

# Cơ Sở Dữ Liệu Phân Tán

## Bài tập về nhà - IS211.M21

Nguyễn Hồ Duy Tri, Nguyễn Thị Kim Yến  
Sinh viên: Phạm Đức Thế - 19522253

Thứ 6, ngày 25 tháng 03 năm 2022

### Bài Tập Phân Mảnh Ngang

#### Bài Tập 1

Một công ty sản xuất có nhiều công nhân. Mỗi công nhân làm việc ở một phân xưởng. Công ty có ba phân xưởng đặt ở Quận 9 (MaPX='Q9'), Thủ Đức (MaPX='TĐ') và Bình Dương (MaPX='BD'). Một công nhân thì thuộc về một tổ. Công ty có hai tổ: tổ giao và nhận (MaTo=1), tổ sản xuất và đóng gói (MaTo=2). Thông tin công nhân được thể hiện qua lược đồ quan hệ sau:

**CONGNHAN**(**MaCN**, **TenCN**, **NgaySinh**, **GioiTinh**, **MaPX**, **MaTo**, **Luong**)

Tên từ: Mỗi công nhân có: mã công nhân, tên công nhân, ngày sinh, giới tính, mã phân xưởng, mã tổ và lương của công nhân đó.

**MATHANG**(**MaMH**, **TenMH**)

Tên từ: Các mặt hàng của công ty: mã mặt hàng, tên mặt hàng

**PHANCONG**(**MaCN**, **MaMH**, **thang**, **nam**, **SoLuong**)

Tên từ: Thông tin về phân công công việc của các công nhân: mã công nhân (MaCN), mã mặt hàng (MaMH), thời gian (tháng, năm) và số lượng (SoLuong).

Cho quan hệ **CONGNHAN** với tập dữ liệu demo như sau:

MaCN	TenCN	NgaySinh	GioiTinh	MaPX	MaTo	Luong
NV1	Nguyễn A	1/1/1975	Nam	Q9	1	5.000.000
NV2	Trần B	1/1/1976	Nữ	TĐ	2	12.000.000
NV3	Bùi C	1/1/1977	Nam	TĐ	2	20.000.000
NV4	Phạm D	1/1/1978	Nữ	BD	1	15.000.000
NV5	Lê E	2/1/1983	Nam	Q9	1	10.000.000
NV6	Bùi T	12/12/1980	Nam	BD	2	7.000.000
NV7	Phạm M	26/09/2003	Nữ	BD	1	3.000.000

Giả sử có ba ứng dụng truy suất đến **CONGNHAN**:

Q1: SELECT \* FROM **CONGNHAN** WHERE **MaPX** = value

Q2: SELECT \* FROM **CONGNHAN** WHERE **MaTo** = value

Q3: SELECT \* FROM **CONGNHAN** WHERE **Luong** < 3.000.000

1. Dùng giải thuật COM\_MIN, tính Pr' thỏa tối tiểu và đầy đủ? (2 điểm)

**Lời giải.**

- Tập vị từ đơn giản dùng để phân hoạch **CONGNHAN**:
  - p1: MaPX = 'Q9'
  - p2: MaPX = 'TĐ'
  - p3: MaPX = 'BD'
  - p4: MaTo = 1
  - p5: MaTo = 2
  - p6: Luong < 3.000.000
  - p7: Luong ≥ 3.000.000

- Khởi tạo  $P_r = \{p1, p2, p3, p4, p5, p6, p7\}$ ,  $P_{r'} = \emptyset$
  - Áp dụng thuật toán COM\_MIN:
    - $i = 1$  làm giá trị khởi đầu, vị từ  $p1$  thỏa quy tắc 1,  $P_{r'} = \{p1\}$ .
    - $i = 2$ : Vị từ  $p2$  thỏa quy tắc 1,  $P_{r'} = \{p1, p2\}$ .
    - $i = 3$ : Ta có vị từ  $p3$  không phân hoạch  $f_2$  (là mảnh giao tối thiểu tạo ra ứng với  $p2$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p2\}$ .
    - $i = 4$ : Vị từ  $p4$  thỏa quy tắc 1,  $P_{r'} = \{p1, p2, p4\}$ .
    - $i = 5$ : Ta có vị từ  $p5$  không phân hoạch  $f_4$  (là mảnh giao tối thiểu tạo ra ứng với  $p4$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p2, p4\}$ .
    - $i = 6$ : Ta có vị từ  $p6$  không phân hoạch  $f_5$  (là mảnh giao tối thiểu tạo ra ứng với  $p5$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p2, p4\}$ .
    - $i = 7$ : Ta có vị từ  $p7$  không phân hoạch  $f_6$  (là mảnh giao tối thiểu tạo ra ứng với  $p6$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p2, p4\}$ .
- Kết luận:  $P_{r'} = \{p1, p2, p4\}$  là đầy đủ và tối thiểu.

2. Sử dụng thuật toán PHORIZONTAL, thiết kế phân mảnh ngang chính cho quan hệ **CONGNHAN**? (2 điểm)

**Lời giải.**

- Các vị từ giao tối thiểu là:
  - m1:  $p1 \wedge p4$ ;
  - m2:  $p1 \wedge \neg p4 = p1 \wedge p5$ ;
  - m3:  $p2 \wedge p4$ ;
  - m4:  $p2 \wedge \neg p4 = p2 \wedge p5$ ;
  - m5:  $(\neg p1 \wedge \neg p2) \wedge p4 = p3 \wedge p4$ ;
  - m6:  $(\neg p1 \wedge \neg p2) \wedge \neg p4 = p3 \wedge p5$ ;
- Vì vậy, ta có tập vị từ giao tối thiểu là:  $M = \{m1, m2, m3, m4, m5, m6\}$
- Phân mảnh ngang chính quan hệ CONGNHAN theo M là:
  - CONGNHAN1 =  $\sigma_{m1}$ (CONGNHAN)
  - CONGNHAN2 =  $\sigma_{m2}$ (CONGNHAN)
  - CONGNHAN3 =  $\sigma_{m3}$ (CONGNHAN)
  - CONGNHAN4 =  $\sigma_{m4}$ (CONGNHAN)
  - CONGNHAN5 =  $\sigma_{m5}$ (CONGNHAN)
  - CONGNHAN6 =  $\sigma_{m6}$ (CONGNHAN)
- Đối với dữ liệu demo mảnh CONGNHAN2 và CONGNHAN3 là rỗng, do đó ta có các mảnh:

CONGNHAN1

MaCN	TenCN	NgaySinh	GioiTinh	MaPX	MaTo	Luong
NV1	Nguyễn A	1/1/1975	Nam	Q9	1	5.000.000
NV5	Lê E	2/1/1983	Nam	Q9	1	10.000.000

CONGNHAN4

MaCN	TenCN	NgaySinh	GioiTinh	MaPX	MaTo	Luong
NV2	Trần B	1/1/1976	Nữ	TĐ	2	12.000.000
NV3	Bùi C	1/1/1977	Nam	TĐ	2	20.000.000

CONGNHAN5

MaCN	TenCN	NgaySinh	GioiTinh	MaPX	MaTo	Luong
NV4	Phạm D	1/1/1978	Nữ	BD	1	15.000.000
NV7	Phạm M	26/09/2003	Nữ	BD	1	3.000.000

CONGNHAN6

MaCN	TenCN	NgaySinh	GioiTinh	MaPX	MaTo	Luong
NV6	Bùi T	12/12/1980	Nam	BD	2	7.000.000

3. Thiết kế phân mảnh ngang dẫn xuất cho bảng **PHANCONG**.

**Lời giải.**

- Phân mảnh ngang dẫn xuất cho bảng **PHANCONG**:  
 $CONGNHAN1 = \sigma_{m1}(CONGNHAN)$   
 $CONGNHAN2 = \sigma_{m2}(CONGNHAN)$   
 $CONGNHAN3 = \sigma_{m3}(CONGNHAN)$   
 $CONGNHAN4 = \sigma_{m4}(CONGNHAN)$   
 $CONGNHAN5 = \sigma_{m5}(CONGNHAN)$   
 $CONGNHAN6 = \sigma_{m6}(CONGNHAN)$

Các mảnh của quan hệ PHANCONG tương ứng là:

$$\begin{aligned} PHANCONG1 &= PHANCONG \bowtie_{MaCN} CONGNHAN1 \\ PHANCONG2 &= PHANCONG \bowtie_{MaCN} CONGNHAN2 \\ PHANCONG3 &= PHANCONG \bowtie_{MaCN} CONGNHAN3 \\ PHANCONG4 &= PHANCONG \bowtie_{MaCN} CONGNHAN4 \\ PHANCONG5 &= PHANCONG \bowtie_{MaCN} CONGNHAN5 \\ PHANCONG6 &= PHANCONG \bowtie_{MaCN} CONGNHAN6 \end{aligned}$$

Các vị từ định tính của phân mảnh là:

$$\begin{aligned} q1: PHANCONG.MaCN &= CONGNHAN.MaCN \wedge MaPX = 'Q9' \wedge MaTo = 1 \\ q2: PHANCONG.MaCN &= CONGNHAN.MaCN \wedge MaPX = 'Q9' \wedge MaTo = 2 \\ q3: PHANCONG.MaCN &= CONGNHAN.MaCN \wedge MaPX = 'TD' \wedge MaTo = 1 \\ q4: PHANCONG.MaCN &= CONGNHAN.MaCN \wedge MaPX = 'TD' \wedge MaTo = 2 \\ q5: PHANCONG.MaCN &= CONGNHAN.MaCN \wedge MaPX = 'BD' \wedge MaTo = 1 \\ q6: PHANCONG.MaCN &= CONGNHAN.MaCN \wedge MaPX = 'BD' \wedge MaTo = 2 \end{aligned}$$

4. Kết quả phân mảnh của lược đồ quan hệ **CONGNHAN** ở câu 2 có đáp ứng được qui tắc đúng đắn của phân mảnh hay không ? Giải thích.

**Lời giải.**

- Kết quả phân mảnh của lược đồ quan hệ **CONGNHAN** ở câu 2 đáp ứng được qui tắc đúng đắn của phân mảnh. Vì nó đảm bảo 3 điều kiện:
  - Điều kiện đầy đủ: Mỗi mục dữ liệu trong quan hệ **CONGNHAN** phải có trong một hoặc nhiều mảnh  $CONGNHAN_i$ .  
 $\forall u \in CONGNHAN, \exists i \in [1, 6] : u \in CONGNHAN_i$
  - Điều kiện tái tạo: Giữa các mảnh có thể áp dụng phép toán hợp để tái tạo lại mảnh ban đầu:  
Ta có:  $CONGNHAN1 \cup CONGNHAN2 \cup CONGNHAN3 \cup CONGNHAN4 \cup CONGNHAN5 \cup CONGNHAN6 = CONGNHAN$   
 $\implies$  Điều kiện tái tạo của phân mảnh này là đúng đắn.
  - Điều kiện tách biệt: Nếu mục dữ liệu  $d_i$  có trong  $R_i$  thì nó không có trong bất kỳ mảnh  $R_k$  khác ( $i \neq k$ ).  
 $\forall i \neq k \text{ và } i, k \in [1, n] : R_i \cap R_k = \emptyset$ .

## Bài Tập 2

Tổng công ty Cảng hàng không Việt Nam hiện muốn quản lý thông tin các sân bay tại Việt Nam. Tùy theo mục đích khai thác mà các sân bay được quy hoạch thành hai nhóm là sân bay Nội địa (LoaiSB = 'Nội địa') và sân bay Quốc tế (LoaiSB = 'Quốc tế'). Ngoài ra, các sân bay còn được phân chia theo loại đường băng để kiểm soát việc máy bay nào có thể cất hay hạ cánh, bao gồm đường băng được làm bê tông (LoaiDB = 'Bê tông') và đường băng được làm bằng nhựa đường (LoaiDB = 'Nhựa đường'). Thông tin của các sân bay được thể hiện qua lược đồ quan hệ sau:

**SANBAY(MaICAO, TenSB, NamTL, LoaiSB, BayDem, LoaiDB)**

Tần từ: Mỗi sân bay bao gồm: Mã ICAO, Tên sân bay, năm thành lập sân bay, loại sân bay, cho phép bay đêm, loại đường băng.

Cho quan hệ **SANBAY** với tập dữ liệu demo như sau:

**SANBAY(MaICAO, TenSB, NamTL, LoaiSB, BayDem, LoaiDB)**

MaICAO	TenSB	NamTL	LoaiSB	BayDem	LoaiDB
VVCM	Cà Mau	1962	Nội địa	Không	Nhựa đường
VVCS	Côn Đảo	1945	Nội địa	Có	Nhựa đường
VVCI	Cát Bi	1985	Quốc tế	Có	Bê tông
VVRG	Rạch Giá	1970	Nội địa	Không	Bê tông
VVDL	Liên Khương	1961	Quốc tế	Có	Nhựa đường
VVCA	Chu Lai	1965	Nội địa	Không	Bê tông
VVPB	Phú Bài	1948	Quốc tế	Có	Nhựa đường
VVTX	Thọ Xuân	1965	Nội địa	Có	Bê tông

Giả sử có ba ứng dụng truy suất đến **SANBAY**:

Q1: SELECT COUNT(\*) FROM **SANBAY** WHERE NamTL < 1965

Q2: SELECT MaICAO, TenSB FROM **SANBAY** WHERE LoaiSB = value

Q3: SELECT \* FROM **SANBAY** WHERE LoaiDB = value

1. Dùng giải thuật COM\_MIN, tính Pr' thỏa tối tiểu và đầy đủ? (2 điểm)

**Lời giải.**

- Tập vị từ đơn giản dùng để phân hoạch SANBAY:  
p1: NamTL < 1965  
p2: NamTL ≥ 1965  
p3: LoaiSB = 'Nội địa'  
p4: LoaiSB = 'Quốc Tế'  
p5: LoaiDB = 'Nhựa đường'  
p6: LoaiDB = 'Bê tông'
- Khởi tạo  $P_r = \{p1, p2, p3, p4, p5, p6\}$ ,  $P_{r'} = \emptyset$
- Áp dụng thuật toán COM\_MIN:
  - i = 1 làm giá trị khởi đầu, vị từ p1 thỏa quy tắc 1,  $P_{r'} = \{p1\}$ .
  - i = 2: Ta có vị từ p2 không phân hoạch  $f_1$  (là mảnh giao tối thiểu tạo ra ứng với p1) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1\}$ .
  - i = 3: Vị từ p3 thỏa quy tắc 1.  $P_{r'} = \{p1, p3\}$ .
  - i = 4: Ta có vị từ p4 không phân hoạch  $f_3$  (là mảnh giao tối thiểu tạo ra ứng với p3) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
  - i = 5: Ta có vị từ p5 không phân hoạch  $f_4$  (là mảnh giao tối thiểu tạo ra ứng với p4) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
  - i = 6: ta có vị từ p6 không phân hoạch  $f_5$  (là mảnh giao tối thiểu tạo ra ứng với p5) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .Kết luận:  $P_{r'} = \{p1, p3\}$  là đầy đủ và tối thiểu.

2. Sử dụng thuật toán PHORIZONTAL, thiết kế phân mảnh ngang chính cho quan hệ **SANBAY**? Lưu ý, sinh viên không cần liệt kê dữ liệu trong các mảnh. (2 điểm)

**Lời giải.**

- Các vị từ giao tối thiểu là:
  - m1:  $p1 \wedge p3$ ;
  - m2:  $p1 \wedge \neg p3 = p1 \wedge p4$ ;
  - m3:  $\neg p1 \wedge p3 = p2 \wedge p3$ ;
  - m4:  $\neg p1 \wedge \neg p3 = p2 \wedge p4$ ;
- Vì vậy, ta có tập vị từ giao tối thiểu là:  $M = \{m1, m2, m3, m4\}$
- Phân mảnh ngang chính quan hệ SANBAY theo M là:
  - $SANBAY1 = \sigma_{m1}(SANBAY)$
  - $SANBAY2 = \sigma_{m2}(SANBAY)$
  - $SANBAY3 = \sigma_{m3}(SANBAY)$
  - $SANBAY4 = \sigma_{m4}(SANBAY)$
- Đối với dữ liệu demo, ta có các mảnh:
  - SANBAY1

MaICAO	TenSB	NamTL	LoaiSB	BayDem	LoaiDB
VVCM	Cà Mau	1962	Nội địa	Không	Nhựa đường
VVCS	Côn Đảo	1945	Nội địa	Có	Nhựa đường

SANBAY2

MaICAO	TenSB	NamTL	LoaiSB	BayDem	LoaiDB
VVDL	Liên Khương	1961	Quốc tế	Có	Nhựa đường
VVPB	Phú Bài	1948	Quốc tế	Có	Nhựa đường

SANBAY3

MaICAO	TenSB	NamTL	LoaiSB	BayDem	LoaiDB
VVRG	Rạch Giá	1970	Nội địa	Không	Bê tông
VVCA	Chu Lai	1965	Nội địa	Không	Bê tông
VVTX	Thọ Xuân	1965	Nội địa	Có	Bê tông

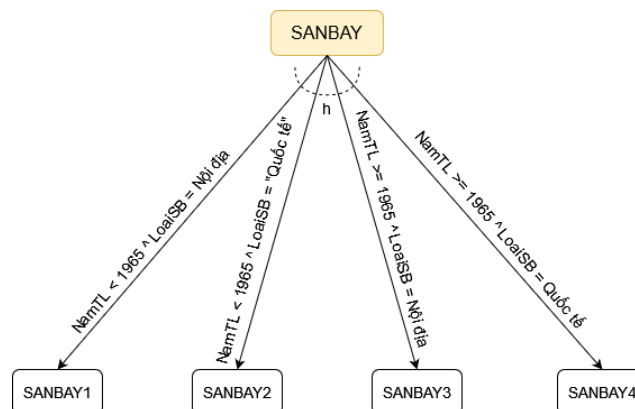
SANBAY4

MaICAO	TenSB	NamTL	LoaiSB	BayDem	LoaiDB
VVCI	Cát Bi	1985	Quốc tế	Có	Bê tông

3. Hãy vẽ cây phân mảnh của lược đồ quan hệ **SANBAY** đã làm. Chứng minh điều kiện tái tạo của phân mảnh này là đúng đắn. (1 điểm)

**Lời giải.**

- Vẽ cây phân mảnh



Hình 1: Cây phân mảnh

- Chứng minh điều kiện tái tạo.
 

Giữa các mảnh có thể áp dụng phép toán hợp để tái tạo lại mảnh ban đầu:  
 Ta có:  $SANBAY1 \cup SANBAY2 \cup SANBAY3 \cup SANBAY4 = SANBAY$   
 $\Rightarrow$  Điều kiện tái tạo của phân mảnh này là đúng đắn.

### Bài Tập 3

Tổng công ty Cảng hàng không Việt Nam hiện muốn quản lý thông tin các sân bay tại Việt Nam. Tùy theo mục đích khai thác mà các sân bay được quy hoạch thành hai nhóm là sân bay có hỗ trợ các chuyến bay vào ban đêm (BayDem = ‘Có’) và sân bay không hỗ trợ các chuyến bay vào ban đêm (BayDem = ‘Không’). Ngoài ra, các sân bay còn được phân chia theo loại đường băng để kiểm soát việc máy bay nào có thể cất hay hạ cánh, bao gồm đường băng được làm bê tông (LoaiDB = ‘Bê tông’) và đường băng được làm bằng nhựa đường (LoaiDB = ‘Nhựa đường’). Thông tin của các sân bay được thể hiện qua lược đồ quan hệ sau:

**SANBAY(MaICAO, TenSB, ChieuDaiDB, LoaiSB, BayDem, LoaiDB)**

Tần từ: Mỗi sân bay bao gồm: Mã ICAO, Tên sân bay, chiều dài đường băng, loại sân bay, cho phép bay đêm, loại đường băng. Cho quan hệ **SANBAY** với tập dữ liệu demo như sau:

**SANBAY(MaICAO, TenSB, ChieuDaiDB, LoaiSB, BayDem, LoaiDB)**

MaICAO	TenSB	ChieuDaiDB	LoaiSB	BayDem	LoaiDB
VVCM	Cà Mau	1500	Nội địa	Không	Nhựa đường
VVCS	Côn Đảo	1830	Nội địa	Có	Nhựa đường
VVCI	Cát Bi	3050	Quốc tế	Có	Bê tông
VVRG	Rạch Giá	3000	Nội địa	Không	Bê tông
VVDL	Liên Khương	2950	Quốc tế	Có	Nhựa đường
VVCA	Chu Lai	3050	Nội địa	Không	Bê tông
VVPB	Phú Bài	2700	Quốc tế	Có	Nhựa đường
VVTX	Thọ Xuân	3200	Nội địa	Có	Bê tông

Giả sử có ba ứng dụng truy suất đến **SANBAY**:

Q1: SELECT COUNT(\*) FROM **SANBAY** WHERE ChieuDaiDB  $\geq$  3000

Q2: SELECT MaICAO, TenSB FROM **SANBAY** WHERE BayDem = value

Q3: SELECT \* FROM **SANBAY** WHERE LoaiDB = value

1. Dùng giải thuật COM\_MIN, tính Pr' thỏa tối tiểu và đầy đủ? (2 điểm)

**Lời giải.**

- Tập vị từ đơn giản dùng để phân hoạch SANBAY:  
p1: ChieuDaiDB  $\geq$  3000  
p2: ChieuDaiDB  $<$  3000  
p3: BayDem = ‘Không’  
p4: BayDem = ‘Có’  
p5: LoaiDB = ‘Nhựa đường’  
p6: LoaiDB = ‘Bê tông’
  - Khởi tạo  $P_r = \{p1, p2, p3, p4, p5, p6\}$ ,  $P_{r'} = \emptyset$
  - Áp dụng giải thuật COM\_MIN:
    - i = 1 làm giá trị khởi đầu, vị từ p1 thỏa quy tắc 1,  $P_{r'} = \{p1\}$ .
    - i = 2: Ta có vị từ p2 không phân hoạch  $f_1$  (là mảnh giao tối thiểu tạo ra ứng với p1) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1\}$ .
    - i = 3: Vị từ p3 thỏa quy tắc 1.  $P_{r'} = \{p1, p3\}$ .
    - i = 4: Ta có vị từ p4 không phân hoạch  $f_3$  (là mảnh giao tối thiểu tạo ra ứng với p3) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
    - i = 5: Ta có vị từ p5 không phân hoạch  $f_4$  (là mảnh giao tối thiểu tạo ra ứng với p4) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
    - i = 6: Ta có vị từ p6 không phân hoạch  $f_5$  (là mảnh giao tối thiểu tạo ra ứng với p5) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
- Kết luận:  $P_{r'} = \{p1, p3\}$  là đầy đủ và tối thiểu.

2. Sử dụng thuật toán PHORIZONTAL, thiết kế phân mảnh ngang chính cho quan hệ **SANBAY**? Lưu ý, sinh viên không cần liệt kê dữ liệu trong các mảnh. (2 điểm)

**Lời giải.**

- Các vị từ giao tối thiểu là:
  - m1:  $p1 \wedge p3$ ;
  - m2:  $p1 \wedge \neg p3 = p1 \wedge p4$ ;
  - m3:  $\neg p1 \wedge p3 = p2 \wedge p3$ ;
  - m4:  $\neg p1 \wedge \neg p3 = p2 \wedge p4$ ;
- Vì vậy, ta có tập vị từ giao tối thiểu là:  $M = \{m1, m2, m3, m4\}$
- Phân mảnh ngang chính quan hệ SANBAY theo M là:  
 $SANBAY1 = \sigma_{m1}(SANBAY)$   
 $SANBAY2 = \sigma_{m2}(SANBAY)$   
 $SANBAY3 = \sigma_{m3}(SANBAY)$   
 $SANBAY4 = \sigma_{m4}(SANBAY)$
- Đối với dữ liệu demo, ta có các mảnh:  
 SANBAY1

MaICAO	TenSB	ChieuDaiDB	LoaiSB	BayDem	LoaiDB
VVRG	Rạch Giá	3000	Nội địa	Không	Bê tông
VVCA	Chu Lai	3050	Nội địa	Không	Bê tông

SANBAY2

MaICAO	TenSB	ChieuDaiDB	LoaiSB	BayDem	LoaiDB
VVCI	Cát Bi	3050	Quốc tế	Có	Bê tông
VVTX	Thọ Xuân	3200	Nội địa	Có	Bê tông

SANBAY3

MaICAO	TenSB	ChieuDaiDB	LoaiSB	BayDem	LoaiDB
VVCM	Cà Mau	1500	Nội địa	Không	Nhựa đường

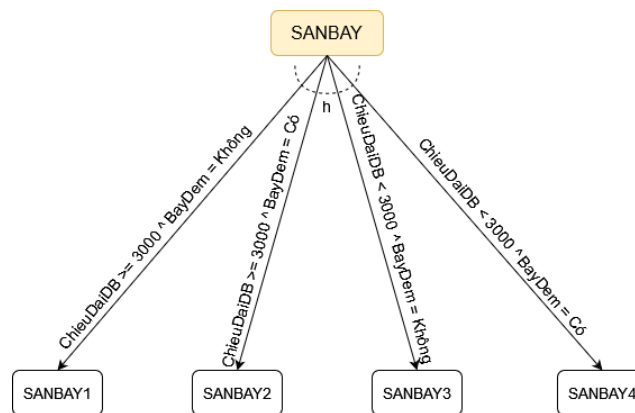
SANBAY4

MaICAO	TenSB	ChieuDaiDB	LoaiSB	BayDem	LoaiDB
VVCS	Côn Đảo	1830	Nội địa	Có	Nhựa đường
VVDL	Liên Khương	2950	Quốc tế	Có	Nhựa đường
VVPB	Phú Bài	2700	Quốc tế	Có	Nhựa đường

3. Hãy vẽ cây phân mảnh của lược đồ quan hệ **SANBAY** đã làm. Chứng minh điều kiện **tách biệt** của phân mảnh này là đúng đắn. (1 điểm)

**Lời giải.**

- Vẽ cây phân mảnh



Hình 2: Cây phân mảnh

- Chứng minh điều kiện tách biệt: Nếu mục dữ liệu  $d_i$  có trong  $R_i$  thì nó không có trong bất kỳ mảnh  $R_k$  khác ( $i \neq k$ ).  
 $\forall i \neq k$  và  $i, k \in [1, n] : R_i \cap R_k = \emptyset$ .

# Cơ Sở Dữ Liệu Phân Tán

## Bài tập về nhà - IS211.M21

Nguyễn Hồ Duy Tri, Nguyễn Thị Kim Yến  
Sinh viên: Phạm Đức Thế - 19522253

Thứ 6, ngày 01 tháng 04 năm 2022

### Bài Tập Phân Mảnh Dọc

#### Bài Tập 1

Cho tập  $Q = \{q1, q2, q3, q4, q5\}$  các truy vấn, tập  $A = \{A1, A2, A3, A4, A5\}$  lần lượt là các thuộc tính **MaCN**, **GioiTinh**, **MaPX**, **MaTo**, **Luong** của quan hệ:

**CONGNHAN**(MaCN, TenCN, NgaySinh, GioiTinh, MaPX, MaTo, Luong)

Tập  $S = \{S1, S2, S3, S4\}$  các vị trí (sites) trong hệ cơ sở dữ liệu phân tán. Giả sử số truy xuất đến các cặp thuộc tính cho mỗi ứng dụng tại các vị trí:  $ref_i(q_j) = 1$ ; với mọi  $i, j$ .

Ma trận truy vấn sử dụng thuộc tính (use) và ma trận tần số sử dụng truy vấn (acc) tại các vị trí như sau:

	A1	A2	A3	A4	A5
q1	1	0	1	0	1
q2	0	1	1	0	0
q3	0	1	0	1	1
q4	0	1	0	0	1
q5	0	0	0	0	1

**Ma trận USE**

	S1	S2	S3	S4
q1	0	5	2	3
q2	5	10	5	0
q3	10	10	0	10
q4	20	0	10	10
q5	10	10	15	15

**Ma trận ACC**

1. Tính ma trận AA. (1 điểm)

**Lời giải.**

- Tính toán ma trận AA:

$$Aff(A_1, A_1) = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10$$

$$Aff(A_1, A_2) = 0$$

$$Aff(A_1, A_3) = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10$$

$$Aff(A_1, A_4) = 0$$

$$Aff(A_1, A_5) = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10$$

$$Aff(A_2, A_2) = acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2)$$

$$+ acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3)$$

$$+ acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4)$$

$$= 5 + 10 + 5 + 0 + 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 = 90$$

$$Aff(A_2, A_3) = acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) = 5 + 10 + 5 + 0 = 20$$

$$Aff(A_2, A_4) = acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) = 10 + 10 + 0 + 10 = 30$$



$$\begin{aligned}
Aff(A_2, A_5) &= acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) \\
&+ acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) \\
&= 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 = 70
\end{aligned}$$

$$\begin{aligned}
Aff(A_3, A_3) &= acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) \\
&+ acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) \\
&= 0 + 5 + 2 + 3 + 5 + 10 + 5 + 0 = 30
\end{aligned}$$

$$Aff(A_3, A_4) = 0$$

$$Aff(A_3, A_5) = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10$$

$$Aff(A_4, A_4) = acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) = 10 + 10 + 0 + 10 = 30$$

$$Aff(A_4, A_5) = acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) = 10 + 10 + 0 + 10 = 30$$

$$\begin{aligned}
Aff(A_5, A_5) &= acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) \\
&+ acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) \\
&+ acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) \\
&+ acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) \\
&= 0 + 5 + 2 + 3 + 10 + 10 + 0 + 10 \\
&+ 20 + 0 + 10 + 10 + 10 + 10 + 15 + 15 = 130
\end{aligned}$$

- Ta có ma trận ái lực AA là:

	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>A5</b>
<b>A1</b>	10	0	10	0	10
<b>A2</b>	0	90	20	30	70
<b>A3</b>	10	20	30	0	10
<b>A4</b>	0	30	0	30	30
<b>A5</b>	10	70	10	30	130

## 2. Sử dụng thuật toán BEA tính ma trận CA. (2 điểm)

### Lời giải.

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_3$  giữa các thuộc tính  $A_1$  và  $A_2$ .

$$\begin{aligned}
Cont(A_0, A_3, A_1) &= 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1) \\
Bond(A_0, A_3) &= Aff(A_1, A_0) \times Aff(A_1, A_3) + Aff(A_2, A_0) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_0) \times Aff(A_3, A_3) + Aff(A_4, A_0) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_0) \times Aff(A_5, A_3) \\
&= 0 \times 10 + 0 \times 20 + 0 \times 30 + 0 \times 0 + 0 \times 10 = 0
\end{aligned}$$

$$\begin{aligned}
Bond(A_3, A_1) &= Aff(A_1, A_3) \times Aff(A_1, A_1) + Aff(A_2, A_3) \times Aff(A_2, A_1) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_1) + Aff(A_4, A_3) \times Aff(A_4, A_1) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_1) \\
&= 0 \times 10 + 20 \times 0 + 30 \times 10 + 0 \times 30 + 10 \times 10 = 500
\end{aligned}$$

$$\begin{aligned}
Bond(A_0, A_1) &= Aff(A_1, A_0) \times Aff(A_1, A_1) + Aff(A_2, A_0) \times Aff(A_2, A_1) \\
&+ Aff(A_3, A_0) \times Aff(A_3, A_1) + Aff(A_4, A_0) \times Aff(A_4, A_1) \\
&+ Aff(A_5, A_0) \times Aff(A_5, A_1) \\
&= 0 \times 10 + 0 \times 0 + 0 \times 10 + 0 \times 0 + 0 \times 10 = 0
\end{aligned}$$

$$\begin{aligned}
\Rightarrow Cont(A_0, A_3, A_1) &= 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1) \\
&= 2 \times 0 + 2 \times 500 - 2 \times 0 = 1000
\end{aligned}$$

$$\begin{aligned}
Cont(A_1, A_3, A_2) &= 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2) \\
Bond(A_1, A_3) &= Aff(A_1, A_1) \times Aff(A_1, A_3) + Aff(A_2, A_1) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_3) + Aff(A_4, A_1) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_3) \\
&= 10 \times 10 + 0 \times 20 + 10 \times 30 + 0 \times 0 + 10 \times 10 = 500 \\
Bond(A_3, A_2) &= Aff(A_1, A_3) \times Aff(A_1, A_2) + Aff(A_2, A_3) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_2) + Aff(A_4, A_3) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_2) \\
&= 10 \times 0 + 20 \times 90 + 30 \times 20 + 0 \times 30 + 10 \times 70 = 3100 \\
Bond(A_1, A_2) &= Aff(A_1, A_1) \times Aff(A_1, A_2) + Aff(A_2, A_1) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_2) + Aff(A_4, A_1) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_2) \\
&= 10 \times 0 + 0 \times 90 + 10 \times 20 + 0 \times 30 + 10 \times 70 = 900 \\
\Rightarrow Cont(A_1, A_3, A_2) &= 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2) \\
&= 2 \times 500 + 2 \times 3100 - 2 \times 900 = 5400
\end{aligned}$$

$$\begin{aligned}
Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\
Bond(A_2, A_3) &= Aff(A_1, A_2) \times Aff(A_1, A_3) + Aff(A_2, A_2) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_2) \times Aff(A_3, A_3) + Aff(A_4, A_2) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_2) \times Aff(A_5, A_3) \\
&= 0 \times 10 + 90 \times 20 + 20 \times 30 + 30 \times 0 + 70 \times 10 = 3100 \\
Bond(A_3, A_0) &= Aff(A_1, A_3) \times Aff(A_1, A_0) + Aff(A_2, A_3) \times Aff(A_2, A_0) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_0) + Aff(A_4, A_3) \times Aff(A_4, A_0) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_0) \\
&= 10 \times 0 + 20 \times 0 + 30 \times 0 + 0 \times 0 + 10 \times 0 = 0 \\
Bond(A_2, A_0) &= Aff(A_1, A_2) \times Aff(A_1, A_0) + Aff(A_2, A_2) \times Aff(A_2, A_0) \\
&+ Aff(A_3, A_2) \times Aff(A_3, A_0) + Aff(A_4, A_2) \times Aff(A_4, A_0) \\
&+ Aff(A_5, A_2) \times Aff(A_5, A_0) \\
&= 0 \times 0 + 90 \times 0 + 20 \times 0 + 30 \times 0 + 70 \times 0 = 0 \\
\Rightarrow Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\
&= 2 \times 3100 + 2 \times 0 - 2 \times 0 = 6200
\end{aligned}$$

$\Rightarrow$  Chèn cột  $A_3$  vào sau cột  $A_1$  và  $A_2$  của CA. Ta có:  $A_1, A_2, A_3$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_4$  giữa các thuộc tính  $A_1, A_2$  và  $A_3$ .

$$\begin{aligned}
Cont(A_0, A_4, A_1) &= 2bond(A_0, A_4) + 2bond(A_4, A_1) - 2bond(A_0, A_1) \\
Bond(A_0, A_4) &= Aff(A_1, A_0) \times Aff(A_1, A_4) + Aff(A_2, A_0) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_0) \times Aff(A_3, A_4) + Aff(A_4, A_0) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_0) \times Aff(A_5, A_4) \\
&= 0 \times 0 + 0 \times 30 + 0 \times 0 + 0 \times 30 + 0 \times 30 = 0 \\
Bond(A_4, A_1) &= Aff(A_1, A_4) \times Aff(A_1, A_1) + Aff(A_2, A_4) \times Aff(A_2, A_1) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_1) + Aff(A_4, A_4) \times Aff(A_4, A_1) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_1) \\
&= 0 \times 10 + 30 \times 0 + 0 \times 10 + 30 \times 0 + 30 \times 10 = 300
\end{aligned}$$

$$\begin{aligned}
Bond(A_0, A_1) &= Aff(A_1, A_0) \times Aff(A_1, A_1) + Aff(A_2, A_0) \times Aff(A_2, A_1) \\
&+ Aff(A_3, A_0) \times Aff(A_3, A_1) + Aff(A_4, A_0) \times Aff(A_4, A_1) \\
&+ Aff(A_5, A_0) \times Aff(A_5, A_1) \\
&= 0 \times 10 + 0 \times 0 + 0 \times 10 + 0 \times 0 + 0 \times 10 = 0 \\
\Rightarrow Cont(A_0, A_4, A_1) &= 2bond(A_0, A_4) + 2bond(A_4, A_1) - 2bond(A_0, A_1) \\
&= 2 \times 0 + 2 \times 300 - 2 \times 0 = 600
\end{aligned}$$

$$\begin{aligned}
Cont(A_1, A_4, A_2) &= 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2) \\
Bond(A_1, A_4) &= Aff(A_1, A_1) \times Aff(A_1, A_4) + Aff(A_2, A_1) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_4) + Aff(A_4, A_1) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_4) \\
&= 10 \times 0 + 0 \times 30 + 10 \times 0 + 0 \times 30 + 10 \times 30 = 300 \\
Bond(A_4, A_2) &= Aff(A_1, A_4) \times Aff(A_1, A_2) + Aff(A_2, A_4) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_2) + Aff(A_4, A_4) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_2) \\
&= 0 \times 0 + 30 \times 90 + 0 \times 20 + 30 \times 30 + 30 \times 70 = 5700 \\
Bond(A_1, A_2) &= Aff(A_1, A_1) \times Aff(A_1, A_2) + Aff(A_2, A_1) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_2) + Aff(A_4, A_1) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_2) \\
&= 10 \times 0 + 0 \times 90 + 10 \times 20 + 0 \times 30 + 10 \times 70 = 900 \\
\Rightarrow Cont(A_1, A_4, A_2) &= 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2) \\
&= 2 \times 300 + 2 \times 5700 - 2 \times 900 = 10200
\end{aligned}$$

$$\begin{aligned}
Cont(A_1, A_4, A_2) &= 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2) \\
Bond(A_1, A_4) &= Aff(A_1, A_1) \times Aff(A_1, A_4) + Aff(A_2, A_1) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_4) + Aff(A_4, A_1) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_4) \\
&= 10 \times 0 + 0 \times 30 + 10 \times 0 + 0 \times 30 + 10 \times 30 = 300 \\
Bond(A_4, A_2) &= Aff(A_1, A_4) \times Aff(A_1, A_2) + Aff(A_2, A_4) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_2) + Aff(A_4, A_4) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_2) \\
&= 0 \times 0 + 30 \times 90 + 0 \times 20 + 30 \times 30 + 30 \times 70 = 5700 \\
Bond(A_1, A_2) &= Aff(A_1, A_1) \times Aff(A_1, A_2) + Aff(A_2, A_1) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_2) + Aff(A_4, A_1) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_2) \\
&= 10 \times 0 + 0 \times 90 + 10 \times 20 + 0 \times 30 + 10 \times 70 = 900 \\
\Rightarrow Cont(A_1, A_4, A_2) &= 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2) \\
&= 2 \times 300 + 2 \times 5700 - 2 \times 900 = 10200
\end{aligned}$$

$$\begin{aligned}
Cont(A_2, A_4, A_3) &= 2bond(A_2, A_4) + 2bond(A_4, A_3) - 2bond(A_2, A_3) \\
Bond(A_2, A_4) &= Aff(A_1, A_2) \times Aff(A_1, A_4) + Aff(A_2, A_2) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_2) \times Aff(A_3, A_4) + Aff(A_4, A_2) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_2) \times Aff(A_5, A_4) \\
&= 0 \times 0 + 90 \times 30 + 20 \times 0 + 30 \times 30 + 70 \times 30 = 5700
\end{aligned}$$

$$\begin{aligned}
Bond(A_4, A_3) &= Aff(A_1, A_4) \times Aff(A_1, A_3) + Aff(A_2, A_4) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_3) + Aff(A_4, A_4) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_3) \\
&= 0 \times 0 + 30 \times 90 + 0 \times 20 + 30 \times 30 + 30 \times 70 = 900 \\
Bond(A_2, A_3) &= Aff(A_1, A_2) \times Aff(A_1, A_3) + Aff(A_2, A_2) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_2) \times Aff(A_3, A_3) + Aff(A_4, A_2) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_2) \times Aff(A_5, A_3) \\
&= 0 \times 10 + 90 \times 20 + 20 \times 30 + 30 \times 0 + 70 \times 10 = 3100 \\
\Rightarrow Cont(A_2, A_4, A_3) &= 2bond(A_2, A_4) + 2bond(A_4, A_3) - 2bond(A_2, A_3) \\
&= 2 \times 5700 + 2 \times 900 - 2 \times 3100 = 7000
\end{aligned}$$

$$\begin{aligned}
Cont(A_3, A_4, A_0) &= 2bond(A_3, A_4) + 2bond(A_4, A_0) - 2bond(A_3, A_0) \\
Bond(A_3, A_4) &= Aff(A_1, A_3) \times Aff(A_1, A_4) + Aff(A_2, A_3) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_4) + Aff(A_4, A_3) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_4) \\
&= 10 \times 0 + 20 \times 30 + 30 \times 0 + 0 \times 30 + 10 \times 30 = 900 \\
Bond(A_4, A_0) &= Aff(A_1, A_4) \times Aff(A_1, A_0) + Aff(A_2, A_4) \times Aff(A_2, A_0) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_0) + Aff(A_4, A_4) \times Aff(A_4, A_0) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_0) \\
&= 0 \times 0 + 30 \times 0 + 0 \times 0 + 30 \times 0 + 30 \times 0 = 0 \\
Bond(A_3, A_0) &= Aff(A_1, A_3) \times Aff(A_1, A_0) + Aff(A_2, A_3) \times Aff(A_2, A_0) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_0) + Aff(A_4, A_3) \times Aff(A_4, A_0) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_0) \\
&= 10 \times 0 + 20 \times 0 + 30 \times 0 + 0 \times 0 + 10 \times 0 = 0 \\
\Rightarrow Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\
&= 2 \times 900 + 2 \times 0 - 2 \times 0 = 1800
\end{aligned}$$

$\Rightarrow$  Chèn cột  $A_4$  vào giữa cột  $A_1$  và  $A_2$  của CA. Ta có:  $A_1, A_4, A_2, A_3$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_5$  giữa các thuộc tính  $A_1, A_4, A_2$  và  $A_3$ .

$$\begin{aligned}
Cont(A_0, A_5, A_1) &= 2bond(A_0, A_5) + 2bond(A_5, A_1) - 2bond(A_0, A_1) \\
Bond(A_0, A_5) &= Aff(A_1, A_0) \times Aff(A_1, A_5) + Aff(A_2, A_0) \times Aff(A_2, A_5) \\
&+ Aff(A_3, A_0) \times Aff(A_3, A_5) + Aff(A_4, A_0) \times Aff(A_4, A_5) \\
&+ Aff(A_5, A_0) \times Aff(A_5, A_5) \\
&= 0 \times 10 + 0 \times 70 + 0 \times 10 + 0 \times 30 + 0 \times 130 = 0 \\
Bond(A_5, A_1) &= Aff(A_1, A_5) \times Aff(A_1, A_1) + Aff(A_2, A_5) \times Aff(A_2, A_1) \\
&+ Aff(A_3, A_5) \times Aff(A_3, A_1) + Aff(A_4, A_5) \times Aff(A_4, A_1) \\
&+ Aff(A_5, A_5) \times Aff(A_5, A_1) \\
&= 10 \times 10 + 70 \times 0 + 10 \times 10 + 30 \times 0 + 130 \times 10 = 1500 \\
Bond(A_0, A_1) &= Aff(A_1, A_0) \times Aff(A_1, A_1) + Aff(A_2, A_0) \times Aff(A_2, A_1) \\
&+ Aff(A_3, A_0) \times Aff(A_3, A_1) + Aff(A_4, A_0) \times Aff(A_4, A_1) \\
&+ Aff(A_5, A_0) \times Aff(A_5, A_1) \\
&= 0 \times 10 + 0 \times 0 + 0 \times 10 + 0 \times 0 + 0 \times 10 = 0 \\
\Rightarrow Cont(A_0, A_5, A_1) &= 2bond(A_0, A_5) + 2bond(A_5, A_1) - 2bond(A_0, A_1) \\
&= 2 \times 0 + 2 \times 1500 - 2 \times 0 = 3000
\end{aligned}$$

$$\begin{aligned}
Cont(A_1, A_5, A_4) &= 2bond(A_1, A_5) + 2bond(A_5, A_4) - 2bond(A_1, A_4) \\
Bond(A_1, A_5) &= Aff(A_1, A_1) \times Aff(A_1, A_5) + Aff(A_2, A_1) \times Aff(A_2, A_5) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_5) + Aff(A_4, A_1) \times Aff(A_4, A_5) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_5) \\
&= 10 \times 10 + 0 \times 70 + 10 \times 10 + 0 \times 30 + 10 \times 130 = 1500 \\
Bond(A_5, A_4) &= Aff(A_1, A_5) \times Aff(A_1, A_4) + Aff(A_2, A_5) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_5) \times Aff(A_3, A_4) + Aff(A_4, A_5) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_5) \times Aff(A_5, A_4) \\
&= 10 \times 0 + 70 \times 30 + 10 \times 0 + 30 \times 30 + 130 \times 30 = 6900 \\
Bond(A_1, A_4) &= Aff(A_1, A_1) \times Aff(A_1, A_4) + Aff(A_2, A_1) \times Aff(A_2, A_4) \\
&+ Aff(A_3, A_1) \times Aff(A_3, A_4) + Aff(A_4, A_1) \times Aff(A_4, A_4) \\
&+ Aff(A_5, A_1) \times Aff(A_5, A_4) \\
&= 10 \times 0 + 0 \times 30 + 10 \times 0 + 0 \times 30 + 10 \times 30 = 300 \\
\Rightarrow Cont(A_1, A_5, A_4) &= 2bond(A_1, A_5) + 2bond(A_5, A_4) - 2bond(A_1, A_4) \\
&= 2 \times 1500 + 2 \times 6900 - 2 \times 300 = 16200
\end{aligned}$$

$$\begin{aligned}
Cont(A_4, A_5, A_2) &= 2bond(A_4, A_5) + 2bond(A_5, A_2) - 2bond(A_4, A_2) \\
Bond(A_4, A_5) &= Aff(A_1, A_4) \times Aff(A_1, A_5) + Aff(A_2, A_4) \times Aff(A_2, A_5) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_5) + Aff(A_4, A_4) \times Aff(A_4, A_5) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_5) \\
&= 0 \times 10 + 30 \times 70 + 0 \times 10 + 30 \times 30 + 30 \times 130 = 6900 \\
Bond(A_5, A_2) &= Aff(A_1, A_5) \times Aff(A_1, A_2) + Aff(A_2, A_5) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_5) \times Aff(A_3, A_2) + Aff(A_4, A_5) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_5) \times Aff(A_5, A_2) \\
&= 10 \times 0 + 70 \times 90 + 10 \times 20 + 30 \times 30 + 130 \times 70 = 16500 \\
Bond(A_4, A_2) &= Aff(A_1, A_4) \times Aff(A_1, A_2) + Aff(A_2, A_4) \times Aff(A_2, A_2) \\
&+ Aff(A_3, A_4) \times Aff(A_3, A_2) + Aff(A_4, A_4) \times Aff(A_4, A_2) \\
&+ Aff(A_5, A_4) \times Aff(A_5, A_2) \\
&= 0 \times 0 + 30 \times 90 + 0 \times 20 + 30 \times 30 + 30 \times 70 = 5700 \\
\Rightarrow Cont(A_4, A_5, A_2) &= 2bond(A_4, A_5) + 2bond(A_5, A_2) - 2bond(A_4, A_2) \\
&= 2 \times 6900 + 2 \times 16500 - 2 \times 5700 = 35400
\end{aligned}$$

$$\begin{aligned}
Cont(A_2, A_5, A_3) &= 2bond(A_2, A_5) + 2bond(A_5, A_3) - 2bond(A_2, A_3) \\
Bond(A_2, A_5) &= Aff(A_1, A_2) \times Aff(A_1, A_5) + Aff(A_2, A_2) \times Aff(A_2, A_5) \\
&+ Aff(A_3, A_2) \times Aff(A_3, A_5) + Aff(A_4, A_2) \times Aff(A_4, A_5) \\
&+ Aff(A_5, A_2) \times Aff(A_5, A_5) \\
&= 0 \times 10 + 90 \times 70 + 20 \times 10 + 30 \times 30 + 70 \times 130 = 16500 \\
Bond(A_5, A_3) &= Aff(A_1, A_5) \times Aff(A_1, A_3) + Aff(A_2, A_5) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_5) \times Aff(A_3, A_3) + Aff(A_4, A_5) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_5) \times Aff(A_5, A_3) \\
&= 10 \times 10 + 70 \times 20 + 10 \times 30 + 30 \times 0 + 130 \times 10 = 3100 \\
Bond(A_2, A_3) &= Aff(A_1, A_2) \times Aff(A_1, A_3) + Aff(A_2, A_2) \times Aff(A_2, A_3) \\
&+ Aff(A_3, A_2) \times Aff(A_3, A_3) + Aff(A_4, A_2) \times Aff(A_4, A_3) \\
&+ Aff(A_5, A_2) \times Aff(A_5, A_3) \\
&= 0 \times 10 + 90 \times 20 + 20 \times 30 + 30 \times 0 + 70 \times 10 = 3100 \\
\Rightarrow Cont(A_2, A_5, A_3) &= 2bond(A_2, A_5) + 2bond(A_5, A_3) - 2bond(A_2, A_3) \\
&= 2 \times 16500 + 2 \times 3100 - 2 \times 3100 = 33000
\end{aligned}$$

$$\begin{aligned}
Cont(A_3, A_5, A_0) &= 2bond(A_3, A_5) + 2bond(A_5, A_0) - 2bond(A_3, A_0) \\
Bond(A_3, A_5) &= Aff(A_1, A_3) \times Aff(A_1, A_5) + Aff(A_2, A_3) \times Aff(A_2, A_5) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_5) + Aff(A_4, A_3) \times Aff(A_4, A_5) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_5) \\
&= 10 \times 10 + 20 \times 70 + 30 \times 10 + 0 \times 30 + 10 \times 130 = 3100 \\
Bond(A_5, A_0) &= Aff(A_1, A_5) \times Aff(A_1, A_0) + Aff(A_2, A_5) \times Aff(A_2, A_0) \\
&+ Aff(A_3, A_5) \times Aff(A_3, A_0) + Aff(A_4, A_5) \times Aff(A_4, A_0) \\
&+ Aff(A_5, A_5) \times Aff(A_5, A_0) \\
&= 10 \times 0 + 70 \times 0 + 10 \times 0 + 30 \times 0 + 130 \times 0 = 0 \\
Bond(A_3, A_0) &= Aff(A_1, A_3) \times Aff(A_1, A_0) + Aff(A_2, A_3) \times Aff(A_2, A_0) \\
&+ Aff(A_3, A_3) \times Aff(A_3, A_0) + Aff(A_4, A_3) \times Aff(A_4, A_0) \\
&+ Aff(A_5, A_3) \times Aff(A_5, A_0) \\
&= 10 \times 0 + 20 \times 0 + 30 \times 0 + 0 \times 0 + 10 \times 0 = 0 \\
\Rightarrow Cont(A_3, A_5, A_0) &= 2bond(A_3, A_5) + 2bond(A_5, A_0) - 2bond(A_3, A_0) \\
&= 2 \times 3100 + 2 \times 0 - 2 \times 0 = 6200
\end{aligned}$$

$\Rightarrow$  Chèn cột  $A_5$  vào giữa cột  $A_4$  và  $A_2$  của CA. Ta có:  $A_1, A_4, A_5, A_2, A_3$

- Ta có ma trận bond là:

	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>A5</b>
<b>A1</b>		900	500	300	1500
<b>A2</b>	900		3100	5700	16500
<b>A3</b>	500	3100		900	3100
<b>A4</b>	300	5700	900		6900
<b>A5</b>	1500	16500	3100	6900	

- Ta có ma trận ái lực tự CA là:

	<b>A1</b>	<b>A4</b>	<b>A5</b>	<b>A2</b>	<b>A3</b>
<b>A1</b>	10	0	10	0	10
<b>A4</b>	0	30	30	30	0
<b>A5</b>	10	30	130	70	10
<b>A2</b>	0	30	70	90	20
<b>A3</b>	10	0	10	20	30

3. Sử dụng thuật toán PARTITION để tìm ra hai phân mảnh dọc gom tụ có dư thừa của quan hệ **CONGNHAN**. (2 điểm)

**Lời giải.**

- $Q = \{q1, q2, q3, q4, q5\}$
- $AQ(q1) = \{A_1, A_3, A_5\}$
- $AQ(q2) = \{A_2, A_3\}$
- $AQ(q3) = \{A_2, A_4, A_5\}$
- $AQ(q4) = \{A_2, A_5\}$
- $AQ(q5) = \{A_5\}$

	<b>A1</b>	<b>A4</b>	<b>A5</b>	<b>A2</b>	<b>A3</b>
<b>A1</b>	10	0	10	0	10
<b>A4</b>	0	30	30	30	0
<b>A5</b>	10	30	130	70	10
<b>A2</b>	0	30	70	90	20
<b>A3</b>	10	0	10	20	30

- $TA = \{A_1, A_4, A_5, A_2\}$   $BA = \{A_3\}$
- $TQ = \{q3, q4, q5\}$   $BQ = \{\emptyset\}$
- $OQ = \{q1, q2\}$
- $CTQ = acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) + acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) + acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) = 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 + 10 + 10 + 15 + 15 = 120$
- $CBQ = 0$
- $COQ = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) + acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) = 0 + 5 + 2 + 3 + 5 + 10 + 5 + 0 = 30$
- $Z = CTQ \times CBQ - COQ^2 = 120 \times 0 - 30^2 = -900$

	A1	A4	A5	A2	A3
A1	10	0	10	0	10
A4	0	30	30	30	0
A5	10	30	130	70	10
A2	0	30	70	90	20
A3	10	0	10	20	30

- $TA = \{A_1, A_4, A_5\}$   $BA = \{A_2, A_3\}$
- $TQ = \{q5\}$   $BQ = \{q2\}$
- $OQ = \{q1, q3, q4\}$
- $CTQ = acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) = 10 + 10 + 15 + 15 = 50$
- $CBQ = acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) = 5 + 10 + 5 + 0 = 20$
- $COQ = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) + acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) + acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) = 0 + 5 + 2 + 3 + 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 = 80$
- $Z = CTQ \times CBQ - COQ^2 = 50 \times 20 - 80^2 = -5400$

	A1	A4	A5	A2	A3
A1	10	0	10	0	10
A4	0	30	30	30	0
A5	10	30	130	70	10
A2	0	30	70	90	20
A3	10	0	10	20	30

- $TA = \{A_1, A_4\}$   $BA = \{A_5, A_2, A_3\}$
- $TQ = \{\emptyset\}$   $BQ = \{q2, q4, q5\}$
- $OQ = \{q1, q3\}$
- $CTQ = 0$
- $CBQ = acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) + acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) + acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) = 5 + 10 + 5 + 0 + 20 + 0 + 10 + 10 + 10 + 10 + 15 + 15 = 110$
- $COQ = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) + acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) = 0 + 5 + 2 + 3 + 10 + 10 + 0 + 10 = 40$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 110 - 40^2 = -1600$

	A1	A4	A5	A2	A3
A1	10	0	10	0	10
A4	0	30	30	30	0
A5	10	30	130	70	10
A2	0	30	70	90	20
A3	10	0	10	20	30

- $TA = \{A_1\}$   $BA = \{A_4, A_5, A_2, A_3\}$
- $TQ = \{\emptyset\}$   $BQ = \{q2, q3, q4, q5\}$
- $OQ = \{q1\}$
- $CTQ = 0$

- $CBQ = acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) + acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) + acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) + acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) = 5 + 10 + 5 + 0 + 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 + 10 + 10 + 15 + 15 = 140$
- $COQ = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 140 - 10^2 = -100$

Áp dụng kết quả phân hoạch ma trận CA vào quan hệ CONGNHAN, ta được các mảnh F: CONGNHAN = {CONGNHAN1, CONGNHAN2}.

- Trong đó:  
 $CONGNHAN1 = \{A_1\}$   
 $CONGNHAN2 = \{A_1, A_4, A_5, A_2, A_3\}$
- Vì thế:  
 $CONGNHAN1 = \{MaCN, TenCN, NgaySinh\}$   
 $CONGNHAN2 = \{MaCN, TenCN, NgaySinh, GioiTinh, MaPX, MaTo, Luong\}$
- MaCN là thuộc tính khóa của quan hệ CONGNHAN.

	<b>A1</b>	<b>A4</b>	<b>A5</b>	<b>A2</b>	<b>A3</b>
<b>A1</b>	10	0	10	0	10
<b>A4</b>	0	30	30	30	0
<b>A5</b>	10	30	130	70	10
<b>A2</b>	0	30	70	90	20
<b>A3</b>	10	0	10	20	30



## Bài Tập 2

Cho tập  $Q = \{q1, q2, q3, q4, q5\}$  các truy vấn, tập  $A = \{A1, A2, A3, A4, A5\}$  lần lượt là các thuộc tính **TenSB**, **NamTL**, **LoaiSB**, **BayDem**, **LoaiDB** của quan hệ:

**SANBAY(MaICAO, TenSB, NamTL, LoaiSB, BayDem, LoaiDB)**

Tập  $S = \{S1, S2, S3, S4\}$  các vị trí (sites) trong hệ cơ sở dữ liệu phân tán. Giả sử số truy suất đến các cặp thuộc tính tại các vị trí được cho theo hàm:

$$ref_j(q_i) = \begin{cases} 1, \forall i \in [1, 4], j \in [1, 3] \\ 2, \forall i \in [1, 4], j \in [4, 5] \end{cases}, \text{ Với } i \text{ là số chỉ vị trí (site), } j \text{ là số chỉ của câu truy vấn.}$$

Ma trận truy vấn sử dụng thuộc tính (use) và ma trận tần số sử dụng truy vấn (acc) tại các vị trí như sau:

	A1	A2	A3	A4	A5
q1	0	1	0	1	0
q2	0	1	1	1	0
q3	1	1	0	0	0
q4	1	0	1	0	0
q5	1	0	1	0	1

**Ma trận USE**

	S1	S2	S3	S4
q1	0	5	0	3
q2	3	0	1	1
q3	7	0	3	0
q4	0	5	0	9
q5	1	0	3	0

**Ma trận ACC**

1. Tính ma trận AA. (1 điểm)

**Lời giải.**

- Ta có ma trận ái lực AA là:

	A1	A2	A3	A4	A5
A1	46	10	36	0	8
A2	10	23	5	13	0
A3	36	5	41	5	8
A4	0	13	5	13	0
A5	8	0	8	0	8

2. Sử dụng thuật toán BEA tính ma trận CA. (2 điểm)

**Lời giải.**

- Ta có ma trận bond là:

	A1	A2	A3	A4	A5
A1		870	3246	310	720
A2	870		745	493	120
A3	3246	745		335	680
A4	310	493	335		40
A5	720	120	680	40	

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_3$  giữa các thuộc tính  $A_1$  và  $A_2$ .

$$\begin{aligned} Cont(A_0, A_3, A_1) &= 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 3246 - 2 \times 0 = 6492 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_3, A_2) &= 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2) \\ &= 2 \times 3246 + 2 \times 745 - 2 \times 870 = 6242 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 745 + 2 \times 0 - 2 \times 0 = 1490 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_3$  vào trước cột  $A_1$  và  $A_2$  của CA. Ta có:  **$A_3, A_1, A_2$**

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_4$  giữa các thuộc tính  $A_3, A_1, A_2$ .

$$\begin{aligned}
Cont(A_0, A_4, A_3) &= 2bond(A_0, A_4) + 2bond(A_4, A_3) - 2bond(A_0, A_3) \\
&= 2 \times 0 + 2 \times 335 - 2 \times 0 = 670 \\
Cont(A_3, A_4, A_1) &= 2bond(A_3, A_4) + 2bond(A_4, A_1) - 2bond(A_3, A_1) \\
&= 2 \times 335 + 2 \times 310 - 2 \times 3246 = -5202 \\
Cont(A_1, A_4, A_2) &= 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2) \\
&= 2 \times 310 + 2 \times 493 - 2 \times 870 = -134 \\
Cont(A_2, A_4, A_0) &= 2bond(A_2, A_4) + 2bond(A_4, A_0) - 2bond(A_2, A_0) \\
&= 2 \times 493 + 2 \times 0 - 2 \times 0 = 986
\end{aligned}$$

$\Rightarrow$  Chèn cột  $A_4$  vào sau cột  $A_1$  và  $A_2$  của ma trận AC. Ta có:  $A_3, A_1, A_2, A_4$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_5$  giữa các thuộc tính  $A_3, A_1, A_2, A_4$ .

$$\begin{aligned}
Cont(A_0, A_5, A_3) &= 2bond(A_0, A_5) + 2bond(A_5, A_3) - 2bond(A_0, A_3) \\
&= 2 \times 0 + 2 \times 680 - 2 \times 0 = 1360 \\
Cont(A_3, A_5, A_1) &= 2bond(A_3, A_5) + 2bond(A_5, A_1) - 2bond(A_3, A_1) \\
&= 2 \times 680 + 2 \times 720 - 2 \times 3246 = -3692 \\
Cont(A_1, A_5, A_2) &= 2bond(A_1, A_5) + 2bond(A_5, A_2) - 2bond(A_1, A_2) \\
&= 2 \times 720 + 2 \times 120 - 2 \times 870 = -60 \\
Cont(A_2, A_5, A_4) &= 2bond(A_2, A_5) + 2bond(A_5, A_4) - 2bond(A_2, A_4) \\
&= 2 \times 120 + 2 \times 40 - 2 \times 493 = -666 \\
Cont(A_4, A_5, A_0) &= 2bond(A_4, A_5) + 2bond(A_5, A_0) - 2bond(A_4, A_0) \\
&= 2 \times 40 + 2 \times 0 - 2 \times 0 = 80
\end{aligned}$$

$\Rightarrow$  Chèn cột  $A_5$  vào trước cột  $A_3$  và  $A_1$  của ma trận AC. Ta có:  $A_5, A_3, A_1, A_2, A_4$

- Ta có ma trận ái lực tự CA là:

	A5	A3	A1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

- Sử dụng thuật toán PARTITION để tìm ra hai phân mảnh dọc của quan hệ **SANBAY**. (2 điểm)

**Lời giải.**

- $Q = \{q1, q2, q3, q4, q5\}$
- $AQ(q1) = \{A_2, A_4\}$
- $AQ(q2) = \{A_2, A_3, A_4\}$
- $AQ(q3) = \{A_1, A_2\}$
- $AQ(q4) = \{A_1, A_3\}$
- $AQ(q5) = \{A_1, A_3, A_5\}$

	A5	A3	A1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

- $TA = \{A_5, A_3, A_1, A_2\}$   $BA = \{A_4\}$
- $TQ = \{q3, q4, q5\}$   $BQ = \{\emptyset\}$
- $OQ = \{q1, q2\}$
- $CTQ = 10 + 2 \times 14 + 2 \times 4 = 46$
- $CBQ = 0$
- $COQ = 8 + 5 = 13$
- $Z = CTQ \times CBQ - COQ^2 = 46 \times 0 - 13^2 = -169$

	A5	A3	A1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

- $TA = \{A_5, A_3, A_1\}$   $BA = \{A_2, A_4\}$
- $TQ = \{q4, q5\}$   $BQ = \{q1\}$
- $OQ = \{q2, q3\}$
- $CTQ = 2 \times 14 + 2 \times 4 = 36$
- $CBQ = 8$
- $COQ = 5 + 10 = 15$
- $Z = CTQ \times CBQ - COQ^2 = 36 \times 8 - 15^2 = 63$

	A5	A3	A1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

- $TA = \{A_5, A_3\}$   $BA = \{A_1, A_2, A_4\}$
- $TQ = \{\emptyset\}$   $BQ = \{q1, q3\}$
- $OQ = \{q2, q4, q5\}$
- $CTQ = 0$
- $CBQ = 8 + 10 = 18$
- $COQ = 5 + 2 \times 14 + 2 \times 4 = 41$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 18 - 41^2 = -1681$

	A5	A3	A1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

- $TA = \{A_5\}$   $BA = \{A_3, A_1, A_2, A_4\}$
- $TQ = \{\emptyset\}$   $BQ = \{q1, q2, q3, q4\}$
- $OQ = \{q5\}$
- $CTQ = 0$
- $CBQ = 8 + 5 + 10 + 2 \times 14 = 51$
- $COQ = 2 \times 4 = 8$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 51 - 8^2 = -64$

Áp dụng kết quả phân hoạch ma trận CA vào quan hệ SANBAY, ta được các mảnh F:  
 $SANBAY = \{SANBAY1, SANBAY2\}$ .

- Trong đó:  
 $SANBAY1 = \{A_5, A_3, A_1\}$   
 $SANBAY2 = \{A_2, A_4\}$
- Vì thế:  
 $SANBAY1 = \{MaICAO, LoaiSB, LoaiDB, TenSB\}$   
 $SANBAY2 = \{MaICAO, NamTL, BayDem\}$
- MaICAO là thuộc tính khóa chính của quan hệ SANBAY.

	<b>A5</b>	<b>A3</b>	<b>A1</b>	<b>A2</b>	<b>A4</b>
<b>A5</b>	8	8	8	0	0
<b>A3</b>	8	41	36	5	5
<b>A1</b>	8	36	46	10	0
<b>A2</b>	0	5	10	23	13
<b>A4</b>	0	5	0	13	13

# Cơ Sở Dữ Liệu Phân Tán

## Ôn Tập Giữa Kỳ - IS211.M21

Nguyễn Hồ Duy Tri, Nguyễn Thị Kim Yến  
Sinh viên: Phạm Đức Thế - 19522253

Thứ 6, ngày 08 tháng 04 năm 2022

### Ôn tập giữa kỳ

#### ĐỀ THI GIỮA KỲ - HỌC KỲ I 2020-2021 - Mã đề 01

Tập đoàn Điện lực Việt Nam hiện muốn quản lý thông tin các nhà máy điện trực tiếp tham gia thị trường phát điện cạnh tranh. Tùy theo nguồn năng lượng được biến đổi thành điện năng mà các nhà máy được phân loại thành hai nhóm là Thủy điện (PhanLoai = “Thủy điện;”) và Tuabin khí (PhanLoai = “Tuabin khí”). Ngoài ra, mỗi nhà máy sẽ được một trong hai cơ quan trực tiếp quản lý, đó là PVPower (CQQL = “PVPower”) và GENCO 3 (CQQL = “GENCO3”). Thông tin của các nhà máy điện được thể hiện qua lược đồ quan hệ sau:

**NHAMAYDIEN** (MaNMĐ, TenNMĐ, PhanLoai, CQQL, CongSuat)

**Tân từ:** Mỗi nhà máy điện có các thông tin bao gồm: Mã nhà máy điện (MaNMĐ), Tên nhà máy điện (TenNMĐ), phân loại nhà máy (PhanLoai), cơ quan quản lý (CQQL), công suất phát điện (CongSuat) được tính theo đơn vị MW.

- (5 điểm) Cho thể hiện của quan hệ **MAYDIEUHOA** như sau:

MaNMĐ	TenNMĐ	PhanLoai	CQQL	CongSuat
NM01	Đak Đrinh	Thủy điện	PVPower	125
NM02	Hủa Na	Thủy điện	PVPower	180
NM03	Nhơn Trạch 2	Tuabin khí	PVPower	750
NM04	Sông Hinh	Thủy điện	GENCO3	70
NM05	Phú Mỹ 1	Tuabin khí	GENCO3	1059
NM06	Buôn Kuốp	Thủy điện	GENCO3	280
NM07	Nhơn Trạch 1	Tuabin khí	PVPower	450
NM08	Thác Bà	Thủy điện	GENCO3	120

Giả sử có ba ứng dụng truy xuất đến **NHAMAYDIEN**:

Q1: SELECT COUNT(\*) FROM **NHAMAYDIEN** WHERE **PhanLoai** = value

Q2: SELECT **MaNMĐ**, **TenNMĐ** FROM **NHAMAYDIEN** WHERE **CQQL** = value

Q3: SELECT \* FROM **NHAMAYDIEN** WHERE **CongSuat** > 400

- Dùng giải thuật COM\_MIN, tính Pr' thỏa tối thiểu và đầy đủ? (2 điểm)

**Lời giải.**

- Tập vị từ đơn giản dùng để phân hoạch **NHAMAYDIEN**:
  - p1: PhanLoai = “Thủy điện”
  - p2: PhanLoai = “Tuabin khí”
  - p3: CQQL = “PVPower”
  - p4: CQQL = “GENCO3”
  - p5: CongSuat > 400
  - p6: CongSuat ≤ 400

- Khởi tạo  $P_r = \{p1, p2, p3, p4, p5, p6\}, P_{r'} = \emptyset$
  - Áp dụng thuật toán COM\_MIN:
    - $i = 1$  làm giá trị khởi đầu, vị từ  $p1$  thỏa quy tắc 1,  $P_{r'} = \{p1\}$ .
    - $i = 2$ : Ta có vị từ  $p2$  không phân hoạch  $f1$  (là mảnh giao tối thiểu tạo ra ứng với  $p1$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1\}$ .
    - $i = 3$ : Vị từ  $p3$  thỏa quy tắc 1,  $P_{r'} = \{p1, p3\}$ .
    - $i = 4$ : Ta có vị từ  $p4$  không phân hoạch  $f3$  (là mảnh giao tối thiểu tạo ra ứng với  $p3$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
    - $i = 5$ : Ta có vị từ  $p5$  không phân hoạch  $f4$  (là mảnh giao tối thiểu tạo ra ứng với  $p4$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
    - $i = 6$ : Ta có vị từ  $p6$  không phân hoạch  $f5$  (là mảnh giao tối thiểu tạo ra ứng với  $p5$ ) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$ .
- Kết luận:  $P_{r'} = \{p1, p3\}$  là đầy đủ và tối thiểu.

(b) Sử dụng thuật toán PHORIZONTAL, thiết kế phân mảnh ngang chính cho quan hệ NHAMAYDIEN? Lưu ý, sinh viên không cần liệt kê dữ liệu trong các mảnh. (2 điểm)

**Lời giải.**

- Các vị từ giao tối thiểu là:
  - $m1: p1 \wedge p3$ ;
  - $m2: p1 \wedge \neg p3 = p1 \wedge p4$ ;
  - $m3: \neg p1 \wedge p3 = p2 \wedge p3$ ;
  - $m4: \neg p1 \wedge \neg p3 = p2 \wedge p4$ ;
- Vì vậy, ta có tập các vị từ giao tối thiểu là:  $M = \{m1, m2, m3, m4\}$
- Phân mảnh ngang chính quan hệ NHAMAYDIEN theo  $M$  là:
  - $NMD1 = \sigma_{m1}(NHAMAYDIEN)$
  - $NMD2 = \sigma_{m2}(NHAMAYDIEN)$
  - $NMD3 = \sigma_{m3}(NHAMAYDIEN)$
  - $NMD4 = \sigma_{m4}(NHAMAYDIEN)$
- Đối với dữ liệu demo, ta có các mảnh:
  - MND1

MaMND	TenNMD	PhanLoai	CQQL	CongSuat
NM01	Đak Đrinh	Thủy điện	PVPower	125
NM02	Hòa Na	Thủy điện	PVPower	180

MND2

MaMND	TenNMD	PhanLoai	CQQL	CongSuat
NM04	Sông Hình	Thủy điện	GENCO3	70
NM06	Buôn Kuốp	Thủy điện	GENCO3	280
NM08	Thác Bà	Thủy điện	GENCO3	120

MND3

MaMND	TenNMD	PhanLoai	CQQL	CongSuat
NM03	Nhơn Trạch 2	Tuabin khí	PVPower	750
NM07	Nhơn Trạch 1	Tuabin khí	PVPower	450

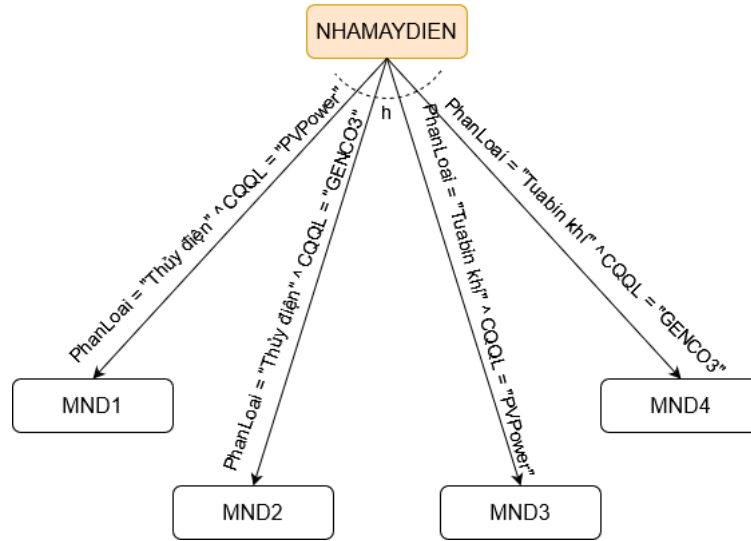
MND4

MaMND	TenNMD	PhanLoai	CQQL	CongSuat
NM05	Phú Mỹ 1	Tuabin khí	GENCO3	1059

- (c) Hãy vẽ cây phân mảnh của lược đồ quan hệ **NHAMAYDIEN** đã làm. Chứng minh điều kiện **tái tạo** của phân mảnh này là đúng đắn. (1 điểm)

**Lời giải.**

- Vẽ cây phân mảnh



Hình 1: Cây phân mảnh

- Chứng minh điều kiện **tái tạo**

Giữa các mảnh có thể áp dụng phép toán hợp để tái tạo lại mảnh ban đầu:

Ta có:  $NMD1 \cup NMD2 \cup NMD3 \cup NMD4 = NHAMAYDIEN$

$\Rightarrow$  Điều kiện tái tạo của phân mảnh này là đúng đắn.

2. (5 điểm) Cho tập  $Q = \{q1, q2, q3, q4, q5\}$  các truy vấn, tập  $A = \{A1, A2, A3, A4, A5\}$  lần lượt là các thuộc tính **MaNMD**, **TenNMD**, **PhanLoai**, **CQQL**, **CongSuat** của quan hệ:

**NHAMAYDIEN** (**MaNMD**, **TenNMD**, **PhanLoai**, **CQQL**, **CongSuat**)

Tập  $S = \{S1, S2, S3, S4\}$  các vị trí (sites) trong hệ cơ sở dữ liệu phân tán. Giả sử số truy xuất đến các cặp thuộc tính tại các vị trí được cho theo hàm:

$$ref_i(q_j) = \begin{cases} 1, \forall i \in [1, 4], j \in \{1, 3, 5\} \\ 2, \forall i \in [1, 4], j \in \{2, 4\} \end{cases}, \text{ Với } i \text{ là số chỉ vị trí (site), } j \text{ là số chỉ của câu truy vấn.}$$

Ma trận truy vấn sử dụng thuộc tính (use) và ma trận tần số sử dụng truy vấn (acc) tại các vị trí như sau:

	A1	A2	A3	A4	A5
q1	0	1	1	0	1
q2	0	0	1	1	0
q3	1	0	1	0	1
q4	1	0	0	1	0
q5	0	1	0	0	1

**Ma trận USE**

	S1	S2	S3	S4
q1	3	0	5	0
q2	0	2	0	6
q3	1	0	3	0
q4	0	3	0	3
q5	2	0	1	3

**Ma trận ACC**

(a) Tính ma trận AA. (1 điểm)

**Lời giải.**

- Ta có ma trận ái lực AA là:

	A1	A2	A3	A4	A5
A1	16	0	4	12	4
A2	0	14	8	0	14
A3	4	8	28	16	12
A4	12	0	16	28	0
A5	4	14	12	0	18

(b) Sử dụng thuật toán BEA tính ma trận CA. (2 điểm)

**Lời giải.**

- Ta có ma trận bond là:

	A1	A2	A3	A4	A5
A1		88	416	592	184
A2	88		504	128	544
A3	416	504		944	680
A4	592	128	944		240
A5	184	544	680	240	

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_3$  giữa các thuộc tính  $A_1, A_2$ .

$$\begin{aligned} Cont(A_0, A_3, A_1) &= 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 416 - 2 \times 0 = 832 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_3, A_2) &= 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2) \\ &= 2 \times 416 + 2 \times 504 - 2 \times 88 = 1664 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 504 + 2 \times 0 - 2 \times 0 = 1008 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_3$  vào giữa cột  $A_1$  và  $A_2$  của CA. Ta có:  $A_1, A_3, A_2$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_4$  giữa các thuộc tính  $A_1, A_3, A_2$ .

$$\begin{aligned} Cont(A_0, A_4, A_1) &= 2bond(A_0, A_4) + 2bond(A_4, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 592 - 2 \times 0 = 1184 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_4, A_3) &= 2bond(A_1, A_4) + 2bond(A_4, A_3) - 2bond(A_1, A_3) \\ &= 2 \times 592 + 2 \times 944 - 2 \times 416 = 2240 \end{aligned}$$

$$\begin{aligned} Cont(A_3, A_4, A_2) &= 2bond(A_3, A_4) + 2bond(A_4, A_2) - 2bond(A_3, A_2) \\ &= 2 \times 944 + 2 \times 128 - 2 \times 504 = 1136 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_4, A_0) &= 2bond(A_2, A_4) + 2bond(A_4, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 128 + 2 \times 0 - 2 \times 0 = 256 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_4$  vào giữa cột  $A_1$  và  $A_3$  của ma trận AC. Ta có:  $A_1, A_4, A_3, A_2$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_5$  giữa các thuộc tính  $A_1, A_4, A_3, A_2$ .

$$\begin{aligned} Cont(A_0, A_5, A_1) &= 2bond(A_0, A_5) + 2bond(A_5, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 184 - 2 \times 0 = 368 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_5, A_4) &= 2bond(A_1, A_5) + 2bond(A_5, A_4) - 2bond(A_1, A_4) \\ &= 2 \times 184 + 2 \times 240 - 2 \times 592 = -336 \end{aligned}$$

$$\begin{aligned} Cont(A_4, A_5, A_3) &= 2bond(A_4, A_5) + 2bond(A_5, A_3) - 2bond(A_4, A_3) \\ &= 2 \times 240 + 2 \times 680 - 2 \times 944 = -48 \end{aligned}$$



$$\begin{aligned}
Cont(A_3, A_5, A_2) &= 2bond(A_3, A_5) + 2bond(A_5, A_2) - 2bond(A_3, A_2) \\
&= 2 \times 680 + 2 \times 544 - 2 \times 504 = 1440 \\
Cont(A_2, A_5, A_0) &= 2bond(A_2, A_5) + 2bond(A_5, A_0) - 2bond(A_2, A_0) \\
&= 2 \times 544 + 2 \times 0 - 2 \times 0 = 1088
\end{aligned}$$

$\implies$  Chèn cột  $A_5$  vào giữa cột  $A_3$  và  $A_2$  của ma trận AC. Ta có:  $A_1, A_4, A_3, A_5, A_2$

- Ta có ma trận ái lực tự CA là:

	A1	A4	A3	A5	A2
A1	16	12	4	4	0
A4	12	28	16	0	0
A3	4	16	28	12	8
A5	4	0	12	18	14
A2	0	0	8	14	14

- (c) Sử dụng thuật toán PARTITION để tìm ra hai phân mảnh dọc của quan hệ **NHAMAY-DIEN**. (2 điểm)

**Lời giải.**

- $Q = \{q1, q2, q3, q4, q5\}$
- $AQ(q1) = \{A_2, A_3, A_5\}$
- $AQ(q2) = \{A_3, A_4\}$
- $AQ(q3) = \{A_1, A_3, A_5\}$
- $AQ(q4) = \{A_1, A_4\}$
- $AQ(q5) = \{A_2, A_5\}$

	A1	A4	A3	A5	A2
A1	16	12	4	4	0
A4	12	28	16	0	0
A3	4	16	28	12	8
A5	4	0	12	18	14
A2	0	0	8	14	14

- $TA = \{A_1, A_4, A_3, A_5\}$   $BA = \{A_2\}$
- $TQ = \{q2, q3, q4\}$   $BQ = \{\emptyset\}$
- $OQ = \{q1, q5\}$
- $CTQ = 2 \times 8 + 4 + 2 \times 6 = 32$
- $CBQ = 0$
- $COQ = 8 + 6 = 14$
- $Z = CTQ \times CBQ - COQ^2 = 32 \times 0 - 14^2 = -196$

	A1	A4	A3	A5	A2
A1	16	12	4	4	0
A4	12	28	16	0	0
A3	4	16	28	12	8
A5	4	0	12	18	14
A2	0	0	8	14	14

- $TA = \{A_1, A_4, A_3\}$   $BA = \{A_5, A_2\}$
- $TQ = \{q2, q4\}$   $BQ = \{q5\}$
- $OQ = \{q1, q3\}$
- $CTQ = 2 \times 8 + 2 \times 6 = 28$
- $CBQ = 6$
- $COQ = 8 + 4 = 12$
- $Z = CTQ \times CBQ - COQ^2 = 28 \times 6 - 12^2 = 24$

	<b>A1</b>	<b>A4</b>	<b>A3</b>	<b>A5</b>	<b>A2</b>
<b>A1</b>	16	12	4	4	0
<b>A4</b>	12	28	16	0	0
<b>A3</b>	4	16	28	12	8
<b>A5</b>	4	0	12	18	14
<b>A2</b>	0	0	8	14	14

- $TA = \{A_1, A_4\}$                        $BA = \{A_3, A_5, A_2\}$
- $TQ = \{q_4\}$                                $BQ = \{q_1, q_5\}$
- $OQ = \{q_2, q_3\}$
- $CTQ = 2 \times 6 = 12$
- $CBQ = 8 + 6 = 14$
- $COQ = 2 \times 8 + 4 = 20$
- $Z = CTQ \times CBQ - COQ^2 = 12 \times 14 - 20^2 = -232$

	<b>A1</b>	<b>A4</b>	<b>A3</b>	<b>A5</b>	<b>A2</b>
<b>A1</b>	16	12	4	4	0
<b>A4</b>	12	28	16	0	0
<b>A3</b>	4	16	28	12	8
<b>A5</b>	4	0	12	18	14
<b>A2</b>	0	0	8	14	14

- $TA = \{A_1\}$                                $BA = \{A_4, A_3, A_5, A_2\}$
- $TQ = \{\emptyset\}$                                $BQ = \{q_1, q_2, q_5\}$
- $OQ = \{q_3, q_4\}$
- $CTQ = 0$
- $CBQ = 8 + 2 \times 8 + 6 = 30$
- $COQ = 2 \times 6 + 4 = 16$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 30 - 16^2 = -256$

Áp dụng kết quả phân hoạch ma trận CA vào quan hệ NHAMAYDIEN, ta được các mảnh F: NHAMAYDIEN = {NMD1, NMD2}.

- Trong đó:  
 $NMD1 = \{A_1, A_4, A_3\}$   
 $NMD2 = \{A_1, A_2, A_5\}$
- Vì thế:  
 $NMD1 = \{MaNMD, CQQL, PhanLoai\}$   
 $NMD2 = \{MaNMD, TenNMD, CongSuat\}$
- MaNMD là thuộc tính khóa chính của quan hệ NHAMAYDIEN.

	<b>A1</b>	<b>A4</b>	<b>A3</b>	<b>A5</b>	<b>A2</b>
<b>A1</b>	16	12	4	4	0
<b>A4</b>	12	28	16	0	0
<b>A3</b>	4	16	28	12	8
<b>A5</b>	4	0	12	18	14
<b>A2</b>	0	0	8	14	14

## ĐỀ THI GIỮA KỲ - HỌC KỲ I 2020-2021 - Mã đề 03

Bộ Giao thông Vận tải hiện muốn quản lý thông tin các tuyến đường cao tốc tại Việt Nam. Tùy theo quy mô mà các tuyến đường cao tốc sẽ được Bộ quản lý khai thác (CapQL = “Trung Ương”) hay được giao về cho các Tỉnh, Thành phố trực tiếp quản lý (CapQL = “Tỉnh TP”). Thông tin của các tuyến đường cao tốc được thể hiện qua lược đồ quan hệ sau:

**DUONGCAOTOC (MaĐCT, TenĐCT, CapQL, SoLanXe, VTToiĐa)**

**Tân từ:** Tân từ: Mỗi tuyến đường cao tốc có các thông tin bao gồm: Mã đường cao tốc (MaĐCT), Tên đường cao tốc (TenĐCT), cấp quản lý (CapQL), quy mô số làn xe (SoLanXe), vận tốc tối đa được cho phép khi chạy trên đường cao tốc (VTToiĐa) được tính bằng đơn vị km/h.

1. (5 điểm) Cho thể hiện của quan hệ **DUONGCAOTOC** như sau:

MaĐCT	TenĐCT	CapQL	SoLanXe	VTToiĐa
CT01	Cầu Giẽ - Ninh Bình	Trung Ương	4	120
CT02	Hồ Chí Minh - Trung Lương	Trung Ương	4	120
CT03	Nội Bài - Cầu Nhật Tân	Tỉnh TP	6	80
CT04	HCM - Long Thành - Dầu Giây	Trung Ương	4	120
CT05	Mai Dịch - Thanh Trì	Tỉnh TP	4	100
CT06	Láng - Hoà Lạc	Tỉnh TP	6	100
CT07	Hà Nội - Hải Phòng	Trung Ương	6	120
CT08	Đà Lạt - Liên Khương	Tỉnh TP	4	80

Giả sử có ba ứng dụng truy xuất đến **DUONGCAOTOC**:

Q1: SELECT COUNT(\*) FROM **DUONGCAOTOC** WHERE CapQL = value

Q2: SELECT MaĐCT, TenĐCT FROM **DUONGCAOTOC** WHERE SoLanXe < 5

Q3: SELECT \* FROM **DUONGCAOTOC** WHERE VTToiĐa > 100

- (a) Dùng giải thuật COM\_MIN, tính Pr' thỏa tối thiểu và đầy đủ? (2 điểm)

**Lời giải.**

- Tập vị từ đơn giản dùng để phân hoạch **DUONGCAOTOC**:

p1: CapQL = “Trung Ương”

p2: CapQL = “Tỉnh TP”

p3: SoLanXe < 5

p4: SoLanXe ≥ 5

p5: VTToiĐa > 100

p6: VTToiĐa ≤ 100

- Khởi tạo  $P_r = \{p1, p2, p3, p4, p5, p6\}$ ,  $P_{r'} = \emptyset$

- Áp dụng thuật toán COM\_MIN:

i = 1 làm vị từ khởi đầu, vị từ p1 thỏa quy tắc 1,  $P_{r'} = \{p1\}$

i = 2: Ta có vị từ p2 không phân hoạch  $f_1$  (là mảnh giao tối thiểu tạo ra ứng với p1) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1\}$

i = 3: Vị từ p3 thỏa quy tắc 1,  $P_{r'} = \{p1, p3\}$

i = 4: Ta có vị từ p4 không phân hoạch  $f_3$  (là mảnh giao tối thiểu tạo ra ứng với p3) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$

i = 5: Ta có vị từ p5 không phân hoạch  $f_4$  (là mảnh giao tối thiểu tạo ra ứng với p4) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$

i = 6: Ta có vị từ p6 không phân hoạch  $f_5$  (là mảnh giao tối thiểu tạo ra ứng với p5) theo quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$

Kết luận:  $P_{r'} = \{p1, p3\}$  là tập vị từ đầy đủ và tối thiểu.

- (b) Sử dụng thuật toán PHORIZONTAL, thiết kế phân mảnh ngang chính cho quan hệ **DUONGCAOTOC**? Lưu ý, sinh viên không cần liệt kê dữ liệu trong các mảnh. (2 điểm)

**Lời giải.**

- Các vị từ giao tối thiểu là:
  - m1:  $p1 \wedge p3$ ;
  - m2:  $p1 \wedge \neg p3 = p1 \wedge p4$ ;
  - m3:  $\neg p1 \wedge p3 = p2 \wedge p3$ ;
  - m4:  $\neg p1 \wedge \neg p3 = p2 \wedge p4$ ;
- Vì vậy, ta có tập vị từ giao thooid thiểu là:  $M = \{m1, m2, m3, m4\}$
- Phân mảnh ngang chính quan hệ **DUONGCAOTOC** theo M là:
  - $DCT1 = \sigma_{m1}(DUONGCAOTOC)$
  - $DCT2 = \sigma_{m2}(DUONGCAOTOC)$
  - $DCT3 = \sigma_{m3}(DUONGCAOTOC)$
  - $DCT4 = \sigma_{m4}(DUONGCAOTOC)$
- Đối với dữ liệu demo, ta có các mảnh:
  - DCT1

MaĐCT	TenĐCT	CapQL	SoLanXe	VTToiĐa
CT01	Cầu Giẽ - Ninh Bình	Trung Ương	4	120
CT02	Hồ Chí Minh - Trung Lương	Trung Ương	4	120
CT04	HCM - Long Thành - Dầu Giây	Trung Ương	4	120

DCT2

MaĐCT	TenĐCT	CapQL	SoLanXe	VTToiĐa
CT07	Hà Nội - Hải Phòng	Trung Ương	6	120

DCT3

MaĐCT	TenĐCT	CapQL	SoLanXe	VTToiĐa
CT05	Mai Dịch - Thanh Trì	Tỉnh TP	4	100
CT08	Đà Lạt - Liên Khương	Tỉnh TP	4	80

DCT4

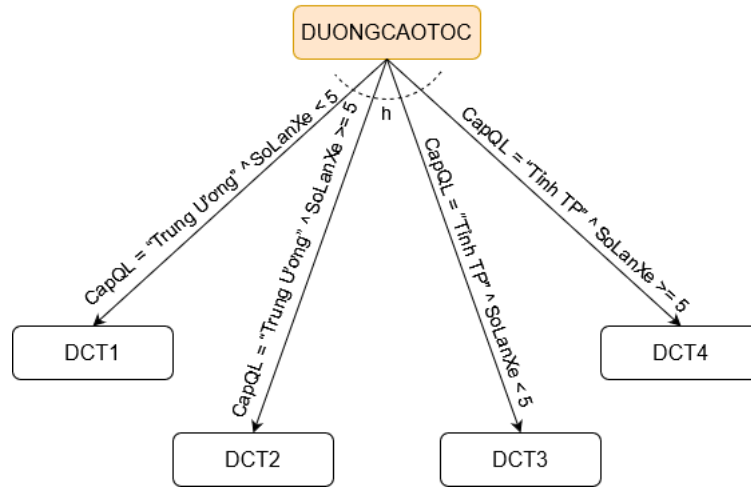
MaĐCT	TenĐCT	CapQL	SoLanXe	VTToiĐa
CT03	Nội Bài - Cầu Nhật Tân	Tỉnh TP	6	80
CT06	Láng - Hòa Lạc	Tỉnh TP	6	100

- (c) Hãy vẽ cây phân mảnh của lược đồ quan hệ **DUONGCAOTOC** đã làm. Chứng minh điều kiện **tái tạo** của phân mảnh này là đúng đắn. (1 điểm)

**Lời giải.**

- Chứng minh điều kiện **tái tạo**
  - Giữa các mảnh có thể áp dụng phép toán hợp để tái tạo lại mảnh ban đầu:
    - Ta có:  $DCT1 \cup DCT2 \cup DCT3 \cup DCT4 = DUONGCAOTOC$
    - $\implies$  Điều kiện tái tạo của phân mảnh này là đúng đắn.

- Vẽ cây phân mảnh



Hình 2: Cây phân mảnh

2. (5 điểm) Cho tập  $Q = \{q1, q2, q3, q4, q5\}$  các truy vấn, tập  $A = \{A1, A2, A3, A4, A5\}$  lần lượt là các thuộc tính **MaĐCT**, **TenĐCT**, **CapQL**, **SoLanXe**, **VTToiĐa** của quan hệ:

**DUONGCAOTOC** (**MaĐCT**, **TenĐCT**, **CapQL**, **SoLanXe**, **VTToiĐa**)

Tập  $S = \{S1, S2, S3, S4\}$  các vị trí (sites) trong hệ cơ sở dữ liệu phân tán. Giả sử số truy xuất đến các cặp thuộc tính tại các vị trí được cho theo hàm:

$$ref_i(q_j) = \begin{cases} 1, \forall i \in [1, 4], j \in \{1, 3, 5\} \\ 2, \forall i \in [1, 4], j \in \{2, 4\} \end{cases}, \text{ Với } i \text{ là số chỉ vị trí (site), } j \text{ là số chỉ của câu truy vấn.}$$

Ma trận truy vấn sử dụng thuộc tính (use) và ma trận tần số sử dụng truy vấn (acc) tại các vị trí như sau:

	A1	A2	A3	A4	A5
q1	1	0	1	1	0
q2	0	1	1	0	0
q3	1	0	0	1	0
q4	0	1	0	0	1
q5	1	0	1	0	0

Ma trận USE

	S1	S2	S3	S4
q1	4	0	2	0
q2	0	1	0	6
q3	2	5	0	0
q4	0	2	4	0
q5	1	0	0	2

Ma trận ACC

- (a) Tính ma trận AA. (1 điểm)

**Lời giải.**

- Ta có ma trận ái lực AA là:

	A1	A2	A3	A4	A5
A1	16	0	9	13	0
A2	0	26	14	0	12
A3	9	14	23	6	0
A4	13	0	6	13	0
A5	0	12	0	0	12

(b) Sử dụng thuật toán BEA tính ma trận CA. (2 điểm)

**Lời giải.**

- Ta có ma trận bond là:

	A1	A2	A3	A4	A5
A1		126	429	431	0
A2	126		686	84	456
A3	429	686		333	168
A4	431	84	333		0
A5	0	456	168	0	

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_3$  giữa các thuộc tính  $A_1, A_2$ .

$$\begin{aligned} Cont(A_0, A_3, A_1) &= 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 429 - 2 \times 0 = 858 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_3, A_2) &= 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2) \\ &= 2 \times 429 + 2 \times 868 - 2 \times 126 = 1978 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 686 + 2 \times 0 - 2 \times 0 = 1372 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_3$  vào giữa cột  $A_1$  và  $A_2$  của CA. Ta có:  $A_1, A_3, A_2$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_4$  giữa các thuộc tính  $A_1, A_3, A_2$ .

$$\begin{aligned} Cont(A_0, A_4, A_1) &= 2bond(A_0, A_4) + 2bond(A_4, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 431 - 2 \times 0 = 862 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_4, A_3) &= 2bond(A_1, A_4) + 2bond(A_4, A_3) - 2bond(A_1, A_3) \\ &= 2 \times 431 + 2 \times 333 - 2 \times 429 = 670 \end{aligned}$$

$$\begin{aligned} Cont(A_3, A_4, A_2) &= 2bond(A_3, A_4) + 2bond(A_4, A_2) - 2bond(A_3, A_2) \\ &= 2 \times 333 + 2 \times 84 - 2 \times 686 = -538 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_4, A_0) &= 2bond(A_2, A_4) + 2bond(A_4, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 84 + 2 \times 0 - 2 \times 0 = 168 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_4$  vào trước cột  $A_1$  và  $A_3$  của ma trận AC. Ta có:  $A_4, A_1, A_3, A_2$

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_5$  giữa các thuộc tính  $A_4, A_1, A_3, A_2$ .

$$\begin{aligned} Cont(A_0, A_5, A_4) &= 2bond(A_0, A_5) + 2bond(A_5, A_4) - 2bond(A_0, A_4) \\ &= 2 \times 0 + 2 \times 0 - 2 \times 0 = 0 \end{aligned}$$

$$\begin{aligned} Cont(A_4, A_5, A_1) &= 2bond(A_4, A_5) + 2bond(A_5, A_1) - 2bond(A_4, A_1) \\ &= 2 \times 0 + 2 \times 0 - 2 \times 431 = -862 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_5, A_3) &= 2bond(A_1, A_5) + 2bond(A_5, A_3) - 2bond(A_1, A_3) \\ &= 2 \times 0 + 2 \times 168 - 2 \times 429 = -522 \end{aligned}$$

$$\begin{aligned} Cont(A_3, A_5, A_2) &= 2bond(A_3, A_5) + 2bond(A_5, A_2) - 2bond(A_3, A_2) \\ &= 2 \times 168 + 2 \times 456 - 2 \times 686 = -124 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_5, A_0) &= 2bond(A_2, A_5) + 2bond(A_5, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 456 + 2 \times 0 - 2 \times 0 = 912 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_5$  vào sau cột  $A_2$  của ma trận AC. Ta có:  $A_4, A_1, A_3, A_2, A_5$

- Ta có ma trận ái lực tự CA là:

	<b>A4</b>	<b>A1</b>	<b>A3</b>	<b>A2</b>	<b>A5</b>
<b>A4</b>	13	13	6	0	0
<b>A1</b>	13	16	9	0	0
<b>A3</b>	6	9	23	14	0
<b>A2</b>	0	0	14	26	12
<b>A5</b>	0	0	0	12	12

- (c) Sử dụng thuật toán PARTITION để tìm ra hai phân mảnh dọc của quan hệ **DUONG-CAOTOC**. (2 điểm)

**Lời giải.**

- $Q = \{q1, q2, q3, q4, q5\}$
- $AQ(q1) = \{A_1, A_3, A_4\}$
- $AQ(q2) = \{A_2, A_3\}$
- $AQ(q3) = \{A_1, A_4\}$
- $AQ(q4) = \{A_2, A_5\}$
- $AQ(q5) = \{A_1, A_3\}$

	<b>A4</b>	<b>A1</b>	<b>A3</b>	<b>A2</b>	<b>A5</b>
<b>A4</b>	13	13	6	0	0
<b>A1</b>	13	16	9	0	0
<b>A3</b>	6	9	23	14	0
<b>A2</b>	0	0	14	26	12
<b>A5</b>	0	0	0	12	12

- $TA = \{A_4, A_1, A_3, A_2\}$                        $BA = \{A_5\}$
- $TQ = \{q1, q2, q3, q5\}$                        $BQ = \{\emptyset\}$
- $OQ = \{q4\}$
- $CTQ = 6 + 2 \times 7 + 7 + 3 = 30$
- $CBQ = 0$
- $COQ = 2 \times 6 = 12$
- $Z = CTQ \times CBQ - COQ^2 = 30 \times 0 - 12^2 = -144$

	<b>A4</b>	<b>A1</b>	<b>A3</b>	<b>A2</b>	<b>A5</b>
<b>A4</b>	13	13	6	0	0
<b>A1</b>	13	16	9	0	0
<b>A3</b>	6	9	23	14	0
<b>A2</b>	0	0	14	26	12
<b>A5</b>	0	0	0	12	12

- $TA = \{A_4, A_1, A_3\}$                        $BA = \{A_2, A_5\}$
- $TQ = \{q1, q3, q5\}$                        $BQ = \{q4\}$
- $OQ = \{q2\}$
- $CTQ = 6 + 7 + 3 = 16$
- $CBQ = 2 \times 6 = 12$
- $COQ = 2 \times 7 = 14$
- $Z = CTQ \times CBQ - COQ^2 = 16 \times 12 - 14^2 = -4$

	<b>A4</b>	<b>A1</b>	<b>A3</b>	<b>A2</b>	<b>A5</b>
<b>A4</b>	13	13	6	0	0
<b>A1</b>	13	16	9	0	0
<b>A3</b>	6	9	23	14	0
<b>A2</b>	0	0	14	26	12
<b>A5</b>	0	0	0	12	12

- $TA = \{A_4, A_1\}$                        $BA = \{A_3, A_2, A_5\}$
- $TQ = \{q3\}$                                $BQ = \{q2, q4\}$
- $OQ = \{q1, q5\}$
- $CTQ = 7$
- $CBQ = 2 \times 6 + 2 \times 7 = 26$
- $COQ = 6 + 3 = 9$
- $Z = CTQ \times CBQ - COQ^2 = 7 \times 26 - 9^2 = 101$

	<b>A4</b>	<b>A1</b>	<b>A3</b>	<b>A2</b>	<b>A5</b>
<b>A4</b>	13	13	6	0	0
<b>A1</b>	13	16	9	0	0
<b>A3</b>	6	9	23	14	0
<b>A2</b>	0	0	14	26	12
<b>A5</b>	0	0	0	12	12

- $TA = \{A_4\}$                                $BA = \{A_1, A_3, A_2, A_5\}$
- $TQ = \{\emptyset\}$                                $BQ = \{q2, q4, q5\}$
- $OQ = \{q1, q3\}$
- $CTQ = 0$
- $CBQ = 2 \times 6 + 2 \times 7 + 3 = 29$
- $COQ = 6 + 7 = 13$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 29 - 13^2 = -169$

Áp dụng kết quả phân hoạch ma trận CA vào quan hệ DUONGCAOTOC, ta được các mảnh F:  $DUONGCAOTOC = \{DCT1, DCT2\}$ .

- Trong đó:  
 $DCT1 = \{A_1, A_4\}$   
 $DCT2 = \{A_1, A_2, A_3, A_5\}$
- Vì thế:  
 $DCT1 = \{MaDCT, SoLanXe\}$   
 $DCT2 = \{MaDCT, TenDCT, CapQL, VTToiDa\}$
- MaDCT là thuộc tính khóa chính của quan hệ DUONGCAOTOC.

	<b>A4</b>	<b>A1</b>	<b>A3</b>	<b>A2</b>	<b>A5</b>
<b>A4</b>	13	13	6	0	0
<b>A1</b>	13	16	9	0	0
<b>A3</b>	6	9	23	14	0
<b>A2</b>	0	0	14	26	12
<b>A5</b>	0	0	0	12	12



## ĐỀ THI GIỮA KỲ - HỌC KỲ II 2020-2021 - Mã đề 01

Một doanh nghiệp chuyên phân phối các sản phẩm điện máy muốn quản lý thông tin các máy điều hòa hiện đang được kinh doanh. Tùy thuộc vào vị trí của kho hàng mà các máy điều hòa sẽ được chia thành hai nhóm, được lưu tại kho Quận 9 (Kho = “Quận 9”) hoặc được lưu tại kho Bình Tân (Kho = “Bình Tân”). Thông tin của các máy điều hòa được thể hiện qua lược đồ quan hệ sau:

**MAYDIEUHOA (MaMĐH, TenMĐH, Kho, CongSuat, DienTichSD, CNBienTan)**

**Tân từ:** Mỗi máy điều hòa có các thông tin bao gồm: Mã máy điều hòa (MaMĐH), tên máy điều hòa (TenMĐH), kho lưu trữ (Kho), công suất làm lạnh (CongSuat), diện tích sử dụng phù hợp (DienTichSD) và có sử dụng công nghệ biến tần hay không (CNBienTan).

1. (5 điểm) Cho thể hiện của quan hệ **MAYDIEUHOA** như sau:

MaMĐH	TenMĐH	Kho	CongSuat	DienTichSD	CNBienTan
M01	SS AR18TY	Quận 9	18000	>20	Có
M02	EL ESV18CR	Bình Tân	12000	15-20	Không
M03	SH AH-X18X	Quận 9	9000	<15	Có
M04	LG V18ENF1	Quận 9	12000	15-20	Không
M05	PN CH22NG	Bình Tân	9000	<15	Có
M06	AQ RT6VTH	Bình Tân	9000	<15	Không
M07	CP W2T6R	Quận 9	12000	15-20	Có
M08	DK Y7WVGF	Bình Tân	18000	>20	Không

Giả sử có ba ứng dụng truy xuất đến **MAYDIEUHOA**:

Q1: SELECT COUNT(\*) FROM **MAYDIEUHOA** WHERE **Kho** = value

Q2: SELECT **MaMĐH**, **TenMĐH** FROM **MAYDIEUHOA** WHERE **CongSuat** > 9000

Q3: SELECT \* FROM **MAYDIEUHOA** WHERE **DienTichSD** = “<15”

- (a) Dùng giải thuật COM\_MIN, tính Pr' thỏa tối thiểu và đầy đủ? (2 điểm)

**Lời giải.**

- Tập vị từ đơn giản dùng để phân hoạch quan hệ **MAYDIEUHOA** là:  
p1: Kho = “Quận 9”  
p2: Kho = “Bình Tân”  
p3: CongSuat > 9000  
p4: CongSuat ≤ 9000  
p5: DienTichSD = “<15”  
p6: DienTichSD ≠ “<15”
- Khởi tạo  $P_r = \{p1, p2, p3, p4, p5, p6\}$ ,  $P_{r'} = \{\emptyset\}$
- Áp dụng thuật toán COM\_MIN:
  - i = 1 làm vị từ khởi đầu, vị từ p1 thỏa quy tắc 1,  $P_{r'} = \{p1\}$
  - i = 2: Ta có vị từ p2 không phân hoạch  $f_1$  (là mảnh giao tối thiểu tạo ra ứng với p1) quy tắc 1. Vì vậy:  $P_{r'} = \{p1\}$
  - i = 3: Vị từ p3 thỏa quy tắc 1,  $P_{r'} = \{p1, p3\}$
  - i = 4: Ta có vị từ p4 không phân hoạch  $f_3$  (là mảnh giao tối thiểu tạo ra ứng với p3) quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$
  - i = 5: Ta có vị từ p5 không phân hoạch  $f_4$  (là mảnh giao tối thiểu tạo ra ứng với p4) quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$
  - i = 6: Ta có vị từ p6 không phân hoạch  $f_5$  (là mảnh giao tối thiểu tạo ra ứng với p5) quy tắc 1. Vì vậy:  $P_{r'} = \{p1, p3\}$
- Kết luận:  $P_{r'} = \{p1, p3\}$  là đầy đủ và tối thiểu.

- (b) Sử dụng thuật toán PHORIZONTAL, thiết kế phân mảnh ngang chính cho quan hệ **MAYDIEUHOA**? Lưu ý, sinh viên không cần liệt kê dữ liệu trong các mảnh. (2 điểm)

**Lời giải.**

- Các vị từ giao tối thiểu là:
  - m1:  $p1 \wedge p3$ ;
  - m2:  $p1 \wedge \neg p3 = p1 \wedge p4$ ;
  - m3:  $\neg p1 \wedge p3 = p2 \wedge p3$ ;
  - m4:  $\neg p1 \wedge \neg p3 = p2 \wedge p4$ ;
- Vì vậy, ta có tập vị từ giao tối thiểu là:  $M = \{m1, m2, m3, m4, m5\}$
- Phân mảnh ngang chính quan hệ MAYDIEUHOA theo M là:
  - $MDH1 = \sigma_{m1}(MAYDIEUHOA)$
  - $MDH2 = \sigma_{m2}(MAYDIEUHOA)$
  - $MDH3 = \sigma_{m3}(MAYDIEUHOA)$
  - $MDH4 = \sigma_{m4}(MAYDIEUHOA)$
- Đối với dữ liệu demo, ta có các mảnh: MDH1

MaMDH	TenMDH	Kho	CongSuat	DienTichSD	CNBientan
M01	SS AR18TY	Quận 9	18000	>20	Có
M04	LG V18ENF1	Quận 9	12000	15-20	Không
M07	CP W2T6R	Quận 9	12000	15-20	Có

MDH2

MaMDH	TenMDH	Kho	CongSuat	DienTichSD	CNBientan
M03	SH AH-X18X	Quận 9	9000	<20	Có

MDH3

MaMDH	TenMDH	Kho	CongSuat	DienTichSD	CNBientan
M02	EL ESV18CR	Bình Tân	12000	15-20	Không
M08	DK Y7WVGF	Bình Tân	18000	>20	Không

MDH4

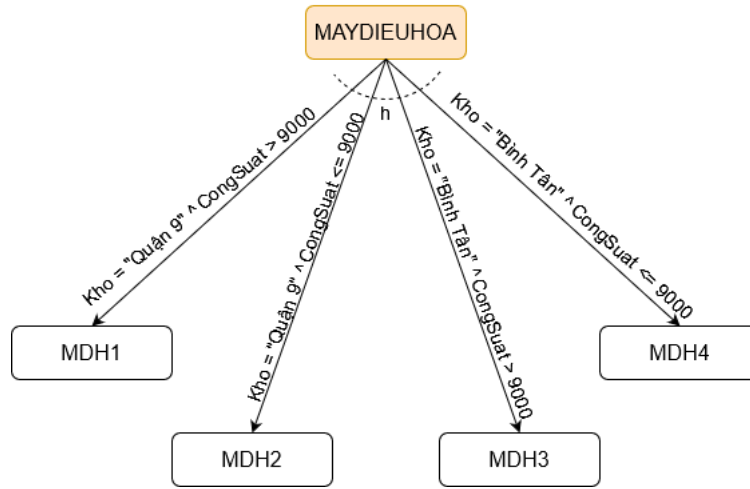
MaMDH	TenMDH	Kho	CongSuat	DienTichSD	CNBientan
M05	PN CH22NG	Bình Tân	9000	<15	Có
M06	AQ RT6VTH	Bình Tân	9000	<15	Không

- (c) Hãy vẽ cây phân mảnh của lược đồ quan hệ **MAYDIEUHOA** đã làm. Chứng minh điều kiện **tái tạo** của phân mảnh này là đúng đắn. (1 điểm)

**Lời giải.**

- Chứng minh điều kiện **tái tạo** Giữa các mảnh có thể áp dụng phép toán hợp để tái tạo lại mảnh ban đầu:  
Ta có:  $MDH1 \cup MDH2 \cup MDH3 \cup MDH4 = MAYDIEUHOA$   
 $\implies$  Điều kiện tái tạo của phân mảnh này là đúng đắn.

- Vẽ cây phân mảnh



Hình 3: Cây phân mảnh

2. (5 điểm) Cho tập  $Q = \{q1, q2, q3, q4, q5\}$  các truy vấn, tập  $A = \{A1, A2, A3, A4, A5\}$  lần lượt là các thuộc tính **TenMDH**, **Kho**, **CongSuat**, **DienTichSD**, **CNBienTan** của quan hệ:

**MAYDIEUHOA** (MaMDH, **TenMDH**, **Kho**, **CongSuat**, **DienTichSD**, **CNBienTan**)

Tập  $S = \{S1, S2, S3, S4\}$  các vị trí (sites) trong hệ cơ sở dữ liệu phân tán. Giả sử số truy xuất đến các cặp thuộc tính tại các vị trí được cho theo hàm:

$$ref_i(q_j) = \begin{cases} 1, \forall i \in [1, 4], j \in \{1, 3, 4\} \\ 2, \forall i \in [1, 4], j \in \{2, 5\} \end{cases}, \text{ Với } i \text{ là số chỉ vị trí (site), } j \text{ là số chỉ của câu truy vấn.}$$

Ma trận truy vấn sử dụng thuộc tính (use) và ma trận tần số sử dụng truy vấn (acc) tại các vị trí như sau:

	A1	A2	A3	A4	A5
q1	0	1	0	0	1
q2	0	0	1	1	0
q3	1	0	1	0	1
q4	0	1	0	1	0
q5	1	0	0	0	1

Ma trận USE

	S1	S2	S3	S4
q1	0	2	6	0
q2	0	3	0	4
q3	2	0	3	1
q4	0	5	1	0
q5	6	0	4	2

Ma trận ACC

- (a) Tính ma trận AA. (1 điểm)

**Lời giải.**

- Ta có ma trận ái lực AA là:

	A1	A2	A3	A4	A5
A1	30	0	6	0	30
A2	0	14	0	6	8
A3	6	0	20	14	6
A4	0	6	14	20	0
A5	30	8	6	0	38

(b) Sử dụng thuật toán BEA tính ma trận CA. (2 điểm)

**Lời giải.**

- Tính toán ma trận bond là:

	<b>A1</b>	<b>A2</b>	<b>A3</b>	<b>A4</b>	<b>A5</b>
<b>A1</b>		240	480	84	2076
<b>A2</b>	240		132	204	416
<b>A3</b>	480	132		560	528
<b>A4</b>	84	204	560		132
<b>A5</b>	2076	416	528	132	

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_3$  giữa các thuộc tính  $A_1, A_2$ .

$$\begin{aligned} Cont(A_0, A_3, A_1) &= 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1) \\ &= 2 \times 0 + 2 \times 480 - 2 \times 0 = 960 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_3, A_2) &= 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2) \\ &= 2 \times 480 + 2 \times 132 - 2 \times 240 = 744 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_3, A_0) &= 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 132 + 2 \times 0 - 2 \times 0 = 264 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_3$  vào trước cột  $A_1$  và  $A_2$  của CA. Ta có:  **$A_3, A_1, A_2$**

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_4$  giữa các thuộc tính  $A_3, A_1, A_2$ .

$$\begin{aligned} Cont(A_0, A_4, A_3) &= 2bond(A_0, A_4) + 2bond(A_4, A_3) - 2bond(A_0, A_3) \\ &= 2 \times 0 + 2 \times 560 - 2 \times 0 = 1120 \end{aligned}$$

$$\begin{aligned} Cont(A_3, A_4, A_1) &= 2bond(A_3, A_4) + 2bond(A_4, A_1) - 2bond(A_3, A_1) \\ &= 2 \times 560 + 2 \times 84 - 2 \times 480 = 328 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_4, A_2) &= 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2) \\ &= 2 \times 84 + 2 \times 204 - 2 \times 240 = 96 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_4, A_0) &= 2bond(A_2, A_4) + 2bond(A_4, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 204 + 2 \times 0 - 2 \times 0 = 408 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_4$  vào trước cột  $A_3$  và  $A_1$  của ma trận AC. Ta có:  **$A_4, A_3, A_1, A_2$**

- Xét ma trận AA, tính toán phần đóng góp khi di chuyển thuộc tính  $A_5$  giữa các thuộc tính  $A_4, A_3, A_1, A_2$ .

$$\begin{aligned} Cont(A_0, A_5, A_4) &= 2bond(A_0, A_5) + 2bond(A_5, A_4) - 2bond(A_0, A_4) \\ &= 2 \times 0 + 2 \times 132 - 2 \times 0 = 264 \end{aligned}$$

$$\begin{aligned} Cont(A_4, A_5, A_3) &= 2bond(A_4, A_5) + 2bond(A_5, A_3) - 2bond(A_4, A_3) \\ &= 2 \times 132 + 2 \times 528 - 2 \times 560 = 200 \end{aligned}$$

$$\begin{aligned} Cont(A_3, A_5, A_1) &= 2bond(A_3, A_5) + 2bond(A_5, A_1) - 2bond(A_3, A_1) \\ &= 2 \times 528 + 2 \times 2076 - 2 \times 480 = 4248 \end{aligned}$$

$$\begin{aligned} Cont(A_1, A_5, A_2) &= 2bond(A_1, A_5) + 2bond(A_5, A_2) - 2bond(A_1, A_2) \\ &= 2 \times 2076 + 2 \times 416 - 2 \times 240 = 4504 \end{aligned}$$

$$\begin{aligned} Cont(A_2, A_5, A_0) &= 2bond(A_2, A_5) + 2bond(A_5, A_0) - 2bond(A_2, A_0) \\ &= 2 \times 416 + 2 \times 0 - 2 \times 0 = 832 \end{aligned}$$

$\Rightarrow$  Chèn cột  $A_5$  vào giữa cột  $A_1$  và  $A_2$  của ma trận AC. Ta có:  **$A_4, A_3, A_1, A_5, A_2$**

- Ta có ma trận ái lực tự CA là:

	<b>A4</b>	<b>A3</b>	<b>A1</b>	<b>A5</b>	<b>A2</b>
<b>A4</b>	20	14	0	0	6
<b>A3</b>	14	20	6	6	0
<b>A1</b>	0	6	30	30	0
<b>A5</b>	0	6	30	38	8
<b>A2</b>	6	0	0	8	14

- (c) Sử dụng thuật toán PARTITION để tìm ra hai phân mảnh dọc của quan hệ **MAY-DIEUHOA**. (2 điểm)

**Lời giải.**

- $Q = \{q1, q2, q3, q4, q5\}$
- $AQ(q1) = \{A_2, A_5\}$
- $AQ(q2) = \{A_3, A_4\}$
- $AQ(q3) = \{A_1, A_3, A_5\}$
- $AQ(q4) = \{A_2, A_4\}$
- $AQ(q5) = \{A_1, A_5\}$

	<b>A4</b>	<b>A3</b>	<b>A1</b>	<b>A5</b>	<b>A2</b>
<b>A4</b>	20	14	0	0	6
<b>A3</b>	14	20	6	6	0
<b>A1</b>	0	6	30	30	0
<b>A5</b>	0	6	30	38	8
<b>A2</b>	6	0	0	8	14

- $TA = \{A_4, A_3, A_1, A_5\}$                        $BA = \{A_2\}$
- $TQ = \{q2, q3, q5\}$                        $BQ = \{\emptyset\}$
- $OQ = \{q1, q4\}$
- $CTQ = 2 \times 7 + 6 + 2 \times 12 = 44$
- $CBQ = 0$
- $COQ = 8 + 6 = 14$
- $Z = CTQ \times CBQ - COQ^2 = 44 \times 0 - 14^2 = -196$

	<b>A4</b>	<b>A3</b>	<b>A1</b>	<b>A5</b>	<b>A2</b>
<b>A4</b>	20	14	0	0	6
<b>A3</b>	14	20	6	6	0
<b>A1</b>	0	6	30	30	0
<b>A5</b>	0	6	30	38	8
<b>A2</b>	6	0	0	8	14

- $TA = \{A_4, A_3, A_1\}$                        $BA = \{A_5, A_2\}$
- $TQ = \{q2\}$                        $BQ = \{q1\}$
- $OQ = \{q3, q4, q5\}$
- $CTQ = 2 \times 7 = 14$
- $CBQ = 8$
- $COQ = 6 + 6 + 2 \times 12 = 36$
- $Z = CTQ \times CBQ - COQ^2 = 14 \times 8 - 36^2 = -1184$

	<b>A4</b>	<b>A3</b>	<b>A1</b>	<b>A5</b>	<b>A2</b>
<b>A4</b>	20	14	0	0	6
<b>A3</b>	14	20	6	6	0
<b>A1</b>	0	6	30	30	0
<b>A5</b>	0	6	30	38	8
<b>A2</b>	6	0	0	8	14

- $TA = \{A_4, A_3\}$   $BA = \{A_1, A_5, A_2\}$
- $TQ = \{q_2\}$   $BQ = \{q_1, q_5\}$
- $OQ = \{q_3, q_4\}$
- $CTQ = 2 \times 7 = 14$
- $CBQ = 8 + 2 \times 12 = 32$
- $COQ = 6 + 6 = 12$
- $Z = CTQ \times CBQ - COQ^2 = 14 \times 32 - 12^2 = 304$

	<b>A4</b>	<b>A3</b>	<b>A1</b>	<b>A5</b>	<b>A2</b>
<b>A4</b>	20	14	0	0	6
<b>A3</b>	14	20	6	6	0
<b>A1</b>	0	6	30	30	0
<b>A5</b>	0	6	30	38	8
<b>A2</b>	6	0	0	8	14

- $TA = \{A_4\}$   $BA = \{A_3, A_1, A_5, A_2\}$
- $TQ = \{\emptyset\}$   $BQ = \{q_1, q_3, q_5\}$
- $OQ = \{q_2, q_4\}$
- $CTQ = 0$
- $CBQ = 8 + 6 + 2 \times 12 = 38$
- $COQ = 2 \times 7 + 6 = 20$
- $Z = CTQ \times CBQ - COQ^2 = 0 \times 38 - 20^2 = -400$

Áp dụng kết quả phân hoạch ma trận CA vào quan hệ MAYDIEUHOA, ta được các mảnh F: MAYDIEUHOA = {MDH1, MDH2}.

- Trong đó:  
MDH1 = {A<sub>4</sub>, A<sub>3</sub>}  
MDH2 = {A<sub>1</sub>, A<sub>2</sub>, A<sub>5</sub>}
- Vì thế:  
MDH1 = {MaMDH, CongSuat, DienTichSD}  
MDH2 = {MaMDH, TenMDH, Kho, CNBienTan}
- MaMDH là thuộc tính khóa chính của quan hệ MAYDIEUHOA.

	<b>A4</b>	<b>A3</b>	<b>A1</b>	<b>A5</b>	<b>A2</b>
<b>A4</b>	20	14	0	0	6
<b>A3</b>	14	20	6	6	0
<b>A1</b>	0	6	30	30	0
<b>A5</b>	0	6	30	38	8
<b>A2</b>	6	0	0	8	14