Tìm bao đóng

Ví dụ: Cho lược đồ quan hệ R(U, F)
Với U = ABCDE và F = { AB -->CD, E --> C, D -->CE, A -->E}. Tìm A+

- Đầu tiên gán A+=A
- Tiếp theo xét xem có PTH nào A->X không? nếu có bỏ X vào A+, ở đây ta có A->E nên A+=AE
- Ta thấy E->C nên A+=ACE
- Cuối cùng ta có A+=ACE

A+ = {AEC}

A+=AEC

Tìm tất cả các khóa trong lược đồ quan hệ

Các khái niệm cơ bản

- **Tập thuộc tính cốt lõi (K_{core})**: Các thuộc tính không xuất hiện ở vế phải của bất kỳ phụ thuộc hàm nào (FD). Những thuộc tính này phải có trong mọi khóa ứng viên.
- **Tập thuộc tính phụ trợ (A_{add})**: Gồm các thuộc tính xuất hiện cả ở vế trái và vế phải của các FD. Đây là các thuộc tính ta sẽ kết hợp với K_{core} để tìm khóa.

Thuật toán hoàn chỉnh

Bước 1: Phân loại thuộc tính

- Xác định tập hợp tất cả các thuộc tính của lược đồ quan hệ.
- Tạo tập K_{core} bằng cách tìm các thuộc tính không xuất hiện ở vế phải của bất kỳ phụ thuộc hàm nào.
- Tạo tập A_{add} bằng cách lấy hiệu của tập tất cả các thuộc tính trừ đi K_{core}.

Bước 2: Tìm các siêu khóa

- Khởi tạo một tập hợp rỗng để lưu trữ các siêu khóa được tìm thấy.
- Lần lượt lấy từng tập hợp con S_i của A_{add} (bao gồm cả tập rỗng).
- Với mỗi S_i, tính bao đóng của tập hợp ($K_{core} \cup S_i$).
- Nếu bao đóng này bằng tất cả các thuộc tính của lược đồ quan hệ, thì ($K_{core} \cup S_i$) là một **siêu khóa**. Lưu siêu khóa này vào một danh sách.

Bước 3: Tìm khóa ứng viên

- Kiểm tra tính tối thiểu của các siêu khóa đã tìm được ở Bước 2.
- Duyệt qua danh sách siêu khóa. Với mỗi siêu khóa SK_j, kiểm tra xem nó có chứa một siêu khóa khác SK_k hay không.
- Nếu SK_j chứa một siêu khóa khác ($SK_k \subset SK_j$), thì SK_j không phải là khóa ứng viên và bị loại bỏ.
- Các siêu khóa còn lại chính là **tất cả các khóa ứng viên** của lược đồ quan hệ.

Ví dụ minh họa

Lược đồ quan hệ: R(A,B,C,D)

Tập phụ thuộc hàm (FDs):

- AB→C
- C→A
- C→D

Bước 1: Phân loại thuộc tính

- **Các thuộc tính ở vế trái:** A,B,C
- **Các thuộc tính ở vế phải:** A,C,D
- Thuộc tính chỉ xuất hiện ở vế trái: B. Vậy, $\mathbf{K_{core}}$**={B}**.
- Các thuộc tính còn lại là A,C,D. Vậy, $\mathbf{A_{add}}$**={A,C,D}**.

Bước 2: Tìm các siêu khóa

- Tìm tất cả các tập con của Aadd={A,C,D}:
    - {},{A},{C},{D},{A,C},{A,D},{C,D},{A,C,D}
- Kết hợp từng tập con với Kcore={B} và tính bao đóng:
    - {B}+={B} → Không phải siêu khóa.
    - {B,A}+={A,B,C,D} → **Siêu khóa** {A,B}.
    - {B,C}+={B,C,A,D} → **Siêu khóa** {B,C}.
    - {B,D}+={B,D} → Không phải siêu khóa.
    - {B,A,C}+={A,B,C,D} → **Siêu khóa** {A,B,C}.
    - {B,A,D}+={A,B,C,D} → **Siêu khóa** {A,B,D}.
    - {B,C,D}+={B,C,A,D} → **Siêu khóa** {B,C,D}.
    - {A,B,C,D}+={A,B,C,D} → **Siêu khóa** {A,B,C,D}.

Bước 3: Tìm khóa ứng viên

- Danh sách siêu khóa đã tìm được: {A,B}, {B,C}, {A,B,C}, {A,B,D}, {B,C,D}, {A,B,C,D}.
- Bây giờ, loại bỏ các siêu khóa không tối thiểu (chứa một siêu khóa khác):
    - {A,B,C} chứa {A,B} và {B,C}. Loại bỏ.
    - {A,B,D} chứa {A,B}. Loại bỏ.
    - {B,C,D} chứa {B,C}. Loại bỏ.
    - {A,B,C,D} chứa tất cả các siêu khóa còn lại. Loại bỏ.
- Các siêu khóa còn lại là {A,B} và {B,C}.

Kết quả: Các khóa ứng viên của lược đồ quan hệ là **{A,B}** và **{B,C}**.

Thuật toán tìm phủ tối thiểu của một tập phụ thuộc hàm

Phủ tối thiểu của một tập phụ thuộc hàm F là một tập phụ thuộc hàm tương đương F+ nhưng thỏa mãn ba điều kiện sau:

1. **Vế phải đơn trị:** Mọi phụ thuộc hàm trong F+ chỉ có một thuộc tính ở vế phải.
2. **Không thuộc tính dư thừa vế trái:** Không thể loại bỏ bất kỳ thuộc tính nào ở vế trái của bất kỳ FD nào trong F+ mà vẫn giữ được tính tương đương.
3. **Loại bỏ phụ thuộc dư thừa:** Loại bỏ PTH dự thừa đến khi không thể loại bỏ bất kỳ FD nào trong F+ mà vẫn giữ được tính tương đương.

Thuật toán tìm phủ tối thiểu

Thuật toán này có ba bước chính, mỗi bước tương ứng với một điều kiện của phủ tối thiểu.

Bước 1: Đơn trị hóa vế phải

- Với mỗi FD có vế phải chứa nhiều hơn một thuộc tính (ví dụ: X→Y1Y2...Yn), ta phân tách nó thành nhiều FD có vế phải đơn trị (ví dụ: X→Y1,X→Y2,...,X→Yn).
- Kết quả thu được là một tập phụ thuộc hàm F1.

Bước 2: Loại bỏ thuộc tính dư thừa ở vế trái

- Lần lượt xét từng FD X→Y trong F1 mà vế trái X chứa nhiều hơn một thuộc tính.
- Lần lượt xét từng thuộc tính A trong X.
- Tính bao đóng của (X−A) với tập F1 hiện tại.
- Nếu bao đóng của (X−A) chứa Y (tức là (X−A)F1+ chứa Y), thì A là thuộc tính dư thừa. Ta loại bỏ A khỏi vế trái của FD đó.
- Lặp lại quá trình này cho đến khi không còn thuộc tính dư thừa nào ở vế trái của bất kỳ FD nào. Kết quả có F2

Bước 3: Loại bỏ phụ thuộc dư thừa

- Lần lượt xét từng FD f trong F2.
- Giả sử ta loại bỏ f khỏi F2 để tạo thành tập F2′.
- Sử dụng thuật toán bao đóng (closure), kiểm tra xem f có được suy ra từ F2′ hay không. Tức là kiểm tra (X)F2′+ có chứa Y hay không (nếu f là X→Y).
- Nếu f có thể được suy ra từ F2′, thì f là một phụ thuộc dư thừa và ta loại bỏ nó vĩnh viễn khỏi tập FĐ.
- Lặp lại quá trình này cho đến khi không còn FĐ dư thừa nào có thể bị loại bỏ.

Kết quả thu được chính là **phủ tối thiểu** của tập phụ thuộc hàm ban đầu.

Ví dụ cụ thể

Cho lược đồ quan hệ R(A,B,C,D,E) và tập phụ thuộc hàm F:

1. A→BC
2. B→E
3. AD→E
4. AC→B

Bước 1: Đơn trị hóa vế phải

- FD 1 có vế phải là BC, ta phân tách thành hai FD:
  - A→B
  - A→C
- Các FD còn lại đã có vế phải đơn trị.
- Tập phụ thuộc hàm sau Bước 1, gọi là F1:
  - A→B
  - A→C
  - B→E
  - AD→E
  - AC→B
-

Bước 2: Loại bỏ thuộc tính dư thừa ở vế trái

- Xét các FD có vế trái chứa nhiều hơn một thuộc tính. Ta có hai FD: AD→E và AC→B.
  - **Xét AD→E**:
    - Kiểm tra thuộc tính A: Ta cần tính bao đóng của $(AD-A)F1^{+}=(D)F1^{+}=\{D\}$. Bao đóng này không chứa E. Vậy A không dư thừa.
    - Kiểm tra thuộc tính D: Ta cần tính bao đóng của $(AD-D)F2^{+}=(A)F1^{+}$.
      - $(A)F2^{+}$ bắt đầu với {A}. Từ A→C, ta có {A,C}. Từ AC→B, ta có {A,C,B}. Từ B→E, ta có {A,C,B,E}.
      - Bao đóng $(A)F1^{+}$ chứa E. Vậy, thuộc tính D là dư thừa. Ta loại bỏ D khỏi vế trái của FD này.
  - FD AD→E trở thành A→E.
- Tập phụ thuộc hàm sau loại bỏ D khỏi vế trái của FD, gọi là F3:
  - A→B
  - A→C
  - B→E
  - A→E
  - AC→B

- **Xét AC→B**:
  - Kiểm tra thuộc tính A: Tính bao đóng của (AC−A)F3+=(C)F3+={C}. Bao đóng này không chứa B. Vậy A không dư thừa.
  - Kiểm tra thuộc tính C: Tính bao đóng của (AC−C)F3+=(A)F3+.
    - (A)F3+={A,C,B,E} (đã tính ở trên).
    - Bao đóng (A)F2+ chứa B. Vậy, thuộc tính C là dư thừa. Ta loại bỏ C khỏi vế trái của FD này.
- FD AC→B trở thành A→B.

- Tập phụ thuộc hàm sau loại bỏ C khỏi vế trái của FD, gọi là F3:
  - A→B
  - A→C
  - B→E
  - A→E
  - A→B
- Có 2 FT A→B nên loại bỏ bới 1 A→B, gọi là F4
  - A→B
  - A→C
  - B→E
  - A→E

Bước 2: Loại bỏ phụ thuộc dư thừa

- Xét từng FD trong F4 xem có dư thừa không.
  - **Xét A→B**: Loại bỏ A→B, tập còn lại là F4′={A→C,B→E,A→E }.
  - Ta cần kiểm tra xem A→B có được suy ra từ F4′ không.
  - Tính bao đóng của A với F4′, (A)F4′+:
    - Bắt đầu với {A}.
    - Từ A→C, ta có {A,C}.
    - Từ A→E, ta có {A,C,E}.
    - Ta thấy bao đóng (A)F4′+ không chứa B. Vậy, A→B là FD không dư thừa.
  - **Xét A→C**: Loại bỏ A→C, tập còn lại là F4′={A→B,B→E,A→E }.
  - Ta cần kiểm tra xem A→C có được suy ra từ F4′ không.
  - Tính bao đóng của A với F4′, (A)F4′+:
    - Bắt đầu với {A}.

        - Từ A→B, ta có {A,B}.
        - Từ B→E, ta có {A,B,E}.
        - Ta thấy bao đóng (A)F4'+ không chứa C. Vậy, A→C là FD không dư thừa.
    - **Xét B→E**: Loại bỏ A→C, tập còn lại là F4'={A→B,A→C,A→E }.
    - Ta cần kiểm tra xem B→E có được suy ra từ F4' không.
    - Tính bao đóng của B với F4', (B)F4'+:
        - Bắt đầu với {B}.
        - Ta thấy bao đóng (B)F4'+ không chứa E. Vậy, B→E là FD không dư thừa.
    - **Xét A→E**: Loại bỏ A→E, tập còn lại là F4'={ A→B,A→C,B→E, }.
    - Ta cần kiểm tra xem A→E có được suy ra từ F4' không.
    - Tính bao đóng của A với F4', (A)F4'+:
        - Bắt đầu với {A}.
        - Từ A→B, ta có {A,B}.
        - Từ A→C, ta có {A,B,C}.
        - Từ B→E, ta có {A,B,C,E}.
        - Ta thấy bao đóng (A)F4'+ có chứa E. Vậy, A→E là FD dư thừa. (loại bỏ **A→E)**

Kết quả cuối cùng

Sau khi hoàn thành tất cả các bước, phủ tối thiểu của tập FD ban đầu là tập sau:

1. A→B
2. A→C
3. B→E