

SƠ ĐỒ LƯỢC ĐỒ CSDL TRUONGHOC

I. CẤU TRÚC CHUNG, KHÓA CHÍNH VÀ KHÓA NGOẠI.

1. Nhận xét

Dựa trên 5 file Excel, em thấy mỗi trường học đều có các thông tin: mã trường, tên trường, địa chỉ, sở, phòng (trừ `GDTX` và `THPT`), loại hình, loại trường.

2. Khóa chính và khóa ngoại

- Do có sự trùng lặp về dữ liệu ở các mục *sở, phòng, loại hình, loại trường* nên em đã tách các dữ liệu trên thành các bảng gồm ID và tên của dữ liệu và đặt ID làm *primary key*.
- Ở bảng *PHONG*, vì với mỗi sở sẽ có các phòng, nên em thêm 1 field *IDSO*. Do đó bảng *PHONG* sẽ gồm:
 - 2 *primary key* là *IDPHG* và *IDSO*.
 - 1 *foreign key* là *IDSO*.
- Ở bảng *TRUONG*, do em gộp cả 5 files Excel vào chung một bảng *TRUONG*, nên để phân biệt các trường học em thêm 1 field là *IDCAP*, và thêm 1 bảng *CAP* gồm 5 cấp GDTX, MN, TH, THCS và THPT. Như vậy, bảng *TRUONG* sẽ có gồm:
 - 2 *primary key* là *MATRG* và *IDCAP*.
 - 5 *foreign key* là *IDSO*, *IDPHG*, *IDLH*, *IDLT*, *IDCAP*.

II. Ràng buộc

- Đối với 5 bảng *LOAIHINH*, *LOAITRUONG*, *CAP*, *SO*, *PHONG*: em đặt ràng buộc là NOT NULL cho tất cả các field (*primary key* bắt buộc là NOT NULL). Vì các bảng này tạo ra từ file Excel và một đối tượng thì phải có tên của nó.
- Đối với bảng *TRUONG*, các ràng buộc như sau:
 - MATRG*, *TENTRG*, *IDCAP*: NOT NULL (giải thích như ý ở trên)
 - DCHI*, *IDSO*, *IDPHG*, *IDLH*, *IDLT*: DEFAULT NULL. Các trường này không bắt buộc phải có thông tin nên mặc định là NULL. Ở cấp GDTX và THPT không có 'phòng' nên ý nghĩa của NULL là: không tồn tại.

III. NHẬN XÉT

Ưu điểm của cách chia này:

- Việc sửa đổi các thông tin bị lặp lại và gán ID giúp dễ dàng tìm kiếm, và có tính nhất quán (giả sử 'Tư thực' có nhiều cách viết từ nhiều người dùng như 'Tu thuc', 'TƯ THỰC',... thì việc đặt *IDLH* = 2, *TENLH* = 'Tư thực' sẽ dễ quản lý hơn).
- Các trường học có 1 cấu trúc chung, và không bị lặp lại cấu trúc bảng. Nếu muốn lấy thông tin của các trường học cấp Tiểu học thì chỉ cần truy vấn điều kiện *IDCAP* = 'TH'.
- Việc tách dữ liệu thành nhiều bảng giúp cho dữ liệu có tính độc lập (giả sử *TENLT* = 'Dân tộc bán trú' bị viết nhầm thành 'Dân tộc nội trú' thì cũng có thể được chỉnh sửa dễ dàng mà không ảnh hưởng nhiều đến dữ liệu khác).