	A
$\beta$	2	B

# شکل نرمال اول First Normal Form

▶ دامنه، اتمیک (atomic) است اگر عناصر آن واحدهای مجزایی باشند.

◦ مثال دامنه های غیر اتمیک:

• مجموعه نامها، خصیصه های ترکیبی.

• شماره های شناسایی مثل CS101 که میتواند به بخشهایش شکسته شود.

▶ یک شمای رابطه ای  $R$  در شکل نرمال اول است اگر دامنه همه خصیصه های  $R$  اتمیک باشد.

▶ مقادیر غیر اتمیک باعث پیچیدگی ذخیره سازی اطلاعات می شوند و باعث افزایش احتمال اطلاعات تکراری می شوند.

◦ مثال: مجموعه حسابها برای هر مشتری ذخیره شود و مجموعه صاحبان حساب برای هر حساب ذخیره شود.

◦ فرض میکنیم همه رابطه ها در شکل نرمال اول باشند.

# هدف: تعیین یک تئوری برای مورد زیر:

- ▶ تصمیم گیری که آیا رابطه  $R$  به شکل «خوبی» طراحی شده؟
- ▶ اگر رابطه  $R$  در شکل خوبی نباشد، آن را به مجموعه ای از روابط  $\{R_1, R_2, \dots, R_n\}$  میشکنیم به طوری که:
  - هر رابطه در شکل خوبی باشد.
  - تجزیه به صورت lossless-join باشد.
- ▶ این تئوری براساس مفاهیم زیر است:
  - وابستگی تابعی functional dependencies
  - وابستگی چند مقداری multivalued dependencies



# وابستگی های تابعی

- ▶ محدودیت بر مجموعه روابط منطقی
- ▶ لازم است که مقدار مجموعه معینی از خصیصه ها به طور یکتا مقدار مجموعه دیگری از خصیصه ها را تعیین کند.
- ▶ وابستگی تابعی، تعمیمی از تعریف کلید است.

# وابستگی های تابعی (ادامه)

► فرض کنید  $r$  شمای رابطه باشد

$$\alpha \subseteq R \text{ and } \beta \subseteq R$$

► وابستگی تابعی  $\alpha \rightarrow \beta$  در  $R$  برقرار است اگر و تنها اگر برای هر رابطه  $r(R)$  اگر در دو رکورد  $t_1$  و  $t_2$  مقدار خصیصه های  $\alpha$  برابر باشد، مقدار خصیصه های  $\beta$  نیز در آنها برابر باشد.

$$t_1[\alpha] = t_2[\alpha] \Rightarrow t_1[\beta] = t_2[\beta]$$

► مثال:  $r(A, B)$  را با نمونه  $r$  زیر در نظر بگیرید:

1	4
1	5
3	7

► در این رابطه  $A \rightarrow B$  برقرار نیست ولی  $B \rightarrow A$  برقرار است.

►  $A \rightarrow C$

$A$	$B$	$C$	$D$
$a_1$	$b_1$	$c_1$	$d_1$
$a_1$	$b_2$	$c_1$	$d_2$
$a_2$	$b_2$	$c_2$	$d_2$
$a_2$	$b_3$	$c_2$	$d_3$
$a_3$	$b_3$	$c_2$	$d_4$

# وابستگی های تابعی (ادامه)

►  $K \rightarrow R$  یک سوپرکلید برای شمای  $R$  است اگر و تنها اگر  $K \rightarrow R$

►  $K$  کلید کاندیدا برای  $R$  است اگر و تنها اگر:

◦  $K \rightarrow R$  و

◦ برای هیچ  $\alpha \subset K$  رابطه  $\alpha \rightarrow R$  را نداشته باشیم.

► شمای زیر را در نظر بگیرید:

*inst\_dept* (ID, name, salary, dept\_name, building, budget).

این وابستگیهای تابعی برقرارند:

$dept\_name \rightarrow building$  و  $ID \rightarrow building$

اما رابطه زیر برقرار نیست:

$dept\_name \rightarrow salary$

# وابستگی های تابعی (ادامه)

► وابستگی تابعی، بدیهی است اگر در همه نمونه های رابطه برقرار باشد.

◦ Example:

- $ID, name \rightarrow ID$
- $name \rightarrow name$

◦ در حالت کلی  $\alpha \rightarrow \beta$  بدیهی است اگر  $\beta \subseteq \alpha$

- ▶ *Room\_number* → *capacity* ممکن است برقرار نباشد
- ▶ *building, room number* → *capacity*

<i>building</i>	<i>room_number</i>	<i>capacity</i>
Packard	101	500
Painter	514	10
Taylor	3128	70
Watson	100	30
Watson	120	50

# بستار مجموعه وابستگیهای تابعی

- ▶ اگر مجموعه وابستگیهای تابعی  $F$  وجود داشته باشند، وابستگیهای دیگری نیز وجود دارند که توسط  $F$  برقرار است.
  - مثال: از  $A \rightarrow B$  و  $B \rightarrow C$  میتوانیم نتیجه بگیریم:  $A \rightarrow C$
- ▶ مجموعه همه وابستگیهای تابعی که توسط  $F$  برقرار میشوند، بستار  $F$  گفته می شود.
- ▶ بستار  $F$  را با  $F^+$  نشان می دهیم که یک superset از  $F$  است.

# شکل نرمال Boyce-Codd

یک شمای رابطه ای  $R$  با توجه به مجموعه وابستگیهای تابعی  $F$  در BCNF است اگر برای همه وابستگیها در  $F^+$  به شکل  $\alpha \rightarrow \beta$  که  $\alpha \subseteq R$  and  $\beta \subseteq R$  ، حداقل یکی از روابط زیر برقرار باشند:

▶  $\alpha \rightarrow \beta$  بدیهی باشد ( $\beta \subseteq \alpha$ )

▶  $\alpha$  یک سوپرکلید برای  $R$  باشد.

مثال: شمای زیر در BCNF نیست:

*instr\_dept* (ID, name, salary, dept\_name, building, budget)

زیرا  $dept\_name \rightarrow building, budget$  برقرار است ولی  $dept\_name$  سوپر کلید نیست.



# تجزیه شما به شکل BCNF

► فرض کنید شما  $R$  و وابستگی غیر بدیهی  $\alpha \rightarrow \beta$  را داریم که باعث نقض BCNF شده است.  
 $R$  را به شکل زیر تجزیه میکنیم:

- $(\alpha \cup \beta)$
- $(R - (\beta - \alpha))$

► در مثال

- $instr\_dept(\underline{ID}, name, salary, dept\_name, building, budget)$
  - $\alpha = dept\_name$
  - $\beta = building, budget$
- $inst\_dept$  با روابط زیر جایگزین می شود.
- $(\alpha \cup \beta) = (dept\_name, building, budget)$
  - $(R - (\beta - \alpha)) = (ID, name, salary, dept\_name)$

# شکل نرمال سوم Third Normal Form

- ▶ شمای رابطه (R) در شکل نرمال سوم (3NF) است اگر برای همه  $\alpha \rightarrow \beta$  در  $F^+$  حداقل یکی از روابط زیر برقرار باشند:
  - $\alpha \rightarrow \beta$  بدیهی باشد ( $\beta \in \alpha$ )
  - $\alpha$  برای R سوپرکاید باشد.
  - هر خصیصه A در  $\beta - \alpha$  در یکی از کلیدهای کاندیدای R وجود داشته باشد.
- ▶ اگر رابطه در BCNF باشد در 3NF هم هست.

# اهداف نرمال سازی

- ▶ فرض کنید  $R$  یک شمای رابطه با مجموعه وابستگی های تابعی  $F$  باشد.
- ▶ آیا شمای رابطه  $R$  در شکل خوبی است؟
- ▶ اگر شمای رابطه  $R$  در شکل خوبی نباشد، به مجموعه شماهای  $\{R_1, R_2, \dots, R_n\}$  شکسته میشود به طوری که:
  - هر شما در شکل خوبی باشد.
  - تجزیه lossless-join decomposition باشد.
  - ترجیحا تجزیه باید وابستگیها را حفظ کند.

# بررسی BCNF

▶ شماهایی هستند که در BCNF هستند ولی به اندازه کافی نرمال نیستند.

▶ این رابطه را در نظر بگیرید:

*inst\_info (ID, child\_name, phone)*

◦ که استاد ممکن است چند فرزند و چند شماره تلفن داشته باشد.

<i>ID</i>	<i>child_name</i>	<i>phone</i>
99999	David	512-555-1234
99999	David	512-555-4321
99999	William	512-555-1234
99999	Willian	512-555-4321

*inst\_info*

## بررسی BCNF (ادامه)

- ▶ همه وابستگیها بدیهی هستند و رابطه در BCNF است.
- ▶ آنومالی درج: اگر شماره 981-992-3443 را به 99999 اضافه کنیم باید دو رکورد اضافه کنیم:  
(99999, David, 981-992-3443)  
(99999, William, 981-992-3443)

<i>ID</i>	<i>child_name</i>	<i>phone</i>
99999	David	512-555-1234
99999	David	512-555-4321
99999	William	512-555-1234
99999	Willian	512-555-4321

*inst\_info*

## بررسی BCNF (ادامه)

► بنابراین بهتر است رابطه به صورت زیر شکسته شود:

<i>inst_child</i>	<i>ID</i>	<i>child_name</i>
	99999	David
	99999	David
	99999	William
	99999	Willian

<i>inst_phone</i>	<i>ID</i>	<i>phone</i>
	99999	512-555-1234
	99999	512-555-4321
	99999	512-555-1234
	99999	512-555-4321

بنابراین شکلهای نرمال بالاتری لازم است.

# مجموعه وابستگی‌های تابعی

► بستار  $F$  که با  $F^+$  نشان می‌دهیم را با بکارگیری قوانین زیر می‌توان به دست آورد:

## اصول آرمسترانگ **Armstrong's Axioms**

- if  $\beta \subseteq \alpha$ , then  $\alpha \rightarrow \beta$  (انعکاسی)
- if  $\alpha \rightarrow \beta$ , then  $\gamma \alpha \rightarrow \gamma \beta$  (افزونی)
- if  $\alpha \rightarrow \beta$ , and  $\beta \rightarrow \gamma$ , then  $\alpha \rightarrow \gamma$  (تعدی)

► خواص این قوانین:

- **درستی:** وابستگی‌هایی را تولید میکنند که واقعا برقرارند.
- **کامل بودن:** همه وابستگی‌هایی را که برقرارند تولید می‌کنند.

## مثال

►  $R = (A, B, C, G, H, I)$

$F = \{ A \rightarrow B$

$A \rightarrow C$

$CG \rightarrow H$

$CG \rightarrow I$

$B \rightarrow H\}$

► برخی اعضا  $F^+$  :

◦  $A \rightarrow H$

• با تعدی از  $A \rightarrow B$  and  $B \rightarrow H$

◦  $AG \rightarrow I$

• با افزودن  $G$  به  $A \rightarrow C$  و تعدی از  $CG \rightarrow I$

◦  $CG \rightarrow HI$

• با افزودن  $I$  به  $CG \rightarrow H$  و دست آوردن  $CG \rightarrow CGI$  و افزودن  $H$  به  $CGI \rightarrow HI$  و سپس تعدی.



# بستار وابستگی‌های تابعی (ادامه)

► قوانین دیگر:

- اگر  $\alpha \rightarrow \beta$  و  $\alpha \rightarrow \gamma$  آنگاه  $\alpha \rightarrow \beta \gamma$  (اجتماع)
- اگر  $\alpha \rightarrow \beta \gamma$  آنگاه  $\alpha \rightarrow \beta$  و  $\alpha \rightarrow \gamma$  برقرارند. (تجزیه)
- اگر  $\alpha \rightarrow \beta$  و  $\gamma \beta \rightarrow \delta$  برقرار باشند، آنگاه  $\alpha \gamma \rightarrow \delta$  (شبه تعدی)