

فصل هفتم: نرمالسازی

نرمال سازی

تبدیل جداول غیر نرمال به صورتهای نرمال اصلی

نرمال سازی رابطه‌ها

- یک مشکل رابطه‌های غیرنرمال، **افزونگی داده (redundancy)** است.
- یک رابطه غیر نرمال ممکن است در عملیات ذخیره‌سازی (درج، حذف و بهنگام‌سازی) مشکلاتی داشته باشد. به این مشکلات **آنومالی** گفته می‌شود.

مثالی از افزونگی و آنومالی در عملیات درج

رابطه T			رابطه C			رابطه STC (جدول نمرات)						
TId	TName	TFamily	CId	CName	Units	SId	SName	SFamily	Field	TId	CId	mark
1	Hadi	Hamidi	01	DB	3	84110	Ali	Ahmadi	Computer	2	01	12
2	Karim	Hassani	02	OS	3	84120	Reza	Rezaei	Math	1	03	8
3	Ali	Omidi				84110	Ali	Ahmadi	Computer	2	02	18

- در این مثال، اطلاعات کلی دانشجو و نمرات او همه در یک جدول قرار می‌گیرند. در هنگام درج اطلاعات یک دانشجو با دو آنومالی ممکن است مواجه باشیم:
 ۱. بازای هر نمره که یک دانشجو می‌گیرد باید هر بار تمام مشخصات او نیز به همراه نمره اش در جدول STC وارد شود و این افزونگی است.
 ۲. برای وارد کردن اطلاعات دانشجوی جدید (که هنوز درسی نگرفته) مجبوریم فقط ۴ فیلد اول در جدول STC را وارد کنیم و بقیه فیلدها Null می‌شوند. وجود مقادیر Null بدلیل هدر دادن بیهوده فضا خوشایند نیست

مثالی از آنومالی در عملیات حذف

رابطه T

TId	TName	TFamily
1	Hadi	Hamidi
2	Karim	Hassani
3	Ali	Omid

رابطه C

CId	CName	Units
01	DB	3
02	OS	3

رابطه STC (جدول نمرات)

SId	SName	SFamily	Field	TId	CId	mark
84110	Ali	Ahmadi	Computer	2	01	12
84120	Reza	Rezaei	Math	1	03	8
84110	Ali	Ahmadi	Computer	2	02	18

- در این مثال هنگام حذف اطلاعات با دو آنومالی ممکن است مواجه باشیم:
 ۱. اگر بخواهیم رکورد مربوط به نمره یکی از دانشجویان را حذف کنیم که فقط همین یک درس را تاکنون گرفته، اطلاعات شخصی او هم از سیستم حذف می شود.
 ۲. اگر بخواهیم اطلاعات یک دانشجو را حذف کنیم باید در تمام جدول STC هر جا که اطلاعات آن دانشجو آمده را یافته و حذف کنیم.

مثالی از آنومالی در عملیات ویرایش

رابطه T

TId	TName	TFamily
1	Hadi	Hamidi
2	Karim	Hassani
3	Ali	Omid

رابطه C

CId	CName	Units
01	DB	3
02	OS	3

رابطه STC (جدول نمرات)

SId	SName	SFamily	Field	TId	CId	mark
84110	Ali	Ahmadi	Computer	2	01	12
84120	Reza	Rezaei	Math	1	03	8
84110	Ali	Ahmadi	Computer	2	02	18

- ویرایش اطلاعات یک دانشجو نیز آنومالی دارد:
 ۱. اگر بخواهیم یک ویژگی یک دانشجو مثلاً نام خ او را تغییر دهیم بجای اینکه در یک جا این تغییر اعمال شود باید در تمام رکورد های مربوط به آن دانشجو در جدول STC این عمل صورت پذیرد.

نرمال سازی رابطه ها

- یعنی شکستن یک رابطه غیر نرمال به چند رابطه کوچکتر با هدف مقابله با:
 ۱. افزونگی داده (redundancy)
 ۲. آنومالی عملیات درج، حذف و بهنگام سازی
 ۳. وجود مقادیر زیاد Null

تجزیه مطلوب چیست؟

۱. همه وابستگیهای موجود در جدول اولیه ، بعد از تجزیه حفظ شوند.
 ۲. در صورت پیوند مجدد جداول حاصل از تجزیه، همان جدول اولیه حاصل شود (بدون رکورد اضافی)
- برای این منظور صفت مشترک بین دو جدول حاصل از تجزیه باید حداقل در یکی کلید کاندید باشد. (ضابطه ریسانن)

قضیه Heath:

اگر در جدول $R(A,B,C)$ وابستگی های $A \rightarrow B$ و $B \rightarrow C$ موجود باشد، تجزیه مطلوب بصورت زیر است:

$R1(\underline{A},B)$, $R2(B,\underline{C})$

مثال:

جدول $R(\underline{X},Y,Z)$ به دو جدول $R1(X,Y)$ و $R2(X,Z)$ تجزیه و نرمال شده است. کدام گزینه از شرط های صحت این عمل می باشد؟ ($X \rightarrow R1$ یعنی X کلید $R1$ است)

(۱) $X \rightarrow R1$ (۲) $X \rightarrow R2$ (۳) $X \rightarrow R1$ و $X \rightarrow R2$ (۴) $X \rightarrow R1$ یا $X \rightarrow R2$

گزینه :

توضیح:

با توجه به اینکه X کلید جدول R است، داریم: $X \rightarrow Y$ و $X \rightarrow Z$. تجزیه R به $R1$ و $R2$ در صورتی صحیح است که این وابستگیها حفظ شوند. پس در $R1$ داریم $X \rightarrow Y$ (پس X در $R1$ کلید است) و در $R2$ داریم $X \rightarrow Z$ (پس X در $R2$ هم کلید است)

صورت های نرمال

۱. صورت نخست نرمال (1NF)
۲. صورت دوم نرمال (2NF)
۳. صورت سوم نرمال (3NF)
۴. صورت نرمال بایس-کاد (BCNF)

- ۵. صورت چهارم نرمال (4NF)
- ۶. صورت پنجم نرمال (5NF)
- ۷. صورت نرمال میدان-کلیدی (DKNF)
- ۸. صورت نرمال تحدید-اجتماع (RUNF)

فرم نرمال 1NF

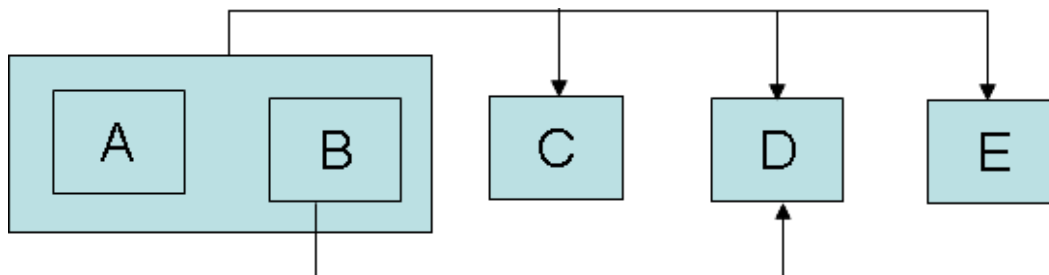
- جدولی در فرم نرمال 1NF است که هیچ فیلد مرکب و چند مقداری نداشته باشد و تمام ویژگیهای ذکر شده برای رابطه ها در آن برقرار باشد.
- از آنجا که هیچیک از نرم افزارهای بانک اطلاعات رابطه ای (مانند Access, Sql Server و ...) اجازه تعریف فیلدهای مرکب و چندمقداری را نمی دهند، پس هر جدولی که ما می سازیم می دانیم که حداقل 1NF هست.

فرم نرمال 2NF

- جدولی در فرم نرمال 2NF است که
 ۱. اولاً 1NF باشد.
 ۲. ثانیاً تمام فیلدهای غیر کلید با کلید کاندید وابستگی تابعی کامل داشته باشند. (یعنی فیلدها به اجزای تشکیل دهنده کلید کاندید وابسته نباشند)
- اگر کلیدهای کاندید یک جدول ترکیبی نباشد (یک فیلد تک باشد)، آن جدول حتما 2NF هم هست.

مثال:

فرض کنیم نمودار وابستگیها در جدول $R(A,B,C,D,E)$ بصورت زیر باشد.



- مشخص است که کلید اصلی جدول (A,B) است.

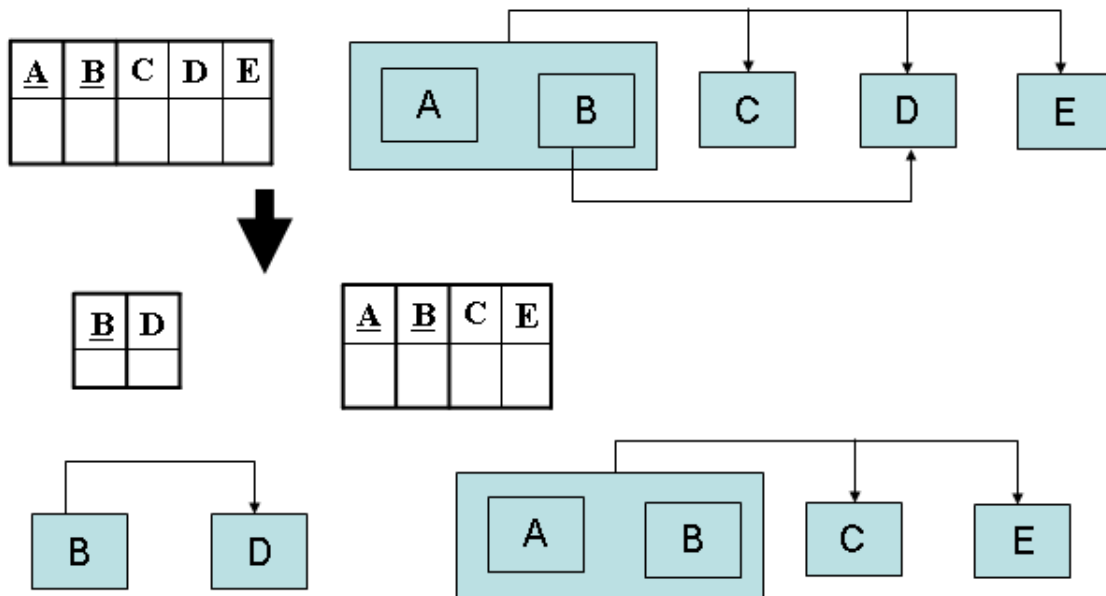
- تنها فیلدی که وابستگی آن به کلید ، کامل نیست D است. پس D باید از این جدول جدا شود.

روش تبدیل جدول از فرم 1NF به فرم 2NF

- جدول را بصورت زیر می شکنیم:

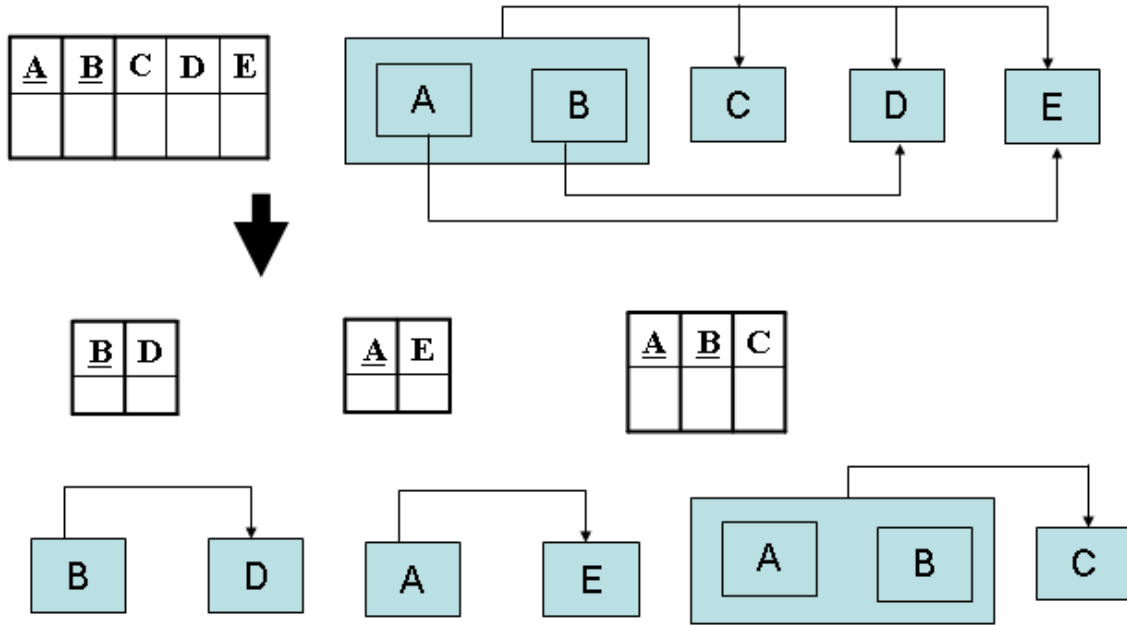
۱. فیلدهایی را که به اجزائی از کلید اصلی وابسته هستند را از جدول جدا کرده، به همراه آن اجزا در یک جدول قرار می دهیم.
۲. کل کلید اصلی به همراه سایر فیلدها در جدول دیگر

مثال: تبدیل جدول مثال قبل به فرم 2NF



دیگر در هیچیک از دو جدول، وابستگی تابعی غیر کامل بین فیلدها با کلید اصلی نداریم

مثالی دیگر:



فرم نرمال 3NF

جدولی در فرم نرمال 3NF است که

– اولاً 2NF باشد.

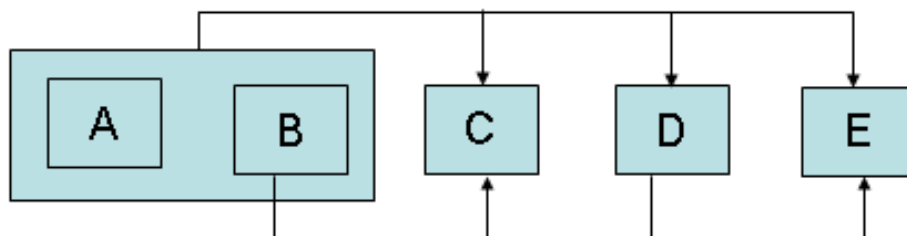
– ثانیاً فیلدهای غیر کلید به هم وابستگی نداشته باشند.

(بعبارت دیگر هر فیلد فقط بطور مستقیم (و نه با رابطه تعدی) به کلید جدول وابسته باشد).

مثال:

فرض کنیم نمودار وابستگیها در جدول $R(A, B, C, D, E)$ بصورت زیر باشد. جدول در چه سطح

نرمالی قرار دارد؟ اگر 3NF نیست آنرا به فرم 3NF تبدیل کنید.



جدول در سطح 1NF است. برای تبدیل به 3NF باید اول آنرا به 2NF و سپس به 3NF تبدیل کنیم.

روش تبدیل جدول از فرم 2NF به فرم 3NF

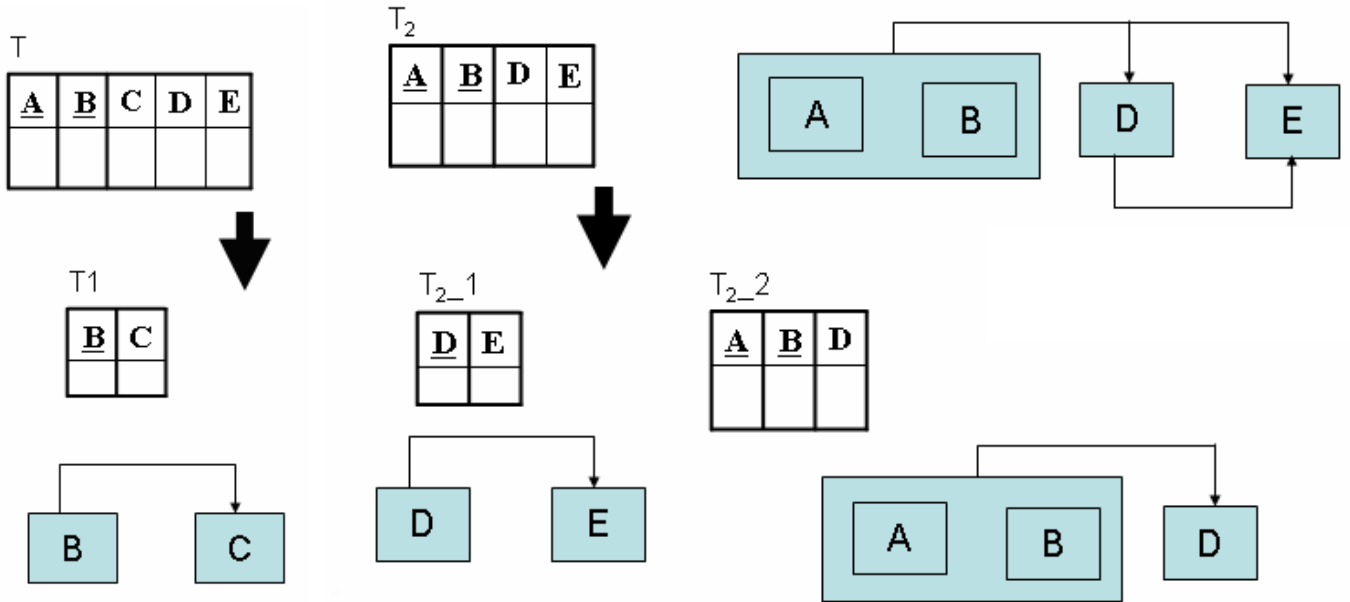
جدول را بصورت زیر می شکنیم:

۱. فیلدهایی (مثل Y) را که به فیلد غیر کلید اصلی (مثل X) وابسته هستند را از جدول جدا

کرده، به همراه X در یک جدول قرار می دهیم.

۲. کلید اصلی به همراه تمام فیلدها بجز Y در جدول دیگر

مثال: تبدیل جدول مثال قبل به فرم 3NF



1: مرحله دوم: تبدیل

جدول T به 3NF

مثال:

فرض کنیم جدولی مشابه جدول STC با اسکیمای زیر طراحی کرده ایم. میخواهیم در صورت لزوم آنرا نرمالسازی کنیم.

مدیر دانشکده دانشکده ارائه دهنده

<u>SId</u>	<u>CId</u>	<u>Sec</u>	<u>Term</u>	Cname	Units	ClgId	Manager	tId	Mark

برای بررسی سطح نرمال جدول، بعنوان طراح وظیفه داریم لیست وابستگی های موجود در جدول را بنویسیم:

Cid → Cname

Cid → Units

ClgId → Manager

(Cid , Sec , term) → Cname

(Cid , Sec , term) → Units

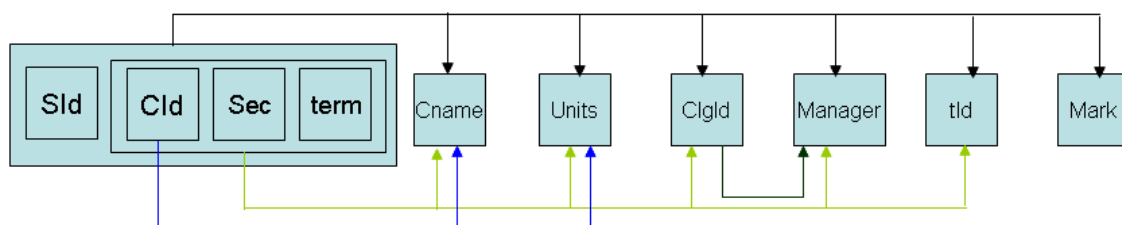
(Cid , Sec , term) → ClgId

(Cid , Sec , term) → manager

(Cid , Sec , term) → tId

(Sid , Cid , Sec , term) → همه فیلدها

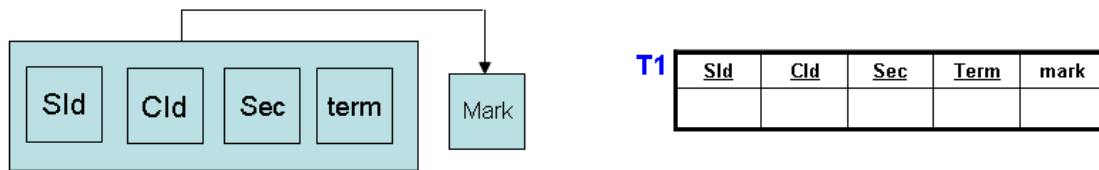
نمودار وابستگی ها:



از نمودار مشخص می شود که جدول در سطح 1NF قرار دارد. برای اینکه 2NF شود باید وابستگی های آبی و سبز از جدول حذف شوند.

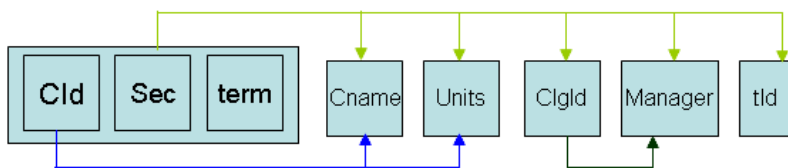
با وابستگی های سبز رنگ کار را شروع می کنیم چون که دترمینان آنها (سمت چپ وابستگی) بزرگتر است.

مرحله اول: تبدیل T به 2NF (شکستن به دو جدول T1 و T2)



T1

SId	CId	Sec	Term	mark

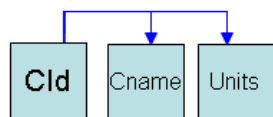


T2

CId	Sec	Term	Cname	Units	clgld	Manager	tld

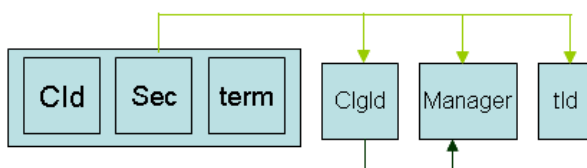
T1 در سطح 2NF و حتی 3NF است اما T2 هنوز 1NF است (بدلیل وجود وابستگی های آبی). پس باید اول T2 2NF شود.

مرحله دوم: تبدیل T2 به 2NF



T2_1

CId	Cname	Units

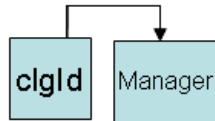


T2_2

CId	Sec	Term	clgld	Manager	tld

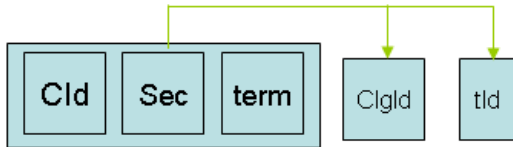
T2_1 در سطح 2NF و حتی 3NF است. T2_2 هم در سطح 2NF است اما بدلیل وجود وابستگی سیاه رنگ 3NF نیست. پس T2_2 تنها جدولی است که باید 3NF شود.

مرحله سوم: تبدیل T2_2 به 3NF



T2_2_1

<u>clgld</u>	Manager



T2_2_2

<u>Cld</u>	<u>Sec</u>	<u>Term</u>	clgld	tld

اکنون تمام جدولها در سطح 3NF هستند.

جمع بندی مثال:

جدول اولیه جهت نرمالسازی تا سطح 3NF به ۴ جدول زیر شکسته شد.

T1

<u>Sld</u>	<u>Cld</u>	<u>Sec</u>	<u>Term</u>	mark

→ جدول نمرات

T2_1

<u>Cld</u>	Cname	Units

→ جدول لیست تمام

T2_2_1

<u>clgld</u>	Manager

→ جدول لیست تمام

T2_2_2

<u>Cld</u>	<u>Sec</u>	<u>Term</u>	clgld	tld

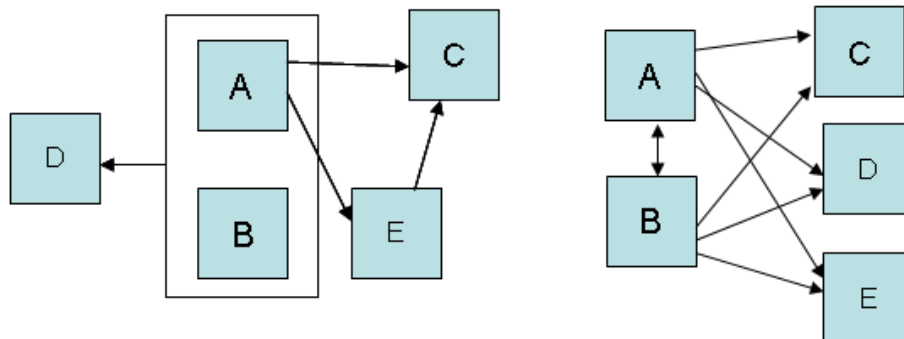
→ جدول لیست گروههای درسی

فرم نرمال BCNF

- این حالت بیشتر زمانی مطرح می شود که:
 - جدول دارای چند کلید کاندید باشد.
 - کلید کاندید ترکیبی داشته باشیم.
 - کلیدهای کاندید در اجزایشان با هم اشتراک داشته باشند. مثلا (A,B) و (A,C)
- اگر در جدولی این ۳ شرط برقرار باشد نیاز به نرمالسازی BCNF دارد.
- جدولی در فرم نرمال BCNF است که
 - اولاً 3NF باشد.
 - ستونهای جدول فقط به کلیدهای کاندید وابستگی داشته باشند و حتی اجزای کلیدهای کاندید مختلف هم به هم وابسته نباشند.

مثال:

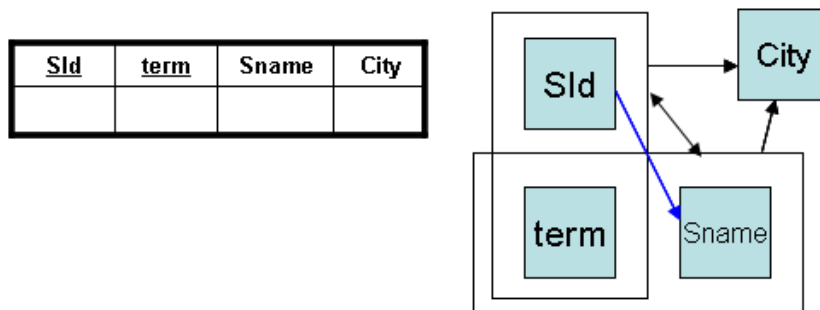
جدول $R(A,B,C,D,E)$ در کدام حالت در سطح BCNF است؟



- در شکل سمت چپ، کلید کاندید (A,B) است اما شاهد وابستگیهایی مثل $A \rightarrow C$ و $E \rightarrow C$ هستیم. پس BCNF نیست.
- در شکل سمت راست A و B هر کدام یک کلید کاندید هستند. همه وابستگیها هم فقط به آنهاست. پس BCNF هست.

مثال:

- با توجه به نمودار وابستگیها (که فرضی است) جدول زیر در چه سطح نرمالی قرار دارد؟ در صورت لزوم آنرا نرمالتر کنید.



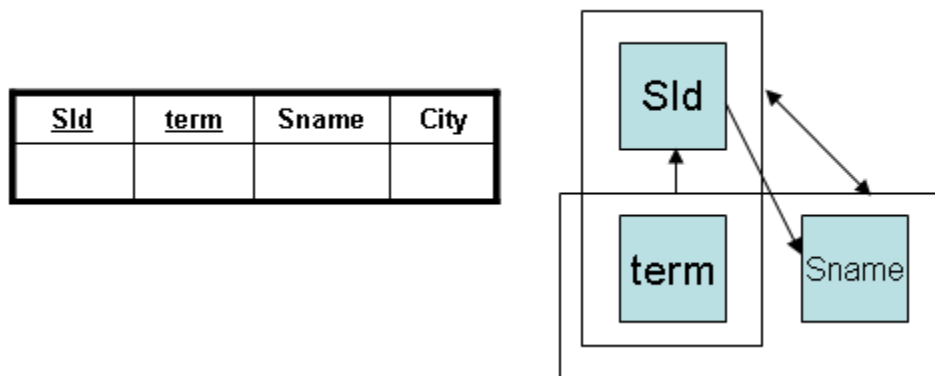
از شکل مشخص است که $(sId,term)$ و $(sname,term)$ هر دو کلید کاندید هستند. جدول در سطح 3NF است اما BCNF نیست (بدلیل وجود وابستگی بین اجزای دو کلید کاندید). پس باید BCNF شود؛ بصورت زیر:

<u>SId</u>	Sname

<u>SId</u>	<u>term</u>	City

نکته مهم:

گاهی نباید جدولی را از یک حد بیشتر نرمال کرد. مثلا جدول زیر در سطح 3NF است اما BCNF نیست. ولی آنرا BCNF نمی کنیم!



از شکل مشخص است که $(sname,term)$ و $(sId,term)$ هر دو کلید کاندید هستند. اگر جدول BCNF شود، بصورت زیر در می آید:

<u>SId</u>	Sname

<u>SId</u>	<u>term</u>

در اثر این تجزیه، وابستگی $SId \rightarrow (Sname, term)$ حذف شده. پس این تجزیه مطلوب نیست.