

Các dạng chuẩn của CSDL quan hệ

Lê Thành Văn

05-12-2023

Khoa Hệ thống thông tin quản lý

Giới thiệu

Giới thiệu

Định nghĩa

Chuẩn hóa cơ sở dữ liệu (csdl) là quá trình tách bảng dữ liệu nhằm:

- giảm thiểu việc dư thừa dữ liệu, và
- hạn chế lỗi có thể phát sinh trong quá trình thay đổi dữ liệu.

Việc chuẩn hóa csdl dựa trên các dạng chuẩn (normal form) được Edgar F. Codd đề ra trong mô hình quan hệ của mình.

Giới thiệu

Các loại lỗi

Lỗi khi thay đổi dữ liệu (thêm, sửa hoặc xóa) là dạng lỗi khi dữ liệu được thay đổi không tương thích với cấu trúc hoặc dữ liệu hiện có.

Giả sử thông tin giảng viên cần bao gồm mã môn học họ giảng dạy. Tuy nhiên, một giảng viên mới có thể chưa có môn, dẫn đến không thể thêm thông tin của họ.

Faculty and Their Courses

Faculty ID	Faculty Name	Faculty Hire Date	Course Code
389	Dr. Giddens	10-Feb-1985	ENG-206
407	Dr. Saperstein	19-Apr-1999	CMP-101
407	Dr. Saperstein	19-Apr-1999	CMP-201

424	Dr. Newsome	29-Mar-2007	?
-----	-------------	-------------	---

Hình 1: Lỗi khi thêm dữ liệu

Một thông tin có thể xuất hiện nhiều lần trong một bảng, nên khi thay đổi có thể bị sót, dẫn đến thông tin không nhất quán trong csdl.

Employees' Skills

Employee ID	Employee Address	Skill
426	87 Sycamore Grove	Typing
426	87 Sycamore Grove	Shorthand
519	94 Chestnut Street	Public Speaking
519	96 Walnut Avenue	Carpentry

Hình 2: Lỗi khi sửa dữ liệu

Lỗi khi xóa dữ liệu là dạng lỗi mà khi xóa một thông tin này có thể xóa những thông tin (không cần xóa) khác.

Faculty and Their Courses

Faculty ID	Faculty Name	Faculty Hire Date	Course Code
389	Dr. Giddens	10-Feb-1985	ENG-206
407	Dr. Saperstein	19-Apr-1999	CMP-101
407	Dr. Saperstein	19-Apr-1999	CMP-201



DELETE

Hình 3: Lỗi khi xóa dữ liệu

Giới thiệu

Các dạng chuẩn

Các dạng chuẩn thường gặp bao gồm (từ thấp đến cao):

- Dạng chuẩn 1 (First Normal Form, 1NF).
- Dạng chuẩn 2 (Second Normal Form, 2NF).
- Dạng chuẩn 3 (Third Normal Form, 3NF).
- Dạng chuẩn Boyce-Codd (Boyce-Codd Normal Form, BCNF, 3.5NF).

Việc đạt được một dạng chuẩn cao có nghĩa là phải thỏa mãn điều kiện của các dạng chuẩn thấp hơn nó.
Tức là, không thể đạt được dạng chuẩn 3 nếu như không đạt được dạng chuẩn 1 hoặc dạng chuẩn 2.

Một csdl quan hệ được gọi là **đã chuẩn hóa** nếu như nó đạt được dạng chuẩn 3.