



Chương 3: PHƯƠNG PHÁP CHUẨN HOÁ LỖCSDL

3.1 DẪN NHẬP:

3.2 CÁC TIÊU CHUẨN CỦA QUÁ TRÌNH CHUẨN HOÁ:

3.3 QUAN ĐIỂM BẢO TOÀN PHỤ THUỘC HÀM:

3.4 QUAN ĐIỂM BẢO TOÀN THÔNG TIN:

3.5 QUAN ĐIỂM BIỂU DIỄN TRỌN VẸN:

3.6 HAI PHƯƠNG PHÁP CHUẨN HÓA MỘT LỖCSDL:





3.1 Dẫn nhập:

- Xuất phát từ giai đoạn phân tích nhu cầu, ta có thể có 1 trong 2 kết quả sau:
 - (i) Dựa trên kinh nghiệm, chúng ta có thể đề nghị một cấu trúc CSDL ban đầu gồm các quan hệ con Q_i cùng các phụ thuộc dữ liệu F_{Q_i} định nghĩa trên các quan hệ con.
 - $C = (\langle Q_i, F_{Q_i} \rangle)_{i=1..n}$
 - (ii) Hoặc chỉ có một quan hệ phổ quát duy nhất Q_0 chứa tất cả các thuộc tính cần được lưu trữ và tập các phụ thuộc F_Q tìm được.
 - $C_0 = \langle Q_0, F_Q \rangle$
- Chúng ta cần kiểm tra và chuẩn hoá các kết quả đầu tiên này, dựa trên một số tiêu chuẩn thiết kế, để có được một cấu trúc quan niệm CSDL được đánh giá tốt hơn, phù hợp hơn với các yêu cầu của môi trường ứng dụng.





3.2 Các Tiêu chuẩn của quá trình chuẩn hoá (1):

- Hầu hết các công trình nghiên cứu về thiết kế CSDL đều thỏa thuận rằng 2 tiêu chuẩn quan trọng cần đạt được qua quá trình chuẩn hóa một CSDL ở mức quan niệm là:
 - (i) CSDL kết quả cần đạt **dạng chuẩn cao nhất**
 - (ii) CSDL kết quả phải **tương đương** với CSDL phân tích **lúc ban đầu**.





3.2 Các Tiêu chuẩn của quá trình chuẩn hoá (2):

1. Tiêu chuẩn dạng chuẩn được đề ra nhằm đáp ứng 2 yêu cầu cụ thể:

- **Cập nhật:** Hạn chế tối đa sự trùng lặp thông tin trong CSDL, do đó sẽ giảm bớt tình huống thông tin bị mâu thuẫn sau những lần cập nhật CSDL.
- **Kiểm tra RBTV:** Tạo điều kiện thuận lợi cho việc kiểm tra RBTV ở dạng phụ thuộc dữ liệu dựa trên **cơ chế khoá** sẵn có bên trong các phần mềm quản trị CSDL.





3.2 Các Tiêu chuẩn của quá trình chuẩn hoá (3):

2. Tiêu chuẩn tương đương:

- Nhằm đáp ứng yêu cầu truy xuất dữ liệu. Với tiêu chuẩn này các **thông tin lưu trữ CSDL ban đầu** đều phải được tìm thấy **đầy đủ trong CSDL kết quả**.
- Có 3 quan niệm khác nhau về tiêu chuẩn tương đương:





3.3 Quan điểm bảo toàn phụ thuộc hàm(1):

- Quan điểm này cho rằng các thông tin được lưu trong CSDL là những thông tin được thể hiện thông qua các phụ thuộc dữ liệu. Do đó cần phải bảo toàn phụ thuộc hàm trong khi biến đổi.
- Tiêu chuẩn tương đương theo quan điểm bảo toàn phụ thuộc hàm được đề ra như sau:
 - Giả sử, $C_1 = \langle Q, F \rangle$ và $C_2 = \{ \langle Q_i, F_i \rangle \}_{i=1..n}$ là một biến đổi từ C_1
 - $C_1 \equiv C_2$ nếu hai điều kiện sau được thỏa:
 - (i.1) $\bigcup_{i=1}^n Q_i^+ = Q^+$ (không được sót thuộc tính)
 - (i.2) $(\bigcup_{i=1}^n F_i)^+ = F^+$. (bảo toàn PTH)





3.3 Quan điểm bảo toàn phụ thuộc hàm(2):

- **Phương pháp Chứng minh Phân rã bảo toàn PTH:**

- Để Chứng minh $(\cup F_i)^+ = F^+$ ta đặt $F' = (\cup F_i)$
- Và chứng minh: $\forall f' \in (F' \setminus F)$ thì $f' \in F^+$ và $\forall f \in (F \setminus F')$ thì $f \in F'^+$

- **Ví dụ:** Cho $Q(ABCD)$ và $F = \{ A \rightarrow C; C \rightarrow A; D \rightarrow C; BD \rightarrow A \}$

- Xét phân rã $Q1(AB); Q2(ACD); Q3(BCD)$
 - a) Xác định tập phụ thuộc hàm chiếu trên từng quan hệ
 - b) Kiểm tra tính bảo toàn phụ thuộc hàm của phân rã trên





3.4 Quan điểm bảo toàn thông tin(1):

- Quan điểm này cho rằng các thông tin lưu trữ trong **CSDL ban đầu** đều phải được tìm thấy **đầy đủ** trong **CSDL kết quả**.
- Tiêu chuẩn tương đương theo quan điểm bảo toàn thông tin được đề ra như sau:
 - Giả sử, $C_1 = \langle Q, F \rangle$ và $C_2 = \{ \langle Q_i, F_i \rangle \}_{i=1,n}$ là một biến đổi từ C_1
 - $C_1 \equiv C_2$ nếu hai điều kiện sau được thỏa:
 - (i.1) $\bigcup_{i=1}^n Q_i^+ = Q^+$ (không được sót thuộc tính)
 - (i.2) $\bigcap_{i=1}^n Q[Q_i^+] = Q$. (bảo toàn thông tin lưu trữ)





3.4 Quan điểm bảo toàn thông tin(2):

- Phương pháp kiểm tra tính chất bảo toàn thông tin của một phân rã:
 - Cho $C = \{Q_i\}$ là 1 phân rã của ldqh Q có tập pth F_{Q_i} .
 - **b1: Xây dựng 1 bảng 2 chiều** mà các cột là các thuộc tính của Q , mỗi dòng là một Q_i trong phân rã nhận được.
 - Mỗi ô ở dòng i cột j chứa ký hiệu:
 - a) aj nếu Q_i có chứa thuộc tính thứ j của Q
 - b) bk nếu ngược lại (trong đó k là số thứ tự xuất hiện b)





3.4 Quan điểm bảo toàn thông tin(3):

- **b2: Biến đổi bảng dựa trên các pth có trong F_Q theo qui tắc sau:**
 - Xét một pth $f : X \rightarrow Y \in F_Q$.
 - Chọn 2 dòng Q_i, Q_j sao cho: $Q_i.X = Q_j.X$
 - Nếu $Q_i.Y \neq Q_j.Y$ thì thực hiện **thay thế** trên Q_i và Q_j ở từng cột A_k thuộc Y theo các trường hợp sau:
 - Nếu cả 2 ô (i,k) và (j,k) **đều không chứa a_k** thì ta **không thay đổi**
 - Ngược lại nếu có **1 ô chứa a_k** thì thay ô kia bằng ký hiệu a_k .

