

TRỊ THỨC DUY TÂN

NGUYỄN XUÂN HUY - LÊ HOÀI BẮC

# BÀI TẬP CƠ SỞ DỮ LIỆU

(Hiệu đính và xuất bản lần thứ 5)



NHÀ XUẤT BẢN THÔNG TIN VÀ TRUYỀN THÔNG

[www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn)

NGUYỄN XUÂN HUY  
LÊ HOÀI BẮC

# BÀI TẬP CƠ SỞ DỮ LIỆU

(Hiệu đính và xuất bản lần thứ 5)

NHÀ XUẤT BẢN THÔNG TIN VÀ TRUYỀN THÔNG

[www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn) - [www.lib.hau.edu.vn](http://www.lib.hau.edu.vn)

Mã số: **GD 03 HM11**

# LỜI GIỚI THIỆU

Cơ sở dữ liệu là một môn học bắt buộc đối với sinh viên chuyên ngành công nghệ thông tin và điện tử viễn thông. Những năm gần đây, trong các kỳ thi chuyển đổi, thi tuyển sinh cao học và nghiên cứu sinh đều có mảng kiến thức về cơ sở dữ liệu. Điều đó khẳng định yếu tố quan trọng của Cơ sở dữ liệu đối với các kỹ sư, thạc sỹ công nghệ thông tin và điện tử viễn thông tương lai.

Nhằm giúp bạn đọc nắm vững phương thức tự kiểm tra và đánh giá tri thức ban đầu về cơ sở dữ liệu, tự mình giải quyết các bài tập cơ sở dữ liệu, nhóm tác giả PGS.TSKH Nguyễn Xuân Huy và TS. Lê Hoài Bắc đã bổ sung chỉnh sửa và phối hợp với Nhà xuất bản Thông tin và Truyền thông tái bản lần thứ 5 "**Bài tập Cơ sở dữ liệu**" giới thiệu đến bạn đọc.

Bố cục cuốn sách được chia thành 4 phần: Phần tóm tắt lý thuyết và bài tập, một số đề thi và bài giải cho các bài tập và đề thi. Trong phần Tóm tắt lý thuyết và bài tập: tác giả đã chọn lọc và giới thiệu một số bài tập thuộc 5 mảng kiến thức: đại số quan hệ, các phép toán trên bộ, ngôn ngữ truy vấn SQL, phụ thuộc hàm và chuẩn hoá.

Ngoài ra để bạn đọc làm quen với các nội dung tổng hợp, tác giả đã tuyển chọn và giới thiệu một số đề thi tuyển sinh cao học và nghiên cứu sinh. Phần cuối sách có các bài giải cụ thể để bạn đọc tự kiểm tra kiến thức và gợi giúp nếu bạn đọc chưa có phương án giải quyết phù hợp.

Mục tiêu cuối cùng của việc ra bài tập là giúp cho người học hiểu sâu và kỹ hơn về các khái niệm đã học. Để đạt được điều này mong bạn đọc không bỏ qua bất kỳ một bài tập nào. Với các bài dễ, bạn có thể giải trong một vài phút. Với các bài có dấu \* là ký hiệu các bài tập khó, để bạn đọc thử sức mình; trong lần luyện tập thứ nhất bạn có thể bỏ qua. Sau một vài

lần thử sức, tin rằng bạn sẽ hoàn toàn làm chủ được các khái niệm liên quan đến cơ sở dữ liệu.

Hy vọng cuốn sách không chỉ là tài liệu học tập hữu ích với các bạn học sinh, sinh viên mà còn cần thiết cho cả các lập trình viên cơ sở dữ liệu và với những bạn đọc yêu thích công nghệ thông tin nói chung và cơ sở dữ liệu nói riêng.

Xin chân thành cảm ơn những đóng góp quý báu của các đồng nghiệp trong quá trình biên soạn. Tác giả mong rằng sẽ tiếp tục nhận được những ý kiến góp ý của bạn đọc gần xa để lần tái bản sau được hoàn thiện hơn.

Trân trọng cảm ơn./.

**TÁC GIẢ**

## Phần 1

# TÓM TẮT LÝ THUYẾT & BÀI TẬP

## Chương 1

# QUAN HỆ VÀ ĐẠI SỐ QUAN HỆ

### TÓM TẮT LÝ THUYẾT

#### Quan hệ

Cho tập hữu hạn  $U = \{A_1, A_2, \dots, A_n\}$  khác trống ( $n \geq 1$ ). Các phần tử của  $U$  được gọi là *thuộc tính*. Ứng với mỗi thuộc tính  $A_i \in U$ ,  $i = 1, 2, \dots, n$  là một tập không rỗng  $dom(A_i)$  được gọi là *miền trị* của thuộc tính  $A_i$  (thậm chí được giả thiết là chứa hơn 1 giá trị).

$$\text{đặt } D = \bigcup_{i=1}^n dom(A_i)$$

Một *quan hệ*  $R$  với các *thuộc tính*  $U = \{A_1, A_2, \dots, A_n\}$ , ký hiệu là  $R(U)$ , là một tập các ánh xạ  $t: U \rightarrow D$  sao cho với mỗi  $A_i \in U$  ta có  $t(A_i) \in dom(A_i)$ . Mỗi ánh xạ được gọi là một *bộ* của quan hệ  $R$ .

Mỗi quan hệ  $R(U)$  có hình ảnh là một bảng, mỗi cột ứng với một thuộc tính, mỗi dòng là một bộ.

Ta ký hiệu  $t(U)$  hoặc  $t/U$  là một bộ trên tập thuộc tính  $U$ .

Một *quan hệ rỗng*, ký hiệu  $\emptyset$ , là quan hệ không chứa bộ nào.

*Chú ý:* Vì mỗi quan hệ là một tập các bộ nên trong quan hệ *không* có hai bộ trùng lặp.

### Các ký hiệu cơ bản

Theo truyền thống của lý thuyết cơ sở dữ liệu chúng ta chấp nhận các quy định sau đây:

- Các thuộc tính được ký hiệu bằng các chữ *LA-TIN HOA* đầu bằng chữ  $A, B, C, \dots$

- Tập thuộc tính được ký hiệu bằng các chữ *LA-TIN HOA* cuối bằng chữ  $X, Y, Z, \dots$

- Các thuộc tính trong một tập được liệt kê như một xâu ký tự, không có các dấu biểu diễn tập, chẳng hạn ta viết  $X = ABC$  thay vì viết  $X = \{A, B, C\}$ .  $XY$  biểu diễn hợp của hai tập  $X$  và  $Y$ ,  $X \cup Y$ . Phép trừ hai tập  $X$  và  $Y$  được ký hiệu là  $X - Y$  hoặc  $X \setminus Y$ .

- Các bộ được biểu diễn bằng các chữ *La-tin* thường có thể kèm chỉ số  $t, u, v, t_1, \dots$

- Với mỗi bộ  $t$  trong quan hệ  $R(U)$  và mỗi tập con các thuộc tính  $X \subseteq U$  ta ký hiệu  $t[X]$  hoặc  $t.X$  là hạn chế của bộ  $t$  trên tập thuộc tính  $X$ .

- Hàm  $Attr(R)$  cho tập thuộc tính của quan hệ  $R$ .

- Hàm  $Card(R)$  cho lực lượng (số bộ) của quan hệ  $R$ .

- Trong trường hợp quan hệ  $R$  với tập thuộc tính  $U$  đã cho trước ta có thể viết đơn giản  $R$  thay cho  $R(U)$ .

- Ký hiệu  $REL(U)$  là tập toàn thể các quan hệ trên tập thuộc tính  $U$ .

Hai quan hệ  $R$  và  $S$  được gọi là *tương thích* nếu chúng có cùng một tập thuộc tính,  $Attr(R) = Attr(S)$ .

Với mỗi bộ  $u$  trong quan hệ  $R(U)$  và mỗi bộ  $v$  trong quan hệ  $S(V)$  ta ký hiệu  $u * v$  là phép *dán bộ*.  $u * v$  cho ta bộ  $t$  trên tập thuộc tính  $UV$  thỏa mãn điều kiện  $t.U = u$  và  $t.V = v$ .

Với mỗi bộ  $u$  trong quan hệ  $R(U)$  và với mỗi quan hệ  $S(V)$  ta ký hiệu  $u * S$  là phép *dán bộ  $u$  với quan hệ  $S$* .  $u * S$  cho ta quan hệ

$$P(UV) = \{u * v \mid v \in S\}$$

Để thể hiện các phép toán quan hệ ta sẽ dùng các ký pháp tựa như ký pháp của hệ ISBL (Information System Base Language).

### Đại số quan hệ

#### Phép chọn (phép lọc, Selection)

Cho quan hệ  $R(U)$  và biểu thức điều kiện (còn gọi là *biểu thức lọc* hay *biểu thức chọn*)  $e$ . Phép chọn trên quan hệ  $R$  theo điều kiện  $e$ , ký hiệu  $R(e)$  cho ta quan hệ:

$$P(U) = R(e) = \{t \in R \mid Sat(t, e)\}$$

trong đó hàm logic  $Sat(t, e)$  kiểm tra bộ  $t$  thỏa mãn điều kiện  $e$  được xác định như sau:

1. Thay mọi xuất hiện của mỗi thuộc tính  $A$  trong biểu thức chọn  $e$  bằng trị tương ứng của  $A$  trong bộ  $t$ ,  $t.A$ , ta thu được một mệnh đề logic  $b$ .
2. Tính trị của  $b$ . Nếu là đúng (*True*) thì bộ  $t$  thỏa mãn điều kiện  $e$ ; ngược lại, nếu trị của  $b$  là sai (*False*) thì bộ  $t$  không thỏa mãn điều kiện  $e$ .

Trong các biểu thức chọn ta sử dụng ký hiệu cho các phép toán logic như sau:

- Tích: & hoặc  $\wedge$  hoặc AND
- Tổng: | hoặc  $\vee$  hoặc OR
- Phủ định: ! hoặc  $\neg$  hoặc NOT
- Kéo theo:  $\Rightarrow$  hoặc IMPLY