

BÀI 3**QUAN HỆ GIỮA CÁC BẢNG VÀ KHOÁ NGOÀI TRONG CƠ SỞ DỮ LIỆU QUAN HỆ***Học xong bài này, em sẽ:*

- ✓ Diễn đạt được khái niệm khoá ngoài của một bảng và mối liên kết giữa các bảng. Giải thích được các khái niệm đó qua ví dụ minh hoạ.
- ✓ Giải thích được ràng buộc khoá ngoài là gì.
- ✓ Biết được các phần mềm quản trị CSDL có cơ chế kiểm soát các cập nhật dữ liệu để đảm bảo ràng buộc khoá ngoài.



Để quản lý sách, người đọc và việc mượn/trả sách của một thư viện (TV) trường học, bạn Anh Thư dự định chỉ dùng một bảng như mẫu ở Hình 1. Theo em, trong trường hợp cụ thể này, việc đưa tất cả dữ liệu cần quản lý vào trong một bảng như Anh Thư thực hiện có ưu điểm và nhược điểm gì?

Số thẻ TV	Họ và tên	Ngày sinh	Lớp	Mã sách	Tên sách	Số trang	Tác giả	Ngày mượn	Ngày trả
...

Hình 1. Cấu trúc bảng SÁCH-NGƯỜI ĐỌC-MƯỢN-TRẢ

Gợi ý: Xét một số trường hợp sau:

- 1) Một học sinh mượn sách nhiều lần, mỗi lần mượn nhiều quyển sách.
- 2) Cần bổ sung dữ liệu về số sách mới mua của thư viện.

1. Tình dư thừa dữ liệu

a) Dư thừa dữ liệu có thể dẫn đến dữ liệu không nhất quán khi cập nhật

Có thể một số người nghĩ rằng nên đưa tất cả dữ liệu cần lưu trữ vào trong một bảng vì khi cần tìm thông tin nào đó thì chỉ phải tìm trong một bảng. Nhưng thực tế cho thấy đa số bài toán quản lý cần dùng nhiều hơn một bảng dữ liệu. Nếu chỉ dùng một bảng thì rất có thể dẫn đến tình trạng dư thừa dữ liệu. Ví dụ, trường hợp nêu ở phần khởi động, giả sử học sinh có số thẻ TV “HS-002” tên là “Lê Bình” sinh ngày “02/3/2007”, học lớp “11A1” đã có 68 lần mượn sách. Như vậy, bộ giá trị (“HS-002”, “Lê Bình”, “02/3/2007”, “11A1”) phải xuất hiện 68 lần (trên 68 bản ghi của bảng). Tình trạng dư thừa dữ liệu có thể dẫn đến sai nhầm, không nhất quán về dữ liệu. Việc gõ nhập 68 lần

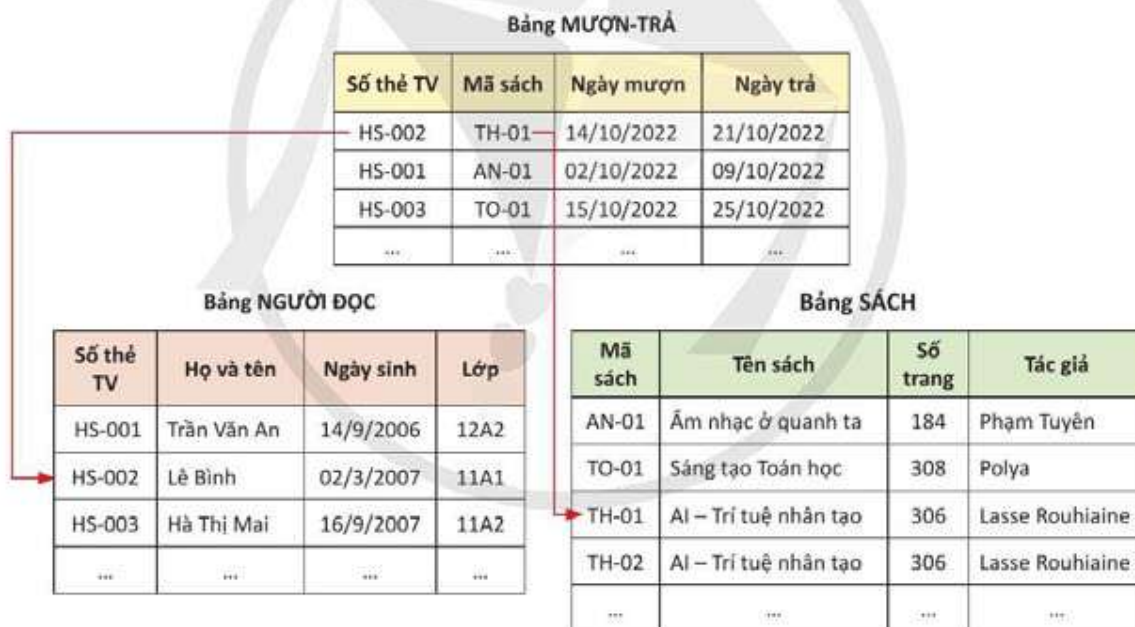
bộ dữ liệu về Lê Bình sẽ dễ xuất hiện sai nhầm hơn so với 68 lần chỉ gõ *Số thẻ TV* của Lê Bình vào bảng. Giải pháp tránh dư thừa là có thể dùng một bảng khác chỉ chứa dữ liệu về người đọc và có khoá chính là trường *Số thẻ TV*.

b) CSDL cần được thiết kế để tránh dư thừa dữ liệu

Dư thừa dữ liệu do trùng lặp dữ liệu có các nhược điểm là tốn nhiều vùng nhớ lưu trữ không cần thiết và dữ liệu có thể không nhất quán (dữ liệu bị mâu thuẫn) khi cập nhật dữ liệu.

Để tránh những nhược điểm do dư thừa dữ liệu gây ra, CSDL quan hệ thường được thiết kế gồm một số bảng, có bảng chứa dữ liệu về riêng một đối tượng (cá thể) cần quản lí, có bảng chứa dữ liệu về những sự kiện liên quan đến các đối tượng được quản lí. Ví dụ, ở một thư viện nhỏ, CSDL có thể gồm 3 bảng (Hình 2):

- Bảng SÁCH chứa dữ liệu về các quyển sách của thư viện.
- Bảng NGƯỜI ĐỌC chứa dữ liệu về những người đọc (có thẻ thư viện).
- Bảng MƯỢN-TRẢ chứa dữ liệu về sự việc một người mượn/trả một quyển sách, sự việc này liên quan đến hai đối tượng quản lí (một người đọc và một quyển sách).



Hình 2. Một CSDL gồm 3 bảng


Với cách tổ chức CSDL như trong ví dụ vừa nêu, mỗi bảng sẽ giảm được dữ liệu lặp lại, tránh thông tin dư thừa và việc cập nhật dữ liệu sẽ bớt được nhiều rủi ro sai nhầm.

2 Liên kết giữa các bảng và khoá ngoài

Để trích xuất thông tin từ CSDL quan hệ, ta có thể cần dữ liệu trong hơn một bảng và phải ghép nối đúng được dữ liệu giữa các bảng với nhau.

Ví dụ: Xét CSDL Thư viện gồm ba bảng như ở Hình 2 và yêu cầu “Cho biết *Họ và tên, Lớp* của những học sinh đã mượn quyển sách có mã TH-01”. Để trả lời yêu cầu này cần dữ liệu ở hai bảng (MUỖN-TRẢ và NGƯỜI ĐỌC). Chú ý rằng giá trị “HS-002” của *Số thẻ TV* trong bảng MUỖN-TRẢ đã “dẫn” ta đến (tham chiếu đến) một bản ghi trong bảng NGƯỜI ĐỌC chứa thông tin cần tìm. Thông qua thuộc tính *Số thẻ TV* mà hai bảng MUỖN-TRẢ và NGƯỜI ĐỌC có được mối liên kết với nhau: mỗi giá trị của *Số thẻ TV* xuất hiện trong MUỖN-TRẢ được giải thích chi tiết hơn trong NGƯỜI ĐỌC. Trong mỗi liên kết đó, bảng MUỖN-TRẢ được gọi là *bảng tham chiếu*, NGƯỜI ĐỌC là *bảng được tham chiếu* của mỗi liên kết. Tương tự, hai bảng MUỖN-TRẢ và SÁCH có mối liên kết với nhau qua thuộc tính *Mã sách*, bảng MUỖN-TRẢ là bảng tham chiếu và bảng SÁCH là bảng được tham chiếu.

Để tham chiếu xác định thì thuộc tính liên kết hai bảng phải là khoá của bảng được tham chiếu, trong ví dụ này *Số thẻ TV* phải là khoá chính của bảng NGƯỜI ĐỌC và còn được gọi là *khoá ngoài* của bảng MUỖN-TRẢ. Liên kết giữa hai bảng trong CSDL được thực hiện thông qua cặp khoá chính – khoá ngoài.

 **Khoá ngoài của một bảng:** một trường (hay một số trường) của bảng này và đồng thời là khoá của một bảng khác.

3 Hệ quản trị CSDL đảm bảo ràng buộc khoá ngoài



Hãy xét tình huống sau đây: CSDL Thư viện có bảng MUỖN-TRẢ liên kết với bảng NGƯỜI ĐỌC qua khoá ngoài *Số thẻ TV*. Hiện tại, bảng NGƯỜI ĐỌC có bốn bản ghi (ghi nhận dữ liệu về bốn học sinh đã làm thẻ thư viện). Người thủ thư đang muốn thêm một bản ghi cho bảng MUỖN-TRẢ (Hình 3). Theo em, cập nhật đó có hợp lý không? Giải thích vì sao?

Bảng MUỖN-TRẢ

Số thẻ TV	Mã sách	Ngày mượn	Ngày trả
HS-002	TH-01	14/10/2022	21/10/2022
HS-001	AN-01	02/10/2022	09/10/2022
HS-007	TO-01	16/9/2022	26/9/2022

Bảng NGƯỜI ĐỌC

Số thẻ TV	Họ và tên	Ngày sinh	Lớp
HS-001	Trần Văn An	14/9/2006	12A2
HS-002	Lê Bình	02/3/2007	11A1
HS-003	Hà Thị Mai	16/9/2007	11A2
HS-004	Nguyễn Lộc	05/10/2006	12A3

Hình 3. Dự kiến bổ sung một bản ghi vào bảng MUỖN-TRẢ

a) Ràng buộc khoá ngoài

Khi hai bảng trong một CSDL có liên kết với nhau, mỗi giá trị khoá ngoài ở bảng tham chiếu sẽ được giải thích chi tiết hơn ở bảng được tham chiếu. Ví dụ, “HS-001”

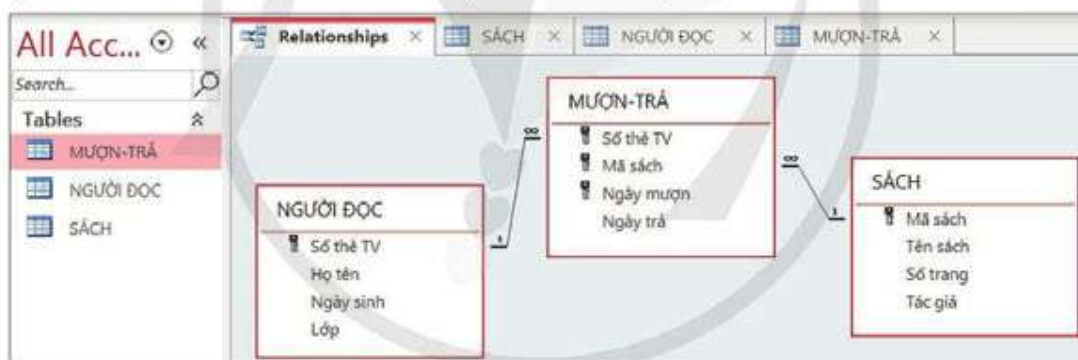
được giải thích bằng thông tin “Họ và tên: Trần Văn An; Ngày sinh: 14/9/2006; Lớp: 12A2”. Nếu có giá trị khoá ngoài nào không xuất hiện trong giá trị khoá ở bảng được tham chiếu thì xảy ra hiện tượng mất tham chiếu. Trong *Hình 3*, “HS-007” không xuất hiện trong *Số thẻ TV* của bảng NGƯỜI ĐỌC. Do vậy, việc bổ sung cho bảng MUỖN-TRẢ một bản ghi mới có giá trị khoá ngoài là “HS-007” sẽ làm cho dữ liệu trong CSDL không còn đúng đắn nữa, không giải thích được “HS-007” là số thẻ thư viện của ai. Muốn cập nhật đó hợp lệ, phải bổ sung bản ghi có giá trị khoá là “HS-007” vào bảng NGƯỜI ĐỌC trước.

Đảm bảo tính tham chiếu đầy đủ giữa các bảng có liên kết với nhau cũng là một phần của việc đảm bảo tính toàn vẹn của dữ liệu. Ràng buộc này áp dụng cho khoá ngoài nên được gọi là *ràng buộc khoá ngoài*. Nói một cách cụ thể hơn, *ràng buộc khoá ngoài* là yêu cầu mọi giá trị của khoá ngoài trong bảng tham chiếu phải xuất hiện trong giá trị khoá ở bảng được tham chiếu.

b) Khai báo liên kết giữa các bảng

Các hệ quản trị CSDL đều cho người tạo lập CSDL được khai báo liên kết giữa các bảng. Phần mềm quản trị CSDL sẽ căn cứ vào các liên kết đó để kiểm soát tất cả thao tác cập nhật, không để xảy ra những vi phạm ràng buộc khoá ngoài.

Hình 4 cho thấy kết quả trực quan của việc khai báo liên kết giữa 3 bảng khi dùng hệ quản trị CSDL Microsoft Access (phiên bản 365).



Hình 4. Kết quả của việc khai báo thành công hai liên kết giữa 3 bảng

4 Thực hành về bảng với khoá ngoài

Yêu cầu:

Khám phá cách khai báo liên kết giữa các bảng trong môi trường Access và nhận biết các cập nhật vi phạm ràng buộc khoá ngoài.

Hướng dẫn thực hiện:

Bước 1. Mở CSDL Thư viện đã có bảng SÁCH (kết quả mục thực hành ở Bài 2). Tạo cấu trúc như ở *Hình 2* cho bảng NGƯỜI ĐỌC và bảng MUỖN-TRẢ. Chọn *Số thẻ TV* làm khoá chính cho bảng NGƯỜI ĐỌC, chọn khoá chính của bảng MUỖN-TRẢ gồm ba thuộc tính: *Số thẻ TV*, *Mã sách* và *Ngày mượn*.

Bước 2. Khám phá cách khai báo liên kết giữa các bảng.

- Trong dải **Database Tools**, chọn **Relationships**.
- Dùng chuột kéo thả các bảng vào cửa sổ khai báo liên kết (vùng trống ở giữa).
- Dùng chuột kéo thả khoá ngoài của bảng tham chiếu thả vào khoá chính của bảng được tham chiếu, làm xuất hiện hộp thoại Edit Relationships.
- Đánh dấu hộp kiểm **Enforce Referential Integrity** và chọn **Create**.

Bước 3. Khám phá báo lỗi của phần mềm quản trị CSDL khi cập nhật vi phạm ràng buộc khoá ngoài.

- Thêm một vài bản ghi trong đó có bản ghi vi phạm lỗi ràng buộc khoá ngoài (tham khảo *Hình 3*), quan sát báo lỗi của phần mềm.
- Chọn xoá một bản ghi trong bảng NGƯỜI ĐỌC nếu giá trị *Số thẻ TV* trong bản ghi này xuất hiện trong bảng MUỖN-TRÀ, quan sát báo lỗi của phần mềm.



Trong việc tạo lập CSDL, sau khi tạo xong cấu trúc cho hai bảng mà ta dự kiến có liên kết với nhau bằng khoá ngoài, ta nên khai báo liên kết trước hay nên nhập dữ liệu cho hai bảng trước? Hãy giải thích vì sao.



Trong các câu sau, những câu nào đúng?

- Một trường là khoá ngoài của một bảng nếu nó là khoá của bảng đó và đồng thời xuất hiện trong một bảng khác.
- Khoá ngoài của một bảng là tập hợp một số trường của bảng đó và đồng thời là khoá của một bảng khác.
- Khi hai bảng có liên kết với nhau qua khoá chính – khoá ngoài, chỉ khi bổ sung bản ghi vào các bảng mới cần thoả mãn ràng buộc khoá ngoài.
- Các hệ quản trị CSDL quan hệ tự động kiểm tra và chỉ chấp nhận các cập nhật thoả mãn ràng buộc khoá ngoài.

Tóm tắt bài học

- ✓ CSDL quan hệ có thể gồm một số bảng, trong đó có những bảng có mối liên kết với nhau. Những liên kết này giúp tìm được những thông tin đúng đắn và đầy đủ.
- ✓ Nếu hai bảng có chung một trường và trường này là khoá của một trong hai bảng thì trường đó là khoá ngoài của bảng còn lại. Hai bảng có thể liên kết với nhau thông qua khoá ngoài.
- ✓ Dữ liệu trong hai bảng liên kết với nhau qua khoá ngoài cần phải thoả mãn ràng buộc khoá ngoài: Mọi giá trị khoá ngoài đều phải xuất hiện trong trường khoá ở bảng được tham chiếu. Mọi hệ quản trị CSDL quan hệ đều có cơ chế đảm bảo cập nhật dữ liệu không vi phạm ràng buộc khoá ngoài đối với các liên kết giữa các bảng.