Cơ Sở Dữ Liệu Phân Tán Bài tập về nhà - IS211.M21

Nguyễn Hồ Duy Tri, Nguyễn Thị Kim Yến Sinh viên: Pham Đức Thể - 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: $ref_i(q_i) = 1$; 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:

$$\begin{array}{lll} Aff(A_2,A_5) & = & acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) \\ & + & acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) \\ & = & 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 = 70 \\ \\ Aff(A_3,A_3) & = & acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) \\ & + & acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) \\ & = & 0 + 5 + 2 + 3 + 5 + 10 + 5 + 0 = 30 \\ \\ Aff(A_3,A_4) & = & 0 \\ Aff(A_3,A_5) & = & acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10 \\ \\ Aff(A_4,A_4) & = & acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) = 10 + 10 + 0 + 10 = 30 \\ Aff(A_4,A_5) & = & acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) = 10 + 10 + 0 + 10 = 30 \\ \\ Aff(A_5,A_5) & = & acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) \\ & + & acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) \\ & + & acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) \\ & + & acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) \\ & = & 0 + 5 + 2 + 3 + 10 + 10 + 0 + 10 \\ & + & 20 + 0 + 10 + 10 + 10 + 10 + 15 + 15 = 130 \\ \end{array}$$

• Ta có ma trận ái lực AA là:

	A 1	A2	A3	A4	A5
A1	10	0	10	0	10
A2	0	90	20	30	70
A3	10	20	30	0	10
A4	0	30	0	30	30
A5	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 .

$$Cont(A_0, A_3, A_1) = 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1)$$

$$Bond(A_0, A_3) = Aff(A_1, A_0) \times Aff(A_1, A_3) + Aff(A_2, A_0) \times Aff(A_2, A_3)$$

$$+ Aff(A_3, A_0) \times Aff(A_3, A_3) + Aff(A_4, A_0) \times Aff(A_4, A_3)$$

$$+ Aff(A_5, A_0) \times Aff(A_5, A_3)$$

$$= 0 \times 10 + 0 \times 20 + 0 \times 30 + 0 \times 0 + 0 \times 10 = 0$$

$$Bond(A_3, A_1) = Aff(A_1, A_3) \times Aff(A_1, A_1) + Aff(A_2, A_3) \times Aff(A_2, A_1)$$

$$+ Aff(A_3, A_3) \times Aff(A_3, A_1) + Aff(A_4, A_3) \times Aff(A_4, A_1)$$

$$+ Aff(A_5, A_3) \times Aff(A_5, A_1)$$

$$= 0 \times 10 + 20 \times 0 + 30 \times 10 + 0 \times 30 + 10 \times 10 = 500$$

$$Bond(A_0, A_1) = Aff(A_1, A_0) \times Aff(A_1, A_1) + Aff(A_2, A_0) \times Aff(A_2, A_1)$$

$$+ Aff(A_5, A_0) \times Aff(A_5, A_1)$$

$$= 0 \times 10 + 0 \times 0 + 0 \times 10 + 0 \times 0 + 0 \times 10 = 0$$

$$\implies Cont(A_0, A_3, A_1) = 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1)$$

$$= 2 \times 0 + 2 \times 500 - 2 \times 0 = 1000$$

```
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
\implies 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
    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
\implies 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
\implies Chèn cột A_3 vào sau cột A_1 và A_2 của CA. Ta có: \overline{A_1, A_2, A_3}
```

- \bullet 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 .

$$Cont(A_0, A_4, A_1) = 2bond(A_0, A_4) + 2bond(A_4, A_1) - 2bond(A_0, A_1)$$

$$Bond(A_0, A_4) = Aff(A_1, A_0) \times Aff(A_1, A_4) + Aff(A_2, A_0) \times Aff(A_2, A_4)$$

$$+ Aff(A_3, A_0) \times Aff(A_3, A_4) + Aff(A_4, A_0) \times Aff(A_4, A_4)$$

$$+ Aff(A_5, A_0) \times Aff(A_5, A_4)$$

$$= 0 \times 0 + 0 \times 30 + 0 \times 0 + 0 \times 30 + 0 \times 30 = 0$$

$$Bond(A_4, A_1) = Aff(A_1, A_4) \times Aff(A_1, A_1) + Aff(A_2, A_4) \times Aff(A_2, A_1)$$

$$+ Aff(A_3, A_4) \times Aff(A_3, A_1) + Aff(A_4, A_4) \times Aff(A_4, A_1)$$

$$+ Aff(A_5, A_4) \times Aff(A_5, A_1)$$

$$= 0 \times 10 + 30 \times 0 + 0 \times 10 + 30 \times 0 + 30 \times 10 = 300$$

```
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
\implies 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
    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
\implies 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
    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
\implies 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
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
```

$$Bond(A_4, A_3) = Aff(A_1, A_4) \times Aff(A_1, A_3) + Aff(A_2, A_4) \times Aff(A_2, A_3) \\ + Aff(A_3, A_4) \times Aff(A_3, A_3) + Aff(A_4, A_4) \times Aff(A_4, A_3) \\ + Aff(A_5, A_4) \times Aff(A_5, A_3) \\ = 0 \times 0 + 30 \times 90 + 0 \times 20 + 30 \times 30 + 30 \times 70 = 900 \\ Bond(A_2, A_3) = Aff(A_1, A_2) \times Aff(A_1, A_3) + Aff(A_2, A_2) \times Aff(A_2, A_3) \\ + Aff(A_3, A_2) \times Aff(A_3, A_3) + Aff(A_4, A_2) \times Aff(A_4, A_3) \\ + Aff(A_5, A_2) \times Aff(A_5, A_3) \\ = 0 \times 10 + 90 \times 20 + 20 \times 30 + 30 \times 0 + 70 \times 10 = 3100 \\ \Rightarrow Cont(A_2, A_4, A_3) = 2bond(A_2, A_4) + 2bond(A_4, A_3) - 2bond(A_2, A_3) \\ = 2 \times 5700 + 2 \times 900 - 2 \times 3100 = 7000 \\ \\ Cont(A_3, A_4, A_0) = 2bond(A_3, A_4) + 2bond(A_4, A_0) - 2bond(A_3, A_0) \\ Bond(A_3, A_4) = Aff(A_1, A_3) \times Aff(A_3, A_4) + Aff(A_2, A_3) \times Aff(A_2, A_4) \\ + Aff(A_5, A_3) \times Aff(A_3, A_4) + Aff(A_4, A_3) \times Aff(A_4, A_4) \\ + Aff(A_5, A_3) \times Aff(A_5, A_4) \\ = 10 \times 0 + 20 \times 30 + 30 \times 0 + 0 \times 30 + 10 \times 30 = 900 \\ Bond(A_4, A_0) = Aff(A_1, A_4) \times Aff(A_3, A_0) + Aff(A_2, A_4) \times Aff(A_2, A_0) \\ + Aff(A_5, A_4) \times Aff(A_5, A_0) \\ = 0 \times 0 + 30 \times 0 + 0 \times 0 + 30 \times 0 + 30 \times 0 = 0 \\ Bond(A_3, A_0) = Aff(A_1, A_3) \times Aff(A_3, A_0) + Aff(A_4, A_3) \times Aff(A_4, A_0) \\ + Aff(A_5, A_4) \times Aff(A_5, A_0) \\ = 10 \times 0 + 20 \times 0 + 30 \times 0 + 0 \times 0 + 30 \times 0 = 0 \\ Bond(A_3, A_0) = Aff(A_1, A_3) \times Aff(A_3, A_0) + Aff(A_4, A_3) \times Aff(A_4, A_0) \\ + Aff(A_5, A_3) \times Aff(A_5, A_0) \\ = 10 \times 0 + 20 \times 0 + 30 \times 0 + 0 \times 0 + 10 \times 0 = 0 \\ \Rightarrow Cont(A_2, A_3, A_0) = 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0) \\ = 2 \times 900 + 2 \times 0 - 2 \times 0 = 1800 \\$$

- \implies 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 .

$$Cont(A_{0}, A_{5}, A_{1}) = 2bond(A_{0}, A_{5}) + 2bond(A_{5}, A_{1}) - 2bond(A_{0}, A_{1})$$

$$Bond(A_{0}, A_{5}) = Aff(A_{1}, A_{0}) \times Aff(A_{1}, A_{5}) + Aff(A_{2}, A_{0}) \times Aff(A_{2}, A_{5})$$

$$+ Aff(A_{3}, A_{0}) \times Aff(A_{3}, A_{5}) + Aff(A_{4}, A_{0}) \times Aff(A_{4}, A_{5})$$

$$+ Aff(A_{5}, A_{0}) \times Aff(A_{5}, A_{5})$$

$$= 0 \times 10 + 0 \times 70 + 0 \times 10 + 0 \times 30 + 0 \times 130 = 0$$

$$Bond(A_{5}, A_{1}) = Aff(A_{1}, A_{5}) \times Aff(A_{1}, A_{1}) + Aff(A_{2}, A_{5}) \times Aff(A_{2}, A_{1})$$

$$+ Aff(A_{3}, A_{5}) \times Aff(A_{3}, A_{1}) + Aff(A_{4}, A_{5} \times Aff(A_{4}, A_{1})$$

$$+ Aff(A_{5}, A_{5}) \times Aff(A_{5}, A_{1})$$

$$= 10 \times 10 + 70 \times 0 + 10 \times 10 + 30 \times 0 + 130 \times 10 = 1500$$

$$Bond(A_{0}, A_{1}) = Aff(A_{1}, A_{0}) \times Aff(A_{1}, A_{1}) + Aff(A_{2}, A_{0}) \times Aff(A_{2}, A_{1})$$

$$+ Aff(A_{3}, A_{0}) \times Aff(A_{3}, A_{1}) + Aff(A_{4}, A_{0}) \times Aff(A_{4}, A_{1})$$

$$+ Aff(A_{5}, A_{0}) \times Aff(A_{5}, A_{1})$$

$$= 0 \times 10 + 0 \times 0 + 0 \times 10 + 0 \times 0 + 0 \times 10 = 0$$

$$\implies Cont(A_{0}, A_{5}, A_{1}) = 2bond(A_{0}, A_{5}) + 2bond(A_{5}, A_{1}) - 2bond(A_{0}, A_{1})$$

$$= 2 \times 0 + 2 \times 1500 - 2 \times 0 = 3000$$

```
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
\implies 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
    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
\implies 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
    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)
                           = \quad 0 \times 10 + 90 \times 20 + 20 \times 30 + 30 \times 0 + 70 \times 10 = 3100
\implies 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
```

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

 \implies Chèn cột A_5 vào giữa cột A_4 và A_2 của CA. Ta có: A_1, A_4, A_5A_2, A_3

• Ta có ma trận bond là:

	A1	A2	A3	A4	A5
A1		900	500	300	1500
A2	900		3100	5700	16500
A3	500	3100		900	3100
A4	300	5700	900		6900
A5	1500	16500	3100	6900	

• Ta có ma trận ái lực tụ CA là:

	A 1	A4	A5	A2	A3
A 1	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

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.

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

	A1	$\mathbf{A4}$	$\mathbf{A5}$	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 = $\{A_1, A_4, A_5, A_2\}$ BA = $\{A_3\}$
- $TQ = \{q3, q4, q5\}$ $BQ = \{\emptyset\}$
- $OQ = \{q1, q2\}$
- CBQ = 0
- COQ = $acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) + acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) = 0 + 5 + 2 + 3 + 5 + 10 + 5 + 0 = 30$
- $Z = CTQ \times CBQ COQ^2 = 120 \times 0 30^2 = -900$

	A 1	A4	A5	A2	A3
A1	10	0	10	0	10
A4	0	30	30	30	0
A 5	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 = $\{A_2, A_3\}$
- $TQ = \{q5\}$ $BQ = \{q2\}$
- $\bullet \ \mathrm{OQ} = \{q1, q3, q4\}$
- $CTQ = acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) = 10 + 10 + 15 + 15 = 50$
- CBQ = $acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) = 5 + 10 + 5 + 0 = 20$
- COQ = $acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) + acc_1(q_3) + acc_2(q_3) + acc_3(q_3) + acc_4(q_3) + acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) = 0 + 5 + 2 + 3 + 10 + 10 + 0 + 10 + 20 + 0 + 10 + 10 = 80$
- $Z = CTQ \times CBQ COQ^2 = 50 \times 20 80^2 = -5400$

	A 1	A4	A5	$\mathbf{A2}$	A3
A 1	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 = $acc_1(q_2) + acc_2(q_2) + acc_3(q_2) + acc_4(q_2) + acc_1(q_4) + acc_2(q_4) + acc_3(q_4) + acc_4(q_4) + acc_1(q_5) + acc_2(q_5) + acc_3(q_5) + acc_4(q_5) = 5 + 10 + 5 + 0 + 20 + 0 + 10 + 10 + 10 + 15 + 15 = 110$
- 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
A 1	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

- $COQ = acc_1(q_1) + acc_2(q_1) + acc_3(q_1) + acc_4(q_1) = 0 + 5 + 2 + 3 = 10$
- $\bullet \ \mathbf{Z} = CTQ \times CBQ COQ^2 = 0 \times 140 10^2 = -100$

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

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

	A 1	A4	A5	A2	A 3
A 1	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

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:

		$\mathbf{A1}$	$\mathbf{A2}$	A3	A4	$\mathbf{A5}$
A	1	46	10	36	0	8
A	2	10	23	5	13	0
A	3	36	5	41	5	8
A	4	0	13	5	13	0
A	5	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 .

$$Cont(A_0, A_3, A_1) = 2bond(A_0, A_3) + 2bond(A_3, A_1) - 2bond(A_0, A_1)$$

$$= 2 \times 0 + 2 \times 3246 - 2 \times 0 = 6492$$

$$Cont(A_1, A_3, A_2) = 2bond(A_1, A_3) + 2bond(A_3, A_2) - 2bond(A_1, A_2)$$

$$= 2 \times 3246 + 2 \times 745 - 2 \times 870 = 6242$$

$$Cont(A_2, A_3, A_0) = 2bond(A_2, A_3) + 2bond(A_3, A_0) - 2bond(A_2, A_0)$$

$$= 2 \times 745 + 2 \times 0 - 2 \times 0 = 1490$$

 \implies Chèn cột A_3 vào trước cột A_1 và A_2 của CA. Ta có: $\overline{A_3}, A_1, A_2$

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

$$Cont(A_0, A_4, A_3) = 2bond(A_0, A_4) + 2bond(A_4, A_3) - 2bond(A_0, A_3)$$

$$= 2 \times 0 + 2 \times 335 - 2 \times 0 = 670$$

$$Cont(A_3, A_4, A_1) = 2bond(A_3, A_4) + 2bond(A_4, A_1) - 2bond(A_3, A_1)$$

$$= 2 \times 335 + 2 \times 310 - 2 \times 3246 = -5202$$

$$Cont(A_1, A_4, A_2) = 2bond(A_1, A_4) + 2bond(A_4, A_2) - 2bond(A_1, A_2)$$

$$= 2 \times 310 + 2 \times 493 - 2 \times 870 = -134$$

$$Cont(A_2, A_4, A_0) = 2bond(A_2, A_4) + 2bond(A_4, A_0) - 2bond(A_2, A_0)$$

$$= 2 \times 493 + 2 \times 0 - 2 \times 0 = 986$$

- \implies 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 .

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

- \implies 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 tụ CA là:

	A5	A3	A 1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A 1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

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

Lời giải.

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

	A5	A3	A 1	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	A 1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A 1	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
A 1	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
- $COQ = 5 + 2 \times 14 + 2 \times 4 = 41$
- $\bullet \ \mathbf{Z} = CTQ \times CBQ COQ^2 = 0 \times 18 41^2 = -1681$

	A5	A 3	A 1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A 1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13

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

$$\begin{split} & SANBAY1 = \{ MaICAO,\, LoaiSB,\, LoaiDB,\, TenSB \} \\ & SANBAY2 = \{ MaICAO,\, NamTL,\, BayDem \} \end{split}$$

• MaICAO là thuộc tính khóa chính của quan hệ SANBAY.

	A5	A3	A 1	A2	A4
A5	8	8	8	0	0
A3	8	41	36	5	5
A 1	8	36	46	10	0
A2	0	5	10	23	13
A4	0	5	0	13	13