

پایگاه داده ها فصل هفتم: نرمالتر سازی رابطه ها Normalisation

دانشجوی دکتری دانشگاه تهران

(کارشناسی و کارشناسی ارشد: کامپیوتر نرم افزار) (دکتری: بیو انفورماتیک)

#### نرمال تر سازی

هر رابطه هر چند نرمال نیز ممکن است در عملیات درج، حذف و بهنگام سازی مشکلاتی داشته باشد.

بنابراین بهتر است یک رابطه نرمال را تا حد امکان نرمال تر کرد. میابراین بهتر است یک رابطه نرمال را تا حد امکان نرمال تر کرد. می است یک رابطه نرمال را تا حد امکان نرمال تر کرد. می است یک رابطه نرمال را تا حد امکان نرمال تر کرد. می است یک رابطه نرمال را تا حد امکان نرمال تر کرد.

#### آنومالي

دشواری و وضع غیرعادی را آنومالی می گویند.

مثلاً وقتی که سطری را حذف می کنیم و پی آمد آن اطلاعات ناخواسته ای نیز حذف شود. یا مقدار صفتی را برای یک سطر تغییر می دهیم در حالیکه در سطرهای دیگر هم نیاز به تغییر داشته باشد.

آنومالی در عملیات ذخیره سازی به هر یک از سه حالت زیر گفته می شود:

- ۱- بروز پیامد بد، بعد از انجام یک عمل
  - ۲- عدم امکان انجام یک عمل
- ۳- بروز اضافه کاری در انجام یک عمل

## آنومالی در بهنگام سازی

در رابطه زیر ، اگر شهر 51 را تغییر دهیم، این عمل باید در تمام سطرهایی که 51 وجود دارد نیز انجام شود (فزونکاری).

در واقع برای جلوگیری از ناسازگاری ، باید بهنگام سازی منتشر شونده انجام شود.

S#	STATUS	CITY	<b>P</b> #	QTY
<b>S1</b>	20	Tehran	P1	300
<b>S</b> 1	20	Tehran	P2	200
<b>S2</b>	10	kish	<b>P1</b>	300
<b>S3</b>	10	yazd	P2	200
<b>S4</b>	20	Tehran	P5	400

#### آنومالی در حذف

در رابطه زیر ، با حذف اطلاع " S3 از P2 به تعداد ۲۰۰ تا تهیه کرده است" این اطلاع که S3 ساکن yazd است نیز حذف می شود.

S#	STATUS	CITY	<b>P</b> #	QTY
<b>S</b> 1	20	Tehran	<b>P</b> 1	300
<b>S</b> 1	20	Tehran	<b>P2</b>	200
<b>S2</b>	10	kish	P1	300
<b>S3</b>	10 0	yazd	<b>P2</b>	200
<b>S4</b>	20	Tehran	<b>P5</b>	400

## آنومالی در درج

در رابطه زیر ، نمی توان اطلاع" تهیه کننده S5 در شهر همدان ساکن است" را درج کرد. چون باید بدانیم چه قطعه ای را تهیه کرده است.

کلید اصلی (S#, P#) است و طبق قاعده جامعیت موجودیتی هیچ جزء کلید اصلی نباید تهی باشد.

S#	STATUS	CITY	<b>P</b> #	QTY
<b>S</b> 1	20	<b>C2</b>	P1	300
<b>S</b> 1	20	<b>C2</b>	P2	200
<b>S2</b>	10	<b>C3</b>	<b>P</b> 1	300
<b>S3</b>	10	<b>C3</b>	<b>P2</b>	200
<b>S4</b>	20	<b>C2</b>	P5	400

تجزیه مطلوب تجزیه مطلوب FaraDars.org

#### ضوابط ریسانن برای تجزیه مطلوب

تجزیه رابطه  ${\bf R}$  به دو رابطه  ${\bf R}$  و  ${\bf R}$  مطلوب است، اگر  ${\bf R}$  و  ${\bf R}$  مستقل از یکدیگر باشند و شرایط زیر برقرار باشند:

۱- صفت مشترک در دو رابطه، حداقل در یکی از آنها کلید کاندید باشد.

این مجموعه R2 های موجود در R در مجموعه FD های R1 و R2 موجود باشند، یا از FD این مجموعه FD منطقاً قابل استنتاج باشند.

#### مثال

کدام یک از تجزیه های رابطه (R (a, b, c) با وابستگی های زیر مطلوب است؟

$$a \rightarrow b$$
 ,  $a \rightarrow c$  ,  $b \rightarrow c$ 

$$R1(\underline{a},b)$$
 ,  $R2(\underline{a},c)$  --

R1 
$$(\underline{a}, c)$$
 , R2  $(\underline{b}, c)$   $-a$ 

تجزیه الف، هر دو شرط قضیه ریسانن را دارد.

در تجزیه  $oldsymbol{ ext{p}}$  ، نمی توان وابستگی  $oldsymbol{ ext{b}} oldsymbol{ ext{c}}$  را از دو وابستگی موجود در روابط  $oldsymbol{ ext{R1}}$  و  $oldsymbol{ ext{R2}}$  منطقاً استنتاج کرد.

در تجزیه a ، نمی توان وابستگی  $a \rightarrow b$  را را از دو وابستگی موجود در روابط a و a منطقاً استنتاج کرد. همچنین شرط اول قضیه ریسانن را نیز ندارد.

#### قضیه هیث

رابطه  $\mathbf{R}(\mathbf{A},\mathbf{B},\mathbf{C})$  که در آن $\mathbf{A}$  و $\mathbf{B}$  و  $\mathbf{B}$  سه مجموعه از صفات هستند، مفروض است.

R2 (B,C) و R1 (A,B) اگر  $A \rightarrow B$  و R1 (B,C) انگاه تجزیه خوب برابر است با:

فرم های نرمال (NORMAL FORMS)

#### فرم های نرمال

1NF

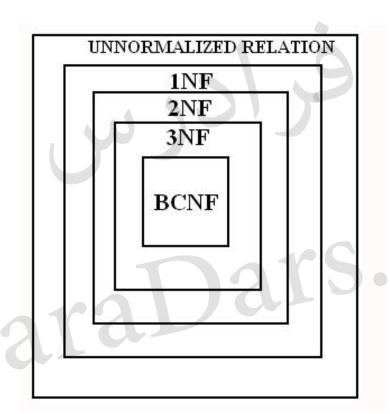
2NF

3NF

**BCNF** 

4NF

**5NF** 



فرم های نرمال عبارتند از:

هر یک از فرم های نرمال، از فرم قبلی نرمالتر هستند.

# تعریف فرم های نرمال

هر صفت خاصه در هر تاپل، تک مقداری باشد.	1NF
1NF باشد و هر صفت خاصه غیر کلید، با کلید اصلی، وابستگی تابعی جزئی نداشته باشد.	2NF
2NF باشد و هر صفت خاصه غیر کلید، با کلید اصلی، وابستگی تابعی با واسطه نداشته باشد.	3NF
هر دترمینان، کلید کاندید باشد.	BCNF
اگر وابستگی تابعی چند مقداری غیر بدیهی $+ - A$ دارد، آنگاه $A$ ابر کلید باشد.	4NF
فقط به کلیدهای کاندیدش وابستگی پیوندی داشته باشد.	5NF

# تبدیل به 1NF

item	colors	price		it
T-shirt	red, blue	12	(3)	7
polo	red, yellow	12.	1	Т
T-shirt	red, blue	12		р
sweatshirt	blue, black	25		p
TaraDa				
	ra.			S



item	color	price
T-shirt	red	12
T-shirt	blue	12
polo	red	12
polo	yellow	12
sweatshirt	blue	25
sweatshirt	black	25

1NF نیست.

1NF

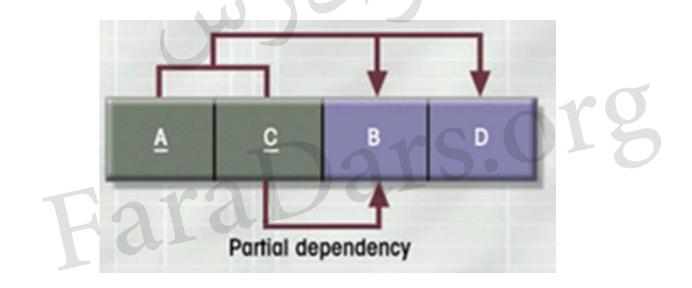
# تبدیل به 1NF

RefNo	Name	Address	AccNo
345	Date	23, High Street	120768 <i>,</i> 348973
543	Rolland	45, The Ash	987654



RefNo	Name	Address	AccNo
345	Date	23, High Street	120768
345	Date	23, High Street	348973
543	Rolland	45, The Ash	987654

# رابطه 1NF ای که وابستگی جزئی ندارد، 2NF است.



S#	P#	City	Quantity
<b>S1</b>	P1	London	1000
<b>S1</b>	<b>P2</b>	London	1500
<b>S1</b>	Р3	London	3400
<b>S1</b>	P4	London	2100
<b>S2</b>	<b>P2</b>	Paris	3400
<b>S2</b>	Р3	Paris	1000
<b>S4</b>	<b>P1</b>	Hamedan	5
<b>S4</b>	P4	Hamedan	7

City
London
Paris
Hamedan

S#	P#	Quantity
<b>S1</b>	P1	1000
S1	P2	1500
<b>S1</b>	Р3	3400
S1	P4	2100
S2	P2	3400
S2	Р3	1000
<b>S4</b>	P1	5
<b>S4</b>	P4	7

#### وابستگی ها :

 $(S\#,P\#) \rightarrow Quantity$ 

 $(S\#,P\#) \rightarrow City$ 

S#→ City

حذف وابستگی جزئی

#### **1NF** به **2NF**

item	color	price
T-shirt	red	12
T-shirt	blue	12
polo	red	12
polo	yellow	12
sweatshirt	blue	25
sweatshirt	black	25

item	color
T-shirt	red
T-shirt	blue
polo	red
polo	yellow
sweatshirt	blue
sweatshirt	black

item	price	
T-shirt	12.00	
polo	12.00	
sweatshirt	25.00	

(item,color) → price

item → price

حذف وابستگی جزئی

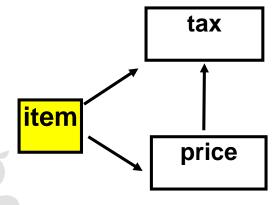
3NF 2NF 2NF earaDars.org

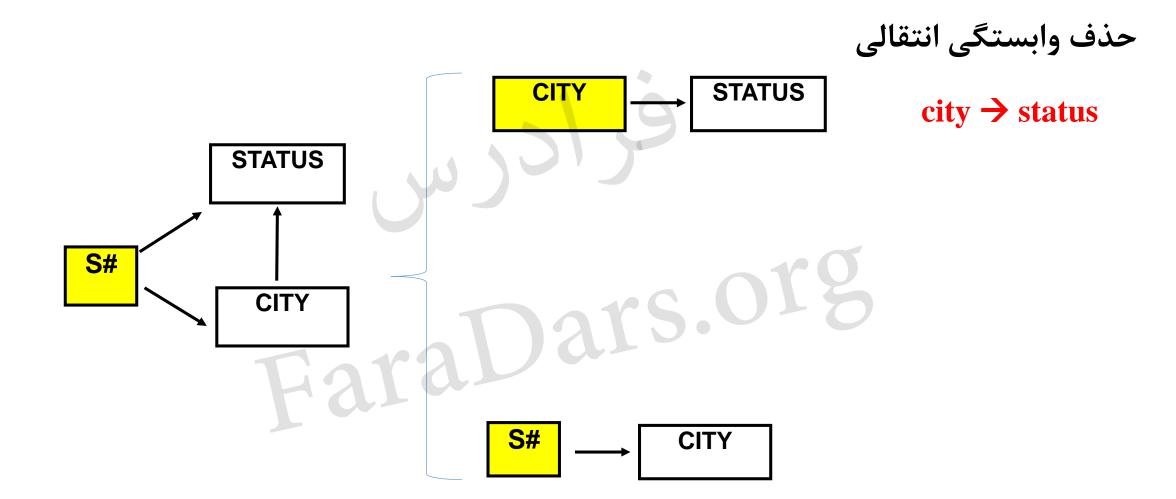
item	price	tax	
T-shirt	12.00	0.60	
polo	12.00	0.60	
sweatshirt	25.00	1.25	

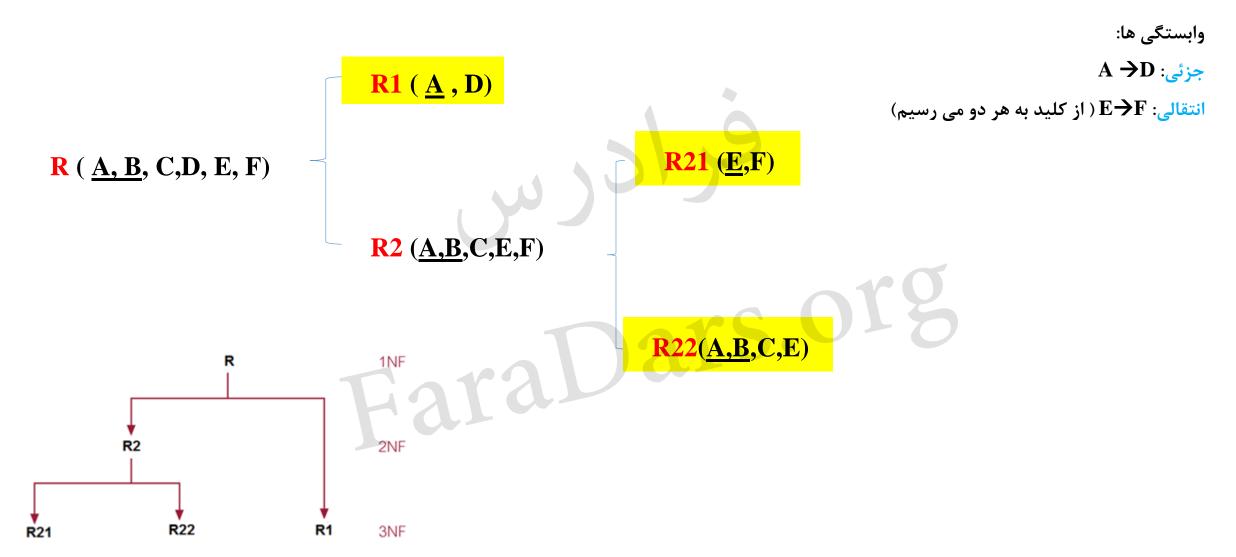
price	tax
12.00	0.60
25.00	1.25

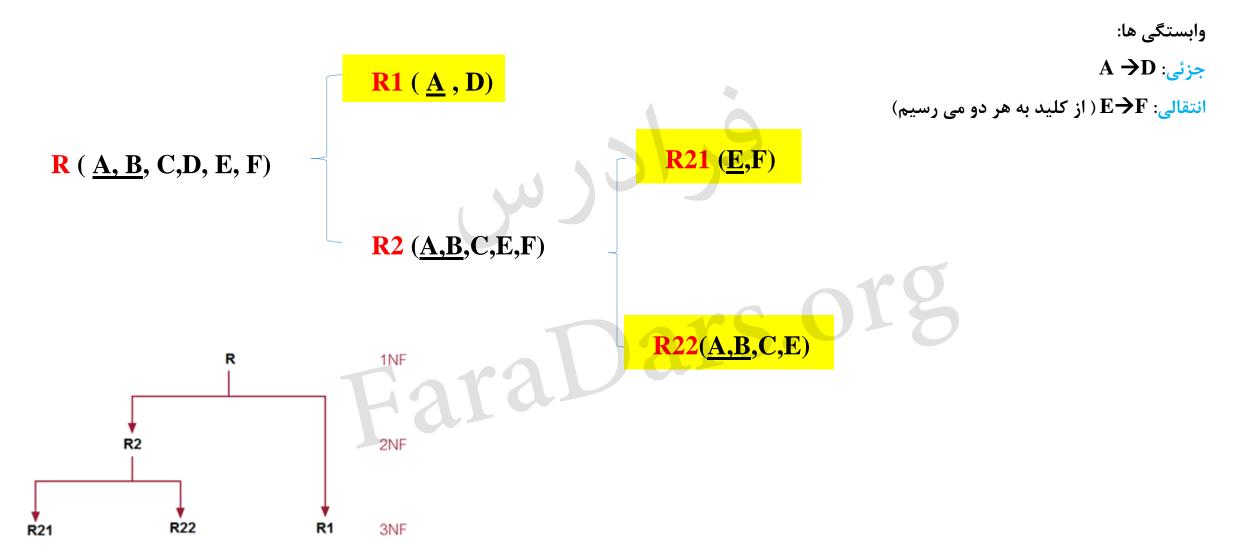
item	price
T-shirt	12.00
polo	12.00
sweatshirt	25.00

# حذف وابستگی انتقالی









وابستگی ها:

جزئى: C,D **→** C

 $\mathbf{C} \rightarrow \mathbf{D}$  انتقالى:

R1(A,B)

 $R21 (\underline{C}, D)$ 

 $\mathbf{R}$  ( $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ , $\mathbf{D}$ )

 $\mathbf{R2}(\underline{\mathbf{A}},\mathbf{C},\mathbf{D})$ 

 $\mathbf{R22} \left( \mathbf{\underline{A}}, \mathbf{C} \right)$ 

R1(A,B,C,F)

وابستگی ها:

 $B \rightarrow E$ 

 $(A,B) \rightarrow D$ 

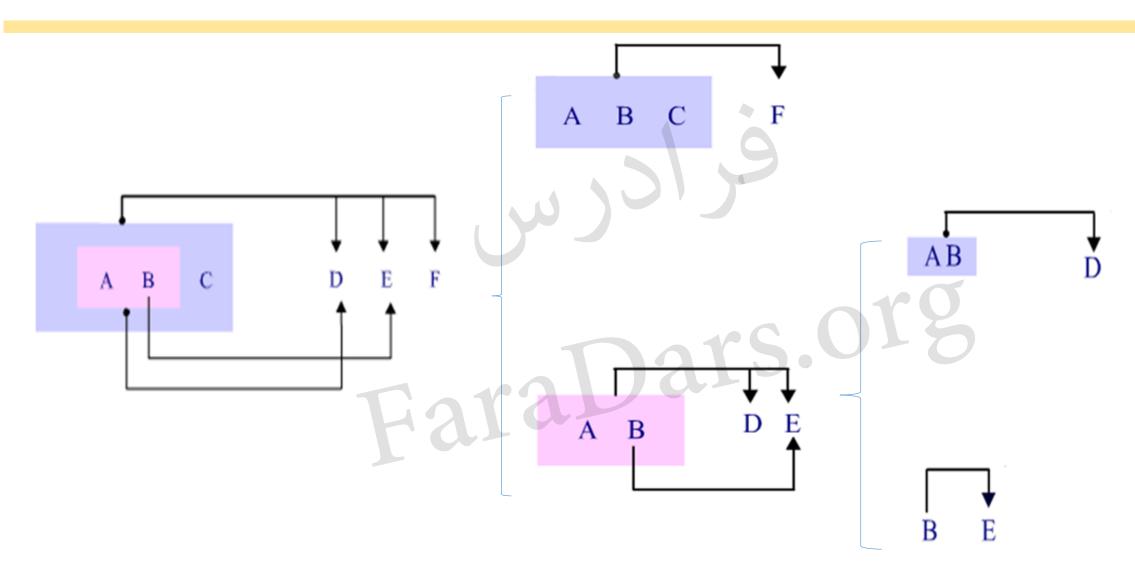
 $\mathbf{R}$  ( $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$ ,  $\mathbf{E}$ ,  $\mathbf{F}$ )

R2 (A, B, D, E)

 $R22(\underline{A},\underline{B},D)$ 

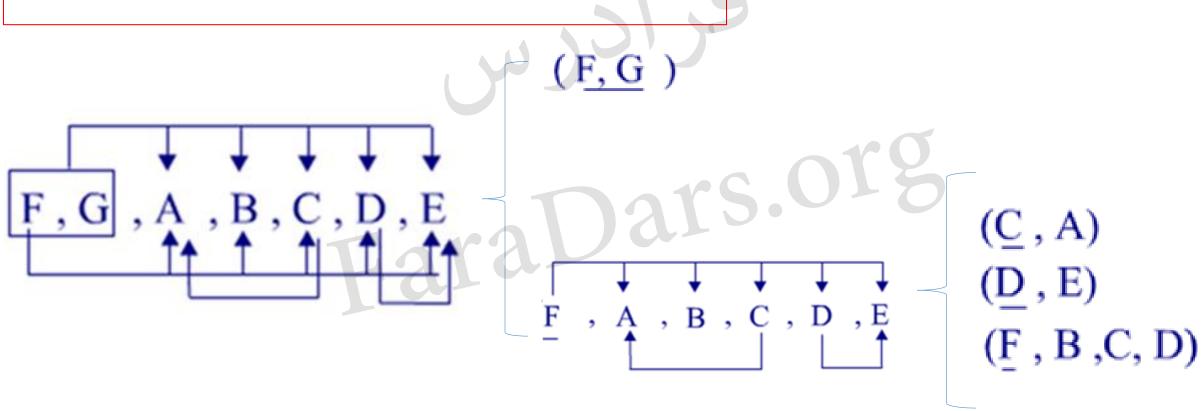
 $R21(\underline{B}, \underline{E})$ 

# نمودار وابستگی مثال قبل



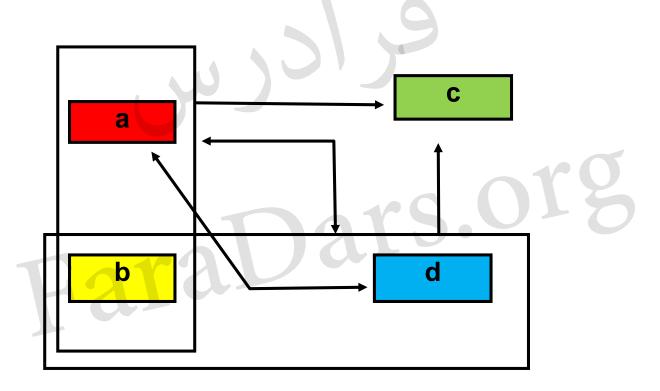
$$R = \{ A, B, C, D, E, F, G \}$$

$$F = \{ AF \rightarrow BE, FC \rightarrow DE, F \rightarrow CD, D \rightarrow E, C \rightarrow A \}$$



## سئوال

# آیا رابطه وابستگی جزئی دارد؟

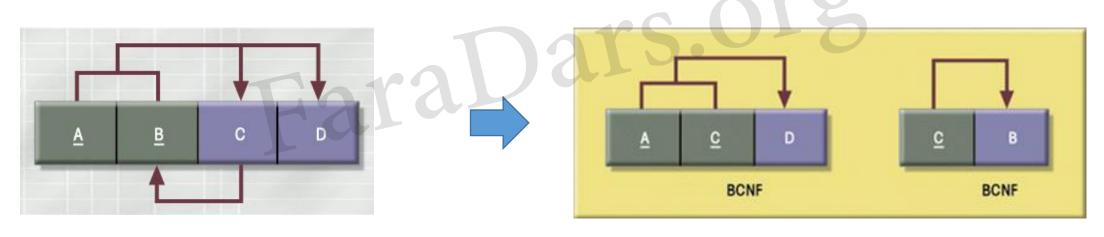


# BCNF. Frg. FaraDars. Frg.

#### رابطه BCNF

تعریف : در  ${f A} 
ightarrow {f B}$ ، می گوییم  ${f A}$  دترمینان است.

رابطه ای در سطح (بایس–کاد)  $\frac{BCNF}{D}$  است اگر و فقط اگر هر دترمینان آن، کلید کاندید باشد.



صفت  ${\bf C}$  دترمینان است ولی کلید کاندید نیست:

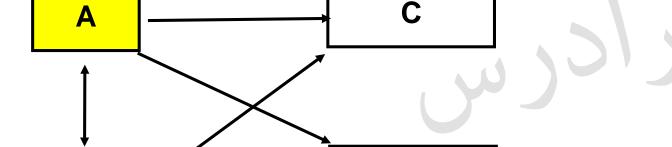
#### نكات

اگر یک رابطه 3NF، تنها یک کلید کاندید داشته باشد و در مجموعه وابستگی های تابعی کاهش ناپذیر آن، وابستگی تابعی دیگری غیر از وابستگی های ناشی از کلید کاندید وجود نداشته باشد، آنگاه رابطه BCNF هم هست.

رابطه تمام کلید، BCNF است.

#### مثال

آیا رابطه 3NF مقابل ، در فرم BCNF است؟



B

شامل دو کلید کاندید: A و B

. حل: بله – چون صفات  ${f B}$  ,  ${f A}$  هم دترمینان هستند، هم کلید کاندید

(A,B) و (A,C) دارای دو کلید کاندید :

 $\mathbf{C} \rightarrow \mathbf{B}$  وابستگی:  $\mathbf{C} \rightarrow \mathbf{B}$  وابستگی

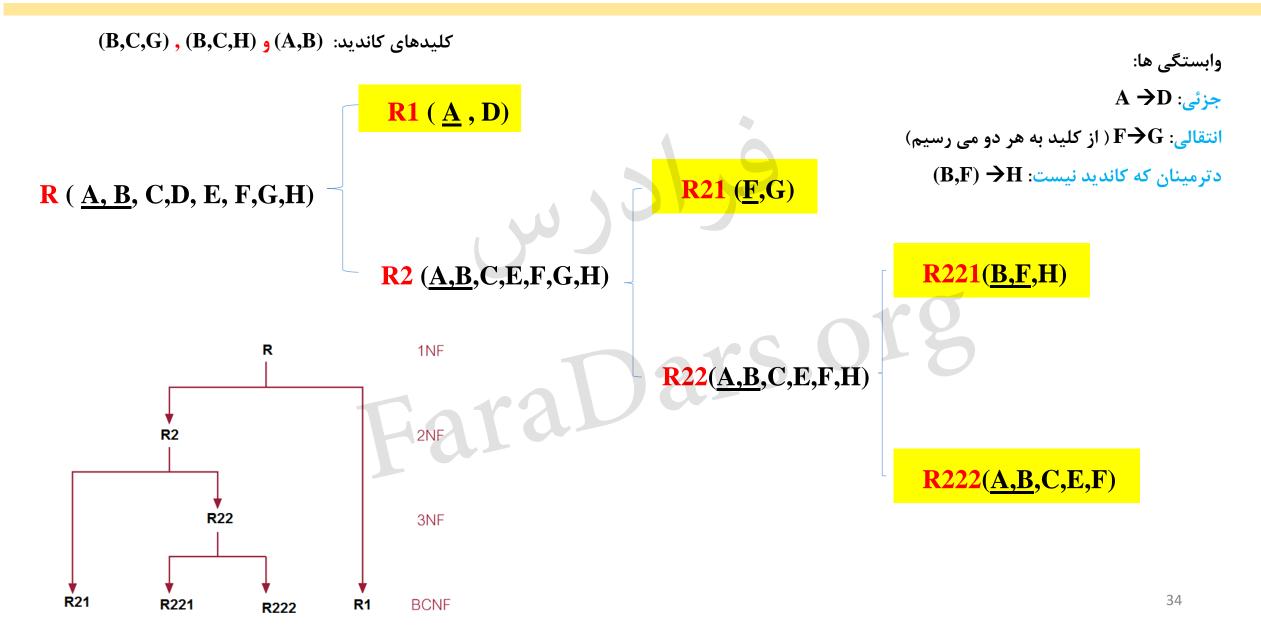
دترمینان است، در حالی که کلید کاندید نمی باشد.  ${f C}$ 

 $\mathbf{R}(\mathbf{A}, \mathbf{B}, \mathbf{C})$ 

**R2** (A, C)

نکته: در حالت وجود صفت مشترک در بین دو کلید کاندید، رابطه ممکن است  $3ar{NF}$  باشد اما BCNF نباشد.

#### **BCNF** به **1NF**



#### مثال واقعى با توجه به اسلايد قبل

#### StaffPropertyInspection

otalii lopolityiilopootioli							
propertyNo	iDate	iTime	pAddress	comments	staffNo	sName	carReg
PG4	18-Oct-00	10.00	6 Lawrence St, Glasgow	Need to replace crockery	SG37	Ann Beech	M231 JGR
PG4	22-Apr-01	09.00	6 Lawrence St, Glasgow	In good order	SG14	David Ford	M533 HDR
PG4	1-Oct-01	12.00	6 Lawrence St, Glasgow	Damp rot in bathroom	SG14	David Ford	N721 HFR
PG16	22-Apr-01	13.00	5 Novar Dr, Glasgow	Replace living room carpet	SG14	David Ford	M533 HDR
PG16	24-Oct-01	14.00	5 Novar Dr, Glasgow	Good condition	SG37	Ann Beech	N721 HFR

#### کلیدهای کاندید:

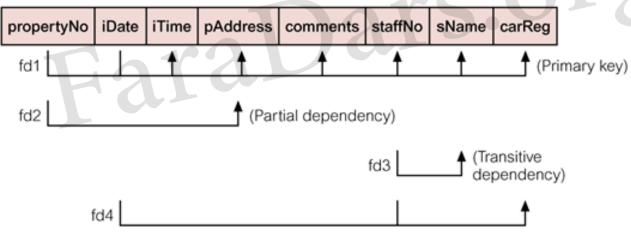
(iDate , iTime , carReg)
(iDate , iTime , staffNo)

كليد اصلى:

(propertyNo , iDate)

(propertyNo , iDate)

StaffPropertyInspection



R: staffPropertyInspection (propertyNo, iDate, iTime, pAddress, comments, staffNo, sName, carReg)

partial dependency (propertyNo  $\rightarrow$  pAddress).

R1: Property (propertyNo, pAddress)

R2: PropertyInspection (propertyNo, iDate, iTime, comments, staffNo, sName, carReg)

Transitive dependency (staffNo→ sName)

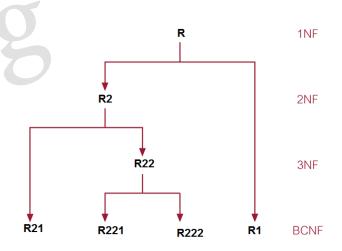
R21: Staff (staffNo, sName)

R22: PropertyInspect (propertyNo, iDate, iTime, comments, staffNo, carReg)

(staffNo, iDate) -> carReg (staffNo, iDate) : Determinant but not candidate key

R221: taffCar (staffNo, iDate, carReg)

R222: Inspection (propertyNo, iDate, iTime, comments, staffNo)



## **BCNF** به **1NF**

```
1NF:

ORDER ( NO , SID, SNAME, DATE )

ORDER_LINE (NO , PID , PNAME , QTY , COST )
```

#### 2NF:

ORDER (NO, SID, SNAME, DATE)
PRODUCT (PID, PNAME, COST)
ORDER\_LINE (NO, PID, QTY)

SNAME, in ORDER table, is transitively dependent on SID.

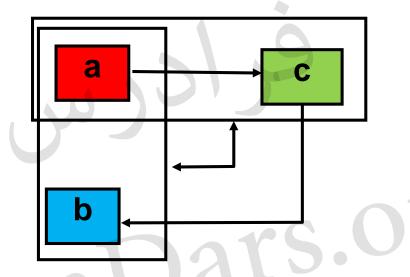
```
3NF:
```

```
ORDER ( NO , SID, DATE )
SUPPLIER ( SID , SNAME )
PRODUCT ( PID , PNAME , COST )
ORDER_LINE ( NO , PID , QTY )
```

All determinants are candidate keys. All relations are in BCNF.

# عدم نیاز به تجزیه

رابطه 3NF داده شده، BCNF نیست ، چون c د ترمینان است، اما کلید کاندید نیست.



اگر این رابطه را به دو رابطه (c, b) R2 (c, b) که هر دو BCNF که هر دو و رابطه (a, b) کنیم، وابستگی  $(a,b) \rightarrow c$ 

## نكته

گاهی تجزیه یک رابطه 3NF به دو رابطه BCNF، ضوابط ریسانن را ندارد. به همین علت اگر در اثر تجزیه، وابستگی هایی حذف شود، در این صورت همان 3NF کفایت می کند و رابطه را نباید تجزیه کرد.

4NB FaraDars.org

# رابطه 4NF

تعریف: وابستگی B op A op B بدیهی است اگر B زیر مجموعه A باشد، یا ستون دیگری به غیر از A و B نداشته باشد.

 $A \! 
ightarrow B$  است اگر وابستگی تابعی چند مقداری غیر بدیهی  $4 {
m NF}$  رابطه ای

در آن وجود داشته باشد، آنگاه A ابر کلید باشد.

## مثال

course	teacher	book
database	shirafkan	Silberschatz
database	shirafkan	C.J. Date
database	rasti	Silberschatz
database	rasti	C.J. Date
database	akbari	Silberschatz
database	akbari	C.J. Date
os	shirafkan	Stallings
os	shirafkan	Tanenbaum
os	hasani	Stallings
OS	hasani	Tanenbaum

## course $\rightarrow \rightarrow$ book course $\rightarrow \rightarrow$ teacher

رابطه R تمام کلید است و حداقل در سطح BCNF است، اما به علت وجود افزونگی دارای آنومالی هایی می باشد.

آنومالی در درج: در این رابطه افزودن اطلاع اینکه sara استادی است در درج: در این رابطه افزودن اطلاع اینکه database که database, Sara, Silberschatz)

(database, Sara, C.J. Date)

# BCNF به

course	teacher	book
database	shirafkan	Silberschatz
database	shirafkan	C.J. Date
database	rasti	Silberschatz
database	rasti	C.J. Date
database	akbari	Silberschatz
database	akbari	C.J. Date
os	shirafkan	Stallings
os	shirafkan	Tanenbaum
os	hasani	Stallings
os	hasani	Tanenbaum

course	teacher
database	shirafkan
database	rasti
database	akbari
os	shirafkan
os	hasani

course	book
database	Silberschatz
database	C.J. Date
os	Stallings
os	Tanenbaum

## **4NF** به **BCNF**

$$\mathbf{R} = (\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{G}, \mathbf{H}, \mathbf{I})$$

$$F = \{ A \rightarrow \rightarrow B , B \rightarrow \rightarrow H , CG \rightarrow \rightarrow H \}$$

رابطه  $A \to A$  در فرم  $A \to A$  نیست، چون با وجود وابستگی چند مقداری  $A \to A$  ، صفت A ابر کلید نیست .

R1 (A, B) R21 (C, G, H) R221(A, C, G, H, I) R222(A, C, G, I)

 $A \rightarrow \rightarrow B$  and  $B \rightarrow \rightarrow HI : A \rightarrow \rightarrow HI$ 

 $A \rightarrow \rightarrow HI : A \rightarrow \rightarrow I$ 

R222(A,C,G)

# 5 Blance Earabara Farabara Far

### 5NF

رابطه ای در سطح 5NF است که در صورت وجود وابستگی پیوندی در آن این وابستگی ها ناشی از کلیدهای کاندید باشند.

به عبارتی فقط به کلیدهای کاندیدش وابستگی پیوندی داشته باشد.

تذكر: به SNF دابطه PJNF نيز مي گويند.

## مثال

رابطه تمام کلید (<u># , P# , J#) SPJ</u> در فرم **5NF نمی باشند**، چون دارای وابستگی پیوندی است که ناشی از کلید کاندید رابطه نمی باشد. یعنی کلید کاندید در همه پرتوهای این رابطه حضور ندارد.

وابستگی پیوندی در این رابطه برابر است با:

\* (SP, PJ, JS)

یعنی اگر رابطه SPJ را به سه رابطه SP, PJ, PJ تجزیه کنیم وسپس دو تجزیه مثلاً SPJ را با هم پیوند کنیم و در نهایت نتیجه را با SPJ پیوند بزنیم، آنگاه حاصل همان SPJ خواهد بود، بدون اینکه تاپلی اضافه یا کم شده باشد.

# مراحل کار

SPJ

510		
S#	<b>P</b> #	J#
S1	P1	<b>J</b> 2
S1	P2	J1
<b>S</b> 2	P1	J1
S1	P1	J1

SP

S#	<b>P</b> #
S1	P1
S1	<b>P</b> 2
S2	P1

 $\Rightarrow$ 

PJ

<b>P</b> #	J#
P1	J2
<b>P</b> 2	J1
P1	_J1

J#	S#
J2	S1
J1	S1
J1	S2

S#	<b>P</b> #	J#
S1	P1	J2
S1	P1	J1
S1	P2	Л1
S2	P2	J2
S2	P1	J1

S#	<b>P</b> #	J#
S1	P1	J2
S1	P2	J1
S2	P1	J1
S1	P1	J1

### مثال

رابطه S (S# , SNAME , STATUS , CITY) در سطح 5NF است.

چون وابستگی های پیوندی موجود در این رابطه ناشی از کلید کاندید است.

یعنی کلید کاندید در همه پرتوهای رابطه وجود دارد.

وابستگی زیر ناشی از کلید کاندید #5 است:

\* ( (S#, SNAME, STATUS, CITY ), ( S#, CITY) )

وابستگی زیر ناشی از کلید کاندید #S یا SNAME است:

\* ( (S#, SNAME ), ( S#, STATUS ), ( SNAME, CITY) )

رابطه هایی که BCNF باشند اما 4NF نباشند و یا 4NF باشند و 5NF باشند و 5NF باشند، نادرند.

اگر رابطه 3NF باشد و تمام کلیدهای کاندید آن صفات ساده باشند، در اینصورت رابطه 5NF است.

# مزایای نرمالترسازی

- ۱- کاهش بعضی از آنومالی ها
  ۲- کاهش بعضی از انواع افزونگی
  ۳- تسهیل اعمال بعضی از قواعد جامعیت FaraDars.org

# معایب نرمالترسازی

- ۱- بروز فزونکاری در سیستم در عمل بازیابی
  - ۲- ایجاد نوعی افزونگی
  - ۳- زمانگیر بودن فرایند نرمالترسازی
- ۴- مشکل شدن تصمیم گیری ها در تعدد تجزیه ها در مواردی

# تمرين

CourseNo	Lecturer	Time	Room
1	Α	9	121
2	В	10	666
5	С	1	343
6	D	11	876
4	С	3	343
3	В	4	666

CourseNo	Lecturer	Time
1	Α	9
2	В	10
5	С	1
6	D	11
4	С	3
3	В	4

Lecturer	Room
A	121
В	666
С	343
D	876



#### Staff Branch

staffNo	sName	position	salary	branchNo	bAddress
SL21	John White	Manager	30000	B005	22 Deer Rd, London
SG37	Ann Beech	Assistant	12000	B003	163 Main St, Glasgow
SG14	David Ford	Supervisor	18000	B003	163 Main St, Glasgow
SA9	Mary Howe	Assistant	9000	B007	16 Argyll St, Aberdeen
SG5	Susan Brand	Manager	24000	B003	163 Main St, Glasgow
SL41	Julie Lee	Assistant	9000	B005	22 Deer Rd, London

#### Staff

staffNo	sName	position	salary	branchNo
SL21	John White	Manager	30000	B005
SG37	Ann Beech	Assistant	12000	B003
SG14	David Ford	Supervisor	18000	B003
SA9	Mary Howe	Assistant	9000	B007
SG5	Susan Brand	Manager	24000	B003
SL41	Julie Lee	Assistant	9000	B005

#### **Branch**

branchNo	bAddress			
B005	22 Deer Rd, London			
B007	16 Argyll St, Aberdeen			
B003	163 Main St, Glasgow			

# تمرین۳

RefNo	Name	Address	Status	AccNo		675	745363
						675	678453
345	C.J. Date	23, High Street	Business	120768		675	348973
345	C.J. Date	23, High Street	Business	348973		0/3	346973
543	F.D. Rolland	45, The Ash	Domestic	987654			400
675	D.R. Howe	17, Low Street	Business	745363	_		
675	D.R. Howe	17, Low Street	Business	678453		S	ノエて
675	D.R. Howe	17, Low Street	Business	348973			
		T	21				
		,	CTT.			RefNo	Name
						345	C.J. Date
						Γ/1 <b>2</b>	ED Dellar

RefNo	AccNo			
345	120768			
345	348973			
543	987654			
675	745363			
675	678453			
675	348973			
	·-			

RefNo	Name	Address	Status
345	C.J. Date	23, High Street	Business
543	F.D. Rolland	45, The Ash	Domestic
675	D.R. Howe	17, Low Street	Business

## مروری بر عناوین تدریس شده

فصل ١: مفاهيم اوليه

فصل ۲: مدل رابطه ای

فصل ۳: نمودار ER

فصل ۴: جبر رابطه ای

فصل ۵: SQL فصل ۶: وابستگی ها فصل ۶: وابستگی ها فصل ۷: نرمال تر سازی

+91719VY+YA مشاوره با مدرس شیرافکن: این اسلاید ها بر مبنای نکات مطرح شده در فرادرس «پایگاه داده ها» تهیه شده است.

برای کسب اطلاعات بیشتر در مورد این آموزش به لینک زیر مراجعه نمایید faradars.org/fvsft105