

به نام خدا

پاسخنامه تکلیف پنجم درس پایگاه داده ها 1

ترم پاییز 1400

1. به سوالات زیر پاسخ دهید:

a. توضیح دهید BCNF چه تفاوتی با 3NF دارد؟ چرا آن را شکل قوی تری از 3NF در نظر می گیرند؟

BCNF شکل قوی تری از 3NF است. همچنین در BCNF در وابستگی تابعی $XY \rightarrow Y$ نیازمند نیست یک prime attribute باشد در حالی که در 3NF باید یک prime attribute باشد. (وابستگی تابعی یا functional dependencies)

BCNF شکل قوی تری از 3NF است زیرا:

- در BCNF هر determinant باید یک کلید کاندید باشد.
- BCNF برخی از وابستگی هایی را که در 3NF مجاز هستند، مجاز نمیداند.
- رابطه ای که در BCNF باشد در 3NF نیز خواهد بود.

b. چگونه شرايطی را توصیف می کنید که در آن یک attribute به attribute دیگری وابسته باشد و هیچ یک از attribute ها جزو کلید اصلی نیستند؟

این شرایط، همان شرایط **transitive dependency** است. یک **transitive dependency** در واقع وابستگی یک **non prime attribute** (غیر کلید اصلی) به یک **non prime attribute** دیگر است.

c. ویژگی های یک رابطه ی bcnf چیست؟ اثبات کنید جدول های با دو attribute همیشه BCNF هستند. (راهنمایی: تمام حالت های وابستگی های تابعی را بررسی کنید.)

اگر R یک رابطه و F مجموعه وابستگی های تابعی تعریف شده روی این رابطه باشد آنگاه رابطه R در BCNF است اگر و فقط اگر به ازای هر یک از اعضای F مثل $A \rightarrow B$ حداقل یکی از دو شرط زیر برقرار باشد:

- $A \rightarrow B$ یک trivial باشد
 - سمت چپ وابستگی تابعی **super key** باشد.
- رابطه R را با دو ویژگی a و b در نظر می گیریم:
- تعداد حالت های وابستگی تابعی در یک رابطه با دو ویژگی 4 تا است.

حالت اول: رابطه هیچ گونه وابستگی تابعی نداشته باشد:

پس کلید مجموعه ی $\{a, b\}$ است. پس طبق تعریف bcnf است.

حالت دوم: وابستگی تابعی $a \rightarrow b$ وجود داشته باشد:

بنابراین a کلید است. پس طبق تعریف bcnf است.

حالت سوم: وابستگی تابعی $b \rightarrow a$ وجود داشته باشد:

بنابراین b کلید است. پس طبق تعریف bcnf است.

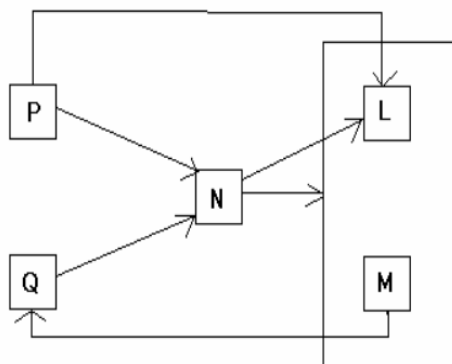
حالت چهارم: وابستگی تابعی $a \rightarrow b$ و $b \rightarrow a$ وجود داشته باشد:

$\{a\}$ و $\{b\}$ هر دو کلید هستند. پس طبق تعریف bcnf است.

2. همان طور که می‌دانیم، گاهی لازم می‌شود اصول نرمال سازی را بخاطر مسائلی مانند کارایی پایگاه، تعدیل کنیم که به آن Denormalization می‌گویند. در مورد هر یک از فیلدهای ستون موضوع جدول، ویژگی‌های دو رویکرد Normalization و Denormalization را توضیح دهید.

موضوع	Normalization	Denormalization
پیاده سازی	افزونگی داده را از پایگاه داده حذف میکند	برای ترکیب جدول های متفاوت به یک جدول استفاده می شود تا پرس و جو ها سریع تر اجرا شوند
تمرکز اصلی روش	نرمال سازی بر پاک کردن پایگاه داده از داده های غیر قابل استفاده تمرکز می کند تا افزونگی داده و ناسازگاری را کاهش دهد	این رویکرد بر روی به دست آوردن اجرای سریع پرس و جوها از طریق معرفی داده های اضافه تمرکز می کند
مصرف حافظه	از حافظه ی کمتری استفاده میکند	به دلیل افزونگی داده از حافظه ی بیشتری استفاده میکند
یکپارچگی داده	نرمال سازی یکپارچگی داده را حفظ می کند. به طور مثال حذف کردن داده از یک جدول باعث ناسازگاری در ارتباط بین جدول ها نمی شود	یکپارچگی داده را حفظ نمی کند
محل استفاده	در جایی استفاده می شود که هزینه ی انجام joinها سنگین است و کویری های پرتکرار روی جدول ها اجرا می شوند.	در جایی استفاده می شود که تعداد زیادی update و insert و delete انجام می شود و هزینه ی join جدول ها سنگین نیست

3. پوشش کانونی برای FD های حاصل از شکل زیر را بدست آورید. (همه ی مراحل را بنویسید)



با توجه به شکل مجموعه ی FD ها به صورت زیر است:

$N \rightarrow LM$
 $N \rightarrow L$
 $M \rightarrow Q$
 $Q \rightarrow N$
 $P \rightarrow N$
 $P \rightarrow L$

عبارت $N \rightarrow LM$ را باید برای پوشش کانونی بررسی کنیم برای این عبارت دو عبارت زیر را داریم :

$N \rightarrow L$
 $N \rightarrow M$

و اینجا عبارت $N \rightarrow L$ مجدداً تکرار شده است و باید حذف شود.

پوشش کانونی :

$N \rightarrow L$
 $N \rightarrow M$
 $M \rightarrow Q$
 $Q \rightarrow N$
 $P \rightarrow N$

$N \rightarrow LM$
 $M \rightarrow Q$
 $Q \rightarrow N$
 $P \rightarrow N$

4. رابطه $R=\{A,B,C,D,E,F,G,H,I,J\}$ و وابستگی های تابعی زیر را در نظر بگیرید:

$F = \{\{A, B\} \rightarrow \{C\}, \{A\} \rightarrow \{D, E\}, \{B\} \rightarrow \{F\}, \{F\} \rightarrow \{G, H\}, \{D\} \rightarrow \{I, J\}\}$

کلید رابطه چیست ؟ رابطه را به روابط 2NF و سپس 3NF تجزیه کنید.(نوشتن تمام مراحل الزامی است.)

تعیین کلید:

یک مجموعه ی کمینه از ویژگی ها که بستار آن شامل همه ی ویژگی های R است یک کلید است.
به طور مثال مجموعه ی {A,B} این ویژگی را دارد.

$$\{A,B\}^+ = R$$

تجزیه ی R به فرم 2nf و سپس 3nf:

وابستگی های جزئی که باعث می شوند R ، نیز 2NF نباشد را پیدا می کنیم. یعنی باید ویژگی هایی را پیدا کنیم که وابستگی تابعی به بخشی از کلید داشته باشند. یعنی ویژگی هایی که فقط به A و یا فقط به B وابسته هستند. پس ما بستار $\{A\}^+$ و $\{B\}^+$ را حساب میکنیم.

$$\{A\}^+ = \{A, D, E, I, J\}, \quad (A \rightarrow A : \text{بدیهی})$$

$$\{B\}^+ = \{B, F, G, H\}, \quad (B \rightarrow B : \text{بدیهی})$$

برای تجزیه به فرم نرمال دوم ویژگی هایی که به بخشی از کلید وابسته هستند را حذف میکنیم و آن ها را در رابطه های جداگانه همراه با جزئی از کلید قرار می دهیم. ویژگی هایی که هم به A و هم به B وابسته هستند در رابطه ی R باقی می مانند.

$$R1 = \{A, D, E, I, J\}$$

$$R2 = \{B, F, G, H\}$$

$$R3 = \{A, B, C\}$$

به این ترتیب جدول های ما به فرم نرمال دوم تبدیل شدند.

برای تبدیل جداول به فرم نرمال سوم باید به دنبال وابستگی های تراگذاری باشیم.

رابطه ی R1 یک رابطه ی تراگذاری دارد .

$$\{A\} \rightarrow \{D\} \rightarrow \{I, J\}$$

به همین ترتیب ویژگی های A و I را همراه با یک کپی از D به یک رابطه ی جدید میبریم. و رابطه ی R1 به دو رابطه ی R11 و R12 تجزیه می شود.

$$R11 = \{D, I, J\}, \quad R12 = \{A, D, E\}$$

به همین ترتیب R2 به R21 و R22 با توجه به رابطه ی تراگذاری زیر تقسیم می شود.

$$\{B\} \rightarrow \{F\} \rightarrow \{G, H\}:$$

$$R21 = \{B, F\}, \quad R22 = \{F, G, H\}$$

به این ترتیب مجموعه رابطه ها به فرم نرمال سوم به صورت زیر خواهد بود:

$$\{\{D, I, J\}, \{A, D, E\}, \{B, F\}, \{F, G, H\}, \{A, B, C\}\}$$

5. رابطه R را در نظر بگیرید که دارای ویژگی هایی است که اطلاعات course ها و section ها را در یک دانشگاه ثبت می کند.

$$R = \{\text{Course_no, Sec_no, Offering_dept, Credit_hours, Course_level, Instructor_ssn, Semester, Year, Days_hours, Room_no, No_of_students}\}$$

فرض کنید وابستگی های تابعی زیر روی R داریم:

$$\{\text{Course_no}\} \rightarrow \{\text{Offering_dept, Credit_hours, Course_level}\}$$

$\{\text{Course_no, Sec_no, Semester, Year}\} \rightarrow \{\text{Days_hours, Room_no, No_of_students, Instructor_ssn}\}$

$\{\text{Room_no, Days_hours, Semester, Year}\} \rightarrow \{\text{Instructor_ssn, Course_no, Sec_no}\}$

تعیین کنید کدام مجموعه از ویژگی ها کلیدهای R را تشکیل می دهند. رابطه را در سطح 3NF نرمال کنید.

$\{\text{Course_no}\}^+ = \{\text{Course_no, Offering_dept, Credit_hours, Course_level}\}$

بستار $\{\text{Course_no, Sec_no, Semester, Year}\}$ را می خواهیم پیدا کنیم. با توجه به سه وابستگی تابعی سوال بستار آن به صورت زیر است:

$\{\text{Course_no, Sec_no, Semester, Year}\}^+ = \{\text{Course_no, Sec_no, Semester, Year, Days_hours, Room_no, No_of_students, Instructor_ssn, Offering_dept, Credit_hours, Course_level}\}$

بستار $\{\text{Course_no, Sec_no, Semester, Year}\}$ برابر R است بنابراین $\{\text{Course_no, Sec_no, Semester, Year}\}$ کلید کاندید رابطه ی R است.

$\{\text{Room_no, Days_hours, Semester, Year}\}^+ = \{\text{Room_no, Days_hours, Semester, Year, Instructor_ssn, Course_no, Sec_no, Offering_dept, Credit_hours, Course_level, No_of_students}\}$

بستار $\{\text{Room_no, Days_hours, Semester, Year}\}$ برابر R است بنابراین $\{\text{Room_no, Days_hours, Semester, Year}\}$ کلیدهای کاندید رابطه ی R است.

بر اساس فرم دوم نرمال نباید وابستگی جزئی وجود داشته باشد. $\{\text{Course_no}\}$ بخشی از کلید است که $\{\text{Offering_dept, Credit_hours, Course_level}\}$ را مشخص میکند که نباید در جدول وجود داشته باشد. بنابراین رابطه ی R را به رابطه ی R1 و R2 تقسیم میکنیم :

$R1 = \{\text{Course_no, Offering_dept, Credit_hours, Course_level}\}$

$R2 = \{\text{Room_no, Days_hours, Semester, Year, Instructor_ssn, Course_no, Sec_no, No_of_students}\}$

رابطه های R1 و R2 در فرم نرمال سوم هستند چون هیچ رابطه ی تراگذاری ندارند.

6. رابطه $\text{REFRIG}(\text{Model\#, Year, Price, Manuf_plant, Color})$ یا به اختصار $\text{REFRIG}(M, Y, P, MP, C)$ و وابستگی های تابعی زیر را در نظر بگیرید:

$F = \{M \rightarrow MP, \{M, Y\} \rightarrow P, MP \rightarrow C\}$

a. هر یک از موارد زیر را به عنوان یک کلید کاندید برای REFRIG ارزیابی کنید و دلایلی را ارائه دهید که چرا می تواند یا نمی تواند کلید باشد.

$\{M\}, \{M, Y\}, \{M, C\}$

b. براساس تعیین کلید بالا، بیان کنید که آیا رابطه REFRIG در 3NF و در BCNF است؟ دلایل مناسب را برای پاسخ خود ارائه دهید.

قسمت اول:

نحوه شناسایی کلیدهای کاندید یک رابطه: (صفحه 15 اسلاید)

اگر R یک رابطه و F مجموعه وابستگی های تابعی موجود روی آن رابطه باشد، در این صورت A (زیرمجموعه ای از خصیصه های R) به عنوان یک کلید کاندید برای R محسوب می شود:

1. همه اعضای A مستقل باشند یعنی هیچ یک از اعضای A قابل استنتاج از روی سایرین نباشد
2. بستار A همه ویژگی های رابطه R را شامل شود.

در اینجا {M} نمیتواند کلید کاندید باشد زیرا بستار آن شامل ویژگی های Y و P نیست.

M -> MP
M,Y ->P
MP ->C
M ->C
{M,Y}+= M , MP , C

در اینجا {M,Y} یک کلید کاندید است زیرا :

{M , Y} ->P

{M,Y} ->M
M ->MP
{M,Y} -> MP

{M,Y} ->MP
MP ->C
{M,Y} ->C

{M,Y}+=M , Y , P , MP , C

در اینجا {M,C} کلید کاندید نمی تواند باشد زیرا بستار آن شامل ویژگی های P و Y نیست.

قسمت دوم:

در اینجا {بنابر صفحه 28 اسلاید} REFRIG در 3NF نیست زیرا

M->MP

همچنین REFRIG در BCNF نیست زیرا M یک super key در M -> MP نیست.

