



CSDL 60TH3 (1) - ôn tập có đáp án có thể tham khảo bài này

Cơ sở dữ liệu (Đại học Kinh tế Quốc dân)

ÔN TẬP CƠ SỞ DỮ LIỆU 60TH3

20h15p – 22h 19/5/2020

Câu	Nội dung	Mentor
1	Đại số quan hệ	Đoàn Khánh Chi
2	Các câu lệnh SQL	Nguyễn Văn Trường
3	Khóa, Phụ thuộc hàm, kiểm tra tính toàn vẹn sau phân rã	Trần Thị Ngọc Bích
4	Các dạng chuẩn và chuẩn hóa	Nguyễn Cao Thiêm

Mục đích buổi ôn tập:

- Ôn tập kiến thức cơ bản (*chủ yếu* dành cho những bạn còn đang cảm thấy khó khăn trong việc học môn CSDL)
- Hệ thống lại kiến thức môn học Cơ sở dữ liệu cho sinh viên 60TH3.
- Trao đổi, bổ sung , chuẩn bị kiến thức để ôn thi hiệu quả.

CÂU 1: Đại số quan hệ (2đ)

Các phép toán đại số quan hệ : 8 phép

1. Phép hợp : $R \cup S = \{ t \mid t \in R \text{ hoặc } t \in S \}$

Phép hợp hai quan hệ R và S thực chất là việc gộp các bản ghi trong hai quan hệ thành 1 quan hệ mà các bản ghi trùng nhau thì chỉ giữ lại một bản ghi.

R	S	R ∪ S																																				
<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a1</td><td>b1</td><td>c1</td></tr> <tr><td>a1</td><td>b2</td><td>c1</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a1	b1	c1	a1	b2	c1	a2	b2	c2	<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a2</td><td>b1</td><td>c2</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a2	b1	c2	a2	b2	c2	<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a1</td><td>b1</td><td>c1</td></tr> <tr><td>a1</td><td>b2</td><td>c1</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> <tr><td>a2</td><td>b1</td><td>c2</td></tr> </table>	A	B	C	a1	b1	c1	a1	b2	c1	a2	b2	c2	a2	b1	c2
A	B	C																																				
a1	b1	c1																																				
a1	b2	c1																																				
a2	b2	c2																																				
A	B	C																																				
a2	b1	c2																																				
a2	b2	c2																																				
A	B	C																																				
a1	b1	c1																																				
a1	b2	c1																																				
a2	b2	c2																																				
a2	b1	c2																																				

2. Phép giao : $R \cap S = \{ t \mid t \in R \text{ và } t \in S \}$

Phép giao giữa R và S thực chất là việc chọn ra trong hai quan hệ R và S những bản ghi trùng nhau.

R	S	$R \cap S$																											
<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a1</td><td>b1</td><td>c1</td></tr> <tr><td>a1</td><td>b2</td><td>c1</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a1	b1	c1	a1	b2	c1	a2	b2	c2	<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a2</td><td>b1</td><td>c2</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a2	b1	c2	a2	b2	c2	<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a2	b2	c2
A	B	C																											
a1	b1	c1																											
a1	b2	c1																											
a2	b2	c2																											
A	B	C																											
a2	b1	c2																											
a2	b2	c2																											
A	B	C																											
a2	b2	c2																											


3. Phép trừ : $R - S = \{ t \mid t \in R \text{ hoặc } t \notin S \}$

Phép trừ R - S thực chất là việc chọn ra các bản ghi chỉ có ở R mà không có ở S.

R	S	$R - S$																														
<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a1</td><td>b1</td><td>c1</td></tr> <tr><td>a1</td><td>b2</td><td>c1</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a1	b1	c1	a1	b2	c1	a2	b2	c2	<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a2</td><td>b1</td><td>c2</td></tr> <tr><td>a2</td><td>b2</td><td>c2</td></tr> </table>	A	B	C	a2	b1	c2	a2	b2	c2	<table border="1"> <tr><th>A</th><th>B</th><th>C</th></tr> <tr><td>a1</td><td>b1</td><td>c1</td></tr> <tr><td>a1</td><td>b2</td><td>c1</td></tr> </table>	A	B	C	a1	b1	c1	a1	b2	c1
A	B	C																														
a1	b1	c1																														
a1	b2	c1																														
a2	b2	c2																														
A	B	C																														
a2	b1	c2																														
a2	b2	c2																														
A	B	C																														
a1	b1	c1																														
a1	b2	c1																														

4. Tích đề-các : $R \times S$

Phép tích đề các của hai quan hệ R và S thực chất là việc lấy mọi bản ghi của R "gắn" với mọi bản ghi của S.

R		S		R x S																																																															
<table><tr><th>A</th><th>B</th><th>C</th></tr><tr><td>a1</td><td>b1</td><td>c1</td></tr><tr><td>a1</td><td>b2</td><td>c1</td></tr><tr><td>a2</td><td>b2</td><td>c2</td></tr></table>	A	B	C	a1	b1	c1	a1	b2	c1	a2	b2	c2		<table><tr><th>D</th><th>E</th><th>F</th></tr><tr><td>d2</td><td>e1</td><td>f2</td></tr><tr><td>d2</td><td>e2</td><td>f2</td></tr></table>	D	E	F	d2	e1	f2	d2	e2	f2		<table><tr><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th><th>F</th></tr><tr><td>a1</td><td>b1</td><td>c1</td><td>d2</td><td>e1</td><td>f2</td></tr><tr><td>a1</td><td>b1</td><td>c1</td><td>d2</td><td>e2</td><td>f2</td></tr><tr><td>a1</td><td>b2</td><td>c1</td><td>d2</td><td>e1</td><td>f2</td></tr><tr><td>a1</td><td>b2</td><td>c1</td><td>d2</td><td>e2</td><td>f2</td></tr><tr><td>a2</td><td>b2</td><td>c2</td><td>d2</td><td>e1</td><td>f2</td></tr><tr><td>a2</td><td>b2</td><td>c2</td><td>d2</td><td>e2</td><td>f2</td></tr></table>	A	B	C	D	E	F	a1	b1	c1	d2	e1	f2	a1	b1	c1	d2	e2	f2	a1	b2	c1	d2	e1	f2	a1	b2	c1	d2	e2	f2	a2	b2	c2	d2	e1	f2	a2	b2	c2	d2	e2	f2
A	B	C																																																																	
a1	b1	c1																																																																	
a1	b2	c1																																																																	
a2	b2	c2																																																																	
D	E	F																																																																	
d2	e1	f2																																																																	
d2	e2	f2																																																																	
A	B	C	D	E	F																																																														
a1	b1	c1	d2	e1	f2																																																														
a1	b1	c1	d2	e2	f2																																																														
a1	b2	c1	d2	e1	f2																																																														
a1	b2	c1	d2	e2	f2																																																														
a2	b2	c2	d2	e1	f2																																																														
a2	b2	c2	d2	e2	f2																																																														

5. Phép chia : $R(z = x+y) \div S(y)$ được kết quả là 1 quan hệ xác định trên tập thuộc tính

(x)

Kết quả trả về là [các bộ với các thuộc tính chỉ có trong R] sao cho sự kết hợp của nó với [các bộ trong S] có mặt trong R

▪ Ví dụ: Đưa ra môn học được dạy ở tất cả các khoá học

Subject		Course		Kết quả	
Name	Course	Course		Name	
Systems	BCS	BCS		Database	
Database	BCS	MCS			
Database	MCS				
Algebra	MCS				

6. Phép chiếu : $R[X]$, kí hiệu $\pi_X(R) \rightarrow$ VD: $\pi_{Name}(Subject)$

Là phép cắt dọc quan hệ để lấy ra dữ liệu trên một số thuộc tính của quan hệ.

Subject

Name	Course	Name
System	BCS	System
Database	BCS	Database
Database	MCS	Database
Algebra	MCS	Algebra

7. Phép chọn : $\sigma_D(R) \rightarrow$ VD: $\sigma_{Name = 'Database'}(Subject)$ Chọn ra các bộ dữ liệu thỏa mãn điều kiện chọn.

Subject

Name	Course	Name	Course
System	BCS	Database	BCS
Database	BCS	Database	MCS
Database	MCS		
Algebra	MCS		

8. Phép kết nối : Phép kết nối giữa hai quan hệ R và S thực chất là việc lấy một bản ghi của R "gắn" với một bản ghi của S sao cho bản ghi kết quả thỏa mãn điều kiện kết nối.

Student $\bowtie_{Id=SID}$ Enrol

Student			Enrol		
Id	Name	Suburb	SID	Course	
1108	Robert	Kew	3936	101	
3936	Glen	Bundoora	1108	113	
8507	Norman	Bundoora	8507	101	
8452	Mary	Balwyn			

Kết quả				
SID	Id	Name	Suburb	Course
1108	1108	Robert	Kew	113
3936	3936	Glen	Bundoora	101
8507	8507	Norman	Bundoora	101

Phép kết nối bằng (Phép kết nối tự nhiên) : kết nối từng bản ghi của R với từng bản ghi của S => được bảng mới lược bỏ bớt đi 1 cột thuộc tính chung

Takes		Enrol				
SID	SNO	SID	Course	SID	SNO	Course
1108	21	3936	101	1108	21	113
1108	23	1108	113	1108	23	113
8507	23	8507	101	8507	23	101
8507	29			8507	29	101

Phép kết nối nội : là kết nối bằng nhưng trong trường hợp 2 thuộc tính cùng tên thì kết quả vẫn giữ lại 2 tên thuộc tính

Phép kết nối ngoài (Phép kết nối trái / Phép kết nối phải) :

Student			Enrol					
ID	Name	Suburb	SID	Course	ID	Name	Suburb	Course
1108	Robert	Kew	3936	101	1108	Robert	Kew	113
3936	Glen	Bundoora	1108	113	3936	Glen	Bundoora	101
8507	Norman	Bundoora	8507	101	8507	Norman	Bundoora	101
8452	Mary	Balwyn			8452	Mary	Balwyn	null

Chữa bài tập 3.2 (Bài tập cuối chương 3) :

- **HANG** (MaH, TenH, SLTon)
- **KHACH** (MaK, TenK, Diachi)
- **HOADON** (SoHD, ngayHD, MaK)
- **CHITIETHD** (SoHD, MaH, SLBan, DGia)

- Đưa ra danh sách địa chỉ của các khách hàng: $\pi_{Diachi}(KHACH)$
- Đưa ra tên hàng và số lượng tồn của những mặt hàng: $\pi_{TenH, SLTon}(HANG)$
- Đưa ra thông tin của các mặt hàng có số lượng tồn > 5: $\sigma_{SLTon > 5}(HANG)$
- Đưa ra các thông tin khách hàng có địa chỉ ở Hà Nội: $\sigma_{Diachi = 'Hà Nội'}(KHACH)$
- Đưa ra tên khách hàng mua hàng ngày 1/1/2013:
 $\pi_{TenK}(\sigma_{ngayHD = '1/1/2013'}(KHACH * HOADON))$
- Đưa ra Mã hàng, Tên hàng có đơn giá bán > 200,000:
 $\pi_{MaH, TenH}(\sigma_{DGia > 200,000}(HANG * CHITIETHD))$
- Đưa ra tên khách hàng ở Hải Phòng mua hàng ngày 2/2/2013:
 $\pi_{TenK}(\sigma_{Diachi = 'Hải Phòng'}(KHACH) * \sigma_{ngayHD = '2/2/2013'}(HOADON))$
- Đưa ra tên hàng được bán trong ngày 2/2/2013:
 $\pi_{TenH}[HANG * (\sigma_{ngayHD = '2/2/2013'}(HOADON * CHITIETHD))]$
- Đưa ra các mã hàng chưa từng được bán:
 $\pi_{MaH}(HANG) - (\pi_{MaH}(CHITIETHD))$
- Đưa ra các mã khách chưa từng mua hàng từ ngày 12/12/2012:

$\pi MaK (KHACH) - [\pi MaK (\sigma ngayHD = '12/12/2012' (HOADON))]$

Ví dụ 1 :

THISINH (sbd , hoten , ngaysinh , noisinh , namduthi , matruong)

TRUONG (matruong , tentruong)

MONTHI (maMT , tenMT)

KETQUA (sbd , maMT , diemthi)

1. Cho biết điểm thi các môn của thí sinh có số báo danh là '080191000001'.
 $\pi diemthi [\pi maMT, tenMT (MONTHI) * \sigma sbd = '080191000001' (KETQUA)]$
2. Cho biết các thí sinh có ít nhất một môn thi nào đó bị điểm 0 ở kỳ thi năm 2010.
 $\sigma diemthi = 0 (KETQUA) * \sigma namduthi = 2010 (THISINH)$
3. Cho biết các thí sinh có điểm tất cả các môn thi đều lớn hơn hoặc bằng 8 trong kỳ thi năm 2010.
 $\sigma namduthi = 2010 (THISINH) * [\pi sbd (THISINH) - \pi sbd (\sigma diemthi < 8 (KETQUA))]$
4. Cho biết các thí sinh cùng trường với thí sinh có số báo danh là '080191000001'.
 $\pi sbd, hoten, matruong (THISINH) * [\pi matruong (\sigma sbd = '080191000001' (THISINH))]$

Ví dụ 2 :

KHACH (maKH , hoten , dchi , sdt , ngsinh , doanhso)

NHANVIEN (maNV , hoten , ngsinh , heso , mucluong)

SANPHAM (maSP , tenSP , DVT , nuocSX , gia)

HOADON (soHD , ngayHD , maKH , maNV , trigia)

CTHD (soHD , maSP , SL)

1. In ra số hóa đơn cùng trị giá của các hóa đơn do nhân viên có tên “Nguyễn Văn A” lập trong ngày 10/10/2005
 $\sigma soHD = trigia [\ HOADON * (\sigma ngayHD = '10/10/2005' (HOADON) * \sigma hoten = 'Nguyễn Văn A' (NHANVIEN))]$
2. In ra danh sách các sản phẩm (mã sản phẩm, tên sản phẩm) được khách hàng có tên “Nguyễn Văn A” mua.
 $\pi maSP, tenSP [SANPHAM * (\sigma hoten = 'Nguyễn Văn A' (KHACH * (HOADON * CTHD)))]$
3. Tìm các số hóa đơn đã mua cùng lúc 2 sản phẩm có mã số “SP01” và “SP02”
 $\pi soHD (\sigma maSP = 'SP01' \vee 'SP02' (CTHD))$
4. In ra danh sách các sản phẩm không bán được trong năm 2005
 $\pi maSP (SANPHAM) - \pi maSP (\sigma ngayHD = '2005' (HOADON * CTHD))$

CÂU 2: Các câu lệnh SQL

1 . Nhóm câu lệnh thao tác với BẢNG

Tạo bảng xong khi mà bạn muốn sửa đổi cấu trúc của bảng:

Thêm 1 ràng buộc

ALTER TABLE<TÊN BẢNG>

ADD CONSTRAINT<TÊN RB>;

Ví dụ tạo ràng buộc :Khóa chính và khóa ngoại

```
ALTER TABLE KHACHHANG  
ADD constraint PK__DONHANG__MaK primary key (MaK)
```

```
ALTER TABLE DONDATHANG  
ADD constraint FK__DONHANG__MaK foreign key (MaK) References [KHACHHANG](MaK)
```

Thêm 1 cột

ALTER TABLE<TÊN BẢNG>

ADD <Tên cột><kiểu dữ liệu>

Ví dụ thêm cột giới tính

```
ALTER TABLE KHACHHANG  
ADD gioitinh nvarchar(5)
```

Xóa bỏ 1 cột

ALTER TABLE<TÊN BẢNG>

DROP COLUMN<Tên cột>

```
ALTER TABLE KHACHHANG  
DROP COLUMN gioitinh
```

2. Nhóm câu lệnh thao tác với bản ghi (dòng dữ liệu)

INSERT INTO

Lệnh:

INSERT INTO <tên bảng>(tencot1, tencot2,...)

VALUES(biểu thức 1, biểu thức 2,...)

Thêm bản ghi thì thêm dấu , và thêm thông tin tiếp ko cần ghi lệnh values

Sửa nội dung của bản ghi

Lệnh

UPDATE Tên bảng

Set cot1=gt1,cot2=gt2,...

Where điều kiện (có hoặc không cũng được)

Ví dụ sửa nội dung tên khách ở bảng khách hàng và mã khách hàng là 002

Update KHACHHANG

Set TENKHACH='HUY'

Where MaK='002';

Xoá bản ghi

DELETE FROM tên bảng

Where dieukien;

Ví dụ Xóa khách hàng mã số là 002 trong bảng khách hàng

DELETE FROM KHACHHANG

Where MaK='002';

3 .Nhóm câu lệnh truy vấn dữ liệu

Câu lệnh Select(Chọn dữ liệu từ 1 cơ sở dữ liệu)

Có 3 kiểu chọn

Chọn Tất cả **Select ***

Chọn không trùng **Select Distinct**

Chọn top đầu ý **Select top(giatri)**

Ví dụ :Chọn 3 nhân viên hiển thị đầu danh sách

Select top 3 manv,tennv

From NHANVIEN

Thường phép chọn này sẽ luôn đi với câu lệnh **Order by**
Order by nhằm để sắp xếp kết quả truy vấn theo 1 hay nhiều cột
Sắp xếp tăng ta dùng thêm **ASC** ngược lại giảm thì dùng **DESC**

Ví dụ:Chọn top 3 nhân viên lương nhiều nhất:

Select top 3 manv,tennv

From NHANVIEN

Order by lương DESC

3.Câu lệnh Where-xác định điều kiện để truy xuất dữ liệu

Có nhiều loại toán tử so sánh(=,>,<,...)

Ví dụ:Chọn ra nhân viên mã nhân viên =002

Select manv,tennv,ngaysinh

From NHANVIEN

Where (manv='002')

Kiểm tra thuộc khoảng –betweenand

Ví dụ Đưa ra thông tin lương của nhân viên thuộc khoảng 2500000 đến 10000000

```
SELECT *  
FROM NHANVIEN  
WHERE lương BETWEEN 2500000 AND 10000000;
```

Câu lệnh Where tác dụng phép kết nối tự nhiên

Phép kết nối như nào?

Ví dụ: cho 2 bảng thông tin: NHANVIEN(manv,hoten,ngaysinh) và CHUCVU(manv,tenchucvu)
In ra thông tin của nhân viên và chức vụ nhân viên đó làm

```
SELECT *  
  
FROM NHANVIEN, CHUCVU  
  
Where NHANVIEN.manv=CHUCVU.manv
```

4.Các thao tác bổ sung.

Có các câu lệnh **Order by**, **Group by**....**Having**

Order by :Cho biết bảng đã được sắp xếp 1 giá trị thuộc tính trong bảng đó

Ví dụ:Cho bảng thông tin NHANVIEN(manv,hoten,ngaysinh,lương)

```
SELECT *  
FROM NHANVIEN  
  
Order by lương asc
```

Chú ý:Sắp xếp tăng ta dùng thêm **ASC** ngược lại giảm thì dùng **DESC**

Group by... Having

Group by để thông kê trên các nhóm dữ liệu trên mỗi nhóm có các phép tính toán cộng,trung bình cộng,....

SUM([ALL/DISTINCT]biểu thức) Cộng

AVG([ALL/DISTINCT]biểu thức) Trung bình

COUNT([ALL/DISTINCT]biểu thức) Đếm

MAX(biểu thức)

MIN(biểu thức)

Đó là những hàm tính toán thường gặp và đi kèm với là phải có **Group by**

Nhóm đó mà cần thêm điều kiện dùng thêm câu lệnh **Having**

Having (điều kiện Group by)

Ví dụ: Cho bảng thông tin NHANVIEN(manv,hoten,ngaysinh,luong)

Đưa ra thông tin nhân viên và tổng số lương của các nhân viên có lương lớn hơn 1000000

```
SELECT manv,hoten, SUM (luong) AS "Tong so luong"
```

```
FROM sanpham
```

```
GROUP BY manv,hoten
```

```
HAVING luong > 1000000;
```

CÂU 3: Phụ thuộc hàm, khóa, kiểm tra tính toàn vẹn thông tin khi phân rã.

- **Điều kiện tiên quyết:** Phải biết tính Bao đóng của Thuộc tính trên tập Phụ thuộc hàm. (X+)

Cách tính bao đóng:

Vd1: Cho Tập thuộc tính: ABCDE, tập PTH $\{CD \rightarrow A, E \rightarrow B, DB \rightarrow C, C \rightarrow D\}$.

Tính A^+, B^+, C^+, CD^+

$A^+ : \{A\}$

$B^+ : \{B\}$

$C^+ : \{C\}$

$CD^+ : \{CDA\}$

Vd2: Cho $Q(ABCDEH)$ với $F = \{AB \rightarrow C, CD \rightarrow E, AH \rightarrow B, B \rightarrow D, A \rightarrow D\}$.

Hãy tính bao đóng của tập thuộc tính AC (Tính $(AC)^+$)

- Ban đầu ta có $ketqua = AC$;
- Ta lần lượt xét tất cả các phụ thuộc hàm có trong F:
- $AB \rightarrow C$ có $AB \not\subseteq ketqua$ nên $ketqua$ vẫn giữ nguyên.
- $CD \rightarrow E$ có $CD \not\subseteq ketqua$ nên $ketqua$ vẫn giữ nguyên.
- $AH \rightarrow B$ có $AH \not\subseteq ketqua$ nên $ketqua$ vẫn giữ nguyên.
- $B \rightarrow D$ có $B \not\subseteq ketqua$ nên $ketqua$ vẫn giữ nguyên.
- $A \rightarrow D$ có $A \subseteq ketqua$ nên $ketqua = ketqua \cup D = ACD$
- **Quay lại từ đầu tập F lần 2:**
- $AB \rightarrow C$ có $AB \not\subseteq ketqua$ nên $ketqua$ vẫn giữ nguyên.

- $CD \rightarrow E$ có $CD \subseteq \text{ketqua}$ nên $\text{ketqua} = \text{ketqua} \cup E = ACDE$ vẫn giữ nguyên.
- $AH \rightarrow B$ có $AH \not\subseteq \text{ketqua}$ nên ketqua vẫn giữ nguyên.
- $B \rightarrow D$ có $B \not\subseteq \text{ketqua}$ nên ketqua vẫn giữ nguyên.
- Tiếp tục các phụ thuộc hàm sau không làm thay đổi kết quả.
- Cuối cùng ta được $(AC)^+ = ACDE$

Chứng minh 1 PTH thuộc F^+ :

- CM $X \rightarrow Y$ thuộc F^+ . Tính X^+ , nếu Y thuộc X^+ thì $X \rightarrow Y$ thuộc F^+ .

vd:

Chứng minh $DE \rightarrow DC$ thuộc tập PTH $\{ CD \rightarrow A, E \rightarrow B, DB \rightarrow C, C \rightarrow D \}$;

$DE^+ : \{DEBCA\}$. DC thuộc $DE^+ \Rightarrow DE \rightarrow DC$ thuộc tập PTH.

XÁC ĐỊNH PHỦ TỐI THIỂU: (Có 2 cách tìm phủ tối thiểu có thể áp dụng 1 trong 2 cách tùy chọn)

(Chọn cách 2 để trình bày ; cái này tùy mỗi người nhé nhưng cách 2 được trình bày theo các bước rõ ràng hơn ^^)

Các bước tìm phủ tối thiểu:

B1: Tách các thuộc tính từ vế phải(nếu vế phải chứa từ 2 thuộc tính trở lên)

B2: Loại bỏ thuộc tính dư thừa ở vế trái

B3: Loại bỏ các PTH dư thừa

B4: Kết luận

Vd1: Cho $R(ABCD)$, $F=\{A \rightarrow BC, B \rightarrow C, AB \rightarrow D\}$, Tìm F_c

B1: Có A BC nên ta tách thành A B, A C

$$F1 = \{A B, A C, B C, AB D\}$$

B2: Loại bỏ thuộc tính dư thừa ở vế trái ABD

Tính $A^+ = ABCD$ chứa B nên B dư thừa

Tính $B^+ = BC$ không chứa A nên A không dư thừa

$$\Rightarrow F2 = \{AB, AC, BC, AD\}$$

B3: Loại bỏ các PTH dư thừa

Xét $A \rightarrow B$: Tính $(A^+ - F - (A \rightarrow B)) = AD$ không chứa B nên $A \rightarrow B$ không dư thừa.

Xét $A \rightarrow C$: Tính $(A^+ - F - (A \rightarrow C)) = ABCD \supseteq C$ nên $A \rightarrow C$ dư thừa, loại $A \rightarrow C$

Xét $B \rightarrow C$: Tính $(B^+ - F - (B \rightarrow C)) = B$ không chứa C nên $B \rightarrow C$ không dư thừa.

Xét AD: Tính $(A^+ - F - (A \rightarrow D)) = A$ không chứa D nên $A \rightarrow D$ không dư thừa.

$$\Rightarrow F3 = \{A \rightarrow B, B \rightarrow C, A \rightarrow D\}$$

Không thể loại bỏ thêm được pth nào nữa

$$\Rightarrow \text{Phủ tối thiểu } F_c = \{A \rightarrow B, B \rightarrow C, A \rightarrow D\}$$

TÌM KHÓA

Có 2 thuật toán tìm **một khóa**: (các bạn tự đọc thêm)

→ *Dùng thuật toán tìm toàn bộ khóa*

F: Tập phụ thuộc hàm, U: Tập thuộc tính.

B1: Tìm giao của các khóa $X = U - (R - L)$. (Lấy thuộc tính trừ vế phải của phụ thuộc hàm)

B2: Tính X^+ , nếu $X^+ = U$, thì quan hệ có một khóa duy nhất là X , ngược lại thì chuyển sang **B3**

B3: Tính X^+ (X hợp với các thuộc tính trong U), nếu $= U$ thì đó là khóa của quan hệ

Mẹo B3 : Chỉ cần kiểm tra X hợp với các thuộc tính nằm trong vế trái của các PTH.

Ví dụ 1: (Trường hợp Có 1 khóa duy nhất): $Q(ABCDEH)$ với $F = \{AB \rightarrow C, CD \rightarrow E, AH \rightarrow B, B \rightarrow D, A \rightarrow D\}$

Cách làm:

-Ta có giao của các khóa X là: $\{ABCDEH\} - \{CEBD\} = \{AH\}$

Tính bao đóng của AH :

$(AH)^+ = AHBDCE = Q$

\Rightarrow Vậy AH là khóa duy nhất của Q

Ví dụ 2: (Trường hợp có nhiều khóa)

$U(ABCDEI), G = \{A \rightarrow C, AB \rightarrow C, C \rightarrow DI, CD \rightarrow I, EC \rightarrow AB, EI \rightarrow C\}$

- Ta có giao của các khóa X là: $\{ABCDEI\} - \{CDIAB\} = \{E\}$

- Tính bao đóng của E :

$E^+ = E \cup U \Rightarrow$ Quan hệ có nhiều hơn 1 khóa.

- Bổ sung thêm các thuộc tính khác vào cùng với giao của các khóa ta sẽ có các khóa khác nhau của lược đồ quan hệ:

- Bổ sung thêm A , ta có $(AE)^+ = ABCDEI = U$

- Bổ sung thêm C , ta có $(CE)^+ = ABCDEI = U$

- Bổ sung thêm I , ta có $(EI)^+ = ABCDEI = U$

\Rightarrow Vậy quan hệ R có 3 khóa AE hoặc CE hoặc EI

KIỂM TRA PHÉP TÁCH BẢO TOÀN THÔNG TIN

- Dùng bảng Tableau

Bỏ qua lý thuyết đọc dài hoang mang → đi luôn vào ví dụ.

Ví dụ

Cho $R(ABCDE)$, $F = \{AB \rightarrow C, C \rightarrow D, D \rightarrow AE\}$

-1) $R_1(ABD)$, $R_2(ACE)$.

-2) $R_1(ABC)$, $R_2(ABDE)$.

Xây dựng bảo Tableau tương ứng với các phép tách R_1 , R_2

A	B	C	D	E
a1	a2	b1	a4	b2
a1	b3	a3	b4	a5

Lần 1:

- Xét $AB \rightarrow C$: Thấy R_1 , R_2 không bằng nhau trên AB, bảng Tableau giữ nguyên

- Xét $C \rightarrow D$: Thấy R_1 , R_2 không bằng nhau trên C, bảng Tableau giữ nguyên

- Xét $D \rightarrow AE$: Thấy R_1 , R_2 không bằng nhau trên D, bảng Tableau giữ nguyên

Lần 2, lần 3 không thay đổi.

→ Vậy phép tách không bảo toàn thông tin

A	B	C	D	E
a1	a2	a3	b1	b2
a1	a2	b3	a4	a5

Lần 1:

- Xét pth: $AB \rightarrow C$: Ta thấy R1, R2 bằng nhau trên AB $\Rightarrow b3 \rightarrow a3$.

Ta thấy R2 toàn các giá trị từ a1 \rightarrow a5. (Chỉ cần 1 hàng chứa toàn a là có thể kết luận bảo toàn thông tin mà không cần xét thêm các pth khác)

\rightarrow Vậy phân rã bảo toàn thông tin.

Câu 4:

Các bước làm câu 4:

- Kiểm tra xem câu 1, câu 2, câu 3 đã làm đúng chưa. ? (ăn chắc 6đ 7đ trước khi mong muốn đạt 8đ 9đ) .
- Xác định dạng chuẩn:
- Chuẩn hóa về 3NF hoặc BCNF

1. Điều kiện tiên quyết:

- *Tìm Phủ tối thiểu của tập phụ thuộc hàm.*
- *Tìm khóa của quan hệ.*

2. Kiểm tra dạng chuẩn của lược đồ quan hệ:

- Thuộc tính đơn là gì ?
- Phụ thuộc hàm nguyên tố là gì ?
- Các dạng chuẩn: 1NF, 2NF, 3NF, BCNF.

→ **Thuộc tính đơn** là thuộc tính không chứa thuộc tính con nào bên trong nó. (Hiểu đơn giản là miền giá trị của nó là đơn).

Ví dụ:

→ Thuộc tính Mã khách: Trong đó chỉ có duy nhất 2 miền giá trị.

→ Thuộc tính Địa chỉ: trong thuộc tính địa chỉ chứa 2 thuộc tính Số nhà và Đường.

→ **Phụ thuộc hàm nguyên tố** : $X \rightarrow A$ là PTH nguyên tố nếu: Không $\exists Y$ là tập con thực sự

của $X, Y \rightarrow A \in F$

Ví dụ:

$F = \{AB \rightarrow C, B \rightarrow C\}$ thì:

$AB \rightarrow C$: không là phụ thuộc hàm nguyên tố vì có $B \rightarrow C$.

$B \rightarrow C$: là phụ thuộc hàm nguyên tố.

Chuẩn 1NF:

- Quan hệ đạt chuẩn 1NF khi các thuộc tính đều là thuộc tính đơn / thuộc tính nguyên tố.
- **Phần lớn** các bài tập về chuẩn hóa (trong nội dung môn học) đều đạt chuẩn **1NF**.

Chuẩn 2NF

- Đạt chuẩn 1NF
- Loại bỏ quan hệ Q ở dạng 2NF nếu **tất cả thuộc tính không khoá** đều phụ thuộc đầy đủ vào **khóa**.

ví dụ:

Cho $Q(ABCD)$, $F = \{A \rightarrow C, B \rightarrow D\}$.

- Tìm khóa: → Khóa là AB
- Các thuộc tính không khóa: C, D

Xét $AB \rightarrow C$: Thấy $AB \rightarrow C$ không phải là PTH nguyên tố vì xuất hiện $A \rightarrow C$. → Quan hệ không đạt 2NF

tương tự có thể xét $AB \rightarrow D$: Thấy $AB \rightarrow D$ không phải là PTH nguyên tố vì có $B \rightarrow D$. Quan hệ không đạt 2NF.

Vậy quan hệ trên đạt chuẩn 1NF

Chuẩn 3NF :

ĐN1:

- Lược đồ quan hệ Q ở dạng 3NF nếu ở 2NF và tất cả các thuộc tính không khoá không phụ thuộc bắc cầu vào khoá.

ĐN2:

- Lược đồ quan hệ Q ở dạng 3NF nếu ở 1NF và tất cả phụ thuộc hàm không hiển nhiên $X \rightarrow Y$ của F^+ thoả một trong hai điều kiện sau:

- (i) X là một siêu khoá (X chứa một khoá nào đó)
- (ii) Mỗi thuộc tính trong tập (Y - X) nằm trong một khoá nào đó.

Ví dụ:

VD: Cho $Q(ABCDE)$, $F = \{AB \rightarrow CDE, B \rightarrow D, DE \rightarrow ABC\}$ ở dạng 3NF vì:

$AB \rightarrow CDE$ có vế trái là một siêu khoá.

$B \rightarrow D$ có $(VP) - (VT) = D$ chứa trong khoá DE.

$DE \rightarrow ABC$ có vế trái là một siêu khoá.

Chuẩn BCNF:

- Lược đồ quan hệ Q ở BCNF nếu ở dạng 1NF và tất cả phụ thuộc hàm không hiển nhiên $X \rightarrow Y$ của F^+ thì X là một siêu khoá (X chứa một khoá nào đó).

Lược đồ CSDL C ở dạng BCNF nếu tất cả các Q_i của C đều ở dạng BCNF.

➔ Bản chất chuẩn BCNF là chuẩn **3NF**, tuy nhiên điều kiện bị khuyết một nửa so với chuẩn 3NF (khắt khe hơn).

Ví dụ:

$Q(ABCD), F=\{A \rightarrow B, C \rightarrow D\}$

- Tìm khóa: Khóa là AC.
 - Xét chuẩn 1NF: **đạt chuẩn 1NF** (các thuộc tính đều là thuộc tính đơn)
 - Xét chuẩn 2NF: không đạt 2NF (các thuộc tính không khóa không phụ thuộc đầy đủ vào khóa).
 - Xét chuẩn 3NF: không đạt 3NF (các thuộc tính về phải không phải siêu khóa, về trái cũng không phải thuộc tính khóa)
- ➔ Vậy Quan hệ đạt chuẩn 1NF

Chuẩn hóa:

- Chuẩn hóa bằng phương pháp Tổng hợp (*Recommender*)
- Chuẩn hóa bằng phương pháp Phân rã (*not Recommender*)

Tại sao ?

- Chuẩn hóa bằng pp Tổng hợp giúp bảo toàn cả Thông tin và bảo toàn PTH, cách làm 1 cách duy nhất đơn giản, tường minh dễ hiểu, tối thiểu đạt 3NF.
- Chuẩn hóa bằng pp Phân rã mặc dù giúp đạt chuẩn BCNF, tuy nhiên cá nhân t thấy cách chuẩn hóa này sẽ khó sử dụng với bạn ít làm (bởi vì có nhiều cách chọn PTH từ một tập PTH) .

Chuẩn hóa bằng phương pháp Tổng hợp.

- Bỏ qua thuật toán tương đối khó tưởng tượng, đi luôn vào cách giải:

B1: Tính phủ tối thiểu Fc

B2: Từ các PTH trong Phủ tối thiểu, tách thành các Q1, Q2, .. Qn tương ứng, mỗi Qi nhận một PTH tương ứng của nó.

B3: Tìm khóa của các Qi đó.

B4: Kiểm tra các Qi, nếu không thấy có khóa của Quan hệ ban đầu (bảng gốc) , thì thêm một Qi nữa, chỉ chứa khóa.

- ➔ Kiểm tra dạng chuẩn của từng Qi trong, ít nhất nó sẽ đạt **3NF**, và khoảng **90%** sẽ đạt **BCNF**.
- ➔ Xong. :)

Vậy điều kiện cần để làm được câu chuẩn hóa:

- Biết tìm Phủ tối thiểu
- Biết tìm Khóa (trình bày cẩn thận tránh mất điểm).
- Hiểu định nghĩa các dạng chuẩn 3NF và BCNF.