پایگاه داده ۲
جلسه۷

فرجیان

IASBS
1992-2012

پایگاه داده‌ها

وابستگی و نرمال‌سازی

مقدمه

- سئوال:
  – اگر همه اطلاعات مربوط به دروس (شامل درس‌ها و گروه‌هاي آنها) را در يک جدول مي‌آورديم، چه مي‌شد؟

جواب : مزيت عمده آن ساده تر شدن درخواست ها و پرس و جوها از بانك اطلاعاتي و عدم نياز به پيوندها (Join) و در نتيجه افزايش سرعت پاسخ گوئي

IASBS
1992-2012

کد دانشجويي	نام	فاميلی	ش. ش	سال تولد	نام پدر	آدرس	تلفن	نام درس	نمره
2	علی	محمدی	11	1360	محمد	رشت	2332	رياضي	13
5	محمد	ناصری	11	1362	رضا	زنجان	3251	رياضي	14
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	فيزيك	18
5	محمد	ناصری	11	1362	رضا	زنجان	3251	فيزيك	15
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	شيمي	17
2	علی	محمدی	11	1360	محمد	رشت	2332	فيزيك	16
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	شيمي	19
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	فارسي	16
3	رضا	رضايی	451	1360	احمد	تبريز	2324	فارسي	9

سه اشكال عمده در رابطه با جداول

1- افزونگي داده‌ها (data redundancy)

– افزونگي يعني تكرار بي‌رويه داده‌ها.

– در بانک اطلاعات رابطه‌اي، تكرار داده‌ها تنها راه برقراري ارتباط بين جداول است و از آن به عنوان كليد خارجي ياد مي‌شود. تكرار بيش از اين، افزونگي نام دارد.

معايب عمده :

1- افزونگي داده ها (Data Redundancy)

کد دانشجويي	نام	فاميلی	ش. ش	سال تولد	نام پدر	آدرس	تلفن	نام درس	نمره
2	علی	محمدی	11	1360	محمد	رشت	2332	رياضي	13
5	محمد	ناصری	11	1362	رضا	زنجان	3251	رياضي	14
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	فيزيك	18
5	محمد	ناصری	11	1362	رضا	زنجان	3251	فيزيك	15
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	شيمي	17
2	علی	محمدی	11	1360	محمد	رشت	2332	فيزيك	16
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	شيمي	19
6	عليرضا	كاظمي	354	1358	حسين	انزلي	2121	فارسي	16
3	رضا	رضايی	451	1360	احمد	تبريز	2324	فارسي	9

سه اشکال عمده در رابطه با جداول

۲- بی‌نظمی (anomaly)

- وجود افزونگی باعث بی‌نظمی در تغییر داده‌ها می‌شود.
- بعنوان مثال در سئوال مطرح شده در ابتدا، اگر بخواهید یک نمره در این جدول وارد نمایید، باید اطلاعات کامل درس را نیز همراه آن وارد نمایید.

- از طرفی updateاطلاعات هم مشکل است. زیرا اگر بخواهیم مثلا" نام درس را تغییر دهیم مجبور به دستکاری در صدها رکورد هستیم.

سه اشکال عمده در رابطه با جداول

۳- مقادیر تهی (NULL values)

- با ادغام جداول، از نشان دادن بعضی از اقلام اطلاعاتی بدون استفاده از مقادیر تهی ناتوانیم.

- مثال: معمولاً دانشکده‌ها همه درس‌های خود را در همه ترم‌ها ارایه نمی‌دهند. این درس‌ها یا باید در جدول نیایند یا برای سایر اطلاعات آن مقدار تهی در نظر بگیریم.
- اگر فرض کنید در مثال قبل بخواهیم ثبت نام اولیه دانشجو (ثبت اطلاعات پایه ای) را انجام دهیم که هنوز انتخاب واحدی را انجام نداده باشد باید مقدار NULLبه جای اسم درس و نمره وارد کنیم.

وابستگي تابعي (functional dependency)

- اگر A و B دو مجموعه صفت در شماي R باشند، آنگاه وابستگي تابعي

$$A \rightarrow B$$

برقرار است اگر براي تمام رابطه‌ها در R، به ازاي هر مقدار A فقط يک مقدار در B داشته باشيم.

يادآوري

- y=f(x)
- به ازاي هر مقدار xيک مقدار براي y داريم.

$$x \xrightarrow{f} y$$

- مثال: آیا S# → avg برقرار است؟
  - برقرار نیست.
- مثال: آیا avg → S# برقرار است؟
  - برقرار است.
- مثال: آیا city → sname,avg برقرار است؟
  - برقرار نیست.
- مثال: s#,sname → city,avg,clg# برقرار است.
  - برقرار است.

- نکته: وابستگي تابعي براي تعريف محدوديت‌هاي بانک اطلاعات نيز به‌کار مي‌رود.

- مثال: avg → S# ؟

  – هر دانشجو یک معدل دارد. (یک محدودیت است)

- مثال: s# → sname ؟

  – درست نیست.

وابستگي تابعي كامل (FFD)
Full Functional Dependency

- اگر داشته باشيم A→B، چنانچه هر صفتي از A را حذف نماييم ديگر وابستگي تابعي وجود نداشته باشد، در اينصورت مي‌گوييم A→B وابستگي تابعي كامل است و مي‌نويسيم:

$$A \xrightarrow{FFD} B$$

وابستگي تابعي كامل مشابه خاصيت كمينگي است.

- مثال: $s\#, sname \xrightarrow{FFD} city$ ؟

ابر کلید

- اگر B شامل تمام صفات شماي R باشد و داشته باشيم A → B، در اينصورت A را ابرکليد ناميده و مي‌نويسيم A → R.

- مثال: s# → stud ابر کلید
- مثال: s#, sname → stud ابر کلید
- مثال: avg → stud

کلید کاندید

- اگر $A \xrightarrow{FFD} R$ آنگاه گوییم، A کلید کانديداي شماي R است.

- مثال:

$$s\#,sname \xrightarrow{FFD} stud$$

$$s\# \xrightarrow{FFD} stud$$

وابستگي تابعي بديهي

- اگر B زيرمجموعه A باشد، A → B را وابستگي تابعي بديهي مي‌ناميم.

- مثال:

$$sname,avg \longrightarrow avg$$

مجموعه پوششي وابستگي

- اگر $\mathcal{F}$ يک مجموعه از وابستگي‌هاي تابعي باشد آنگاه مجموعه تمام وابستگي‌هاي تابعي که از آن منتج مي‌شود را مجموعه پوششي $\mathcal{F}$ مي‌ناميم و با $\mathcal{F}^+$ نمايش مي‌دهيم.

قواعد استخراج وابستگي تابعي

1. قاعده بازتاب (Reflexivity):

$B \subset A \quad \Rightarrow \quad A \rightarrow B$

همان وابستگی تابعی بدیهی است

2. قاعده افزایش (Augmentation):

$A \rightarrow B \quad \Rightarrow \quad AC \rightarrow BC$

قواعد استخراج وابستگي تابعي

3. قاعده انتقال (Transfinite):

$$\left.\begin{array}{l} A \rightarrow B \\ \wedge \\ B \rightarrow C \end{array}\right\} \Rightarrow A \rightarrow C$$

4. قاعده اجتماع (Union):

$$\left.\begin{array}{l} A \rightarrow B \\ \wedge \\ A \rightarrow C \end{array}\right\} \Rightarrow A \rightarrow BC$$

قواعد استخراج وابستگي تابعي

5. قاعده تجزيه (Decomposition):

$$A \rightarrow BC \Rightarrow \begin{cases} A \rightarrow B \\ \wedge \\ A \rightarrow C \end{cases}$$

6. تركيب (Composition):

$$\left.\begin{array}{l} A \rightarrow B \\ \wedge \\ C \rightarrow D \end{array}\right\} \Rightarrow AC \rightarrow BD$$

قواعد استنتاج آرمسترانگ

فرض: A، B، C و D زیرمجموعه‌هایی از صفات رابطه R باشند. قواعد زیر برقرارند:

1. قاعده انعکاس: اگر $B \subseteq A$ آنگاه: $A \to B$
2. قاعده تعدی(تراگذری): اگر $A \to B$ و $B \to C$ آنگاه $A \to C$
3. قاعده افزایش: اگر $A \to B$ آنگاه $(A,C) \to (B,C)$
4. قاعده تجزیه: اگر $A \to (B,C)$ آنگاه $A \to C$ و $A \to B$
5. قاعده اجتماع: اگر $A \to C$ و $A \to B$ آنگاه $A \to (B,C)$
6. قاعده ترکیب: اگر $C \to D$ و $A \to B$ آنگاه $(A,C) \to (B,D)$
7. قاعده شبه تعدی: اگر $(C,B) \to D$ و $A \to B$ آنگاه $(A,C) \to D$
8. قاعده یگانگی عمومی: اگر $C \to D$ و $A \to B$ آنگاه $A \cup (C-B) \to (B,D)$