



# BÀI GIẢNG

## CƠ SỞ DỮ LIỆU

### 7. Mô hình quan hệ: Các phép toán

Nguyễn Hải Châu

Khoa Công nghệ Thông tin  
Trường Đại học Công nghệ, ĐHQGHN

# Các lược đồ và trạng thái quan hệ

- Lược đồ:  $DONVI(\underline{MaDV}, TenDV, MaNV, NgayBatDau)$  và  $DUAN(\underline{MaDA}, TenDA, DiaDiem, MaDV)$
- Thuộc tính  $MaDV$  của  $DUAN$  là khóa ngoài tham chiếu tới khóa chính  $\underline{MaDV}$  của  $DONVI$
- Quan hệ  $DONVI$ :

<u>MaDV</u>	TenDV	MaNV	NgayBatDau
DV1	Tài chính	NV1	2020-01-09
DV2	Nhân sự	NV5	2017-07-15
DV3	Kinh doanh	NV2	2018-09-10
DV4	Quản trị	NV8	2018-09-10

- Quan hệ  $DUAN$ :

<u>MaDA</u>	TenDA	DiaDiem	MaDV
DA1	Tuyển dụng	Cần Thơ	DV2
DA2	Quảng cáo trực tuyến	Toàn quốc	DV3
DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DA4	Quy trình nghiệp vụ	Hà Nội	<i>null</i>

# Các phép toán cập nhật

# Phép chèn (insert)

- **Phép chèn (insert)** đưa một bộ mới  $t$  vào một quan hệ  $R$ , với tham số là bộ giá trị của các thuộc tính. Phép chèn có thể vi phạm các ràng buộc miền, khóa và toàn vẹn tham chiếu:
- Lược đồ  $DUAN(\underline{MaDA}, TenDA, DiaDiem, MaDV)$ :
  - Phép chèn bộ  $\langle null, 'Phát tở rơi', 'Đà Nẵng', 'DV3' \rangle$  vào  $DUAN$  vi phạm ràng buộc toàn vẹn thực thể: giá trị  $null$  cho khóa chính
  - Phép chèn bộ  $\langle 'DA2', 'Phát tở rơi', 'Đà Nẵng', 'DV3' \rangle$  vào  $DUAN$  vi phạm ràng buộc khóa: giá trị khóa 'DA2' đã tồn tại
  - Phép chèn bộ  $\langle 'DA5', 'Phát tở rơi', 'Đà Nẵng', 'DV7' \rangle$  vào  $DUAN$  vi phạm ràng buộc tham chiếu: không có giá trị khóa 'DV7' trong quan hệ bị tham chiếu  $DONVI$
  - Phép chèn bộ  $\langle 'DA5', 'Phát tở rơi', 'Đà Nẵng', 'DV3' \rangle$  vào  $DUAN$  được chấp nhận vì thoả mãn tất cả các ràng buộc

# Phép chèn (insert)

Nếu chèn bộ < 'DA5', 'Phát tờ rơi', 'Đà Nẵng', 'DV3' > vào *DUAN*, quan hệ này trở thành:

<b>MaDA</b>	<b>TenDA</b>	<b>DiaDiem</b>	<b>MaDV</b>
DA1	Tuyển dụng	Cần Thơ	DV2
DA2	Quảng cáo trực tuyến	Toàn quốc	DV3
DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DA4	Quy trình nghiệp vụ	Hà Nội	<i>null</i>
DA5	Phát tờ rơi	Đà Nẵng	DV3

# Phép xóa (delete)

**Phép xóa (delete)** được sử dụng để xóa một hoặc nhiều bộ giá trị của một quan hệ. Phép xóa có thể vi phạm ràng buộc tham chiếu trong trường hợp bộ bị xóa được tham chiếu từ các bộ khác trong cơ sở dữ liệu

- Xóa bộ <'DV3', 'Kinh doanh', 'NV2', '2018-09-10'> trong *DONVI* vi phạm toàn vẹn tham chiếu, do giá trị khóa chính 'DV3' được tham chiếu từ bộ <'DA2', 'Quảng cáo trực tuyến', 'Toàn quốc', 'DV3' > của *DUAN*, do đó không được chấp nhận
- Xóa bộ <'DV4', 'Quản trị', 'NV8', '2018-09-10' > trong *DONVI* được chấp nhận vì không vi phạm ràng buộc tham chiếu: không có tham chiếu đến giá trị khóa chính 'DV4'

# Phép xóa (delete)

Nếu xóa bộ < 'DV4', 'Quản trị', 'NV8', '2018-09-10' > từ *DONVI*, quan hệ này trở thành:

<u>MaDA</u>	<u>TenDV</u>	<u>MaNV</u>	<u>NgayBatDau</u>
DV1	Tài chính	NV1	2020-01-09
DV2	Nhân sự	NV5	2017-07-15
DV3	Kinh doanh	NV2	2018-09-10

# Phép sửa (update)

- **Phép sửa (update)** được dùng để thay đổi các giá trị của một hoặc nhiều thuộc tính trong một (hoặc nhiều) bộ của một quan hệ
- Việc sửa một thuộc tính không phải là khóa hoặc khóa ngoài thường không gây ra các vi phạm ràng buộc, hệ quản trị cơ sở dữ liệu chỉ kiểm tra giá trị mới thuộc miền
- Sửa giá trị khóa chính tương đương với xóa bộ cũ và chèn bộ mới: Các vi phạm ràng buộc tuân theo các ràng buộc của phép chèn và xóa
- Sửa giá trị khóa ngoài: cần đảm bảo giá trị mới thỏa mãn ràng buộc tham chiếu, tức là giá trị mới hoặc là *null* hoặc phải tham chiếu đến một bộ tồn tại trong quan hệ bị tham chiếu



# Các phép toán đại số quan hệ

# Phép chọn (select)

- **Phép chọn (select)** được sử dụng để chọn một tập hợp các bộ thỏa mãn điều kiện chọn từ một quan hệ
- Có thể xem phép chọn như một bộ lọc, nó chỉ giữ lại các bộ thỏa mãn điều kiện đặt ra
- Phép chọn được ký hiệu là  $\sigma_{\langle cond \rangle}(R)$ , trong đó  $\langle cond \rangle$  là điều kiện chọn
- Danh sách dự án có địa điểm ở Hà Nội:  $\sigma_{\langle DiaDiem='Hà Nội' \rangle}(DUAN)$ , kết quả:

MaDA	TenDA	DiaDiem	MaDV
DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DA4	Qui trình nghiệp vụ	Hà Nội	<i>null</i>

# Phép chiếu (project)

- Nếu ta xem một quan hệ như một bảng thì **phép chiếu (project)** là phép toán chọn một số cột của bảng
- Phép chiếu được ký hiệu là:  $\pi_{\langle attr \rangle}(R)$  trong đó  $\langle attr \rangle$  là tập con các thuộc tính của  $R$ .
- Liệt kê tên và địa điểm các dự án:  $\pi_{\langle TenDA, DiaDiem \rangle}(DUAN)$ , kết quả:

TenDA	DiaDiem
Tuyển dụng	Cần Thơ
Quảng cáo trực tuyến	Toàn quốc
Xây dựng phần mềm tài chính	Hà Nội
Quy trình nghiệp vụ	Hà Nội

# Phép đổi tên (rename)

- Phép đổi tên được ký hiệu là  $\rho_{S(B_1, B_2, \dots, B_n)}(R)$ ,  $\rho_S(R)$  hoặc  $\rho(B_1, B_2, \dots, B_n)(R)$ , trong đó:
  - $\rho_{S(B_1, B_2, \dots, B_n)}(R)$  đổi tên quan hệ  $R(A_1, A_2, \dots, A_n)$  thành  $S(B_1, B_2, \dots, B_n)$  với tên các thuộc tính mới là  $B_1, B_2, \dots, B_n$ ,
  - $\rho_S(R)$  chỉ đổi tên quan hệ  $R$  thành  $S$  và
  - $\rho(B_1, B_2, \dots, B_n)(R)$  chỉ đổi tên các thuộc tính của  $R$  từ  $A_1, A_2, \dots, A_n$  thành  $B_1, B_2, \dots, B_n$

# Phép đổi tên (rename)

- Để tìm mã và tên các dự án ở Hà Nội trong quan hệ *DUAN*, chúng ta có thể làm một trong hai cách:
  - $Tenduan \leftarrow \pi_{\langle MaDA, TenDA \rangle}(\sigma_{\langle DiaDiem='Hà\ Nội' \rangle}(DUAN))$ , hoặc
  - $HN \leftarrow \sigma_{\langle DiaDiem='Hà\ Nội' \rangle}(DUAN)$   
 $Tenduan \leftarrow \pi_{\langle MaDA, TenDA \rangle}(HN)$
- Hai cách trên đều cho kết quả:

MaDA	TenDA
DA3	Xây dựng phần mềm tài chính
DA4	Quy trình nghiệp vụ

- Đổi tên quan hệ kết quả:  $\rho_{Project}(PrCode, PrName)(Tenduan)$

# Các phép toán lý thuyết tập hợp

- Các phép toán lý thuyết tập hợp chỉ áp dụng được cho các quan hệ  $R(A_1, A_2, \dots, A_n)$  và  $S(B_1, B_2, \dots, B_n)$  thỏa mãn điều kiện *tương thích đồng nhất*:  $Dom(A_i) = Dom(B_i) \forall i = \overline{1, n}$
- **Phép hợp**: Hợp của hai quan hệ  $R$  và  $S$ , được ký hiệu là  $R \cup S$ , cho kết quả là một quan hệ chứa tất cả các bộ có trong  $R$  hoặc ở trong  $S$  hoặc ở trong cả hai. Các bộ trùng lặp bị loại bỏ
- **Phép giao**: Giao của hai quan hệ  $R$  và  $S$ , được ký hiệu là  $R \cap S$ , cho kết quả là một quan hệ chứa tất cả các bộ có trong cả hai quan hệ  $R$  và  $S$
- **Phép trừ**: Phép trừ  $R$  và  $S$ , được ký hiệu là  $R - S$ , cho kết quả là một quan hệ chứa tất cả các bộ có trong  $R$  nhưng không có trong  $S$

# Tích Đề-các ( $\times$ )

- Tích Đề-các của hai quan hệ  $R(A_1, A_2, \dots, A_n)$  và  $S(B_1, B_2, \dots, B_m)$ , ký hiệu là  $R(A_1, A_2, \dots, A_n) \times S(B_1, B_2, \dots, B_m)$ , được dùng để kết hợp mọi bộ của  $R$  với mọi bộ của  $S$
- Kết quả của tích Đề-các là quan hệ  $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$  có  $m + n$  thuộc tính
- Giả sử  $R$  có  $n_R$  bộ,  $S$  có  $n_S$  bộ, khi đó  $Q$  có  $n_R \times n_S$  bộ

# Tích Đề-các (×)

Tích Đề-các của *DONVI* và *DUAN* có 16 bộ như sau:

MaDV	TenDV	MaNV	NgayBatDau	MaDA	TenDA	DiaDiem	DUAN.MaDV
DV1	Tài chính	NV1	2020-01-09	DA1	Tuyển dụng	Cần Thơ	DV2
DV2	Nhân sự	NV5	2017-07-15	DA1	Tuyển dụng	Cần Thơ	DV2
DV3	Kinh doanh	NV2	2018-09-10	DA1	Tuyển dụng	Cần Thơ	DV2
DV4	Quản trị	NV8	2018-09-10	DA1	Tuyển dụng	Cần Thơ	DV2
DV1	Tài chính	NV1	2020-01-09	DA2	Quảng cáo trực tuyển	Toàn quốc	DV3
DV2	Nhân sự	NV5	2017-07-15	DA2	Quảng cáo trực tuyển	Toàn quốc	DV3
DV3	Kinh doanh	NV2	2018-09-10	DA2	Quảng cáo trực tuyển	Toàn quốc	DV3
DV4	Quản trị	NV8	2018-09-10	DA2	Quảng cáo trực tuyển	Toàn quốc	DV3
DV1	Tài chính	NV1	2020-01-09	DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DV2	Nhân sự	NV5	2017-07-15	DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DV3	Kinh doanh	NV2	2018-09-10	DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DV4	Quản trị	NV8	2018-09-10	DA3	Xây dựng phần mềm tài chính	Hà Nội	DV1
DV1	Tài chính	NV1	2020-01-09	DA4	Quy trình nghiệp vụ	Hà Nội	<i>null</i>
DV2	Nhân sự	NV5	2017-07-15	DA4	Quy trình nghiệp vụ	Hà Nội	<i>null</i>
DV3	Kinh doanh	NV2	2018-09-10	DA4	Quy trình nghiệp vụ	Hà Nội	<i>null</i>
DV4	Quản trị	NV8	2018-09-10	DA4	Quy trình nghiệp vụ	Hà Nội	<i>null</i>



# Phép nối (join)

- **Phép nối (join)** Phép nối được ký hiệu là  $\bowtie$  và được dùng để kết hợp các bộ có liên hệ với nhau từ hai quan hệ thành một bộ. Phép toán này rất quan trọng đối với cơ sở dữ liệu quan hệ có nhiều bảng bởi vì nó cho phép ta xử lý các mối liên kết giữa các quan hệ
- Dạng tổng quát của phép nối trên hai quan hệ  $R(A_1, A_2, \dots, A_n)$  và  $S(B_1, B_2, \dots, B_m)$  là  $R \underset{\langle cond \rangle}{\bowtie} S$  trong đó  $\langle cond \rangle$  là điều kiện nối.  
Kết quả của phép nối là một quan hệ  $Q(A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m)$  có  $n + m$  thuộc tính.
- Mỗi bộ của  $Q$  là một kết nối giữa một bộ của  $R$  và một bộ của  $S$  thoả mãn điều kiện nối
- $R \underset{\langle cond \rangle}{\bowtie} S = \sigma_{\langle cond \rangle}(R \times S)$
- Hầu hết các phép nối đều có  $\langle cond \rangle$  là so sánh bằng; các phép nối đó được gọi là phép nối bằng (equi-join)

# Phép nối (join)

Phép nối  $Q \leftarrow DONVI \bowtie_{<DONVI.MaDV=DUAN.MaDV>} DUAN$  cho kết quả:

MaDV	TenDV	MaNV	NgayBatDau	MaDA	TenDA	DiaDiem	DUAN.MaDV
DV1	Tài chính	NV1	2020-01-09	DA3	Xây dựng phần mềm	Hà Nội	DV1
DV2	Nhân sự	NV5	2017-07-15	DA1	Tài chính Tuyển dụng	Cần Thơ	DV2
DV3	Kinh doanh	NV2	2018-09-10	DA2	Quảng cáo trực tuyến	Toàn quốc	DV3

trong đó ký hiệu  $DONVI.MaDV$  nhằm chỉ rõ thuộc tính  $MaDV$  là của quan hệ  $DONVI$ .

- Kết quả của phép nối là một quan hệ  $Q$  có 8 thuộc tính, trong đó 4 thuộc tính đầu được lấy từ các thuộc tính của  $DONVI$ , 4 thuộc tính còn lại lấy từ  $DUAN$
- Ở mỗi bộ của  $Q$ , chúng ta thấy điều kiện nối được thỏa mãn:  $MaDV = DUAN.MaDV$ , với  $MaDV$  được hiểu là  $DONVI.MaDV$

# Phép nối tự nhiên (\*)

- Trong ví dụ phép nối, có hai thuộc tính của hai quan hệ *DONVI* và *DUAN* trùng tên, đó là *MaDV*, do đó chúng ta chỉ rõ *DUAN.MaDV* trong cột cuối để tránh nhầm lẫn
- Quan hệ *Q* là kết quả của phép nối bằng, nên cả hai cột *MaDV* có mặt trong *Q* là dư thừa → người ta đề nghị phép nối tự nhiên, ký hiệu là  $*$ , nhằm loại bỏ các thuộc tính thừa
- Phép nối tự nhiên đòi hỏi các quan hệ tham gia phải có các thuộc tính trùng tên

Phép nối tự nhiên  $Q_* \leftarrow DONVI \underset{\langle MaDV \rangle}{*} DUAN$  cho kết quả:

MaDV	TenDV	MaNV	NgayBatDau	MaDA	TenDA	DiaDiem
DV1	Tài chính	NV1	2020-01-09	DA3	Xây dựng phần mềm tài chính	Hà Nội
DV2	Nhân sự	NV5	2017-07-15	DA1	Tuyển dụng	Cần Thơ
DV3	Kinh doanh	NV2	2018-09-10	DA2	Quảng cáo trực tuyến	Toàn quốc

*MaDV* dư thừa đã bị loại bỏ

# Tập hợp đầy đủ các phép toán quan hệ

- Tập các phép toán  $P = \{\sigma, \pi, \cup, -, \times\}$  là một tập đầy đủ: Mọi phép toán đại số quan hệ khác đều có thể được biểu diễn qua các phép toán trong  $P$
- Ví dụ:
  - Phép nối:  $R \bowtie_{\langle cond \rangle} S = \sigma_{\langle cond \rangle}(R \times S)$
  - Phép giao:  $R \cap S = (R \cup S) - ((R - S) \cup (S - R))$

# Các phép toán nối mở rộng

- Có ba phép nối mở rộng là **nối ngoài trái (left outer join)**, **nối ngoài phải (right outer join)** và **nối ngoài đầy đủ (full outer join)**, ký hiệu tương ứng là:  $\bowtie$ ,  $\ltimes$ ,  $\rhd$
- **Phép nối ngoài trái** của  $R(A_1, A_2, \dots, A_n)$  và  $S(B_1, B_2, \dots, B_m)$  được định nghĩa như sau:  
$$R \underset{\langle cond \rangle}{\bowtie} S = (R \underset{\langle cond \rangle}{\times} S) \cup T$$
 trong đó  $T$  là một quan hệ gồm các bộ  $\langle a_1, a_2, \dots, a_n, null, \dots, null \rangle$  ( $m$  giá trị  $null$ ), trong đó  $\langle a_1, a_2, \dots, a_n \rangle \in R$  không thỏa mãn điều kiện  $\langle cond \rangle$
- Tương tự như vậy đối với các **phép nối ngoài phải** và các **phép nối ngoài đầy đủ**

# Phép nối ngoài trái $\bowtie$

- Kết quả của  $DONVI \bowtie_{DONVI.MaDV=DUAN.MaDV} DUAN$ :

MaDV	TenDV	MaNV	NgayBatDau	MaDA	TenDA	DiaDiem
DV1	Tài chính	NV1	2020-01-09	DA3	Xây dựng phần	Hà Nội
DV2	Nhân sự	NV5	2017-07-15	DA1	mềm tài chính	Cần Thơ
DV3	Kinh doanh	NV2	2018-09-10	DA2	Tuyển dụng	Toàn quốc
DV4	Quản trị	NV8	2018-09-10	null	Quảng cáo trực tuyển null	null

- Các đơn vị không giám sát dự án nào:  $MaDA$  nhận giá trị  $null$  trong kết quả phép nối ngoài trái

- Kết quả của  $DONVI \bowtie_{DONVI.MaDV=DUAN.MaDV} DUAN$ :

MaDV	TenDV	MaNV	NgayBatDau	MaDA	TenDA	DiaDiem
<i>null</i>	<i>null</i>	<i>null</i>	<i>null</i>	DA4	Qui trình nghiệp vụ	Hà Nội
DV1	Tài chính	NV1	2020-01-09	DA3	Xây dựng phần mềm tài chính	Hà Nội
DV2	Nhân sự	NV5	2017-07-15	DA1	Tuyển dụng	Cần Thơ
DV3	Kinh doanh	NV2	2018-09-10	DA2	Quảng cáo trực tuyến	Toàn quốc

- Các dự án chưa có đơn vị nào quản lý:  $MaDV$  nhận giá trị *null* trong kết quả phép nối ngoài phải

# Phép nối ngoài đầy đủ $\bowtie$

- Kết quả của  $DONVI \bowtie DUAN$ :  
 $DONVI.MaDV = DUAN.MaDV$

MaDV	TenDV	MaNV	NgayBatDau	MaDA	TenDA	DiaDiem
<i>null</i>	<i>null</i>	<i>null</i>	<i>null</i>	DA4	Qui trình nghiệp vụ	Hà Nội
DV1	Tài chính	NV1	2020-01-09	DA3	Xây dựng phần mềm tài chính	Hà Nội
DV2	Nhân sự	NV5	2017-07-15	DA1	Tuyển dụng	Cần Thơ
DV3	Kinh doanh	NV2	2018-09-10	DA2	Quảng cáo trực tuyến	Toàn quốc
DV4	Quản trị	NV8	2018-09-10	<i>null</i>	<i>null</i>	<i>null</i>

- Các đơn vị không giám sát dự án nào:  $MaDA$  nhận giá trị *null* trong kết quả phép nối ngoài trái
- Các dự án chưa có đơn vị nào quản lý:  $MaDV$  nhận giá trị *null* trong kết quả phép nối ngoài phải



# Các hàm nhóm và phép nhóm

- Phép nhóm:  $\langle attr \rangle \mathcal{J}_{\langle f \rangle}(R)$  (ký hiệu  $\mathcal{J}$  trong tiếng Anh là "calligraphic J")
- $\mathcal{J}$  là ký hiệu phép nhóm,  $\langle attr \rangle$  là danh sách các thuộc tính nhóm,  $\langle f \rangle$  là danh sách các cặp ( $\langle hàm \rangle \langle thuộc tính \rangle$ ) với  $\langle hàm \rangle$  là một trong các hàm SUM, AVERAGE, MIN, MAX, COUNT và  $\langle thuộc tính \rangle$  là một trong các thuộc tính của  $R$
- Kết quả của  $DiaDiem \mathcal{J}_{COUNT()}(DUAN)$ :

<b>DiaDiem</b>	<b>Tongso</b>
Cần Thơ	1
Hà Nội	2
Toàn quốc	1

Ý nghĩa: Đếm tổng số dự án ở mỗi địa điểm