NGUYỄN XUÂN HUY - LÊ HOÀI BẮC

Bài tập
CƠ SỞ DỮ LIỆU

NHÀ XUẤT BẢN THỐNG KÊ - 2003

NGUYỄN XUÂN HUY - LÊ HOÀI BẮC

BÀI TẬP
CƠ SỞ DỮ LIỆU

NHÀ XUẤT BẢN THỐNG KÊ
HÀ NỘI - 2003

MỤC LỤC

LỜI NÓI ĐẦU

Khác với toán học, trong tủ sách tin học nước nhà, ta chỉ thấy một số sách bài tập lập trình. Đó chắc chắn là một thiệt thòi cho sinh viên và các bạn tự học.

Cuốn Bài tập cơ sở dữ liệu này là một thử nghiệm nhằm trợ giúp các bạn trẻ một phương thức tự kiểm tra và đánh giá tri thức ban đầu, mức nhập môn, về một lĩnh vực chiếm vị trí đáng nói trong quá trình phát triển của công nghệ thông tin.

Những năm gần đây, trong các kỳ thi tốt nghiệp đại học, thi chuyển đổi, thi tuyển cao học và nghiên cứu sinh đều có mảng về cơ sở dữ liệu. Đó là điều dễ hiểu, vì cơ sở dữ liệu là phần không thể thiếu trong các hệ thống tin học hoá.

Trong phương án đầu tiên của cuốn sách chúng tôi chọn lọc và đề xuất một số bài tập thuộc năm mảng tri thức sau đây: đại số quan hệ, các phép toán trên bộ, ngôn ngữ hỏi SQL, phụ thuộc hàm và chuẩn hoá. Mỗi mảng tri thức được trình bày thành ba phần: Phần thứ nhất bao gồm một số điều tóm tắt về lý thuyết. Phần tiếp theo là các bài tập, cuối cùng là các bài giải. Dấu * được dùng để ghi chú các bài tập ở mức nâng cao.

Phần cuối sách chúng tôi tuyển chọn và giới thiệu một số đề thi tuyển cao học và nghiên cứu sinh để bạn đọc làm quen với các nội dung tổng 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 đừng bỏ qua bài tập nào. Với các bài dễ, bạn có thể giải trong một vài phút. Với các bài khó, 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.

Chúng tôi cho rằng các tài liệu sau đây sẽ giúp ích bạn đọc tra cứu các nguồn tri thức cơ sở:

1. Date C. J., *Nhập môn các hệ cơ sở dữ liệu*, Những người dịch: Hồ Thuần, Nguyễn Quang Vinh, Nguyễn Xuân Huy, NXB Thống Kê, Hà Nội, Tập I (1985), Tập II (1986).

2. Nguyễn Xuân Huy, *Thuật toán*, NXB Thống Kê, Hà Nội, 1987.

3. Vũ Đức Thi, *Cơ sở dữ liệu:Kiến thức và thực hành*, NXB Thống Kê, Hà Nội, 1997.

4. Lê Tiến Vương, *Nhập môn cơ sở dữ liệu quan hệ*, Tái bản lần thứ 4, NXB Thống Kê, Hà Nội, 1999.

5. Garcia-Molina H., Ullman J., Widom J., *Database System: The Complete Book*, Prentice Hall, 2002.

6. Maier D., *The Theory of Relational Database*, Computer Science Press, Rockville, Md, 1983.

7. Ullman, J., *Principles of Data-base and Knowledge-base Systems*, (Second Edition), Computer Science Press, Potomac, Md., 1982, (Có bản dịch tiếng Việt của Trần Đức Quang.)

Người đầu tiên định hướng cho chúng tôi tìm hiểu về cơ sở dữ liệu và luôn luôn khuyến khích chúng tôi học tập và trao đổi kiến thức là giáo sư Hồ Thuần, Viện Công nghệ Thông tin.

Cuốn sách này được khởi thảo và hoàn thành theo phương án đầu tiên là nhờ nhiệt tình đóng góp về ý tưởng, nội dung và thẩm định của các đồng nghiệp của chúng tôi. Giáo sư Lê Tiến Vương, Tổng cục Địa chính, giáo sư Hoàng Kiếm, giáo sư Trần Vĩnh Phước, Đại học Quốc gia thành phố Hồ Chí Minh đã thảo luận chi tiết về những nội dung cơ bản và kiến trúc cho tập sách.

Đặc biệt, các đồng nghiệp trẻ, giáo sư Vũ Ngọc Loãn, Đại học Quốc gia Hà Nội, giáo sư Nguyễn Thanh Thuỷ, Đại học Bách khoa Hà Nội, tiến sỹ Trịnh Đình Thắng, Đại học Sư phạm Hà Nội II, tiến sỹ Dương Anh Đức, tiến sỹ Đỗ Văn Nhơn, thạc sỹ Nguyễn Tấn Trần Minh Khang, Đại học Quốc gia thành phố Hồ Chí Minh, thạc sỹ Nguyễn Xuân Tùng, Trung tâm Tin học Bưu điện Hà Nội, thạc sỹ Nguyễn Ngọc Hà, Trung tâm Tin học Bưu điện Hải Phòng, thạc sỹ Trịnh Thanh Lâm, Intel, thạc sỹ Nguyễn Xuân Hoàng, Misa Group đã có những góp ý cụ thể về nội dung chương trình đào tạo và các yêu cầu thực tiễn của cơ sở dữ liệu. Các cử nhân Bùi Thuý Hằng và Trần Quốc Dũng, Viện Công nghệ Thông tin đã giúp chúng tôi đọc lại và chỉnh sửa các trang bản thảo.

Chúng tôi chân thành cảm ơn những đóng góp vô giá của các đồng nghiệp.

Chúng tôi mong rằng sẽ tiếp tục nhận được những ý kiến chỉ giáo của bạn đọc gần xa về nội dung và cấu trúc của tập sách.

Cát Bà, Mùa Hoa Phượng, 2003

Các tác giả

NGUYỄN XUÂN HUY - LÊ HOÀI BẮC

1

TÓM TẮT LÝ THUYẾT VÀ BÀI TẬP

Chương 1

QUAN HỆ VÀ ĐẠI SỐ QUAN HỆ

TÓM TẮT LÝ THUYẾT

Quan hệ

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

Đặt $D = \bigcup_{i=1}^{n} dom(A_i)$

Một *quan hệ R* với các *thuộc tính* $U = \{ A_1, A_2, \ldots, A_n \}$, 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 là một bộ trên tập thuộc tính U.

Một *quan hệ trống*, ký hiệu ∅, 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 *LATIN HOA* đầu bảng chữ *A*, *B*, *C*,...
- Tập thuộc tính được ký hiệu bằng các chữ *LATIN HOA* cuối bảng chữ *X*, *Y*, *Z*,...
- 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 thuộc tính *X* và *Y*, $X \cup Y$. Phép trừ hai tập hợp X và Y được ký hiệu là X-Y hoặc $X \setminus Y$.
- Các bộ được biểu diễn bằng các chữ *Latin thường* có thể kèm chỉ số $t, u, v, t_1 ...$
- 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 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, tức là nếu *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* thoả 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)

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 thoả 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* thoả điều kiện *e*; ngược lại, nếu trị của *b* là *sai* (*False*) thì bộ *t* không thoả điều kiện *e*.

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

- Tích: & hoặc AND
- Tổng: | hoặc OR

- Phủ định: ! hoặc NOT
- Kéo theo: $\Rightarrow$ hoặc IMPLY

Phép chiếu

Phép chiếu quan hệ $R(U)$ trên tập con thuộc tính $X \subseteq U$, ký hiệu $R[X]$, cho ta quan hệ

$$P(X) = R[X] = \{ t.X \mid t \in R \}$$

$R[X]$ được tính theo 2 bước như sau:

1. Xoá các cột không thuộc X của bảng R,
2. Xoá bớt các dòng giống nhau trong bảng kết quả: chỉ giữ lại một dòng trong số các dòng giống nhau.

Phép kết nối tự nhiên

Phép kết nối (*tự nhiên*) hai quan hệ $R(U)$ và $S(V)$, ký hiệu $R*S$, cho ta quan hệ chứa các bộ được dán từ các bộ u của quan hệ R với mỗi bộ v của quan hệ S sao cho các trị trên miền thuộc tính chung (nếu có) của hai bộ này giống nhau.

$$P(UV) = R*S = \{ u*v \mid u \in R, v \in S, u.M = v.M, M = U \cap V \}$$

Nếu $M = U \cap V = \varnothing$, $R*S$ sẽ cho ta tích Descartes trong đó mỗi bộ của quan hệ R sẽ được ghép với mọi bộ của quan hệ S.

Phép cộng (hợp)

Phép cộng (*hợp* theo lý thuyết tập hợp hoặc *kết nối dọc*) hai quan hệ tương thích $R(U)$ và $S(U)$, ký hiệu $R+S$, cho ta quan hệ chứa các bộ của mỗi quan hệ thành phần,

$$P(U) = R+S = \{ t \mid t \in R \vee t \in S \}$$

Phép trừ

Phép trừ (theo lý thuyết tập hợp hoặc *lấy phần riêng*) hai quan hệ tương thích $R(U)$ và $S(U)$, ký hiệu R-S, cho ta quan hệ chứa các bộ của quan hệ R không có trong quan hệ S,

$$P(U) = R\text{-}S = \{ t \mid t \in R, t \notin S \}$$

Phép giao

Phép giao (theo lý thuyết tập hợp hoặc *lấy phần chung*) hai quan hệ tương thích $R(U)$ và $S(U)$, ký hiệu $R\&S$, cho ta quan hệ chứa các bộ xuất hiện đồng thời trong cả hai quan hệ thành phần,

$$P(U) = R\&S = \{ t \mid t \in R, t \in S \}$$

Các phép toán cộng, trừ và giao được gọi là các *phép toán tập hợp* trên các quan hệ (tương thích).

Phép chia

Cho hai quan hệ $R(U)$ và $S(V)$. *Phép chia* quan hệ R cho quan hệ S, ký hiệu $R : S$, cho ta quan hệ

$$P(M) = R : S = \{ t.M \mid t \in R, (t.M) * S \subseteq R, M = U - V \}$$

Thứ tự thực hiện các phép toán quan hệ

Trong một biểu thức quan hệ các phép toán *một ngôi* có độ *ưu tiên cao* hơn (do đó được thực hiện sớm hơn) các phép toán hai ngôi. Tiếp đến là nhóm các phép toán kết nối, giao và chia, cuối cùng là nhóm các phép toán cộng và trừ. Thứ tự ưu tiên từ cao đến thấp của các phép toán quan hệ được liệt kê như sau:

( ), [ ]

$*$, & , :

$+$, -

Dãy các phép toán cùng thứ tự ưu tiên được thực hiện lần lượt từ trái qua phải. Nếu biểu thức quan hệ có chứa các cặp ngoặc ( ) thì các biểu thức con trong các cặp ngoặc được thực hiện trước.

Một số hàm tiện ích

1. *Sum*(*R*,*A*): cho *tổng* các giá trị số trong thuộc tính (cột) *A* của quan hệ *R*, $Sum(R,A) = \Sigma\ (\ t.A \mid t \in R\ )$
2. *Avg*(*R*,*A*): cho *trung bình cộng* các giá trị trong thuộc tính (cột) *A* của quan hệ *R*, $Avg(R,A) = Sum(R,A)\ /\ Card(R)$ nếu $Card(R) \neq 0$
3. *Max*(*R* ,*A*): cho *giá trị lớn nhất* trong thuộc tính (cột) *A* của quan hệ *R*.
4. *Min*(*R*,*A*): cho *giá trị nhỏ nhất* trong thuộc tính (cột) *A* của quan hệ *R*.

Nếu trong biểu thức quan hệ có chứa các hàm tiện ích thì các hàm này được thực hiện *sớm nhất* trong ngữ cảnh cho phép.

Thí dụ

Biểu thức quan hệ $P = S*R(A > Avg(S,A))[AB]$ sẽ được thực hiện theo trật tự sau đây:

1. Tính hàm $c = Avg(S,A)$
2. Thực hiện phép chọn $P_1 = R(A > c)$
3. Thực hiện phép chiếu $P_2 = P_1\ [AB]$

4. Thực hiện phép kết nối $P = S*P_2$

Chú ý: Trong một số tài liệu có sử dụng ký pháp khác cho các phép toán quan hệ như sau

Phép toán	Ký hiệu	Ký hiệu khác
chọn	$R(e)$	$\sigma_e(R)$
chiếu	$R[X]$	$\pi_X(R)$
kết nối tự nhiên	$R * S$	$R \bowtie S$
cộng	$R + S$	$R \cup S$
giao	$R \& S$	$R \cap S$
trừ	$R - S$	$R \setminus S$
chia	$R : S$	$R \div S$

Cơ sở dữ liệu minh họa: CSDL Thực tập

Hầu hết bài tập trong chương này liên quan đến CSDL **Thực tập** gồm ba quan hệ sau đây:

SV(SV#, HT, NS, QUE, HL)

DT(DT#, TDT, CN, KP)

SD(SV#, DT#, NTT, KM, KQ)

- Quan hệ SV(SV#, HT, NS, QUE, HL) chứa thông tin về các sinh viên trong một lớp của một trường đại học,

  SV - tên quan hệ sinh viên

  SV# - mã số sinh viên

  HT - họ và tên sinh viên

NS - năm sinh của sinh viên

QUE - quê (tỉnh)

HL - học lực thể hiện qua điểm trung bình

- Quan hệ DT(DT#, TDT, CN, KP) chứa thông tin về các đề tài nhà trường quản lý,

DT - tên quan hệ đề tài

DT# - mã số đề tài

TDT - tên đề tài

CN - họ và tên chủ nhiệm đề tài

KP - kinh phí cấp cho đề tài (triệu đồng).

- Quan hệ SD(SV#, DT#, NTT, KM, KQ) chứa thông tin về tình hình thực tập của các sinh viên theo các đề tài,

SD - tên quan hệ sinh viên - đề tài

SV# - mã số sinh viên

DT# - mã số đề tài mà sinh viên đó tham gia

NTT - nơi thực tập để triển khai đề tài (tỉnh)

KM - khoảng cách từ nơi thực tập đến trường

KQ - kết quả thực tập theo đề tài đã chọn

- Giả thiết là một sinh viên có thể tham gia nhiều đề tài, mỗi đề tài sinh viên đó thực tập tại một địa điểm.
- Với mỗi câu hỏi, yêu cầu trả lời bằng một biểu thức của đại số quan hệ. Tuổi được tính đến năm 2003. Thí dụ,

Câu hỏi

Cho danh sách các sinh viên trẻ (dưới 18 tuổi tính đến năm 2003), học và thực tập đều đạt loại khá/giỏi (điểm không dưới 8.5)

Trả lời

(SD(KQ >= 8.5)[SV#]*SV)(2003-NS < 18 & HL >= 8.5)[HT]

BÀI TẬP

1.1. Cho hai quan hệ *R*(*A*,*B*,*C*,*D*) và *S*(*C*,*D*) như sau

R(A	B	C	D)
a	1	x	2
a	1	y	2
b	2	x	1
b	2	y	1
a	4	x	2
c	5	y	7

S(C	D)
x	1
y	2

Hãy xác định:

a) *R*[*AB*]

b) *R*(3-*B*+*D* >1)

c) *R*(*B* < 4) + *R*(*D* > 3)

d) *R*(*B* >= 1 & *B* <= 5)

e) *R***S*[*C*]

f) *R*(*B* < 4) - *R*(*D* > 3)

g) *R*(*B* < 4) & *R*(*D* > 3).

h) *R* : *S*

1.2. Cho thông tin về những sinh viên sinh trước năm 1983, quê ở Hải Phòng?

1.3. Cho danh sách các tỉnh có sinh viên đến thực tập?

1.4. Cho biết các địa điểm thực tập xa trường (KM > 100) của đề tài số 7?

1.5. Cho thông tin về việc thực tập tại Nha Trang của các sinh viên?

***1.6.** Cho danh sách sinh viên thực tập tại quê nhà?

1.7. Cho thông tin về các đề tài có sinh viên thực tập?

1.8. Cho biết mã của các đề tài không có sinh viên nào tham gia?

1.9. Cho biết mã của những đề tài có kinh phí 1.5 triệu và những đề tài có kinh phí trên 2 triệu?

1.10. Cho biết mã của những sinh viên dưới 20 tuổi, thực tập khá (có điểm kết quả trên 7)?

***1.11.** Cho biết mã của những đề tài có địa bàn thực tập ít ra là như đề tài 1.

***1.12.** Cho danh sách những đề tài được triển khai thực tập ở tất cả các tỉnh có sinh viên thực tập.

***1.13.** Cho danh sách những sinh viên thực tập theo đề tài có kinh phí lớn hơn một phần năm tổng kinh phí cấp cho các đề tài.

***1.14.** Cho danh sách các sinh viên có điểm học tập cao hơn điểm thực tập trung bình của đề tài mã số 4.

1.15. Một phép toán 2 ngôi θ có tính chất *giao hoán* nếu:

$$(\forall R, S):\ R\,\theta\, S = S\,\theta\, R$$

hoặc cả hai vế đồng thời không có nghĩa.

Chứng minh rằng các phép toán quan hệ *kết nối*, *cộng* và *giao* có tính *giao hoán*.

1.16. Tìm thí dụ chứng tỏ các phép toán *trừ* và *chia không có tính giao hoán*.

1.17. Cho quan hệ $R(U)$ và hai biểu thức chọn e và h. Chứng minh

a) $R(e \,\&\, h) = R\,(h \,\&\, e)$

b) $R(e \,\&\, h) = R\,(e) \,\&\, R(h) \subseteq R(e)$

c) $R(e \,\&\, h) = R(h) \,\&\, R(e) \subseteq R(h)$

d) $R(e \,\&\, h) = R(e)(h)$

e) $R(e \mid h) = R(h \mid e)$

f) $R(e \mid h) = R(e) + R(h)$

g) $R(!\, e) = R - R(e)$

h) $R(\text{True}) = R$

i) $R(\text{False}) = \varnothing$

1.18. Cho quan hệ $R(U)$, các biểu thức chọn e, h trên U và tập con các thuộc tính $X \subseteq U$. Ký hiệu Attr(e) là tập các thuộc tính của U có mặt trong e. Chứng minh, nếu Attr(e) $\subseteq X$ thì

a) $R(e)[X] = R[X](e)$

b) $R(h \,\&\, e)[X] = R(h)(e)[X] = R(h)[X](e)$

***1.19.** Chứng minh rằng phép chia có thể được biểu diễn qua các phép chiếu, kết nối và trừ như sau,

$$R : S = R[M] - (R[M] * S - R)[M]$$

trong đó $M = Attr(R) - Attr(S)$.

1.20. Phép toán quan hệ được gọi là *đóng* nếu với mọi quan hệ đầu vào ta đều thu được đầu ra là một quan hệ. Cho biết tính đóng (ghi *có / không*) của các phép toán quan hệ

Phép toán	Ký hiệu	Tính đóng
chọn	( )	
chiếu	[ ]	
kết nối tự nhiên	*	
cộng	+	
giao	&	
trừ	-	
chia	:	

1.21. Phép toán 2 ngôi θ có tính chất *kết hợp* nếu:

$$(\forall R, S, T):\ (R\,\theta\, S)\,\theta\, T = R\,\theta\,(S\,\theta\, T)$$

hoặc cả hai vế đồng thời không có nghĩa.

Chứng minh rằng các phép toán *kết nối*, *cộng* và *giao* của đại số quan hệ có tính chất *kết hợp*.

1.22. Tìm thí dụ chứng tỏ các phép toán *trừ* và *chia* *không có tính kết hợp*.

1.23. Chứng minh rằng với mọi cặp quan hệ tương thích R và S ta có

$$R - (R - S) = R \,\&\, S$$

1.24. Chứng minh rằng với mọi quan hệ $R(U)$, mọi tập con X trong U và mọi biểu thức điều kiện e ta có

a) $R(e)(e) = R(e)$

b) $R[X][X] = R[X]$

1.25. Chứng minh rằng với mọi quan hệ $R(U)$ ta có

a) $R * R = R$

b) $R + R = R$

c) $R \& R = R$

d) $R - R = \varnothing$

e) $R : R = \varnothing$, $\mathrm{Attr}(R : R) = \varnothing$.

1.26. Phép toán quan hệ được gọi là *nở* (*co*) *ngang* nếu quan hệ kết quả có số thuộc tính *nhiều hơn* (*ít hơn*) các quan hệ đầu vào, được gọi là *nở* (*co*) *dọc* nếu quan hệ kết quả có số bộ *nhiều hơn* (*ít hơn*) các quan hệ đầu vào. Hãy đánh dấu (+), (-) hoặc (=) để khẳng định tính nở hoặc co hoặc không nở/co của mỗi phép toán tương ứng.

Phép toán	**Ký hiệu**	**Nở/Co ngang**	**Nở/Co dọc**
chọn	( )		
chiếu	[ ]		
kết nối tự nhiên	*		
cộng	+		
giao	&		
trừ	-		
chia	:		

1.27. Cho quan hệ $R(U)$ và e và h là hai biểu thức chọn trên U. Chứng minh, nếu $e \Rightarrow h$ thì:

a) $R(e)(h) = R(e)$

b) $R(e) \subseteq R(h)$

1.28. Gọi T và F lần lượt là các công thức logic hằng đúng và hằng sai. Chứng minh rằng với mọi quan hệ R ta có:

a) $R(T) = R$

b) $R(F) = \varnothing$

***1.29.** Cho quan hệ $R(U)$. Hãy dùng một phép toán quan hệ để sinh ra quan hệ rỗng $S(U)$.

1.30. Chứng minh rằng với mọi quan hệ $R(U)$ và mọi tập thuộc tính X và Y thoả điều kiện $X \subseteq Y \subseteq U$ thì

$$R[Y][X] = R[X]$$

1.31. Cho hai quan hệ $R(U)$ và $S(V)$ và hai biểu thức chọn e trên U, h trên V. Chứng minh

$$(R*S)(e \,\&\, h) = R(e)*S(h)$$

1.32. Cho các quan hệ $R(U)$, $S(V)$ và các tập thuộc tính $X \subseteq U$, $Y \subseteq V$. Biết $Z = U \cap V$. Chứng minh

$$(R*S)[XZY] = R[XZ]*S[ZY]$$

***1.33.** Cho quan hệ $R(U)$ và các tập con X_1, X_2, ..., X_k thoả điều kiện $X_1 \cup X_2 \ldots \cup X_k = U$. Chứng minh

$$R[X_1] * R[X_2] * \ldots * R[X_k] \supseteq R$$

Tìm thí dụ chứng tỏ $R[X_1] * R[X_2] * \ldots * R[X_k] \neq R$

***1.34.** Cho quan hệ $R(U)$ và hai tập con thuộc tính $X \subseteq Y \subseteq U$.

Chứng minh

a) $R[U] = R$

b) $R[X] * R[X] = R[X]$

c) $R[Y] * R[X] = R[Y]$

d) $R * R[X] = R$

***1.35.** Cho tập hữu hạn các phần tử M. Ký hiệu Poset(M) là tập các tập con của M. Ánh xạ f: Poset(M) → Poset(M) được gọi là đóng nếu f thoả ba tính chất sau:

$\forall X, \forall Y \in$ Poset(M):

($C1$) Tính phản xạ: $f(X) \supseteq X$

($C2$) Tính đồng biến: nếu $X \subseteq Y$ thì $f(X) \subseteq f(Y)$

($C3$) Tính luỹ đẳng: $f(f(X)) = f(X)$

Cho tập thuộc tính U. Cố định các tập con $X_1, X_2, \ldots, X_k$ của U thoả điều kiện $X_1 \cup X_2 \ldots \cup X_k = U$.

Xét ánh xạ ρ: $REL(U) \rightarrow REL(U)$

$$\forall R \in REL(U): \rho(R) = R[X_1] * R[X_2] * \ldots * R[X_k]$$

Chứng minh ρ là một ánh xạ đóng trên REL(U).

***1.36.** (*Nghịch lý giao hoán-kết hợp*) Ta đã biết phép kết nối tự nhiên (*) có tính kết hợp và giao hoán. Xét các quan hệ SV, DT và SD trong cơ sở dữ liệu THỰC TẬP. Hai quan hệ SV và DT không có thuộc tính chung, trong khi hai quan hệ *SV* và *SD* chung nhau thuộc tính *SV#* và hai quan hệ *SD* và *DT* chung nhau thuộc tính *DT#*. Giải thích vì sao

$$SV * (SD * DT) = (SV * SD) * DT = (SV * DT) * SD$$

*Cho biết kết quả của phép kết nối (SV * DT) trong biểu thức phải.*

Chương 2

CÁC THAO TÁC TRÊN BỘ VÀ QUAN HỆ

TÓM TẮT LÝ THUYẾT

Trong phần trình bày cú pháp của các cấu trúc điều khiển (câu lệnh) và các hàm thao tác trên bộ và quan hệ dưới đây phần viết trong [ ] là tuỳ chọn.

(T1) `for each t in R [with e] do P endfor`

Thực hiện toán tử *P* trên những bộ *t* của quan hệ *R* [nếu thoả điều kiện e].

(T2) `if e then P[else Q] endif`

Nếu điều kiện *e* thoả thì thực hiện *P* [nếu không, thực hiện *Q*]

(T3) `Create(R,X)`

Tạo lập quan hệ rỗng *R* (không chứa bộ nào) với tập thuộc tính *X* cho trước.

(T4) `Attr(R)`

Hàm cho tập thuộc tính của quan hệ *R* .

(T5) `Card(R)`

Hàm cho biết lực lượng (số bộ) của quan hệ *R*.

(T6) `t [not_]in R`

Hàm cho giá trị True nếu bộ t [không] có trong quan hệ R, ngược lại hàm cho giá trị False.

(T7) add t to R

Nạp bộ *t* vào quan hệ *R*.

(T8) t[X] hoặc t.X

Tạo bộ mới từ bộ *t* với tập thuộc tính *X*. Bộ mới bao gồm các giá trị của mỗi thuộc tính *A* trong *t*, $A \in X$.

(T9) t/X

Tạo bộ mới từ bộ t bằng cách bỏ đi những giá trị của mỗi thuộc tính *A* trong *t*, $A \in X$.

(T10) u*v

Tạo bộ mới bằng phép dán bộ *v* với bộ *u*. Các giá trị trên miền thuộc tính chung của hai bộ *u* và *v* phải bằng nhau và chỉ lấy một trong hai trị bằng nhau trên mỗi thuộc tính chung.

Các thuật toán được diễn đạt thông qua ngôn ngữ quy ước sau đây:

```
Algorithm <tên thuật toán>
Format
Input
Output
Method
// Chú thích được viết sau hai gạch nghiêng
End.
```

BÀI TẬP

2.1. Viết thuật toán thực hiện phép chọn trên quan hệ: $R(e)$.

2.2. Viết thuật toán thực hiện phép chiếu trên quan hệ: $R[X]$.

2.3. Viết thuật toán thực hiện phép kết nối tự nhiên hai quan hệ: $R*S$.

2.4. Viết thuật toán thực hiện phép hợp hai quan hệ tương thích: $R+S$.

2.5. Viết thuật toán thực hiện phép giao hai quan hệ tương thích: $R\&S$.

2.6. Viết thuật toán thực hiện phép trừ hai quan hệ tương thích: $R-S$.

2.7. Viết thuật toán thực hiện phép chia hai quan hệ: $R:S$.

2.8. Phép chọn_chiếu quan hệ $R(U)$ theo biểu thức chọn e và trên tập con thuộc tính $X \subseteq U$ cho ta quan hệ

$$P(X) = R(e,X) = R(e)[X] = \{ t.X \mid t \in R,\ Sat(t,e) \}$$

Viết thuật toán thực hiện trực tiếp phép chọn_chiếu.

2.9. Phép kết_nối_chọn_chiếu hai quan hệ $R(U)$ và $S(V)$ theo biểu thức chọn e và trên tập con thuộc tính $X \subseteq UV$ cho ta quan hệ

$P(X) = (R*S)(e,X) = (R*S)(e)[X] =$

$= \{ (u*v).X \mid u \in R,\ v \in S,\ u.M = v.M,\ M = U \cap V,\ Sat(u*v,e) \}$

Viết thuật toán thực hiện trực tiếp phép kết_nối_chọn_chiếu.

2.10. Cài đặt hàm $Card(R)$: cho lực lượng (số bộ) của quan hệ R.

2.11. Cài đặt hàm $Sum(R,A)$: cho tổng các giá trị trong thuộc tính (cột) A của quan hệ R: $Sum(R,A) = \sum_{t \in R} t.A$.

2.12. Cài đặt hàm $Avg(R,A)$: cho trung bình cộng các giá trị trong thuộc tính (cột) A của quan hệ R,

$$Avg(R,A) = Sum(R,A) \,/\, Card(R), \text{ nếu } Card(R) \neq 0$$

2.13. Cài đặt hàm *Max*(*R*, *A*): cho giá trị lớn nhất trong thuộc tính (cột) *A* của quan hệ *R*.

2.14. Cài đặt hàm *Min*(*R*, *A*): cho giá trị nhỏ nhất trong thuộc tính (cột) *A* của quan hệ *R*.

2.15. Cài đặt hàm *Sum*(*R*,*A*,*e*): cho tổng các giá trị trong thuộc tính (cột) *A* của quan hệ *R* xét trên những bộ thoả điều kiện *e*,

$$Sum(R,A,e) = Sum(R(e),A) = \sum_{\substack{t \in R \\ sat(t,e)}} t.A .$$

2.16. Cài đặt hàm *Count*(*R*,*e*): cho số lượng các bộ thoả điều kiện *e* trong quan hệ *R*.

2.17. Cài đặt hàm *Avg*(*R*,*A*,*e*): cho trung bình cộng của các giá trị không nhất thiết khác nhau trong thuộc tính (cột) *A* của quan hệ *R* xét trên những bộ thoả điều kiện *e*.

2.18. Cài đặt hàm *Maxe*(*R*, *A*, *e*): cho giá trị lớn nhất trong thuộc tính (cột) *A* của quan hệ *R* xét trên những bộ thoả điều kiện *e*.

2.19. Cài đặt hàm *Mine*(*R*, *A*, *e*): cho giá trị nhỏ nhất trong thuộc tính (cột) *A* của quan hệ *R* xét trên những bộ thoả điều kiện *e*.

Chương 3
NGÔN NGỮ HỎI SQL
(Structured Query Language)

TÓM TẮT LÝ THUYẾT

SQL là ngôn ngữ hỏi thuộc lớp ngôn ngữ các phép tính trên bộ.

Cấu trúc chính của SQL có dạng

```
SELECT [DISTINCT/UNIQUE] X
FROM R1, R2, . . . ,Rk
WHERE e;
```

trong đó X là danh sách các thuộc tính của quan hệ ra, $R_1, R_2,...,R_k$ là các quan hệ, e là điều kiện. Cấu trúc trên tương đương với biểu thức đại số quan hệ sau đây

$$(R_1 * R_2 * . \quad . \quad . * R_k)\ (e)\ [X]$$

trong đó phép $*$ được hiểu là phép kết nối có điều kiện. Điều kiện này được xác định trong mục WHERE.

Khi liệt kê tên các thuộc tính, để tránh hiện tượng trùng tên của hai thuộc tính trong hai quan hệ khác nhau ta có thể chỉ rõ quan hệ chứa thuộc tính đó. Thí dụ, *R.A* nói về thuộc tính *A* của quan hệ *R*. Từ khoá DISTINCT là chỉ thị lược bớt các bộ trùng lặp ở kết quả.

Các từ khóa và ký hiệu

** - danh sách đầy đủ các thuộc tính*

IN - là phần tử của

NOT IN - không phảI là phần tử của

ANY - một phần tử, một xuất hiện

ALL - với mọi

AND - phép nhân logic

OR - phép cộng logic

NOT - phép phủ định

ORDER BY DESC/ASC - sắp giảm/tăng

GROUP BY - nhóm theo

CONTAINS - chứa

UNION - hợp hai quan hệ tương thích

INTERSECTION - giao hai quan hệ tương thích

DIFFERENCE/MINUS - trừ hai quan hệ tương thích

EXISTS - cho giá trị True nếu biểu thức sau nó chứa ít nhất một bộ, ngược lại cho giá trị False.

Các hàm trên cột

COUNT - cho số lượng phần tử của cột

SUM - cho tổng các trị trong cột

MIN - cho giá trị nhỏ nhất trong cột

MAX - cho giá trị lớn nhất trong cột

AVG - cho giá trị trung bình cộng của cột

***Bí danh** là các định danh đặt thêm cho một quan hệ để tiện dùng*

BÀI TẬP

Các bài tập liên quan đến CSDL ***Thực tập*** gồm ba quan hệ như đã mô tả trong chương về đại số quan hệ.

SV(*SV#*, *HT*, *NS*, *QUE*,*HL*)

DT(*DT#*,*TDT*,*CN*,*KP*)

SD(*SV#*,*DT#*,*NTT*,*KM*, *KQ*)

Với mỗi câu hỏi, yêu cầu trả lời bằng *một biểu thức* SQL.

3.1. Danh sách các sinh viên trẻ (dưới 18 tuổi) và học khá/giỏi (*HL* > 8.5)

3.2. Thông tin về các đề tài được cấp kinh phí trên 10 triệu đồng

3.3. Danh sách các sinh viên trẻ (dưới 18 tuổi), học và thực tập đều đạt loại khá/giỏi (*HL* > 8.5 và KQ > 8.5)

3.4. Danh sách các chủ nhiệm đề tài có các sinh viên quê ở Hà Nội tham gia.

3.5. Danh sách các sinh viên học giỏi hơn các sinh viên Hà Nội.

3.6. Điểm trung bình của các sinh viên Hà Nội.

3.7. Tổng số đoạn đường thực tập theo đề tài 5.

3.8. Tổng số sinh viên đi thực tập.

3.9. Số tỉnh có sinh viên đến thực tập theo đề tài 5.

3.10. Danh sách các tỉnh và số sinh viên quê ở tỉnh đó, nhóm theo *QUE*.

3.11. Các đề tài có trên 10 sinh viên đăng ký tham gia:

***3.12.** Dùng SQL để biểu thị các phép toán của đại số quan hệ:

a) *R*(*e*)

b) *R*[*X*]

c) *R***S*

d) *R+S*

e) *R&S*

f) *R-S*

3.13. Cho thông tin về những sinh viên sinh trước năm 1973 và quê ở Hải Phòng?

3.14. Cho danh sách các tỉnh có sinh viên đến thực tập?

3.15. Cho biết các địa điểm thực tập xa trường (*KM > 100*) của đề tài số 7?

3.16. Cho thông tin về việc thực tập tại Nha Trang của các sinh viên?

***3.17.** Cho danh sách sinh viên thực tập tại quê nhà?

3.18. Cho thông tin về các đề tài có sinh viên thực tập?

3.19. Cho biết mã của các đề tài không có sinh viên nào tham gia?

3.20. Cho biệt mã của những đề tài có kinh phí 1.5 triệu và những đề tài có kinh phí trên 2 triệu?

3.21. Cho biết mã của những sinh viên dưới 24 tuổi, thực tập khá (có điểm kết quả trên 6)?

***3.22.** Cho danh sách các đề tài có sinh viên học giỏi nhất lớp tham gia.

***3.23.** Cho danh sách các đề tài không có sinh viên học kém nhất lớp tham gia.

***3.24.** Cho danh sách những sinh viên thực tập theo đề tài có kinh phí lớn hơn một phần năm tổng kinh phí cấp cho các đề tài.

***3.25.** Cho danh sách các sinh viên có điểm học tập cao hơn điểm thực tập trung bình của đề tài mã số 4.

***3.26.** Cho quan hệ *R(U)*. Hãy dùng SQL để sinh ra quan hệ rỗng *S(U)*.

Chương 4
PHỤ THUỘC HÀM

TÓM TẮT LÝ THUYẾT

Cho tập thuộc tính *U*. Một *phụ thuộc hàm* (*PTH*) trên *U* là công thức dạng

$$f: X \rightarrow Y; \quad X, Y \subseteq U$$

Nếu $f: X \rightarrow Y$ là một phụ thuộc hàm trên *U* thì ta nói tập thuộc tính *Y* *phụ thuộc* vào tập thuộc tính *X*, hoặc tập thuộc tính *X* *xác định hàm* tập thuộc tính *Y*.

Cho quan hệ *R*(*U*) và một *PTH* $f: X \rightarrow Y$ trên *U*. Ta nói quan hệ *R* *thoả PTH f* và viết R(f), nếu hai bộ tuỳ ý trong *R* giống nhau trên *X* thì chúng cũng giống nhau trên *Y*,

$$R(X\rightarrow Y) \Leftrightarrow (\forall u,v\in R): (u.X=v.X) \Rightarrow (u.Y=v.Y$$

Nếu *Y* không phụ thuộc hàm vào *X* thì ta viết $X \mathrel{!\!\rightarrow} Y$.

Cho tập *PTH F* trên tập thuộc tính *U*. Ta nói quan hệ *R*(*U*) *thoả tập PTH F*, và viết *R*(*F*), nếu *R* thoả mọi *PTH* trong *F*,

$$R(F) \Leftrightarrow (\forall f \in F): R(f)$$

Cho trước tập thuộc tính *U*, ký hiệu *SAT*(*F*) là *tập toàn thể các quan hệ trên U thoả tập PTH F*.

Cho tập $\mathfrak{R}$ các quan hệ trên *U*, ký hiệu *FD*($\mathfrak{R}$) là *tập các PTH trên U thoả trong mọi quan hệ của* $\mathfrak{R}$.

Cho tập *PTH F* trên tập thuộc tính *U*, *Bao đóng* của *F*, ký hiệu F^+ là tập nhỏ nhất các *PTH* trên *U* chứa *F* và thoả các tính chất *F1 - F3* của hệ tiên đề Armstrong A^o sau đây: $\forall X, Y, Z \subseteq U$:

F1. Tính phản xạ: Nếu $X \supseteq Y$ thì $X \rightarrow Y \in F^+$

F2. Tính gia tăng: Nếu $X \rightarrow Y \in F^+$ thì $XZ \rightarrow YZ \in F^+$

F3. Tính bắc cầu: Nếu $X \rightarrow Y \in F^+$ và $Y \rightarrow Z \in F^+$ thì $X \rightarrow Z \in F^+$

Chú ý

Các *PTH* có vế trái chứa vế phải như mô tả trong (*F1*) được gọi là *tầm thường*. Các *PTH* tầm thường thoả trong mọi quan hệ.

Suy dẫn theo tiên đề

Ta nói *PTH f* được *suy dẫn theo tiên đề* (hoặc *suy dẫn logic*) từ tập *PTH F* và ký hiệu là $F \models f$, nếu $f \in F^+$.

$$F \models f \Leftrightarrow f \in F^+$$

Nói cách khác *f* được suy dẫn theo tiên đề từ tập *PTH F* nếu xuất phát từ *F*, áp dụng các luật *F1*, *F2* và *F3* của hệ tiên đề Armstrong sau hữu hạn lần ta sẽ thu được *PTH f*.

Ta viết $F \not\models f$ để biểu thị tập *PTH F không dẫn theo logic* ra được *PTH f*.

Bao đóng của tập thuộc tính

Cho tập *PTH F* trên tập thuộc tính *U* và một tập con các thuộc tính *X* trong *U*. *Bao đóng của tập thuộc tính X*, ký hiệu X^+ là tập thuộc tính

$$X^+ = \{ A \in U \mid X \to A \in F^+ \}$$

Thuật toán tìm bao đóng của một tập thuộc tính

Cho tập *PTH F* trên tập thuộc tính *U* và một tập con các thuộc tính *X* trong *U*. Để xác định bao đóng của tập thuộc tính *X*, X^+ ta xuất phát từ tập *X* và bổ sung dần cho *X* các thuộc tính thuộc vế phải *R* của các *PTH* $L \to R \in F$ thỏa điều kiện $L \subseteq X$. Thuật toán sẽ dừng khi không thể bổ sung thêm thuộc tính nào cho *X*.

Suy dẫn theo quan hệ

Cho tập *PTH F* trên tập thuộc tính *U* và *f* là một *PTH* trên *U*. Ta nói *PTH f được suy dẫn theo quan hệ* từ tập *PTH F* và viết $F \models f$, nếu mọi quan hệ *R*(*U*) thoả *F* thì *R* cũng thoả *f*.

$$F \models f \Leftrightarrow SAT(F) \subseteq SAT(f)$$

Cho tập thuộc tính *U* và tập *PTH F* trên *U*, ta định nghĩa F^* là tập toàn bộ các *PTH f* được suy dẫn theo quan hệ từ tập *PTH F*

$$F^* = \{ f: X \to Y \mid X, Y \subseteq U,\ F \models f \}$$

Ta viết $F \mathrel{!\models} f$ để biểu thị tập *PTH F không dẫn theo quan hệ* ra được *PTH f* .

Định lý (*Tính đủ và chặt của hệ tiên đề Armstrong*)

$$F^+ = F^*$$

Nói cách khác, suy dẫn theo quan hệ và suy dẫn theo tiên đề là một, tức là

$$F \vdash f \Leftrightarrow F \models f$$

Quy ước giản lược

Ta thường viết $X \to Y$ thay vì viết $X \to Y \in F^+$ hoặc $F \models X \to Y$.

Bài toán thành viên

Cho tập thuộc tính *U*, một tập các *PTH F* trên *U* và một *PTH* $X \rightarrow Y$ trên *U*. Hỏi rằng $X \rightarrow Y \in F^+$ hay không ?

Định lý

$$X \rightarrow Y \in F^+ \text{ khi và chỉ khi } Y \subseteq X^+.$$

Lược đồ quan hệ

Lược đồ quan hệ (*LĐQH*) là một cặp $p = (U,F)$, trong đó *U* là tập hữu hạn các thuộc tính, *F* là tập các phụ thuộc hàm trên *U*.

Quy ước

Trong trường hợp không chỉ rõ tập *PTH F*, ta xem *LĐQH* chỉ là một tập hữu hạn các thuộc tính *U*.

BÀI TẬP

4.1. Cho quan hệ *R*(*A*,*B*,*C*,*D*) như sau

R(A	B	C	D)
a	1	x	2
a	1	y	2
b	2	x	1
b	2	y	1

Cho biết các phụ thuộc hàm nào liệt kê dưới đây được thoả trong quan hệ *R*?

f_1: $A \rightarrow A$

f_2: $A \rightarrow B$

f_3: $A \rightarrow C$

f_4: $AC \rightarrow C$

f_5: $A \rightarrow D$

f_6: $D \rightarrow A$

4.2. Viết thuật toán tìm bao đóng của một tập thuộc tính.

4.3. Chứng minh *tính đúng đắn* (chặt) của hệ tiên đề Armstrong A^o, tức là chứng minh

$$F^+ \subseteq F^*$$

theo sơ đồ sau:

Cho tập thuộc tính U, các tập con thuộc tính $X, Y, Z \subseteq U$. Chứng minh với mọi quan hệ $R \in REL(U)$, ta có

F1. Nếu $Y \subseteq X$ thì $R(X \rightarrow Y)$ (*tính phản xạ*)

F2. Nếu $R(X \rightarrow Y)$ thì $R(XZ \rightarrow YZ)$ (*tính gia tăng*)

F3. Nếu $R(X \rightarrow Y)$ và $R(Y \rightarrow Z)$ thì $R(X \rightarrow Z)$ (*tính bắc cầu*)

4.4. Chứng minh *tính đủ* của hệ tiên đề Armstrong A^o, tức là chứng minh

$$F^* \subseteq F^+$$

theo sơ đồ phản chứng sau đây:

Chứng minh rằng nếu f: $X \rightarrow Y \notin F^+$ thì $f \notin F^*$ bằng cách chỉ ra một quan hệ $R(U)$ thoả các PTH trong tập F (thậm chí trong F^+) nhưng không thoả PTH f.

Chú ý: Quan hệ R xây dựng như trên được gọi là quan hệ Armstrong.

4.5. Cho tập thuộc tính U và các tập phụ thuộc hàm F, G trên U. Chứng minh

a) Nếu $F \subseteq G$ thì $SAT(F) \supseteq SAT(G)$

b) $SAT(FG) = SAT(F) \cap SAT(G)$

4.6. Cho tập thuộc tính U và các quan hệ R và S trên U. Chứng minh

a) $FD(R+S) \subseteq FD(R) \cap FD(S)$

b) $R \subseteq S \Rightarrow FD(R) \supseteq FD(S)$

Tìm thí dụ chứng tỏ $FD(R+S) \subset FD(R) \cap FD(S)$

4.7. Cho tập thuộc tính U, tập phụ thuộc hàm F trên U và tập các quan hệ $\mathfrak{R}$ trên U. Chứng minh

a) $F \subseteq FD(SAT(F))$

b) $\mathfrak{R} \subseteq SAT(FD(\mathfrak{R}))$

c) $SAT(FD(SAT(F))) = SAT(F)$

d) $FD(SAT(FD(\mathfrak{R}))) = FD(\mathfrak{R})$

4.8. Sử dụng ba tiên đề Armstrong để chứng minh các tính chất F4 - F11 sau đây:

Với mọi tập con X, Y, Z, V của U và với mọi thuộc tính A trong U:

F4. *Tính tựa bắc cầu:* Nếu $X \rightarrow Y$, $YZ \rightarrow V$ thì $XZ \rightarrow V$

F5. *Tính phản xạ chặt:* $X \rightarrow X$

F6. *Mở rộng vế trái và thu hẹp vế phải:*

Nếu $X \rightarrow Y$ thì $XZ \rightarrow Y \setminus V$

F7. *Cộng tính đầy đủ:*

Nếu $X \rightarrow Y$ và $Z \rightarrow V$ thì $XZ \rightarrow YV$

F8. *Mở rộng vế trái:*

Nếu $X \rightarrow Y$ thì $XZ \rightarrow Y$

F9. *Cộng tính ở vế phải:*

Nếu $X \to Y$ và $X \to Z$ thì $X \to YZ$

F10. *Bộ phận ở vế phải:*

Nếu $X \to YZ$ thì $X \to Y$

F11. *Tính tích luỹ:*

Nếu $X \to YZ$, $Z \to AV$ thì $X \to YZA$

4.9. Cho ánh xạ đóng f: Poset(U) $\to$ Poset(U) trên tập hữu hạn U. Ngoài ba tính chất ($C1$)-($C3$) được sử dụng làm tiên đề cho các ánh xạ đóng, chứng minh các tính chất sau đây của ánh xạ đóng (xem bài 1.35):

$\forall X, Y \in$ Poset(U)

($C4$) $f(XY) \supseteq f(X)f(Y)$

($C5$) $f(X \cap Y) \subseteq f(X) \cap f(Y)$

(C6) $f(f(X)Y) = f(Xf(Y)) = f(XY)$

Chứng minh phép toán lấy bao đóng của tập thuộc tính là một ánh xạ đóng thỏa các tính chất sau:

1. *Tính phản xạ:* $X^+ \supseteq X$
2. *Tính đơn điệu:* nếu $X \subseteq Y$ thì $X^+ \subseteq Y^+$
3. *Tính luỹ đẳng:* $X^{++} = X^+$
4. $(XY)^+ \supseteq X^+Y^+$
5. $(X^+Y)^+ = (XY^+)^+ = (XY)^+$
6. $X \to Y$ khi và chỉ khi $Y \subseteq X^+$
7. $X \to Y$ khi và chỉ khi $Y^+ \subseteq X^+$

8. $X \to X^+$ và $X^+ \to X$
9. $X^+ = Y^+$ khi và chỉ khi $X \to Y$ và $Y \to X$

Tổng kết các tính chất của PTH

Các bài tập 4.10-4.13 dưới đây liên quan đến các tính chất F1-F11 của các PTH.

Cho tập thuộc tính U. Với mọi tập con các thuộc tính X, Y, Z và V trong U và với mọi thuộc tính A trong U ta có:

F1. *Tính phản xạ*: Nếu $Y \subseteq X$ thì $X \to Y$

F2. *Tính gia tăng*: Nếu $X \to Y$ thì $XZ \to YZ$

F3. *Tính bắc cầu*: Nếu $X \to Y$ và $Y \to Z$ thì $X \to Z$

F4. *Tính tựa bắc cầu*: Nếu $X \to Y$, $YZ \to V$ thì $XZ \to V$

F5. *Tính phản xạ chặt*: $X \to X$

F6. *Mở rộng vế trái và thu hẹp vế phải*: Nếu $X \to Y$ thì $XZ \to Y \setminus V$

F7. *Cộng tính đầy đủ*: Nếu $X \to Y$ và $Z \to V$ thì $XZ \to YV$

F8. *Mở rộng vế trái*: Nếu $X \to Y$ thì $XZ \to Y$

F9. *Cộng tính ở vế phải*: Nếu $X \to Y$ và $X \to Z$ thì $X \to YZ$

F10. *Bộ phận ở vế phải*: Nếu $X \to YZ$ thì $X \to Y$

F11. *Tính tích luỹ*: Nếu $X \to YZ$, $Z \to AV$ thì $X \to YZA$

4.10. Chứng minh rằng hệ tiên đề $B°$ sau đây tương đương với hệ tiên đề Armstrong $A°$

$$B° = \{F5, F10, F11\}$$

4.11. Chứng minh rằng hệ tiên đề $S°$ sau đây tương đương với hệ tiên đề Armstrong $A°$

$$S° = \{F1, F4\}$$

4.12. Chứng minh rằng hệ tiên đề D^o sau đây tương đương với hệ tiên đề Armstrong A^o

$$D^o = \{F3, F5, F6, F7\}$$

4.13. Chứng minh rằng hệ tiên đề M^o sau đây tương đương với hệ tiên đề Armstrong A^o

$$M^o = \{F4, F5, F8\}$$

4.14. Cho LĐQH $p = (U,F)$ với $U = ABCDE$, $F = \{ A \rightarrow C, BC \rightarrow D, D \rightarrow E, E \rightarrow A \}$. Tính

a) $(AB)^+$

b) $(BD)^+ - D^+$

4.15. Cho LĐQH $p = (U,F)$ với $U = ABCDEG$, $F = \{ B \rightarrow C, AC \rightarrow D, D \rightarrow G, AG \rightarrow E \}$. Cho biết

a) $AB \rightarrow G \in F^+$?

b) $BD \rightarrow AD \in F^+$?

TÓM TẮT LÝ THUYẾT

Phủ

Cho hai tập PTH F và G trên cùng một tập thuộc tính U. Ta nói F *suy dẫn* ra được G, ký hiệu $F \models G$, nếu $(\forall g \in G)$: $(F \models g)$.

Ta nói F *tương đương* với G, ký hiệu $F \equiv G$, nếu $F \models G$ và $G \models F$.

Nếu $F \equiv G$ ta nói G là một *phủ* của F.

Ký hiệu

- $F\ !\models G$: *F không suy dẫn ra được G.*
- $F\ !\equiv G$: *F và G không tương đương.*

Cho tập PTH F trên tập thuộc tính U và X là tập con của U, ta dùng ký hiệu X_F^+ trong trường hợp cần chỉ rõ bao đóng của tập thuộc tính X lấy theo tập PTH F.

Phủ thu gọn tự nhiên

Cho hai tập PTH F và G trên cùng một tập thuộc tính U. G là *phủ thu gọn tự nhiên* của F nếu

1) G là một phủ của F, và

2) G có dạng thu gọn tự nhiên theo nghĩa sau:

   a) Hai vế trái và phải của mọi PTH trong G rời nhau (không giao nhau)

   b) Các vế trái của mọi PTH trong G khác nhau đôi một.

4.16 Chứng minh rằng nếu F và G là hai tập PTH trên cùng một tập thuộc tính U thì $F \equiv G$ khi và chỉ khi $(\forall X \subseteq U)$: $(X_F^+ = X_G^+)$.

4.17 Cho tập PTH F trên tập thuộc tính U. Chứng minh $F \equiv F^+$.

4.18 Cho các tập con X, Y, Z, V và một thuộc tính A của tập thuộc tính U. Xác định một trong các quan hệ cao nhất các cặp tập PTH sau đây bằng cách đặt dấu $\models$ hoặc dấu $\equiv$ vào chỗ dấu ? Giải thích vì sao.

a) $\{X \to Y, Z \to V\}$? $\{XZ \to YV\}$

b) $\{X \to Y\}$? $\{X \to Y\text{-}X\}$

c) $\{X \to Y\}$? $\{XZ \to Y\}$

d) $\{X \to Y, Y \to Z\}$? $\{X \to Z\}$

e) $\{X \to Y, YZ \to V\}$? $\{XZ \to V\}$

f) $\{X \to Y\}$? $\{XZ \to Y - V\}$

g) $\{X \to Y, X \to Z\}$? $\{X \to YZ\}$

h) $\{X \to YZ\}$? $\{X \to Y\}$

i) $\{X \to YZ, Z \to AV\}$? $\{X \to YZA\}$

4.19 Xây dựng thuật toán tìm phủ thu gọn tự nhiên của tập PTH *F*.

***Phủ không dư*.**

Cho hai tập PTH *F* và *G* trên tập thuộc tính *U*. *G* được gọi là *phủ không dư* của *F* nếu

1) *G* là một phủ của *F*, và

2) *G* có dạng không dư theo nghĩa sau:

$(\forall g \in G)$: $G - \{g\}$ $!\equiv$ G

4.20 Xây dựng thuật toán tìm phủ không dư của tập PTH *F*.

Phủ thu gọn

a) Cho hai tập PTH *F* và *G* trên tập thuộc tính *U*. *G* được gọi là *phủ thu gọn trái* của *F* nếu

1) *G* là một phủ của *F*, và

2) $(\forall X \to Y \in G, \forall A \in X)$: $G - \{X \to Y\} \cup \{(X-\{A\}) \to Y\}$ $!\equiv G$

b) *G* được gọi là *phủ thu gọn phải* của *F* nếu

1) *G* là một phủ của *F*, và

2) $(\forall X \to Y \in G, \forall A \in Y)$: $G - \{X \to Y\} \cup \{X \to (Y-\{A\})\}$ $!\equiv G$

c) *G* được gọi là *phủ thu gọn* của *F* nếu *G* đồng thời là phủ thu gọn trái và thu gọn phải của *F*.

4.21. Xây dựng thuật toán tìm phủ thu gọn trái của tập PTH *F*.

4.22. Xây dựng thuật toán tìm phủ thu gọn phải của tập PTH *F*.

4.23. Xây dựng thuật toán tìm phủ thu gọn của tập PTH *F*.

4.24. Xây dựng thí dụ chứng tỏ với tập PTH *F* sau khi thực hiện sơ đồ dưới đây thì *F* vẫn chưa phải là phủ thu gọn phải:

a) Thu gọn phải *F*

b) Thu gọn trái *F*.

Phủ tối tiểu (Ullman J.)

Cho hai tập PTH *F* và *G* trên tập thuộc tính *U*. *G* được gọi là *phủ tối tiểu* của *F* nếu

1) *G* là một phủ thu gọn của *F*,

2) Vế phải của mọi PTH trong G chỉ chứa một thuộc tính,

4.25. Xây dựng thuật toán tìm phủ tối tiểu của tập PTH *F*.

Phụ thuộc đầy đủ

Tập thuộc tính $Y \subseteq U$ được gọi là *phụ thuộc đầy đủ* vào tập thuộc tính $X \subseteq U$, và được ký hiệu là $X \mathrel{+\!\!\rightarrow} Y$ nếu

1) $X \rightarrow Y$, và

2) $(\forall A \in X)$: $X\text{-}\{A\} \;!\!\!\rightarrow Y$

4.26. Chứng minh rằng với mọi tập PTH *F* trên *U* luôn tồn tại một phủ *G* của *F* sao cho mọi PTH trong *G* đều là phụ thuộc đầy đủ.

4.27. Xây dựng thuật toán tìm một phủ đầy đủ của tập PTH *F*.

Phụ thuộc bắc cầu

Tập thuộc tính $Y \subseteq U$ được gọi là *phụ thuộc bắc cầu* vào tập thuộc tính $X \subseteq U$, và được ký hiệu là $X \,\%\!\!\rightarrow Y$ nếu

$$(\exists Z \subseteq U): Y\text{-}Z \neq \varnothing,\ X \to Z, Z\ ! \to X, Z \to Y.$$

Nếu $X \to Y$ và Y không phụ thuộc bắc cầu vào X thì ta nói *Y phụ thuộc trực tiếp* vào X và ký hiệu là $X \overset{*}{\to} Y$.

Phụ thuộc mạnh, yếu và đối ngẫu

Cho tập thuộc tính U và hai tập con các thuộc tính $X, Y \subseteq U$.

Quan hệ $R(U)$ thỏa *phụ thuộc mạnh* $X\ (s){\to}\ Y$ nếu với hai bộ tùy ý u và v trong R giống nhau tại một thuộc tính A nào đó trong X thì hai bộ đó giống nhau trên Y.

$$\forall u, v \in R: (\exists A \in X: u.A = v.A \Rightarrow u.Y = v.Y)$$

Quan hệ $R(U)$ thỏa *phụ thuộc yếu* $X\ (w){\to}\ Y$ nếu với hai bộ tùy ý u và v trong R giống nhau trên X thì hai bộ đó giống nhau tại một thuộc tính B nào đó của Y.

$$\forall u, v \in R: (u.X = v.X) \Rightarrow (\exists B \in Y: u.B = v.B)$$

Quan hệ $R(U)$ thỏa *phụ thuộc đối ngẫu* $X\ (d){\to}\ Y$ nếu với hai bộ tùy ý u và v trong R giống nhau tại một thuộc tính A nào đó của X thì hai bộ đó giống nhau tại một thuộc tính B nào đó của Y.

$$\forall u, v \in R: (\exists A \in X: u.A = v.A) \Rightarrow (\exists B \in Y: u.B = v.B)$$

4.28. Cho quan hệ R trên U và các tập con thuộc tính X, Y của U. Chứng minh:

a) $R(X\ (s){\to}\ Y) \Rightarrow R(X \to Y)$

b) $R(X \to Y) \Rightarrow R(X\ (w){\to} Y)$

c) $R(X\ (d){\to}\ Y) \Rightarrow R(X\ (w){\to} Y)$

d) $R(X\ (s){\to}\ Y) \Rightarrow R(X\ (d){\to}\ Y)$

4.29. Với mọi tập con thuộc tính X, Y, Z và thuộc tính A của tập thuộc tính U. Điền các ký hiệu f, s, w hoặc d thay cho dấu ? để khẳng định các tính chất cao nhất của các phụ thuộc: *hàm*, *mạnh*, *yếu* hoặc *đối ngẫu* sau đây:

1. *Tính phản xạ*: Nếu $X \supseteq Y$ thì $X (?)\rightarrow Y$
2. *Tính gia tăng*: Nếu $X (?)\rightarrow Y$ thì $XZ\ (?)\rightarrow YZ$
3. *Tính bắc cầu*: Nếu $X (?)\rightarrow Y$ và $Y (?)\rightarrow Z$ thì $X\ (?)\rightarrow Z$
4. *Tính tựa bắc cầu*: Nếu $X (?)\rightarrow Y$, $YZ (?)\rightarrow V$ thì $XZ(?)\rightarrow V$
5. *Tính phản xạ chặt*: $X (?)\rightarrow X$
6. *Mở rộng vế trái và thu hẹp vế phải: Nếu* $X (?)\rightarrow Y$ *thì* $XZ (?)\rightarrow Y \setminus V$
7. *Cộng tính đầy đủ*: Nếu $X (?)\rightarrow Y$ và $Z (?)\rightarrow V$ thì $XZ (?)\rightarrow YV$
8. *Mở rộng vế trái*: Nếu $X (?)\rightarrow Y$ thì $XZ (?)\rightarrow Y$
9. *Cộng tính ở vế phải*: Nếu $X (?)\rightarrow Y$ và $X (?)\rightarrow Z$ thì $X (?)\rightarrow YZ$
10. *Bộ phận ở vế phải*: Nếu $X (?)\rightarrow YZ$ thì $X (?)\rightarrow Y$.
11. *Tính tích luỹ*: Nếu $X (?)\rightarrow YZ$, $Z (?)\rightarrow AV$ thì $X (?)\rightarrow YZA$

Khóa của lược đồ quan hệ

Cho LĐQH $p = (U,F)$. Tập thuộc tính $K \subseteq U$ được gọi là *khoá* của LĐ p nếu

(i) $K^+ = U$

(ii) $\forall A \in K$: $(K - \{A\})^+ \neq U$

Hai điều kiện trên tương đương với

(i') $K \rightarrow U$

(ii') $\forall A \in K$: $(K - \{A\})\ !\rightarrow\ U$

Thuộc tính $A \in U$ được gọi là *thuộc tính khoá* (*nguyên thuỷ hoặc cơ sở*) nếu *A* có trong một khoá nào đấy. *A* được gọi là thuộc tính *không khoá* (*phi nguyên thuỷ hoặc thứ cấp*) nếu *A* không có trong bất kỳ khoá nào.

Nếu *K* thoả điều kiện (i) thì *K* được gọi là một *siêu khoá*.

Chú ý: Trong một số tài liệu thuật ngữ *khoá* được dùng theo nghĩa *siêu khoá* và thuật ngữ *khoá tối tiểu* được dùng theo nghĩa *khoá*.

BÀI TẬP

4.30. Xây dựng thuật toán tìm một khóa của LĐQH.

4.31. Cho LĐQH *p*. Biết p có một khóa *K*. Hãy xây dựng thuật toán tìm một khóa thứ hai *M* của *p*. Nếu *p* không có khóa thứ hai, thuật toán cho kết quả là một tập rỗng.

4.32. Xây dựng một LĐQH có 5 thuộc tính ABCDE, mỗi thuộc tính là một khóa.

4.33. LĐQH có 5 thuộc tính có thể có tối đa bao nhiêu khóa. Cho thí dụ.

4.34. Xây dựng một LĐQH có 5 thuộc tính ABCDE và chỉ có một khóa duy nhất.

4.35. (*Nguyễn Xuân Huy*) Cho *K* là một khóa của LĐQH $p = (U,F)$. Chứng minh rằng với mọi tập con *X* của *K* ta có: $X^+ \cap K = X$.

4.36. (*Lê Văn Bào, Nguyễn Xuân Huy, Hồ Thuần*) Cho LĐQH $p = (U,F)$. Gọi *M* là giao của các khóa của *p*. Chứng minh rằng

$$M = U - \bigcup_{L \to R \in F} (R - L)$$

4.37. (*Lê Văn Bào, Hồ Thuần*)Cho LĐQH $p = (U,F)$. Gọi M là giao của các khóa của p. Chứng minh rằng p có một khóa duy nhất khi và chỉ khi $M^+ = U$.

4.38. Cho tập thuộc tính U với n phần tử. Chứng minh rằng có thể xây dựng tập PTH F sao cho LĐQH $p = (U,F)$ có

$$C_n^{\lfloor \frac{n}{2} \rfloor} \text{ khóa,}$$

trong đó toán tử $\lfloor x \rfloor$ cho ta cận nguyên dưới của số nguyên x, C_n^m là tổ hợp chặp m của n phần tử.

4.39. Tìm tập thuộc tính nguyên thuỷ của LĐQH sau:

$$p = (U,F),\ U = ABCDE,$$
$$F = \{ AB \rightarrow C, AD \rightarrow B, B \rightarrow D \}.$$

4.40. Khẳng định hoặc bác bỏ tính đúng của các mệnh đề sau:

a) $K \subseteq U$ là một khoá khi và chỉ khi $K \overset{+}{\rightarrow} U$ (phụ thuộc đầy đủ)

b) Hai khoá khác nhau của một LĐQH không giao nhau.

c) Hai khoá khác nhau của một LĐQH khổng bao nhau.

d) Mọi LĐQH đều có ít nhất một khoá.

e) Tồn tại một LĐQH không có khoá nào.

f) Số khoá của một LĐQH không thể lớn hơn số thuộc tính.

g) U không thể là khoá của LĐQH (U,F).

h) Mọi LĐQH không thể có hai khoá đơn tức là khoá chỉ gồm một thuộc tính.

Chương 5

CHUẨN HÓA

TÓM TẮT LÝ THUYẾT

Phép tách

Cho lược đồ quan hệ $p = (U,F)$. Một *phép tách* trên tập thuộc tính U là một họ các tập con của U, $\rho = (X_1, X_2,...,X_k)$ thỏa tính chất: $\bigcup_{i=1}^{k} X_i = U$.

Phép tách $\rho = (X_1, X_2,...,X_k)$ trên tập thuộc tính U được gọi là *không tổn thất* (hoặc *không mất thông tin)* đối với tập PTH F nếu

$$\forall\ R(U) \in \text{SAT}(F): R[X_1] * R[X_2] * ... * R[X_k] = R$$

Ngược lại, nếu không tồn tại đẳng thức thì ta gọi ρ là phép tách *tổn thất*.

Kiểm tra tính tổn thất của phép tách bằng kỹ thuật bảng

Input

- LĐQH $p = (U,F)$
- Phép tách $\rho = (X_1, X_2,...,X_k)$

Output

- True, nếu ρ là một phép tách không tổn thất
- False, ngoài ra.

Method

1. *Khởi trị:* Lập bảng T với các cột là các thuộc tính trong U và k dòng, mỗi dòng ứng với một thành phần của X_i trong ρ: Dòng i chứa các ký hiệu phân biệt (KHPB) a_j ứng với các thuộc tính A_j trong X_i và các ký hiệu không phân biệt (KHKPB) b_{ij} ứng với các thuộc tính A_j trong U-X_i. Chú ý rằng mọi KHPB trong cột j của T là giống nhau và bằng a_j còn mọi KHKPB trong toàn bảng T là khác nhau.

2. *Sửa bảng:* Lặp đến khi bảng T không còn thay đổi:

2.1. Vận dụng các F-luật để biến đổi bảng như sau:

Với mỗi PTH $L \rightarrow R$ trong F, nếu trong bảng T có chứa hai dòng u và v giống nhau trên L thì sửa các ký hiệu của chúng cho giống nhau trên mọi cột A trong T như sau:

a) nếu $u.A = v.A$: không sửa,

b) nếu chỉ một trong hai ký hiệu $u.A$ hoặc $v.A$ là KHPB thì sửa mọi xuất hiện trong bảng của KHKPB thành KHPB đó,

c) nếu cả hai ký hiệu $u.A$ và $v.A$ đều là KHKPB thì sửa mọi xuất hiện trong bảng của ký hiệu có chỉ số thứ nhất lớn hơn thành ký hiệu thứ hai.

3. *Kết luận:*

Nếu trong bảng chứa một dòng toàn KHPB thì return True

nếu không return False.

end.

BÀI TẬP

5.1. Chứng minh rằng nếu $X \rightarrow Y$ và $X \cap Y = \varnothing$ thì phép tách (XY, U-Y) là không tổn thất.

5.2. Dùng kỹ thuật bảng để kiểm tra tính tổn thất của các phép tách sau:

a) $p = (U,F)$, $U = ABCD$,

$F = \{ A \rightarrow B, AC \rightarrow D \}$

$\rho = (AB, ACD)$.

b) $p = (U,F)$, $U = ABCDE$,

$F = \{ A \rightarrow y, B \rightarrow C, C \rightarrow D, DE \rightarrow C, CE \rightarrow A \}$

$\rho = (AD, AB, BE, CDE)$.

Các dạng chuẩn

LĐQH $p = (U,F)$ được gọi là

a) ở dạng chuẩn 1 (1*NF*) nếu mọi thuộc tính trong U đều không phải là thuộc tính phức hợp,

b) ở dạng chuẩn 2 (2*NF*) nếu p ở 1*NF* và mọi thuộc tính không khoá đều phụ thuộc đầy đủ vào mọi khoá,

c) ở dạng chuẩn 3 (3*NF*) nếu p ở 1*NF* và mọi thuộc tính không khoá đều phụ thuộc trực tiếp vào mọi khoá,

d) ở dạng chuẩn 3 mạnh hoặc Boyce-Codd (3*SNF*, *BCNF*) nếu p ở 1*NF* và mọi thuộc tính đều phụ thuộc trực tiếp vào mọi khoá,

e) ở dạng chuẩn 3 mạnh hoặc Boyce-Codd (3*SNF*, *BCNF*) nếu p ở 1*NF* và mọi PTH không tầm thường $X \rightarrow Y$ đều cho ta X là một siêu khóa.

Thuật toán chuẩn hoá 3NF không tổn thất và bảo toàn PTH

Algorithm 3NF

Input: LĐQH $p = (U, F)$

Output: Các LĐQH 3NF

$(U_1, K_1), (U_2, K_2), \ldots, (U_s, K_s)$ thoả

a) $\forall R \in REL(U): R[U_1]*R[U_2]*\ldots*R[U_s] = R$

b) $K_1, K_2, \ldots, K_s$ là các khoá của các lược đồ tương ứng

Method

1. Tìm một khoá K của p

2. Tìm một phủ tối tiểu

$$G = \{K_1 \to A_1, K_2 \to A_2, \ldots, K_m \to A_m\}$$

của F.

3. Ghép các PTH có cùng vế trái trong G để thu được phủ

$$G = \{K_1 \to X_1, K_2 \to X_2, \ldots, K_s \to X_s\}$$

4. Xét phép tách $\rho = (K_1X_1, K_2X_2, \ldots, K_sX_s)$. Nếu khoá K không có mặt trong thành phần nào của ρ thì thêm thành phần K vào ρ.

5. return $(K_1X_1, K_1), (K_2X_2, K_2), \ldots, (K_sX_s, K_s)$

End 3NF.

5.3. Một lược đồ quan hệ ở dạng chuẩn 3 và có một khoá duy nhất có thể ở dạng chuẩn Boyce-Codd không? Vì sao?

5.4. Xác định và giải thích dạng chuẩn cao nhất của LĐQH sau:

$p = (U, F)$; $U = ABCD$, $F = \{ A \to C, D \to B, C \to ABD \}$

5.5. Chứng minh rằng một LĐQH ở dạng chuẩn 3 thì đồng thời ở dạng chuẩn 2.

5.6. Chứng minh rằng một LĐQH ở dạng chuẩn BC thì đồng thời ở dạng chuẩn 3.

5.7. Chuẩn hoá 3*NF* CSDL **Thực tập**:

SV(*SV#*, *HT*, *NS*, *QUE*, *HL*)

DT(*DT#*, *TDT*, *CN*, *KP*)

SD(*SV#*, *DT#*, *NTT*, *KM*, *KQ*)

với các tập PTH sau:

$F_{SV} = \{SV\# \to HT, NS, QUE, HL\}$

$F_{DT} = \{DT\# \to TDT, CN, KP\}$

$F_{SD} = \{SV\#, DT\# \to NTT, KQ ; NTT \to KM\}$

5.8. Chuẩn hoá 3*NF* LĐQH $p = (U,F)$ sau:

y = MLTGSDP,

$y = \{ M \to T, GP \to M, GT \to P, MS \to D, GS \to P\}$

với ngữ nghĩa sau:

M: Môn học chuyên đề

L: Lớp chuyên đề

T: Thày - giáo viên phụ trách chuyên đề

G: Giờ học chuyên đề

S: Sinh viên theo học chuyên đề

D: Số đăng ký của sinh viên trong chuyên đề đó

P: Phòng học dành cho chuyên đề

$M \rightarrow T$: Mỗi chuyên đề có một thày phụ trách

$GP \rightarrow M$: Tại mỗi thời điểm, mỗi phòng học được dành cho không quá một môn

$GT \rightarrow P$: Tại mỗi thời điểm, mỗi thày dạy trong không quá một phòng học

$MS \rightarrow D$: Mỗi sinh viên tham gia chuyên đề nào thì được cấp một mã số ghi danh theo chuyên đề đó

$GS \rightarrow P$: Tại mỗi thời điểm, mỗi sinh viên có mặt trong không quá một phòng học

5.9. Chứng minh sự tương đương của hai định nghĩa *d* và *e* về BCNF.

MỘT SỐ ĐỀ THI

ĐỀ THI TUYỂN CAO HỌC ĐẠI HỌC THÁI NGUYÊN

NĂM 2002

Môn: Cơ sở công nghệ thông tin,

Thời gian 180 phút, không tham khảo tài liệu

Đề bài gồm 6 câu, trong đó 2 câu 5 và 6 có nội dung về cơ sở dữ liệu quan hệ

Câu 5. a) Định nghĩa quan hệ, phụ thuộc hàm.

b) Phát biểu hệ tiên đề Armstrong và tính đúng của hệ tiên đề này.

c) Định nghĩa lược đồ quan hệ và khoá của lược đồ quan hệ.

d) Phát biểu bài toán thành viên đối với lược đồ quan hệ và kết quả liên quan tới bài toán thành viên.

Câu 6. a) Trình bày thuật toán tính bao đóng của một tập thuộc tính trong một lược đồ quan hệ.

b) Trình bày thuật toán tìm một khoá của lược đồ quan hệ.

c) Cho lược đồ quan hệ s = (R,F) với R = abcdefg, F = { abc → de, bcd → g, abf → eg, ce → fg }

Tìm một khoá của lược đồ s.

ĐỀ THI TUYỂN NGHIÊN CỨU SINH
VIỆN CÔNG NGHỆ THÔNG TIN
NĂM 2002

Đề số 1

Thời gian 180 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, $XY = X \cup Y$. Thuật ngữ khoa được hiểu là khoá tối tiểu, LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1. Định nghĩa LĐQH. Định nghĩa bao đóng của tập thuộc tính trong LĐQH. Thuật toán tìm bao đóng của tập thuộc tính trong LĐQH.

Câu 2. Cho LĐQH p = (U, F) với tập thuộc tính U = ABCDEGH và tập phụ thuộc hàm $F = \{ CD \to H,\ E \to B,\ D \to G,\ BH \to E,\ CH \to DG,\ C \to A \}$.

a. *Tìm tập M là giao của toàn bộ các khoá của p. Cho biết p có đúng 1 khoá hay không?*
b. *Tập ABD có phải là khoá của p không? Vì sao?*
c. *Tập CH có phải là khoá của p không? Vì sao?*
d. *Tính $Z = (X^+ Y)^+ \cap (K^+ - Y)$ biết X = CD, Y = CH, K là một siêu khoá của p.*

Câu 3. Phép phân tách một LĐQH: định nghĩa phép phân tách và phép phân tách không mất thông tin. Thuật toán bảng kiểm tra một phép phân tách không mất thông tin. Cho LĐQH p như trong câu 2. Vận dụng thuật toán kiểm tra phép phân tách w dưới đây có mất thông tin hay không?

w = [ABCDE,BCH,CDEGH]

Câu 4. Định nghĩa các dạng chuẩn 2NF, 3NF và BCNF của LĐQH. Sơ đồ tương quan giữa các dạng chuẩn. Xác định dạng chuẩn cao nhất của LĐQH h sau đây:

$h = (U, F);\ U = ABCD,\ F = \{ D \rightarrow B,\ C \rightarrow A,\ B \rightarrow ACD \}$

Giải thích vì sao?

ĐỀ THI TUYỂN NGHIÊN CỨU SINH
VIỆN CÔNG NGHỆ THÔNG TIN
NĂM 2002

Đề số 2

Thời gian 180 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, $XY = X \cup Y$. Thuật ngữ khoá được hiểu là khoá tối tiểu, LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1. Định nghĩa PTH và LĐQH. Phát biểu bài toán thành viên trên LĐQH. Điều kiện cần và đủ liên quan đến bài toán thành viên.

Câu 2. Định nghĩa khoá của LĐQH. Thuật toán tìm 1 khoá của LĐQH.

Câu 3. Định nghĩa thuộc tính khoá (thuộc tính cơ bản hay nguyên thuỷ), thuộc tính không khoá (thuộc tính thứ cấp). Cho LĐQH

s = (U, F), trong đó U = ABCD, F = {AD → BC, B → A, D → C}.

a) Tìm các khoá của s.

b) Cho biết C có phải là thuộc tính khoá hay không?

Câu 4. Định nghĩa các dạng chuẩn 2NF, 3NF và BCNF của LĐQH. Sơ đồ tương quan giữa các dạng chuẩn. Xác định dạng chuẩn cao nhất của LĐQH h sau:

$h = (U, F)$; $U = ABCD$, $F = \{CD \to B, A \to C, B \to ACD\}$

Giải thích vì sao?

Câu 5. Cho LĐQH $p = (U, F)$ với tập thuộc tính $U = ABCDEHG$ và tập phụ thuộc hàm $F = \{DE \to G, H \to C, E \to A, CG \to H, DG \to EA, D \to B\}$.

a) Tìm tập M là giao của toàn bộ các khoá của p. Cho biết p có đúng 1 khoá hay không?

b) Tìm 1 khoá của p.

c) Tập BCE có phải là khoá của p không? Vì sao?

d) Hãy thêm hoặc bớt 1 phụ thuộc hàm cho F để LĐQH có đúng 1 khoá.

ĐỀ THI TUYỂN NGHIÊN CỨU SINH

ĐẠI HỌC BÁCH KHOA HÀ NỘI

NĂM 2002

(Tuyển nghiên cứu sinh học trong nước và nước ngoài)

Thời gian làm bài 180 phút, kể từ khi phát xong đề thi.

Thí sinh không được tham khảo tài liệu

Bài 1. Định nghĩa: lược đồ quan hệ, khoá của lược đồ quan hệ. Thuật toán tìm một khoá của lược đồ quan hệ.

Bài 2. Cho lược đồ quan hệ p = (U,F) với tập thuộc tính U = ABCDEH và tập phụ thuộc hàm $F = \{BC \to E,\ D \to A,\ C \to A,\ AE \to D,\ BE \to CH\}$

- a) *Tìm một khoá K của lược đồ p.*
- b) *Ngoài khoá K, lược đồ p còn khoá nào khác không? Vì sao?*
- c) *Tập BCH có phải là khoá của p không? Vì sao?*
- d) *Tập BD có phải là khoá của p không? Vì sao?*
- e) *Tính* $Z = (X^+ \cup Y)^+ \cap K^+ - (X \cup Y)$ *với* $X = AB$, $Y = D$, *K là một siêu khoá của p.*
- f) *Hãy thêm cho F một phụ thuộc hàm để p có đúng một khoá. Giải thích cách làm.*

Bài 3. Một lược đồ quan hệ ở dạng chuẩn 3 và có một khoá duy nhất có thể ở dạng chuẩn Boyce-Codd không? Vì sao?

ĐỀ THI TUYỂN CAO HỌC NGÀNH GIS NĂM 2003

ĐẠI HỌC BÁCH KHOA TP HỒ CHÍ MINH

Môn: Cơ sở dữ liệu, đề số 1

Thời gian 120 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, XY = X∪Y. Các từ viết tắt: LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1.

Phép toán trừ hai quan hệ: Định nghĩa. Thuật toán.

Câu 2.

a) *Định nghĩa LĐQH.*

b) *Định nghĩa bao đóng của tập thuộc tính trong LĐQH.*

c) *Thuật toán tìm bao đóng của tập thuộc tính trong LĐQH.*

Câu 3. *Cho LĐQH p = (U, F) với tập thuộc tính U = ABCDE và tập PTH F = { DE → A, B → C, E → AD }.*

a) Tìm một khoá của lược đồ p.

b) Tập BCE có phải là khoá của p không? Vì sao?

c) Tập AD có phải là khoá của p không? Vì sao?

d) Lược đồ p còn khoá nào nữa không? Vì sao?

e) Tính $Z = (X^+Y)^+ \cap (K^+ - Y)$ biết $X = DE$, $Y = AD$, K là một siêu khoá của p.

f) Có thể thêm vào F một PTH để thu được lược đồ có đúng 2 khoá không. Giải thích cách làm?

ĐỀ THI TUYỂN CAO HỌC NGÀNH GIS NĂM 2003

ĐẠI HỌC BÁCH KHOA TP HỒ CHÍ MINH

Môn: Cơ sở dữ liệu, đề số 2

Thời gian 120 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, XY = X∪Y. Các từ viết tắt: LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1.

Phép toán kết nối tự nhiên hai quan hệ: Định nghĩa. Thuật toán.

Câu 2.

a) Định nghĩa LĐQH.

b) Định nghĩa khoá của LĐQH.

c) Thuật toán tìm một khoá của LĐQH.

Câu 3. *Cho LĐQH p = (U, F) với tập thuộc tính U = ABCDE và tập PTH F = { EA → B, C → D, A → BE }.*

a) Tìm một khoá của lược đồ p.

b) Tập CDA có phải là khoá của p không? Vì sao?

c) Tập BE có phải là khoá của p không? Vì sao?

d) Lược đồ p còn khoá nào nữa không? Vì sao?

e) Tính $Z = (X^+ Y)^+ \cap K^+$ biết $X = AE$, $Y = BE$, K là một khoá của p.

f) Có thể thêm vào F một PTH để thu được lược đồ có đúng 2 khoá không. Giải thích cách làm?

ĐỀ THI TUYỂN CAO HỌC NGÀNH GIS NĂM 2003

ĐẠI HỌC BÁCH KHOA TP HỒ CHÍ MINH

Môn: Cơ sở dữ liệu, đề số 3

Thời gian 120 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, $XY = X \cup Y$. Các từ viết tắt LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1.

Phép toán lấy giao hai quan hệ: Định nghĩa. Thuật toán.

Câu 2.

a). Định nghĩa bao đóng của tập thuộc tính trong LĐQH.

b) Thuật toán tìm bao đóng của tập thuộc tính trong LĐQH.

Câu 3. *Cho LĐQH $p = (U, F)$ với tập thuộc tính $U = ABCDE$ và tập PTH $F = \{ AB \rightarrow C,\ D \rightarrow E,\ B \rightarrow CA \}$.*

a) Tìm một khoá của lược đồ p.

b) Tập DEB có phải là khoá của p không? Vì sao?

c) Tập CA có phải là khoá của p không? Vì sao?

d) Lược đồ p còn khoá nào nữa không? Vì sao?

e) Tính $Z = (X^+ Y)^+ \cap K^+$ biết $X = AB$, $Y = CA$, K là một siêu khoá của p.

f) Có thể thêm vào F một PTH để thu được lược đồ có đúng 2 khoá không. Giải thích cách làm?

2

BÀI GIẢI

Bài giải Chương 1

QUAN HỆ VÀ ĐẠI SỐ QUAN HỆ

Cơ sở dữ liệu minh họa: CSDL Thực tập

SV(SV#, HT, NS, QUE, HL)

DT(*DT#, TDT, CN, KP*)

SD(SV#, DT#, NTT, KM, KQ)

SV **- tên quan hệ sinh viên**

SV# - mã số sinh viên

HT - họ và tên sinh viên

NS - năm sinh của sinh viên

QUE - quê (tỉnh)

HL - học lực thể hiện qua điểm trung bình

DT **- tên quan hệ đề tài**

DT# - mã số đề tài

TDT - tên đề tài

CN - họ và tên chủ nhiệm đề tài

KP - kinh phí cấp cho đề tài (triệu đồng).

SD **- tên quan hệ sinh viên - đề tài**

SV# - mã số sinh viên

DT# - mã số đề tài mà sinh viên đó tham gia

NTT - nơi thực tập để triển khai đề tài (tỉnh)

KM - khoảng cách từ nơi thực tập đến trường

KQ - kết quả thực tập theo đề tài đã chọn

Giả thiết là một sinh viên có thể tham gia nhiều đề tài, mỗi đề tài sinh viên đó thực tập tại một địa điểm.

1.1.

```
a) R[AB]=(A B)
          a 1
          b 2
          a 4
          c 5
```

```
b) R(3-B+D>1) =
   (A B C D)
    a 1 x 2
    a 1 y 2
    b 2 x 1
    b 2 y 1
    c 5 y 7
```

```
c) R(B<4)+R(D>3) =
   (A B C D)
    a 1 x 2
    a 1 y 2
    b 2 x 1
    b 2 y 1
    c 5 y 7
```

```
d) R(B>=1&B<=5)=
   (A B C D)     = R
    a 1 x 2
    a 1 y 2
    b 2 x 1
    b 2 y 1
    a 4 x 2
    c 5 y 7
```

```
e) R*S[C]
=(A B C D)  = R
   a 1 x 2
   a 1 y 2
   b 2 x 1
   b 2 y 1
   a 4 x 2
   c 5 y 7
```

```
f) R(B<4)-R(D>3)
=  (A B C D)
    a 1 x 2
    a 1 y 2
    b 2 x 1
    b 2 y 1
```

```
h) R:S =(A B)
       = Ø
```

```
g) R(B<4)& R(D>3)
=(A B C D)
= Ø
```

1.2. *SV(NS<1983 & QUE="Hai Phòng")*

1.3. *SD[NTT]*

1.4. *SD(DT#=7 & KM > 100)[NTT]*

1.5. *SD(NTT="Nha Trang")*

***1.6.** *((SV[SV#,HT,QUE])*(SD[SV#,NTT]))(QUE=NTT)[HT]*

1.7. *DT*SD[DT#]*

1.8. *(DT[DT#])-(SD[DT#])*

1.9. *DT(KP=1.5 | KP>2)[DT#]*

1.10. *SV(2003-NS<20)[SV#]*(SD(KQ>7)[SV#])*

***1.11.** *SD[DT#,NTT] : (SD(DT#=1)[NTT])*

***1.12.** *(DT[DT#,TDT]*(SD[DT#,NTT]:SD[NTT]))[TDT]*

***1.13.** *(SV*(SD*(DT(KP*5>Sum(DT,KP))[DT#]))[SV#])[HT]*

***1.14.** *SV(HL>Avg(SD(DT#=4),KQ))[HT]*

1.15. Suy trực tiếp từ định nghĩa

1.16. Thí dụ chứng tỏ các phép toán *trừ* và *chia không có tính giao hoán.*

```
R =(A)    S(A)
    a        a
    b
R-S =(A)
      b
S-R = ∅
```

```
R =(A B)   S(B)
    a 1       1
    b 2
R:S =(A)
      a
S:R = ∅
```

1.17.

a) $R(e \,\&\, h) = R(h \,\&\, e)$, vì $e\&h = h\&e$.

b) $R(e \,\&\, h) = R(e) \,\&\, R(h) \subseteq R(e)$, vì $e \,\&\, h \Rightarrow e$.

c) $R(e \,\&\, h) = R(h) \,\&\, R(e) \subseteq R(h)$, vì $e \,\&\, h \Rightarrow h$.

d) $R(e \,\&\, h) = R(e)(h)$, vì $\forall\, t \in R$: $Sat(t, e\&h) \Leftrightarrow Sat(t,e) \,\&\, Sat(t,h)$.

e) $R(e \mid h) = R(h \mid e)$, vì $\forall\, t \in R$: $Sat(t, e \mid h) \Leftrightarrow Sat(t,e) \mid Sat(t,h)$.

f) $R(e \mid h) = R(e) + R(h)$, xem e.

g) $R(!\, e) = R\text{-}R(e)$, vì $\forall t \in R$: $Sat(t, !\, e) \Leftrightarrow !\, Sat(t,e)$

h) $R(True) = R$, vì $\forall t \in R$: $Sat(t, True)$

i) $R(False) = \varnothing$, vì $\forall t \in R$: $!\, Sat(t, False)$

***1.19.** $\forall t \in R : S \Leftrightarrow t \in R[M] \,\&\, t{*}S \subseteq R \Leftrightarrow \exists\, u \in R$: $u.M = t \,\&\, t{*}S \subseteq R \Leftrightarrow t{=}u.M \in R[M] \,\&\, u \notin R[M]{*}S - R \Leftrightarrow t \in R[M] \,\&\, t \notin (R[M]{*}S - R)[M] \Leftrightarrow t \in R[M] - (R[M]{*}S - R)[M]$.

1.20.

Phép toán	Ký hiệu	Tính đóng
chọn	( )	có
chiếu	[ ]	có
kết nối tự nhiên	*	có
cộng	+	không *
giao	&	không *
trừ	-	không *
chia	:	không

(*) Các phép toán quan hệ cộng, giao, trừ không đóng với hai quan hệ không tương thích.

1.21. Suy trực tiếp từ định nghĩa.

1.22. a) Thí dụ chứng tỏ phép toán trừ không có tính kết hợp:

Cho tập thuộc tính tuỳ ý $U \neq \varnothing$. Chọn 3 bộ khác nhau tuỳ ý trên U là t, u, v. Xây dựng các quan hệ $R = \{ t, u \}$; $S = R$; $T = \{ u, v\}$. Ta có

R-(S-T) = R-(R-T) = $\{u\} \neq$ (R-S)-T = (R-R)-T = $\varnothing$ - T = $\varnothing$.

a) Thí dụ chứng tỏ phép chia không có tính kết hợp:

Xét các quan hệ $R(A,B)$, $S(B,C)$ và $T(C)$. Ta có

R:(S:T) cho ta quan hệ có thuộc tính A, trong khi biểu thức (R:S):T không có nghĩa.

1.26.

Phép toán	Ký hiệu	Nở/Co ngang	Nở/Co dọc
Chọn	( )	=	-
Chiếu	[ ]	-	-
Kết nối tự nhiên	*	+	+ -
Cộng	+	=	+
Giao	&	=	-
Trừ	-	=	-
Chia	:	-	-

***1.29.** Chẳng hạn, R(False) hoặc $R(A \neq A)$ với A là một thuộc tính bất kỳ trong U.

*1.33. Thí dụ chứng tỏ $R[X_1] * R[X_2] * \ldots * R[X_k] \neq R$

R = (A B C)	P = R[AB] = (A B)	Q = R[BC] = (B C)	P*Q = (A B C) ≠ R
a 1 x	a 1	1 x	a 1 x
b 1 y	b 1	1 y	a 1 y
			b 1 x
			b 1 y

*1.36. *SV* * *DT* cho ta tích Descartes.

Bài giải Chương 2

CÁC THAO TÁC TRÊN BỘ VÀ QUAN HỆ

2.1.

```
Algorithm  Selection
Format: P = R(e)
Input:          - Quan hệ R(U)
                - Biểu thức chọn e trên U.
Output: - Quan hệ
        P(U) = R(e) = {t∈R|Sat(t,e)}
Method
// Tạo lập quan hệ P với tập thuộc  tính
// của quan hệ R
        Create(P,Attr(R));
        for each tuple t in R
            with Sat(t,e) do
               add t to P ;
        endfor;
        return P;
end Selection;
```

2.2.

```
Algorithm  Projection
Format: P = R[X]
Input:        - Quan hệ R(U)
              - Tập con thuộc tính X của U.
Output: - Quan hệ R[X]={t.X | t∈R}
Method
    Create(P,X);
      for each tuple t in R
          with t.X not_in P do
              add t to P ;
       endif;
      endfor;
      return P;
end    Projection;
```

2.3.

```
Algorithm Join
Format: P = R*S
Input:        - Quan hệ R(U)
              - Quan hệ S(V)
Output: - Quan hệ
 R*S={u*v|u∈R,v∈S,u.M = v.M, M = U ∩ V}
```

```
Method
        X = Attr(R) ∪ Attr(S);
        M = Attr(R) ∩ Attr(S);
      Create(P,X);
        for each tuple u in R do
              for each tuple v in S do
                  if u.M=v.M then
                          add u*v to P ;
                  endif;
              endfor;
        endfor;
        return P;
end Join;
```

2.4.

```
Algorithm Union
Format: P = R+S
Input:        - Quan hệ R(U)
              - Quan hệ S(U)
Output: - Quan hệ R+S={t|t∈R ∨ t∈S}
Method
      Create(P,Attr(R));
        for each tuple u in R do
              add u to P;
```

```
        endfor;
        for each tuple v in S
              with v not_in R do
                    add v to P ;
        endfor;
        return P;
end Union;
```

2.5.

```
Algorithm Intersection
Format: P = R&S
Input:        - Quan hệ R(U)
              - Quan hệ S(U)
Output: - Quan hệ R&S={t|t∈R ∧ t∈S}
Method
    Create(P,Attr(R));
      for each tuple u in R
            with u in S do
                  add u to P;
      endfor;
      return P;
end Intersection;
```

2.6.

```
Algorithm Substraction
Format: P = R-S
Input:          - Quan hệ R(U)
                - Quan hệ S(U)
Output: - Quan hệ R-S={t|t∈R ∧ t∉S}
Method
    Create(P,Attr(R));
      for each tuple u in R
            with u not_in S do
                add u to P;
      endfor;
      return P;
 end  Substraction;
```

2.7.

```
Algorithm Division
      Format: P = R:S
      Input:          - Quan hệ R(U)
                      - Quan hệ S(V)
      Output:
      - Quan hệ R:S={t.M|t∈R,(t.M)*S⊆R,M = U-V}
      Method
            M = Attr(R)-Attr(S);
```

```
        c := Card(S);//số bộ của S
    Create(P,M);
        for each tuple t in R
            with t.M not_in P do
                d:=0;//khởi tạo biến đếm
                for each tuple v in S
                    if (t.M)*v in R then
                        d:=d+1
                    else breakfor;
                    endif;
                endfor;
                if d=c then
                    add t.M to P;
                endif;
        endfor;
        return P;
end Division;
```

2.8.

```
Algorithm  Selection_Projection
Format: P = R(e,X)
Input:        - Quan hệ R(U)
              - Biểu thức chọn e trên U.
              - Tập con X của U.
```

```
      Output: - Quan hệ
P(X) = R(e,X) = {t.X | t∈R, Sat(t,e)}
      Method
            Create(P,X);
            for each tuple t in R
                with Sat(t,e) do
                    if t.X not_in P then
                        add t.X to P ;
                    endif;
            endfor;
            return P;
      end Selection_Projection;
```

2.9.

```
      Algorithm  Join_Selection_Projection
      Format: P = (R*S)(e,X)
      Input:          - Quan hệ R(U)
                      - Quan hệ S(V)
                      - Biểu thức chọn e trên UV.
                      - Tập con X của UV.
      Output: - Quan hệ
P(X) = (R*S)(e,X) = {(u*v).X | u∈R,v∈S, u.M = v.M,
Sat(u*v,e), M = U ∩ V }
      Method
```

```
        M := Attr(R) ∩ Attr(S);
        Create(P,X);
        for each tuple u in R do
              for each tuple v in S
                 with u.M = v.M do
                    if Sat(u*v,e) then
                          construct t=(u*v).X;
                          if t not_in P then
                          add t to P ;
                          endif;
                    endif;
              endfor;
        endfor;
        return P;
end Join_Selection_Projection;
```

2.10.

```
Algorithm Card
Format: c = Card(R)
Input:  - Quan hệ R(U)
Output: - số bộ của quan hệ R.
Method
      c:=0;
        for each tuple u in R do
```

```
            c:=c+1;
        endfor;
        return c;
    end Card;
```

2.11.

```
Algorithm Sum
Format: s = Sum(R,A)
Input:        - Quan hệ R(U)
              - Thuộc tính kiểu số A trong U
Output: - Tổng s các trị trên cột A trong
quan hệ R
Method
    s:=0;
      for each tuple u in R do
          s:=s+u.A;
      endfor;
      return s;
  end Sum;
```


2.12.

```
Algorithm Avg
Format: s = Avg(R,A)
Input:        - Quan hệ R(U)
              - Thuộc tính kiểu số A trong U
```

Output: - Trị trung bình của cột A trong quan hệ R

```
Method
    s:=0;c:=0;
      for each tuple u in R do
            s:=s+u.A;
            c:=c+1;
      endfor;
      if c > 0 then return s/c
      else return null;
      endif;
end Avg;
```


2.13.

Algorithm Max

Format: smax = Max(R,A)

Input: - Quan hệ R(U)

- Thuộc tính kiểu sánh được A trong U

Output: - Trị lớn nhất của cột A trong quan hệ R

```
Method
    smax:= -∞;
      for each tuple u in R do
```

```
            if smax < u.A then
                smax:=u.A;
            endif;
        endfor;
        return smax;
    end Max;
```

2.14.

```
    Algorithm Min
    Format: smin = Min(R,A)
    Input:        - Quan hệ R(U)
                  - Thuộc tính kiểu sánh được A
trong U
    Output: - Trị nhỏ nhất của cột A trong quan
hệ R
    Method
        smin:= +∞;
          for each tuple u in R do
              if smin > u.A then
                  smin:=u.A;
              endif;
          endfor;
          return smin;
    end Min;
```

2.15.

```
Algorithm Sume
Format: s = Sume(R,A,e)
Input:        - Quan hệ R(U)
              - Thuộc tính kiểu số A trong U
              - Biểu thức e trên U
Output: - Tổng s các trị trên cột A của các
bộ thỏa điều kiện e trong quan hệ R
Method
    s:=0;
      for each tuple u in R
            with sat(u,e) do
              s:= s+u.A;
      endfor;
      return s;
end Sume;
```

Bài giải Chương 3

NGÔN NGỮ HỎI SQL

CSDL ***Thực tập***

SV(SV#, HT, NS, QUE,HL)

DT(DT#,TDT,CN,KP)

SD(SV#,DT#,NTT,KM, KQ)

3.1. Danh sách các sinh viên trẻ (dưới 18 tuổi) và học khá/giỏi (HL>8.5):

```
SELECT HT
FROM SV
WHERE 2003-NS < 18 AND HL > 8.5;
```

3.2. Thông tin về các đề tài được cấp kinh phí trên 10 triệu đồng:

```
SELECT *
FROM DT
WHERE KP > 10;
```

3.3. Danh sách các sinh viên trẻ (dưới 18 tuổi), học và thực tập đều đạt loại khá/giỏi (HL > 8.5 và KQ > 8.5):

```
SELECT HT
FROM SV
```

```
WHERE 2003-NS < 18 AND HL > 8.5
AND SV# IN
    (SELECT SV#
     FROM SD
     WHERE KQ > 8.5);
```

3.4. Danh sách các chủ nhiệm đề tài có các sinh viên quê ở Hà Nội tham gia:

```
SELECT CN
FROM DT
WHERE DT# IN
    (SELECT DT#
     FROM SD
     WHERE SV# IN
         (SELECT SV#
          FROM SV
          WHERE QUE = 'Ha Noi'));
```

3.5. Danh sách các sinh viên học giỏi hơn các sinh viên Hà Nội:

```
SELECT HT
FROM SV
WHERE HL > ALL
    (SELECT HL
```

```
    FROM SV
    WHERE QUE = 'Ha Noi');
```

3.6. Điểm trung bình của các sinh viên Hà Nội:

```
SELECT AVG(HL)
FROM SV
WHERE QUE = 'Ha Noi';
```

3.7. Tổng số đoạn đường thực tập theo đề tài 5:

```
SELECT SUM(KM)
FROM SD
WHERE DT# = 5;
```

3.8. Tổng số sinh viên đi thực tập:

```
SELECT COUNT(DISTINCT SV#)
FROM SD;
```

3.9. Số tỉnh có sinh viên đến thực tập theo đề tài 5:

```
SELECT COUNT(DISTINCT NTT)
FROM SD
WHERE DT# = 5;
```

3.10. Danh sách các tỉnh và số sinh viên quê ở tỉnh đó, nhóm theo QUE:

```
SELECT QUE, COUNT(*)
FROM SV
```

```
GROUP BY QUE;
```

3.11. Các đề tài có trên 10 sinh viên đăng ký tham gia:

```
SELECT TDT
FROM DT
WHERE 10 < (SELECT COUNT(*)
            FROM SD
            WHERE DT.DT# = SD.DT#);
```

*3.12. Dùng SQL để biểu thị các phép toán của đại số quan hệ:

a) $R(e)$:

```
SELECT *
FROM R
WHERE e;
```

b) $R[X]$:

```
SELECT DISTINCT X
FROM R;
```

c) $R*S$: Giả sử $Attr(R) \cap Attr(S) = \{A_1, A_2, \ldots, A_k\}$

```
SELECT *
FROM R, S
WHERE R.A1 = S.A1
AND R.A2 = S.A2
AND...
```

```
AND R.Ak = S.Ak;
```

d) *R+S*:

```
R
UNION
S;
```

e) *R&S*:

```
R
INTERSECT
S;
```

f) *R-S*:

```
R
MINUS
S;
```

3.13. Cho thông tin về những sinh viên sinh trước năm 1973 và quê ở Hải Phòng:

```
SELECT *
FROM SV
WHERE NS < 1973 AND QUE = 'Hai Phong';
```

3.14. Cho danh sách các tỉnh có sinh viên đến thực tập:

```
SELECT DISTINCT NTT
FROM SD;
```

3.15. Cho biết các địa điểm thực tập xa trường (KM > 100) của đề tài số 7:

```
SELECT DISTINCT NTT
FROM SD
WHERE DT# = 7
    AND KM > 100;
```

3.16. Cho thông tin về việc thực tập tại Nha Trang của các sinh viên:

```
SELECT *
FROM SD
WHERE NTT = 'Nha Trang'
```

*3.17. Cho danh sách sinh viên thực tập tại quê nhà:

```
SELECT HT
FROM SV, SD
WHERE SV.SV# = SD.SV#
      AND SV.QUE = SD.NTT;
```

3.18. Cho thông tin về các đề tài có sinh viên thực tập:

```
SELECT DISTINCT *
FROM DT
WHERE EXISTS
        (SELECT *
         FROM SD
         WHERE DT.DT# = SD.DT#);
```

3.19. Cho biết mã của các đề tài không có sinh viên nào tham gia:

```
SELECT DT#
FROM DT
MINUS
SELECT DISTINCT DT#
FROM SD;
```

3.20. Cho biết mã của những đề tài có kinh phí 1.5 triệu và những đề tài có kinh phí trên 2 triệu:

```
SELECT DT#
FROM DT
WHERE KP = 1.5
    OR KP > 2;
```

3.21. Cho biết mã của những sinh viên dưới 24 tuổi, thực tập khá (có điểm kết quả trên 6):

```
SELECT SV#
FROM SV
WHERE 2003 - NS < 24
   AND SV# IN
         (SELECT SV#
          FROM SD
          WHERE KQ > 6);
```

*3.22. Cho danh sách các đề tài có sinh viên học giỏi nhất lớp tham gia:

```
SELECT TDT
FROM DT
WHERE DT# IN
    ( SELECT DISTINCT DT#
      FROM SD
      WHERE SV# IN
         ( SELECT SV#
           FROM SV
           WHERE HL = ( SELECT MAX(HL)
                        FROM SV )));
```

*3.23. Cho danh sách các đề tài không có sinh viên học kém nhất lớp tham gia:

```
SELECT TDT
FROM DT
WHERE DT# IN
        (( SELECT DT#
          FROM DT )
        MINUS
       ( SELECT DISTINCT DT#
```

```
        FROM SD
        WHERE SV# IN
          ( SELECT SV#
            FROM SV
            WHERE HL = ( SELECT MIN(HL)
                         FROM SV ))));
```

Chú ý: biểu thức SQL sau đây

```
SELECT TDT
FROM DT
WHERE DT# IN
  ( SELECT DISTINCT DT#
    FROM SD
    WHERE SV# IN
      ( SELECT SV#
        FROM SV
        WHERE HL > ( SELECT MIN(HL)
                     FROM SV )));
```

cho biết danh sách các đề tài có sinh viên tham gia, nhưng những sinh viên này không phải là những người học kém nhất lớp.

***3.24.** Cho danh sách những sinh viên thực tập theo đề tài có kinh phí lớn hơn một phần năm tổng kinh phí cấp cho các đề tài:

```
SELECT HT
FROM SV
```

```
WHERE SV# IN
   ( SELECT SV#
     FROM SD
     WHERE DT# IN
        ( SELECT DT#
          FROM DT
          WHERE 5*KP > ( SELECT SUM(KP)
                             FROM DT )));
```

*3.25. Cho danh sách các sinh viên có điểm học tập cao hơn điểm thực tập trung bình của đề tài mã số 4:

```
SELECT HT
FROM SV
WHERE HL > ( SELECT AVG(KQ)
             FROM SD
             WHERE DT# = 4);
```

*3.26. Cho quan hệ *R(U)*. Hãy dùng SQL để sinh ra quan hệ rỗng *S(U)*:

```
SELECT *
FROM R
WHERE False;
```

Bài giải Chương 4

PHỤ THUỘC HÀM

4.1.

f_1: $A \to A$: *thỏa,*

f_2: $A \to B$: *thỏa,*

f_3: $A \to C$: *không thỏa,*

f_4: $AC \to C$: *thỏa,*

f_5: $A \to D$: *thỏa,*

f_6: $D \to A$: *thỏa.*

4.2.

```
Algorithm Closure
Format: Y = X⁺
Input:       - LĐQH p =(U,F)
             - Tập thuộc tính X ⊆ U
Output: - Y = X⁺ = {A ∈ U | X → A ∈ F⁺}
Method
        Y:=X;
     repeat
             Z:=Y;
             for each FD L → R in F do
```

```
                    if L ⊆ Y then
                         Y:=Y ∪ R;
                    endif;
               endfor;
          until Y=Z;
          return Y;
     end Closure;
```

4.3.

F1. Nếu $Y \subseteq X$ thì $R(X \to Y)$ (*tính phản xạ*)

$\forall u,v \in R$: $u.X=v.X \Rightarrow u.Y=v.Y$, vì $Y \subseteq X$.

F2. Nếu $R(X \to Y)$ thì $R(XZ \to YZ)$ (*tính gia tăng*)

$\forall u,v \in R$: $u.XZ=v.XZ \Rightarrow u.X=v.X$ & $u.Z=v.Z \Rightarrow u.Y=v.Y$ & $u.Z=v.Z \Rightarrow u.YZ=v.YZ$

F3. Nếu $R(X \to Y)$ và $R(Y \to Z)$ thì $R(X \to Z)$ (*tính bắc cầu*)

$\forall u,v \in R$: $u.X=v.X \Rightarrow u.Y=v.Y \Rightarrow u.Z=v.Z$

4.4. Giả sử $f: X \to Y \notin F^+$. Ta chứng minh $X \to Y \notin F^+$ bằng cách **chỉ ra** một quan hệ Armstrong $R(U)$ thoả các PTH trong tập F (thậm chí trong F^+) nhưng không thoả PTH f .

Quan hệ Armstrong R được xây dựng như sau:

Giả sử $U = \{A_1, A_2,..., A_n\}$ và a_i và b_i là hai phần tử khác nhau của $dom(A_i)$, $i = 1..n$. Quan hệ R chứa 2 bộ u và v như sau:

$u = (a_1, a_2,..., a_n)$.

$v.A_i=a_i$, nếu $A_i \in X^+$, nếu không ta đặt $v.A_i=b_i$, $i = 1..n$.

Trước hết ta chứng minh R không thoả PTH $X \rightarrow Y$. Theo cách xây dựng R, ta có hai bộ u và v giống nhau trên miền lớn duy nhất là X^+, $u.X^+ = v.X^+$ và do $X^+ \supseteq X$ nên $u.X = v.X$. Giả sử $u.Y = v.Y$. Thế thì $Y \subseteq X^+$. Theo định nghĩa bao đóng ta suy ra $X \rightarrow Y \in F^+$, mâu thuẫn với giả thiết. Vậy R không thoả PTH $X \rightarrow Y$.

Ta chứng minh R thoả mọi PTH trong F^+. Giả sử $W \rightarrow Z \in F^+$ và $u.W = v.W$. Từ đây rút ra, do đặc điểm của R, $W \subseteq X^+$. Theo định nghĩa bao đóng, $X \rightarrow W \in F^+$, theo tính chất bắc cầu cho các PTH $X \rightarrow W$ và $W \rightarrow Z$ ta suy ra $X \rightarrow Z \in F^+$. Lại theo định nghĩa bao đóng ta có $Z \subseteq X^+$ và do đó, theo đặc điểm của R ta có $u.Z = v.Z$. Vậy R thoả $W \rightarrow Z$ *đpcm*.

Chú ý: *Chứng minh được dựa trên giả thiết là miền trị của các thuộc tính trong quan hệ chứa ít nhất 2 trị phân biệt. Giả thiết này là khá tự nhiên, vì nếu trong bảng có một cột chỉ chứa một trị duy nhất thì ta có thể xoá cột đó.*

4.5. a. Nếu $F \subseteq G$ thì $SAT(F) \supseteq SAT(G)$: $R(G) \Rightarrow R(F)$.

b. $SAT(FG) = SAT(F) \cap SAT(G)$: $R(FG) \Leftrightarrow R(F) \wedge R(G)$.

4.6. ***Chú ý:*** *Những chỗ có dấu ? là gợi ý bạn đọc giải thích vì sao.*

Chứng minh mệnh đề b trước sau đó suy ra mệnh đề a.

b. $R \subseteq S \Rightarrow FD(R) \supseteq FD(S)$:

Giả sử $X \rightarrow Y \in FD(S)$ và u và v là hai bộ trong R thỏa $u.X = v.X$. Ta có $u, v \in S$ (?). Do đó $u.Y = v.Y$ (?), từ đó suy ra $X \rightarrow Y \in FD(R)$.

a. $FD(R+S) \subseteq FD(R) \cap FD(S)$

Vì $R+S \supseteq R$ và $R+S \supseteq S$ nên, theo câu b ta có

$FD(R+S) \subseteq FD(R)$ và $FD(R+S) \subseteq FD(S)$. Từ đó suy ra a.

Thí dụ chứng tỏ $FD(R+S) \subset FD(R) \cap FD(S)$.

$U=AB$; R chứa một bộ duy nhất $u = (1,x)$; S chứa một bộ duy nhất $v = (1,y)$, $x \neq y$. R và S thỏa mọi PTH trên U(?). Quan hệ $P=R+S$ chứa 2 bộ u và v. P không thỏa PTH $A \rightarrow B$(?).

4.7.

Nhận xét : Các toán tử *SAT* và *FD* có tính nghịch biến(?) và

$$SAT(F) = \bigcap_{f \in F} SAT(f) \quad \text{và}$$

$$FD(\Re) = \bigcap_{R \in \Re} FD(R)$$

4.8.

Với mọi tập con X, Y, Z, V của U và với mọi thuộc tính A trong U:

F4. Tính tựa bắc cầu: Nếu $X \rightarrow Y$, $YZ \rightarrow V$ thì $XZ \rightarrow V$

$X \rightarrow Y$ (gt)

$YZ \rightarrow V$ (gt)

$XZ \rightarrow YZ$ (F2)

$XZ \rightarrow V$ (F3)

F5. Tính phản xạ chặt: $X \rightarrow X$ (F1)

F6. Mở rộng vế trái và thu hẹp vế phải: Nếu $X \rightarrow Y$ thì $XZ \rightarrow Y \setminus V$

$X \rightarrow Y$ (gt)

$Y \rightarrow Y \setminus V$ (F1)

$X \to Y \setminus V$ (F3)

$XZ \to X$ (F1)

$XZ \to Y \setminus V$ (F3)

F7. Cộng tính đầy đủ: Nếu $X \to Y$ và $Z \to V$ thì $XZ \to YV$

$X \to Y$ (gt)

$Z \to V$ (gt)

$XZ \to YZ$ (F2)

$YZ \to YV$ (F2)

$XZ \to YV$ (F3)

F8. Mở rộng vế trái: Nếu $X \to Y$ thì $XZ \to Y$

$XZ \to X$ (F1)

$X \to Y$ (gt)

$XZ \to Y$ (F3)

F9. Cộng tính ở vế phải: Nếu $X \to Y$ và $X \to Z$ thì $X \to YZ$

$X \to Y$ (gt)

$X \to XY$ (F2)

$X \to Z$ (gt)

$XY \to YZ$ (F2)

$X \to YZ$ (F3)

F10. Bộ phận ở vế phải: Nếu $X \to YZ$ thì $X \to Y$

$X \to YZ$ (gt)

$YZ \to Y$ (F1)

$X \rightarrow Y$ (F3)

F11. Tính tích luỹ: Nếu $X \rightarrow YZ$, $Z \rightarrow AV$ thì $X \rightarrow YZA$

$Z \rightarrow AV$ (gt)

$YZZ \rightarrow YZAV$ (F2)

$YZ \rightarrow YZAV$

$YZAV \rightarrow YZA$ (F1)

$YZ \rightarrow YZA$ (F3)

$X \rightarrow YZ$ (gt)

$X \rightarrow YZA$ (F3)

4.9.

$\forall X, Y \in Poset(U)$

(C4) $f(XY) \supseteq f(X)f(Y)$:

$XY \supseteq X$ (ltth- theo lý thuyết tập hợp)

$f(XY) \supseteq f(X)$ (C2)

$XY \supseteq Y$ (ltth)

$f(XY) \supseteq f(Y)$ (C2)

$f(XY) \supseteq f(X)f(Y)$ (Cộng tính của ltth)

(C5) $f(X \cap Y) \subseteq f(X) \cap f(Y)$:

$X \cap Y \subseteq X$ (ltth)

$f(X \cap Y) \subseteq f(X)$ (C2)

$X \cap Y \subseteq Y$ (ltth)

$f(X \cap Y) \subseteq f(Y)$ (C2)

$f(X \cap Y) \subseteq f(X) \cap f(Y)$ (ltth)

(C6) $f(f(X)Y) = f(XY)$:

$X \subseteq f(X)$ (C1)

$XY \subseteq f(X)Y$ (ltth)

$f(XY) \subseteq f(f(X)Y)$ (C2)

$X \subseteq XY$ (ltth)

$f(X) \subseteq f(XY)$ (C2)

$Y \subseteq XY$ (ltth)

$XY \subseteq f(XY)$ (C1)

$Y \subseteq f(XY)$ (Tính chất bắc cầu của ltth)

$f(X)Y \subseteq f(XY)$ (Cộng tính của ltth)

$f(f(X)Y) \subseteq f(f(XY)) = f(XY)$ (C2,C3)

$f(Xf(Y)) = f(XY)$: tương tự

Chứng minh phép toán lấy bao đóng của tập thuộc tính là một ánh xạ đóng thỏa các tính chất sau:

1. Tính phản xạ $X^+ \supseteq X$: (đnbđ - Theo định nghĩa bao đóng)

2. Tính đơn điệu nếu $X \subseteq Y$ thì $X^+ \subseteq Y^+$ (ttbđ - thuật toán tìm bao đóng)

3. Tính luỹ đẳng $X^{++} = X^+$: (ttbđ)

4. $(XY)^+ \supseteq X^+Y^+$: (C4)

5. $(X^+Y)^+ = (XY^+)^+ = (XY)^+$: (C6)

6. $X \to Y$ khi và chỉ khi $Y \subseteq X^+$: (đnbđ)

7. $X \rightarrow Y$ khi và chỉ khi $Y^+ \subseteq X^+$: (đnbđ, C2,C3)

8. $X \rightarrow X^+$ và $X^+ \rightarrow X$: (đnbđ,C1)

9. $X^+ = Y^+$ khi và chỉ khi $X \rightarrow Y$ và $Y \rightarrow X$: (đnbđ)

4.10.

$A^o = \{F1, F2, F3\} \Leftrightarrow B^o = \{F5, F10, F11\}$

$[A^o \Rightarrow B^o]$: *xem bài 4.8*

$[B^o \Rightarrow A^o]$

$[B^o \Rightarrow F1]$:

$X \supseteq Y$ (gt)

$X = MY$ (gt)

$MY \rightarrow MY$ (F5)

$X \rightarrow MY$

$X \rightarrow Y$ (F10), đpcm.

Trước hết chứng minh $[B^o \Rightarrow F11']$ với

F11' (*Tính tích luỹ mở rộng*): Nếu $X \rightarrow YZ$ và $Z \rightarrow MV$ thì $X \rightarrow YZM$ trong đó M là một tập con của U.

Giả sử $M = A_1A_2...A_k$. Ta ký hiệu $V_i = (M \backslash A_i)V$, $i = 1..k$. Ta có, với mọi $i = 1..k$: $MV = A_iV_i$.

$X \rightarrow YZ$ (gt)

$Z \rightarrow A_1V_1$ (gt)

$X \rightarrow YZA_1$ (F11)

$X \rightarrow (YA_1)Z$

$Z \rightarrow A_2V_2$ (gt)

$X \rightarrow YA_1ZA_2$ (F11)

$X \rightarrow (YA_1A_2)Z$

...

$X \rightarrow (YA_1A_2 \ldots A_k)Z$ (F11), hay

$X \rightarrow YZM$, đpcm.

[$B^o \Rightarrow$ F3]:

$X \rightarrow \varnothing Y$ (gt)

$Y \rightarrow \varnothing Z$(gt)

$X \rightarrow \varnothing YZ$ (F11')

$X \rightarrow Z$ (F10), đpcm.

[$B^o \Rightarrow$ F2]:

$XZ \rightarrow XZ$ (F5)

$X \rightarrow Y$ (gt)

$XZ \rightarrow XYZ$ (F11')

$XYZ \rightarrow YZ$ (F1 đã chứng minh)

$XZ \rightarrow YZ$ (F3 đã chứng minh), đpcm.

4.11. A^o = {F1, F2, F3} $\Leftrightarrow$ S^o = {F1, F4}

[$A^o \Rightarrow S^o$]: *xem bài 4.8*

[$S^o \Rightarrow A^o$]:

[$S^o \Rightarrow$ F2]:

$X \rightarrow Y$ (gt)

$YZ \rightarrow YZ$ (F1)

$XZ \rightarrow YZ$ (F4), đpcm

$[S^o \Rightarrow$ F3]:

$X \rightarrow Y$ (gt)

$Y\varnothing \rightarrow Z$ (gt)

$X\varnothing \rightarrow Z$ (F4)

$X \rightarrow Z$, đpcm.

4.12. A^o = {F1, F2, F3} $\Leftrightarrow$ D^o = {F3, F5, F6, F7}

$[A^o \Rightarrow D^o]$: *xem bài 4.8*

$[D^o \Rightarrow A^o]$:

$[D^o \Rightarrow$ F1]:

$X \supseteq Y$ (gt)

$X = MY$ (gt)

$Y \rightarrow Y$ (F5)

$MY \rightarrow Y - \varnothing$ (F6)

$X \rightarrow Y$, đpcm.

$[D^o \Rightarrow$ F2]:

$X \rightarrow Y$ (gt)

$Z \rightarrow Z$ (F5)

$XZ \rightarrow YZ$ (F7), đpcm.

4.13. $A^o = \{F1, F2, F3\} \Leftrightarrow M^o = \{F4, F5, F8\}$

$[A^o \Rightarrow M^o]$: *xem bài 4.8*

$[M^o \Rightarrow A^o]$:

$[M^o \Rightarrow F1]$:

$X \supseteq Y$ (gt)

$X = MY$ (gt)

$Y \to Y$ (F5)

$MY \to Y$ (F8)

$X \to Y$,đpcm.

$[M^o \Rightarrow F2]$:

$X \to Y$ (gt)

$YZ \to YZ$ (F5)

$XZ \to YZ$ (F4), đpcm.

$[M^o \Rightarrow F3]$:

$X \to Y$ (gt)

$Y \to Z$ (gt)

$Y\varnothing \to Z$ (gt)

$X\varnothing \to Z$ (F4)

$X \to Z$, đpcm.

4.14. $p = (U,F)$, $U = ABCDE$, $F = \{ A \to C, BC \to D, D \to E, E \to A \}$.

a) $(AB)^+ = ABCDE = U$.

b) $(BD)^+ - D^+ = ABCDE\text{-}ACDE = B$.

4.15. $p = (U,F)$, $U = ABCDEG$, $F = \{ B \to C, Y \to D, D \to G, AG \to E \}$.

a) $AB \to G \in F^+$ vì $(AB)^+ = ABCDEG \supseteq G$

b) $BD \to AD \notin F^+$ vì $AD \not\subset (BD)^+ = BCDG$.

4.16 $F \equiv G \Leftrightarrow (\forall X \subseteq U): (X_F^+ = X_G^+)$?

$X \to Y \in F^+ \Leftrightarrow Y \subseteq X_F^+ = X_G^+ \Leftrightarrow X \to Y \in G^+$.

4.17 $F \equiv F^+$?

$F \models F^+$ (theo định nghĩa)

$F^+ \models F$ (vì $F^+ \supseteq F$)

4.18

a) $F=\{ X \to Y, Z \to V \} \models G=\{ XZ \to YV \}$, vì $(XZ)^+_F = XYZV \supseteq YV$.

b) $F=\{ X \to Y \} \equiv G=\{ X \to Y\text{-}X \}$, vì $X^+_F = X_G^+ = XY$.

c) $F=\{ X \to Y \} \models G=\{ XZ \to Y \}$, vì $(XZ)^+_F = XYZ \supseteq Y$.

d) $F=\{ X \to Y, Y \to Z \} \models G=\{ X \to Z \}$, vì $X^+_F = XYZ \supseteq Z$.

e) $F=\{ X \to Y, YZ \to V \} \models G=\{ XZ \to V \}$, vì $(XZ)^+_F = XYZV \supseteq V$.

f) $F=\{ X \to Y \} \models G=\{ XZ \to Y\text{-}V \}$, vì $(XZ)^+_F = XYZ \supseteq Y \supseteq Y\text{-}V$.

g) $F=\{X \to Y, X \to Z\} \equiv G=\{X \to YZ\}$, vì $X^+_F = X^+_G = XYZ \supseteq Y, Z, YZ$.

h) $F=\{ X \to YZ \} \models G=\{ X \to Y \}$, vì $X^+_F = XYZ \supseteq Y$.

i) $F=\{ X \to YZ, Z \to AV \} \models G=\{ X \to YZA \}$, vì $X^+_F = XYZAV \supseteq YZA$.

4.19 Thuật toán tìm phủ thu gọn tự nhiên của tập PTH F.

```
Algorithm Natural_Reduced
```

```
Format: Natural_Reduced(F)
Input:  - Tập PTH F
Output: - Một phủ thu gọn tự nhiên G của F
    •  G ≡ F
    •  ∀ L→R ∈ G: L∩G = ∅
    •  ∀Li→Ri,∀Lj→Rj∈G: i≠j ⇒ Li≠Lj
Method
      G:=∅;
    for each FD L→R in F do
          Z:=R-L;
        if Z ≠ ∅ then
          if there is an FD L→Y in G then
              replace L→Y in G by L→YZ
             else add L→Z to G;
           endif;
        endif;
    endfor;
    return G;
  end Natural_Reduced;
```

4.20 *Thuật toán tìm phủ không dư của tập PTH F.*

```
Algorithm Nonredundant
```

```
Format: Nonredundant(F)
Input:        - Tập PTH F
Output: - Một phủ không dư G của F
    • G ≡ F
    • ∀ g ∈G: G-{g} !≡ G
Method
    G:=F;
  for each FD g:L → R in F do
      if R ⊆ L+G-{g}  then
         G:= G -{g};
      endif;
   endfor;
   return G;
 end Nonredundant;
```

4.21. *Xây dựng thuật toán tìm phủ thu gọn trái của tập PTH F.*

Để ý rằng ta luôn có

$\forall\ g{:}L\to R \in G, \forall A\in L{:}\ G-\{g\}\cup\{L-\{A\}\to R\} \models L\to R,$

vì $L-\{A\}\subseteq L$, do đó ta chỉ cần kiểm tra

$G \models (L-\{A\})\to R.$

```
Algorithm Left_Reduced
      Format: Left_Reduced(F)
```

```
Input:        - Tập PTH F
Output: - Một phủ thu gọn trái G của F
    • G ≡ F
    • ∀ g:L→R ∈G,∀A∈L: G-{g}∪{L-{A}→R} !≡G
Method
    G:=F;
  for each FD g:L→R in F do
     X:=L;
     for each attribute A in X do
              if R ⊆(L-{A})+G then
            delete A from L in G;
         endif;
     endfor;
  endfor;
  return G;
 end Left_Reduced;
```

4.22. Xây dựng thuật toán tìm phủ thu gọn phải của tập PTH F.

Để ý rằng ta luôn có

$\forall g:L\rightarrow R \in G, \forall A\in R: G \models L\rightarrow R-\{A\}$,

vì $R-\{A\}\subseteq R$, do đó ta chỉ cần kiểm tra

$G-\{L\rightarrow R\}\cup\{L\rightarrow R-\{A\}\} \models L\rightarrow R$.

```
Algorithm Right_Reduced
    Format: Right_Reduced(F)
    Input:      - Tập PTH F
    Output: - Một phủ thu gọn phải G của F
        • G ≡ F
        • ∀ g:L→R ∈G,∀A∈R: G-{g}∪{L→R-{A}} !≡G
    Method
        G:=F;
        for each FD g:L→R in F do
            X:=R;
            for each attribute A in X do
                if A in L+G-{L→R}∪{L→R-{A}}  then
                    delete A from R in G;
                endif;
            endfor;
        endfor;
        return G;
    end Right_Reduced;
```

4.23. Xây dựng thuật toán tìm phủ thu gọn của tập PTH F.

```
Algorithm Reduced
    Format: Reduced(F)
```

```
Input:          - Tập PTH F
Output: - Một phủ thu gọn G của F
Method
        G:=Right_Reduced(Left_Reduced(F));
     return G;
end Reduced;
```

4.24. Xây dựng thí dụ chứng tỏ với tập PTH F sau khi thực hiện

G:=Left_Reduced(Right_Reduced(F));

G lại có thể trở thành phủ chưa thu gọn phải.

Xét tập PTH

$F = \{DA \to BC, B \to CD, A \to D\}$

H:=Right_Reduced$(F) = \{DA \to B, B \to CD, A \to D\}$

G:=Left_Reduced$(H) = \{A \to B, B \to CD, A \to D\}$

G không phải là tập PTH thu gọn phải vì ta còn có thể bỏ thuộc tính D ở PTH thứ ba trong G,

Right_Reduced$(G) = \{A \to B, B \to CD, A \to \emptyset\} \equiv \{A \to B, B \to CD\}$

Chú ý rằng các PTH có vế phải là tập rỗng thì không chứa thông tin do đó có thể loại bỏ chúng khỏi tập PTH. Từ đó suy ra rằng nếu tập PTH có các vế phải là một thuộc tính thì việc tìm thu gọn phải tương đương với việc tìm phủ không dư.

4.25. Thuật toán tìm phủ tối tiểu của tập PTH F.

Algorithm MinCover

```
Format: MinCover(F)
Input:      - Tập PTH F
Output: - Một phủ tối tiểu G của F
Method
      // Tách mỗi PTH L→R trong F thành
      // các PTH L→A, A ∈ R
       G := Ø;
      for each FD L → R in F do
        for each attribute A in R do
       if L→A not_in G then
          add L→A to G;
       endif;
     endfor;
   endfor;
  G:=Reduced(G);
  return G;
end MinCover;
```

4.26. Chứng minh rằng với mọi tập PTH F trên *U* luôn tồn tại một phủ *G* của *F* sao cho mọi PTH trong *G* đều là phụ thuộc đầy đủ:

Đó chính là phủ thu gọn trái của F(?).

4.27. Xây dựng thuật toán tìm một phủ đầy đủ của tập PTH *F*: *Xem bài 4.21.*

4.28. Cho quan hệ R trên U và các tập con thuộc tính X, Y của U.

Chứng minh:

a) $R(X\ (s)\!\!\rightarrow Y) \Rightarrow R(X \rightarrow Y)$

$\forall\ u,v \in R,\ u.X = v.X$ (gt)

$u.Y = v.Y$ (vì $R(X\ (s)\!\!\rightarrow Y)$)

$R(X\rightarrow Y)$, đpcm

b) $R(X \rightarrow Y) \Rightarrow R(X\ (w)\!\!\rightarrow Y)$

$\forall\ u,v \in R$: $u.X = v.X$ (gt)

$u.Y = v.Y$ (vì $R(X \rightarrow Y)$)

$R(X\ (w)\!\!\rightarrow Y)$, đpcm

c) $R(X\ (d)\!\!\rightarrow Y) \Rightarrow R(X\ (w)\!\!\rightarrow Y)$

$\forall\ u,v \in R,\ u.X = v.X$ (gt)

$\exists\ B \in Y$: $u.B = v.B$ (vì $R(X\ (d)\!\!\rightarrow Y)$)

$R(X\ (w)\!\!\rightarrow Y)$, đpcm

d) $R(X\ (s)\!\!\rightarrow Y) \Rightarrow R(X\ (d)\!\!\rightarrow Y)$

$\forall\ u,v \in R, \exists\ A \in X$: $u.A = v.A$ (gt)

$u.Y = v.Y$ (vì $R(X\ (s)\!\!\rightarrow Y)$)

$R(X\ (d)\!\!\rightarrow Y)$, đpcm

4.29. Với mọi tập con thuộc tính X, Y, Z và thuộc tính A trên tập thuộc tính U. Điền các ký hiệu f, s, w hoặc d thay cho dấu ? để khẳng định các tính chất cao nhất của các phụ thuộc mạnh, yếu hoặc đối ngẫu sau đây:

1. Tính phản xạ: Nếu $Y \subseteq X$ thì $X\ (f)\!\!\rightarrow Y$

2. Tính gia tăng: Nếu $X\,(f)\!\to Y$ thì $XZ\ (f)\!\to YZ$

3. Tính bắc cầu: Nếu $X\,(s)\!\to Y$ và $Y\,(s)\!\to Z$ thì $X\,(s)\!\to Z$

4. Tính tựa bắc cầu: Nếu $X\,(s)\!\to Y$, $YZ\,(s)\!\to V$ thì $XZ\,(s)\!\to V$

5. *Tính phản xạ chặt:* $X\,(d)\!\to X$ và $X\,(f)\!\to X$

6. *Mở rộng vế trái và thu hẹp vế phải:*

   Nếu $X\,(f)\!\to Y$ *thì* $XZ\,(f)\!\to Y \setminus V$

7. Cộng tính đầy đủ: Nếu $X\,(f)\!\to Y$ và $Z\,(f)\!\to V$ thì $XZ\,(f)\!\to YV$

8. Mở rộng vế trái: Nếu $X\,(f)\!\to Y$ thì $XZ\,(f)\!\to Y$

9. Cộng tính ở vế phải: Nếu $X\,(s)\!\to Y$ và $X\,(s)\!\to Z$ thì $X\,(s)\!\to YZ$

10. Bộ phận ở vế phải: Nếu $X\,(s)\!\to YZ$ thì $X\,(s)\!\to Y$.

11. Tính tích luỹ: Nếu $X\,(s)\!\to YZ$, $Z\,(s)\!\to AV$ thì $X\,(s)\!\to YZA$

4.30. Xây dựng thuật toán tìm một khóa của LĐQH.

Tư tưởng: *Xuất phát từ một siêu khóa K tùy ý của LĐQH, duyệt lần lượt các thuộc tính A của K, nếu bất biến $(K-\{A\})^+ = U$ được bảo toàn thì loại A khỏi K. Có thể thay kiểm tra $(K-\{A\})^+ = U$ bằng kiểm tra*

$A \in (K-\{A\})^+$ *(?).*

Algorithm Key

Format: *Key(U,F)*

Input: - *Tập thuộc tính U*

- *Tập PTH F*

Output: - *Khóa K* $\subseteq U$ *thỏa*

- $K^+ = U$
- $\forall A \in K$: $(K-\{A\})^+ \neq U$

```
Method
        K:=U;
      for each attribute A in U do
            if A ∈(K-{A})+ then
                K := K-{A}
            endif;
     endfor;
     return K;
end Key;
```

4.31. Cho LĐQH p. Biết p có một khóa K. Hãy xây dựng thuật toán tìm một khóa thứ hai M của p. Nếu p không có khóa thứ hai thuật toán cho kết quả là một tập rỗng.

Tư tưởng: *Xuất phát từ một siêu khóa M tùy ý, trước hết duyệt các thuộc tính A của K, nếu bất biến $(M-\{A\})^+ = U$ được bảo toàn thì loại A khỏi M. Sau đó duyệt tương tự với các thuộc tính trong U-K.*

```
Algorithm Key2
     Format: Key2(U,F)
     Input:        - Tập thuộc tính U
                   - Tập PTH F
                   - Khóa K ⊆U
```

```
Output: - Khóa thứ hai, nếu có, M ⊆ U thỏa
    •  M+ = U
    •  ∀ A ∈ M: (M-{A})+ ≠ U
    Nếu không có khóa thứ hai: Ø.
Method
      M:=U;
     for each attribute A in K do
          if A ∈(M -{A})+ then
             M := M -{A}
          endif;
    endfor;
    for each attribute A in U-K do
          if A ∈(M -{A})+ then
             M := M -{A}
          endif;
    endfor;
    if M = K then  return Ø
      else return M;
    endif
end Key2;
```

4.32. Xây dựng một LĐQH có 5 thuộc tính ABCDE, mỗi thuộc tính là một khóa:

$p = (U,F)$; $U = ABCDE$; $F = \{A \to B, B \to C, C \to D, D \to E, E \to A\}$.

Ta có $\forall x \in U$: $x^+ = U$ (?), do đó x là một khóa của p.

Chú ý: *Tập PTH F có dạng như trên được gọi là tập phụ thuộc hàm dạng vòng.*

4.33. LĐQH có 5 thuộc tính có thể có tối đa bao nhiêu khóa. Cho thí dụ.

10 khóa, thí dụ, p = (U,F); U = ABCDE; F = {AB→CDE, AC→BDE, AD→BCE, AE→BCD, BC→ADE, BD→ACE, BE→ACD, CD→ABE, CE→ABD, DE→ABC}. Dễ thấy mỗi vế trái tạo ra một khóa (?)

4.34. Xây dựng một LĐQH có 5 thuộc tính ABCDE và chỉ có một khóa duy nhất. p = (U,F); U = ABCDE; F = {A→BCDE}.

4.35. (*Nguyễn Xuân Huy*) Cho K là một khóa của LĐQH p = (U,F). Chứng minh rằng với mọi tập con X của K ta có: $X^+ \cap K = X$.

Vì $X \subseteq X^+$ và $X \subseteq K$ nên $X \subseteq X^+ \cap K$. Ta cần chứng minh $X^+ \cap K \subseteq X$. Giả sử $A \in X^+ \cap K$ và $A \notin X$. Ta xét tập $M = K - \{A\}$. Dễ thấy $X \subseteq M$ (?). Ta có, theo tính chất đồng biến của bao đóng: $A \in X^+ \subseteq M^+$

Từ đây suy ra $M^+ \supseteq K$ nên $M^+ = U$(?), tức là M là bộ phận thực sự của khóa K lại đồng thời là siêu khóa, trái với định nghĩa khóa. Vậy $A \in X$. đpcm.

Chú ý: *Tính chất trên là một đặc trưng của các thuộc tính nguyên thủy (thuộc tính khóa).*

4.36. (*Lê Văn Bào, Nguyễn Xuân Huy, Hồ Thuần*) Cho LĐQH p = (U,F). Gọi M là giao của các khóa của p. Chứng minh rằng

$$M = U - \bigcup_{L \to R \in F} (R - L)$$

Chứng minh

Trước hết để ý rằng các PTH L→R và L→(R-L) là tương đương do đó ta có thể giả thiết rằng mọi PTH trong F đều có dạng L→R, L∩ R=∅ (xem thêm bài 4.19). Do giả thiết này ta có R-L=R. Dễ nhận thấy M là tập các thuộc tính không có mặt trong vế phải của mọi PTH trong F (?) do đó chúng phải có mặt trong mọi khóa (?). Giả sử A là một thuộc tính có trong vế phải của PTH L→AR' của F. Ta chứng minh A sẽ không xuất hiện trong một khóa K nào đấy của p. Thật vậy, xét tập X = U-A. Dễ thấy X ⊇ L (?) và X là siêu khóa (?). Từ siêu khóa X không chứa A (?) ta lấy ra được một khóa K. K không chứa A (?).

4.37. (*Lê Văn Bào, Hồ Thuần*)Cho LĐQH p = (U,F). Gọi M là giao của các khóa của p. Chứng minh rằng p có một khóa duy nhất khi và chỉ khi $M^+ = U$.

Chứng minh:

Nếu $M^+ = U$ thì M là siêu khóa và M không thể chứa thực sự một khóa (?), tức là M là một khóa. Vì M là giao của các khóa đồng thời lại là khóa nên p không thể còn khóa nào khác ngoài M. Ngược lại, nếu p chỉ có một khóa duy nhất M thì giao của các khóa đương nhiên là M, và do đó, theo tính chất của khóa $M^+ = U$.

4.38. Cho tập thuộc tính U với n phần tử. Chứng minh rằng có thể xây dựng tập PTH F sao cho LĐQH $p = (U,F)$ có

$$C_n^{\lfloor \frac{n}{2} \rfloor} \text{ khóa,}$$

trong đó toán tử $\lfloor x \rfloor$ cho ta cận nguyên dưới của số nguyên x,

C_n^m là tổ hợp chặp m của n phần tử.

Gợi ý: Giả sử $U = \{A_1, A_2,...,A_n\}$, $n > 0$. Xây dựng tập PTH $F = \{L_i \rightarrow U\text{-}L_i \mid i = 1..m\}$, trong đó mỗi L_i là một tổ hợp chặp $\lfloor n / 2 \rfloor$ của n. Chứng ming rằng mỗi vế trái L_i là một khóa.

4.39. Tìm tập thuộc tính nguyên thuỷ của LĐQH sau:

$p = (U,F)$, $U = ABCDE$,

$F = \{ AB \rightarrow C, AD \rightarrow B, B \rightarrow D \}$.

Lược đồ p có giao các khóa là: AE, mà $(AE)^+ = AE \neq U$ nên p có trên 1 khóa, cụ thể là 2 khóa: ABE và ADE. Vậy tập thuộc tính nguyên thủy là $U_k = ABE \cup ADE = ABDE$.

4.40. Khẳng định hoặc bác bỏ tính đúng của các mệnh đề sau:

a) $K \subseteq U$ là một khoá khi và chỉ khi $K \xrightarrow{+} U$ (phụ thuộc đầy đủ): *đúng*.

b) Hai khoá khác nhau của một LĐQH không giao nhau: *sai, xem bài 4.39*.

c) Hai khoá khác nhau của một LĐQH không bao nhau: *đúng*.

d) Mọi LĐQH đều có ít nhất một khoá: *đúng*.

e) Tồn tại một LĐQH không có khoá nào: *sai.*

f) Số khoá của một LĐQH không thể lớn hơn số thuộc tính: *sai, xem bài 4.33.*

g) U không thể là khoá của LĐQH (U,F): *sai; khi F = ∅ hoặc F chỉ chứa các PTH tầm thường thì U là khóa.*

h) Mọi LĐQH không thể có hai khoá đơn tức là khoá chỉ gồm một thuộc tính: *sai, xem bài 4.32.*

Bài giải chương 5

CHUẨN HÓA

5.1. Chứng minh rằng nếu $X \to Y$ và $X \cap Y = \varnothing$ thì phép tách $(XY, U\text{-}Y)$ là không tổn thất.

Chứng minh

Đặt $M=XY$, $Z=U\text{-}Y$, ta có $MZ=U$ và $M \cap Z = X$. Ta cần chứng minh $\forall R \in SAT(X \to Y)$: $R[M]*R[Z]=R$. Theo bài 1.33 ta có $R[M]*R[Z] \supseteq R$, ta còn phải chứng minh $R[M]*R[Z] \subseteq R$. Giả sử $u \in R[M]$, $v \in R[Z]$ và u và v thoả điều kiện kết nối tự nhiên, tức là $u.X = v.X$, ta chứng minh $u*v \in R$. Vì $u \in R[M]$ nên $\exists\ u' \in R$: $u = u'.M$. Vì $v \in R[Z]$ nên $\exists\ v' \in R$: $v = v'.Z$. Vì $X = M \cap Z$ nên $X \subseteq M$ và $X \subseteq Z$, từ đó suy ra $u.X = u'.M.X = u'.X = v.X = v'.Z.X = v'.X$. Vì R thỏa PTH $X \to Y$ nên từ $u'.X = v'.X$ ta suy ra $u'.Y = v'.Y$ và do đó $u'.M = u'.XY = v'.XY = v'.M$. Từ đây suy ra $u*v = v' \in R$, *đpcm.*

5.2. Dùng kỹ thuật bảng để kiểm tra tính tổn thất của các phép tách sau:

a) $p = (U,F)$, $U = ABCD$,

$F = \{ A \to B, AC \to D \}$

$\rho = (AB, ACD)$.

$T \Rightarrow T^*$

	A	B	C	D
AB	a_1	a_2	b_{13}	b_{14}
ACD	a_1	b_{22} / a_2	a_3	a_4

Vì T^* chứa dòng thứ hai gồm toàn ký hiệu *phân biệt* nên phép tách đã cho là *không tổn thất*.

b) $p = (U,F)$, $U = ABCDE$,

$F = \{ A \to C, B \to C, C \to D, DE \to C, CE \to A \}$

$p = (AD, AB, BE, CDE)$.

$T \Rightarrow T^*$

	A	B	C	[illegible]	E
AD	a_1	b_{12}	b_{13}/ a_3	a_4	b_{15}
AB	a_1	a_2	b_{23}/b_{13}/a_3	b_{24} / a_4	b_{25}
BE	b_{31}	a_2	b_{33}/b_{13}/a_3	b_{34} / a_4	a_5
CDE	b_{41} / b_{31}	b_{42}	a_3	a_4	a_5

Vì T* *không chứa* một dòng *toàn ký hiệu phân biệt* nên phép tách đã cho là *tổn thất*.

5.3. Một lược đồ quan hệ ở dạng chuẩn 3 và có một khoá duy nhất có thể ở dạng chuẩn Boyce-Codd không? Vì sao?

Có.

Chứng minh: Giả sử LĐ p ở 3NF, có một khóa duy nhất K nhưng không ở BCNF. Vậy phải có một PTH không tầm thường X→A,

$A \notin X$ sao cho X không phải là siêu khóa. Ta có $K \to X$ vì K là khóa, X $! \to K$ vì X không phải là siêu khóa. Nếu A không phải là thuộc tính khóa thì phụ thuộc bắc cầu $K \to X$, $X ! \to K$, $X \to A$, $A \notin X$ chứng tỏ LĐ đã cho không ở 3NF. Vậy A phải là thuộc tính khóa, tức là A nằm trong khóa duy nhất K. Xét tập $M = (K\text{-}A)X$. Vì $M \supseteq X$ nên theo tiên đề phản xạ Armstrong $M \to X$, mặt khác $X \to A$ theo giả thiết nên theo tiên đề bắc cầu Armstrong ta suy ra $M \to A$. Khi đó M^+ chứa A và do đó chứa K. Tức là M là siêu khóa không chứa A. Từ siêu khóa này ta có thể tìm được một khóa $K' \neq K$ (vì K chứa A mà K' không chứa A). Hai khóa này mâu thuẫn với điều kiện khóa K là duy nhất. Vậy giả thiết về sự tồn tại PTH không tầm thường $X \to A$, $A \notin X$ sao cho X không phải là siêu khóa là sai. LĐ phải ở dạng BCNF, đpcm.

5.4. Xác định và giải thích dạng chuẩn cao nhất của LĐQH sau:

p = (U, F); $U = ABCD$, $F = \{A \to C, D \to B, C \to ABD\}$.

LĐ p có 2 khoá là A và C vì $A^+ = C^+ = ABCD = U$. Tập thuộc tính khoá là $U_k = AC$, tập thuộc tính không khoá là $U_0 = BD$. Ta có, $C \to D$ (?), $D ! \to C$ (?), $D \to B$. Đây là phụ thuộc bắc cầu của thuộc tính không khoá B vào khoá C. Vậy p không phải là 3NF và đương nhiên p không phải là BCNF(?). Vì hai khoá A và C đều chỉ có một thuộc tính nên các thuộc tính không khoá khác là B và D đều phụ thuộc đầy đủ vào hai khoá này. Vậy p là 2NF.

5.5. Chứng minh rằng một LĐQH ở dạng chuẩn 3 thì đồng thời ở dạng chuẩn 2.

Chứng minh: Nếu A là thuộc tính không khoá và A không phụ thuộc đầy đủ vào khoá K thì ta có $\exists M \subset K$: $M \to A$. Khi đó $M ! \to K$ (?), $A \notin M$

(?). Đây là một phụ thuộc bắc cầu, mâu thuẫn với giả thiết về 3NF. Vậy mọi thuộc tính không khoá phải phụ thuộc đầy đủ vào mọi khoá, tức là lược đồ phải là 2NF.

5.6. Chứng minh rằng một LĐQH ở dạng chuẩn BC thì đồng thời ở dạng chuẩn 3.

Chứng minh: Suy trực tiếp từ định nghĩa.

5.7. Chuẩn hoá 3NF CSDL **Thực tập**:

SV(SV#, HT, NS, QUE, HL)

DT(DT#, TDT, CN, KP)

SD(SV#, DT#, NTT, KM, KQ)

với các tập PTH sau:

F_{SV} = {SV# → HT, NS, QUE, HL}

F_{DT} = {DT# → TDT, CN, KP}

F_{SD} = {SV#, DT# → NTT, KQ ; NTT → KM}

Giải

Hai quan hệ SV và DT chỉ có 1 khoá đơn (1 thuộc tính) tương ứng là SV# và DT# và hai tập PTH tương ứng chỉ chứa 1 PTH nên hai quan hệ này ở BCNF(?).

Quan hệ SD có 1 khoá là K = SV#DT#. SD không ở 3NF vì có phụ thuộc bắc cầu của thuộc tính không khoá KM vào khoá K. Ta chuẩn hoá SD như sau:

1. *Tách*

SV#, DT# → NTT

$SV\#, DT\# \rightarrow KQ$

$NTT \rightarrow KM$

2. *Tìm phủ tối tiểu*

$G = \{SV\#, DT\# \rightarrow NTT; SV\#, DT\# \rightarrow KQ; NTT \rightarrow KM\}$

3. *Gộp: ta thu lại được F*

4. *Kết quả: ρ = (SV#,DT#,NTT,KQ ; NTT, KM). Ta thu được hai quan hệ*

SD(SV#,DT#,NTT,KQ) với khoá K=SV#,DT# và

NK(NTT,KM) với khoá NTT

5.8. Chuẩn hoá 3NF LĐQH p = (U,F) sau:

U = MLTGSDP,

$F = \{M \rightarrow T, GP \rightarrow M, GT \rightarrow P, MS \rightarrow D, GS \rightarrow P\}$

với ngữ nghĩa sau:

M: Môn học chuyên đề

L: Lớp chuyên đề

T; Thày - giáo viên phụ trách chuyên đề

G: Giờ học chuyên đề

S: Sinh viên theo học chuyên đề

D: Số đăng ký của sinh viên trong chuyên đề đó

P: Phòng học dành cho chuyên đề

$M \rightarrow T$: Mỗi chuyên đề có một thày phụ trách

$GP \rightarrow M$: Tại mỗi thời điểm, mỗi phòng học được dành cho không quá một môn

$GT \to P$: Tại mỗi thời điểm, mỗi thày dạy trong không quá một phòng học

$MS \to D$: Mỗi sinh viên tham gia chuyên đề nào thì được cấp một mã số ghi danh theo chuyên đề đó

$GS \to P$: Tại mỗi thời điểm, mỗi sinh viên có mặt trong không qúa một phòng học

1. Tìm một khoá: Tập thuộc tính có trong mọi khoá $K = LGS$. Ta có $K^+ = (LGS)^+ = LGSPMTD = U$ do đó lược đồ p có một khoá duy nhất là K. Tập các thuộc tính khoá: $U_k = LGS$, tập các thuộc tính không khoá: $U_0 = PMTD$.

2. Vì bộ phận thực sự của khoá K là GS và $GS \to P$, $P \in U_0$ nên p không ở 2*NF* và do đó không ở 3*NF*.

3. Để ý rằng F đã ở dạng tối tiểu, ta có:

$\rho = (MT, GPM, GTP, MSD, GSP)$

Khoá $K = LGS$ không có trong thành phần nào của ρ nên ta thêm K vào để thu được kết quả sau:

$\rho = (MT, GPM, GTP, MSD, GSP, LGS)$

5.9. Chứng minh sự tương đương của hai định nghĩa *d* và *e* về *BCNF*.

LĐQH $\rho = (U,F)$ được gọi là

d) ở dạng chuẩn 3 mạnh hoặc Boyce-Codd (3*SNF*, *BCNF*) nếu *p* ở 1*NF* và mọi thuộc tính đều phụ thuộc trực tiếp vào mọi khoá,

e) ở dạng chuẩn 3 mạnh hoặc Boyce-Codd (3*SNF*, *BCNF*) nếu *p* ở 1NF và mọi PTH không tầm thường $X \to Y$ đều cho ta X là một siêu khóa.

Chứng minh

[$d \Rightarrow e$] Giả sử lược đồ p ở 1*NF* và mọi thuộc tính đều phụ thuộc trực tiếp vào mọi khoá và $X \rightarrow Y$ là một PTH không tầm thường. Ta phải chứng minh X là một siêu khoá. Vì PTH $X \rightarrow Y$ không tầm thường nên ta phải có $X \rightarrow A$ với $A \in Y \backslash X$. Nếu X không phải là siêu khoá thì với mọi khoá K ta có $X \,!\!\rightarrow K$, $K \rightarrow X$, $X \rightarrow A$, $A \notin X$, tức là A phụ thuộc bắc cầu vào khoá K, mâu thuẫn với định nghĩa d. Vậy X phải là siêu khoá, tức là ta có e, đpcm.

[$e \Rightarrow d$] Giả sử lược đồ p ở 1*NF* và mỗi *PTH* không tầm thường $X \rightarrow Y$ đều cho ta X là một siêu khoá. Cho khoá K của p và một thuộc tính A trong U. Giả sử A phụ thuộc bắc cầu vào K, tức là $\exists\, X \subseteq U$: $K \rightarrow X$, $X \,!\!\rightarrow K$, $X \rightarrow A$, $A \notin X$. Như vậy $X \rightarrow A$ là PTH không tầm thường nên ta phải có X là một siêu khoá của p và do đó $X \rightarrow K$, mâu thuẫn với giả thiết về phụ thuộc bắc cầu. Vậy mọi thuộc tính A phải phụ thuộc trực tiếp vào mọi khoá của p, tức là ta có d, đpcm.

BÀI GIẢI

CÁC ĐỀ THI

BÀI GIẢI

ĐỀ THI TUYỂN CAO HỌC ĐẠI HỌC THÁI NGUYÊN NĂM 2002

Môn: Cơ sở công nghệ thông tin

Câu 5.

a.1) Định nghĩa quan hệ:

Cho tập hữu hạn $U = \{A_1, A_2, \ldots, A_n\}$ khác trống ($n \geq 1$). Các phần tử của U được gọi là *thuộc tính*. Ứng với mỗi thuộc tính $A_i \in U$, $i = 1,2, \ldots, n$ có một tập không rỗng $dom(A_i)$ được gọi là *miền trị* (miền biến thiên) của thuộc tính A_i. Ta đặt $D = \bigcup_{A_i \in U} dom(A_i)$.

Một *quan hệ R* với các thuộc tính $U = \{A_1, A_2, \ldots, A_n\}$, 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.

a.2) Định nghĩa phụ thuộc hàm:

Cho tập thuộc tính U. Một phụ thuộc hàm (PTH) trên U là công thức dạng $f: X \rightarrow Y$; $X, Y \subseteq U$.

Cho quan hệ $R(U)$ và một PTH $f: X \rightarrow Y$ trên U. Ta nói quan hệ R thoả PTH f và viết $R(f)$, nếu hai bộ tuỳ ý trong R giống nhau trên X thì chúng cũng giống nhau trên Y,

$(\forall\, u,v \in R): (u.X = v.X) \Rightarrow (u.Y = v.Y)$

b.1) Phát biểu hệ tiên đề Armstrong:

Cho tập PTH *F* trên tập thuộc tính *U*. Bao đóng của *F*, ký hiệu F^+ là tập nhỏ nhất các PTH trên *U* chứa *F* và thoả các tính chất F1-F3 của hệ tiên đề Armstrong $A°$ sau đây:

F1. Tính phản xạ: Nếu $X \supseteq Y$ thì $X \rightarrow Y \in F^+$

F2. Tính gia tăng: Nếu $X \rightarrow Y \in F^+$ thì $XZ \rightarrow YZ \in F^+$

F3. Tính bắc cầu: Nếu $X \rightarrow Y \in F^+$ và $Y \rightarrow Z \in F^+$ thì $X \rightarrow Z \in F^+$

Nếu $f \in F^+$ ta nói tập PTH *F* dẫn ra được PTH *f* theo tiên đề Armstrong.

Cho tập PTH *F* trên tập thuộc tính *U* và *f* là một PTH trên *U*. Ta nói PTH *f* được suy dẫn theo quan hệ từ tập PTH *F*, nếu mọi quan hệ *R*(*U*) thoả *F* thì *R* cũng thoả *f*. Ta ký hiệu F^* là tập toàn bộ các PTH *f* được suy dẫn theo quan hệ từ tập PTH *F*.

b.2) Phát biểu tính đúng của hệ tiên đề Arrnstrong:

Định lý: *Cho tập PTH F trên tập thuộc tính U và một PTH f trên U. Nếu f được dẫn từ F theo tiên đề Armstrong thì f cũng được dẫn từ F theo quan hệ, tức là* $F^+ \subseteq F^*$.

c.1) Định nghĩa lược đồ quan hệ: Lược đồ quan hệ (LĐQH) là một cặp $p = (U,F)$, trong đó *U* là tập hữu hạn các thuộc tính, *F* là tập các phụ thuộc hàm trên *U*.

c.2) Định nghĩa khoá của lược đồ quan hệ: Cho LĐQH $p = (U,F)$. Tập thuộc tính $K \subseteq U$ được gọi là khoá của LĐ *p* nếu

(i) $K^+ = U$

(ii) $\forall A \in K: (K - \{A\})^+ \neq U$

d.1) Phát biểu bài toán thành viên đối với lược đồ quan hệ:

Cho tập thuộc tính U, một tập các PTH F trên U và một PTH $f: X \rightarrow Y$ trên U. Hỏi rằng $X \rightarrow Y \in F^+$ hay không ?

d.2) Kết quả liên quan tới bài toán thành viên:

Định lý *: $X \rightarrow Y \in F^+$ khi và chỉ khi $Y \subseteq X^+$.*

Vì đã có thuật toán tìm bao đóng của tập thuộc tính với độ phức tạp thời gian đa thức nên định lý trên cho ta thuật toán giải bài toán thành viên với độ phức tạp thời gian đa thức.

Câu 6.

a.1) Thuật toán tính bao đóng của một tập thuộc tính trong một lược đồ quan hệ:

Ý tưởng: Cho tập PTH F trên tập thuộc tính U và một tập con các thuộc tính X trong U. Để xác định bao đóng của tập thuộc tính X, X^+ ta xuất phát từ tập X và bổ sung dần cho X các thuộc tính thuộc vế phải R của các PTH $L \rightarrow R \in F$ thỏa điều kiện $L \subseteq X$. Thuật toán sẽ dừng khi không thể bổ sung thêm thuộc tính nào cho X.

Algorithm Closure

Format: $C = X^+$

Input: - *LĐQH* $p = (U, F)$

- *Tập thuộc tính* $X \subseteq U$

Output: - $Y = X^+ = \{ A \in U \mid X \rightarrow A \}$

Method

$Y := X;$

```
repeat
        Z:=Y;
        for each FD L→R in F do
                if L ⊆ Y then
                        Y:=Y ∪ R;
                endif;
        endfor;
    until Y=Z;
    return Y;
end Closure;
```


b) Thuật toán tìm một khoá của lược đồ quan hệ:

Tư tưởng: Xuất phát từ một siêu khóa K, loại dần các thuộc tính trong K, bảo toàn bất biến $K^+ = U$.

```
Algorithm Key
    Format: Key(U,F)
    Input:        - Tập thuộc tính U
                  - Tập PTH F
    Output: - Khóa K ⊆ U thỏa
        •  K⁺ = U
        •  ∀ A ∈ K: (K-{A})⁺ ≠ U
    Method
        K:=U;
```

```
    for each attribute A in U do
        if (K-{A})⁺ = U then
            K := K-{A}
        endif;
    endfor;
    return K;
end Key;
```

c) Giả thiết:

LĐQH $s = (R,F)$;

$R = abcdefg$,

$F = \{ abc \to de, bcd \to g, abf \to eg, ce \to fg \}$

Tìm một khoá của lược đồ s.

LĐQH s có khóa $K = abc$, vì:

$(abc)^+ = abcdefg = U$ và $(bc)^+ = bc \neq U$, $(ac)^+ = ac \neq U$, $(ab)^+ = ab \neq U$.

BÀI GIẢI

ĐỀ THI TUYỂN NGHIÊN CỨU SINH VIỆN CÔNG NGHỆ THÔNG TIN NĂM 2002

Đề số 1

Câu 1.

1.a) Định nghĩa lược đồ quan hệ:

Lược đồ quan hệ (LĐQH) là một cặp $p = (U,F)$, trong đó U là tập hữu hạn các thuộc tính, F là tập các phụ thuộc hàm trên U.

1.b) Định nghĩa bao đóng của tập thuộc tính trong LĐQH:

Cho tập PTH F trên tập thuộc tính U và một tập con các thuộc tính X trong U. Bao đóng của tập thuộc tính X, ký hiệu X^+ là tập thuộc tính $X^+ = \{ A \in U \mid X \to A \in F^+ \}$

1.c) Thuật toán tìm bao đóng của tập thuộc tính trong LĐQH:

Ý tưởng: Cho tập PTH F trên tập thuộc tính U và một tập con các thuộc tính X trong U. Để xác định bao đóng của tập thuộc tính X, X^+ ta xuất phát từ tập X và bổ sung dần cho X các thuộc tính thuộc vế phải R của các PTH $L \to R \in F$ thỏa điều kiện $L \subseteq X$. Thuật toán sẽ dừng khi không thể bổ sung thêm thuộc tính nào cho X.

```
Algorithm Closure
```

```
Format: C = X⁺
Input:        - LĐQH p =(U,F)
              - Tập thuộc tính X ⊆ U
Output: - Y = X⁺ = { A ∈ U  | X → A ∈ F⁺ }
Method
      Y:=X;
    repeat
          Z:=Y;
          for each FD L→R in F do
                if L ⊆ Y then
                      Y:=Y ∪ R;
                endif;
          endfor;
      until Y=Z;
      return Y;
  end Closure;
```

Câu 2.

Giả thiết:

LĐQH $p = (U, F)$;

$U = ABCDEGH$;

$F = \{CD \to H,\ E \to B,\ D \to G,\ BH \to E,\ CH \to DG,\ C \to A\}$.

a) Tìm tập M là giao của toàn bộ các khoá của p:

$$M = U - \bigcup_{L \to R \in F}(R - L) = C$$

b) Cho biết *p* có đúng 1 khoá hay không?

Vì $C^+ = CA \neq U$ nên LĐQH p có hơn 1 khóa.

c) Tập *ABD* có phải là khoá của *p* không? Vì sao?

ABD không phải là khóa của p vì nó không chứa C là phần có mặt trong mọi khóa.

d) Tập *CH* có phải là khoá của *p* không? Vì sao?

Tập *CH* không phải là khoá của *p* vì $(CH)^+ = ACDGH \neq U$.

e) Tính $Z = (X^+ Y)^+ \cap (K^+ - Y)$ *biết* $X = CD$, $Y = CH$, K là một siêu khoá của p.

Vì K là siêu khóa nên $K^+ = U$, do đó $K^+ - Y = ABCDEGH - CH = ABDEG$. Ta lại biết $(X^+ Y)^+ = (X Y)^+ = (CDH)^+ = ACDGH$, nên

$Z = ACDGH \cap ABDEG = ADG$.

Câu 3. Phép phân tách một LĐQH:

3.a) Cho lược đồ quan hệ p = (U,F). Một phép tách trên tập thuộc tính U là một họ các tập con của U $\rho = (X_1, X_2,...,X_k)$ thỏa tính chất:

$$\bigcup_{i=1}^{k} X_i = U.$$

Phép tách $\rho = (X_1, X_2,...,X_k)$ trên tập thuộc tính U được gọi là *không tổn thất* (hoặc *không mất thông tin)* đối với tập PTH F nếu

$\forall R(U) \in SAT(F)$: $R[X_1]*R[X_2]*...*R[X_k] = R$. Ngược lại, nếu không tồn tại đẳng thức thì ta gọi ρ là phép tách *tổn thất*.

Kiểm tra tính tổn thất của phép tách bằng kỹ thuật bảng

Input

- LĐQH $p = (U,F)$
- Phép tách $\rho = (X_1, X_2,...,X_k)$

Output

- True, nếu ρ là một phép tách không tổn thất
- False, ngoài ra.

Method

1. *Khởi trị:* Lập bảng T với các cột là các thuộc tính trong U và k dòng, mỗi dòng ứng với một thành phần của X_i trong ρ: Dòng i chứa các ký hiệu phân biệt (KHPB) a_j ứng với các thuộc tính A_j trong X_i và các ký hiệu không phân biệt (KHKPB) b_{ij} ứng với các thuộc tính A_j trong $U\text{-}X_i$. Chú ý rằng mọi KHPB trong cột j của T là giống nhau và bằng a_j còn mọi KHKPB trong bảng T lúc đầu là khác nhau.

2. *Sửa bảng:* Lặp đến khi bảng T không còn thay đổi:

2.1. Vận dụng các F-luật để biến đổi bảng như sau:

Với mỗi PTH $L \rightarrow R$ trong F, nếu trong bảng T có chứa hai dòng u và v giống nhau trên L thì sửa các ký hiệu của chúng cho giống nhau trên mọi cột A trong bảng T như sau:

a) nếu u.A = v.A: không sửa,

b) nếu chỉ một trong hai ký hiệu u.A hoặc v.A là KHPB thì sửa ký mọi xuất hiện trong bảng của KHKPB thành KHPB đó,

c) nếu cả hai ký hiệu u.A và v.A đều là KHKPB thì sửa mọi xuất hiện trong bảng của ký hiệu có chỉ số thứ nhất lớn hơn thành ký hiệu thứ hai.

3. *Kết luận:*

Gọi bảng kết quả là T*.

Nếu T* chứa một dòng toàn KHPB thì return True nếu không return False.

end.

3.c) Cho LĐQH *p* như trong câu 2. Vận dụng thuật toán kiểm tra phép phân tách *w* có mất thông tin hay không?

Giả thiết

$p = (U, F)$,

$U = ABCDEGH$,

$F = \{ CD \rightarrow H,\ E \rightarrow B,\ D \rightarrow G,\ BH \rightarrow E,\ CH \rightarrow DG,\ C \rightarrow A \}$.

$w = [ABCDE, BCH, CDEGH]$

$T \Rightarrow T^*$

	A	B	C	D	E	G	H
ABCDE	a_1	a_2	a_3	a_4	a_5	b_{16} / a_6	b_{17} / a_7
BCH	b_{21}/a_1	a_2	a_3	b_{24}/a_4	b_{25} / a_5	b_{26} / a_6	a_7
CDEGH	b_{31}/a_1	b_{32}/a_2	a_3	a_4	a_5	a_6	a_7

Do T có chứa dòng toàn ký hiệu phân biệt nên phép tách w đã cho là không tổn thất.

Câu 4.

4.a) Định nghĩa các dạng chuẩn *2NF*, *3NF* và *BCNF* của LĐQH:

LĐQH p = (U,F) được gọi là

- ở dạng chuẩn 1 (1NF) nếu mọi thuộc tính trong U đều không phải là thuộc tính phức hợp,
- ở dạng chuẩn 2 (2NF) nếu p ở 1NF và mọi thuộc tính không khoá đều phụ thuộc đầy đủ vào mọi khoá,
- ở dạng chuẩn 3 (3NF) nếu p ở 1NF và mọi thuộc tính không khoá đều phụ thuộc trực tiếp vào mọi khoá,
- ở dạng chuẩn 3 mạnh hoặc Boyce-Codd (3SNF, BCNF) nếu p ở 1NF và mọi PTH không tầm thường X→Y đều cho ta X là một siêu khóa.

4.b) Sơ đồ tương quan giữa các dạng chuẩn:

BCNF ⇒ 3NF ⇒ 2NF

4.c) Xác định dạng chuẩn cao nhất của LĐQH *h* sau đây:

$h = (U, F)$; $U = ABCD$, $F = \{ D \to B,\ C \to A,\ B \to ACD \}$

Giải thích vì sao?

Lược đồ h có 2 khóa $K_1 = D$; $K_2 = B$ *vì* $D^+ = ABCD$ và $B^+ = ABCD$.

Vì có PTH $C \to A$ mà *C* không phải là siêu khóa nên h không thể ở dạng chuẩn BCNF.

Tập thuộc tính khóa là $U_K = BD$; Tập thuộc tính không khóa là $U_0 = AC$;

Ta có:

$B \to C$ (vì *B* là khóa),

C !→ *B* (vì *C* không là khóa)

C → *A* (giả thiết)

và A là thuộc tính không khóa. Như vậy thuộc tính không khóa *A* phụ thuộc bắc cầu vào khóa *B* nên *h* không ở dạng chuẩn 3NF.

Hai thuộc tính không khóa *A* và *C* đều phụ thuộc đầy đủ vào các khóa *B* và *D* do đó lược đồ *h ở dạng chuẩn 2NF*.

BÀI GIẢI

ĐỀ THI TUYỂN NGHIÊN CỨU SINH VIỆN CÔNG NGHỆ THÔNG TIN NĂM 2002

Đề số 2

Câu 1.

1.a) Định nghĩa PTH: Cho tập thuộc tính U. Một phụ thuộc hàm (PTH) trên U là công thức dạng f: $X \rightarrow Y$; $X, Y \subseteq U$.

Cho quan hệ R(U) và một PTH f: $X \rightarrow Y$ trên U. Ta nói quan hệ R thoả PTH f và viết R(f), nếu hai bộ tuỳ ý trong R giống nhau trên X thì chúng cũng giống nhau trên Y,

$$(\forall u,v \in R): (u.X = v.X) \Rightarrow (u.Y = v.Y)$$

1.b) Định nghĩa LĐQH: Lược đồ quan hệ (LĐQH) là một cặp p = (U,F), trong đó U là tập hữu hạn các thuộc tính, F là tập các phụ thuộc hàm trên U.

1.c) Phát biểu bài toán thành viên trên LĐQH: Cho tập thuộc tính U, một tập các PTH F trên U và một PTH f: $X \rightarrow Y$ trên U. Hỏi rằng

$X \rightarrow Y \in F^+$ hay không ?

1.d) Điều kiện cần và đủ liên quan đến bài toán thành viên:

Định lý: $X \rightarrow Y \in F^+$ khi và chỉ khi $Y \subseteq X^+$.

Câu 2.

2.a) Định nghĩa khoá của LĐQH: Cho LĐQH p = (U,F). Tập thuộc tính $K \subseteq U$ được gọi là khoá của LĐ p nếu

(i) $K^+ = U$

(ii) $\forall A \in K: (K - \{A\})^+ \neq U$

2.b) Thuật toán tìm 1 khoá của LĐQH:

Tư tưởng: Xuất phát từ một siêu khóa K, loại dần các thuộc tính trong K, bảo toàn bất biến $K^+ = U$.

```
Algorithm Key
    Format:  Key(U,F)
    Input:          - Tập thuộc tính U
                    - Tập PTH F
    Output:  - Khóa K ⊆ U thỏa
         •  K+=U
         •  ∀ A ∈ K: (K-{A})+ ≠ U
    Method
         K:=U;
       for each attribute A in U do
            if (K-{A})+ = U then
                K := K-{A}
            endif;
       endfor;
       return K;
    end Key;
```

Câu 3.

3.a) Định nghĩa thuộc tính khoá (thuộc tính cơ bản hay nguyên thuỷ), thuộc tính không khoá (thuộc tính thứ cấp): Cho LĐQH p = (U,F). Thuộc tính A trong U được gọi là thuộc tính *khóa* nếu A có trong một khóa của p. A được gọi là thuộc tính *không khóa* nếu A không có trong bất kỳ khóa nào của p.

3.b)

Giả thiết:

LĐQH $s = (U, F)$

$U = ABCD$,

$F = \{AD \to BC,\ B \to A,\ D \to C\}$.

a) Tìm các khoá của *s:* Lược đồ s có 2 khóa:

$K_1 = BD$, vì $(BD)^+ = U$; $B^+ = AB \neq U$; $D^+ = CD \neq U$.

$K_2 = AD$; vì $(AD)^+ = U$; $A^+ = A \neq U$ và $D^+ \neq U$;

b) Cho biết *C* có phải là thuộc tính khoá hay không? *C* không phải là thuộc tính khóa của s vì *C* không có trong khóa nào.

Câu 4.

4.a) Định nghĩa các dạng chuẩn *2NF, 3NF* và *BCNF* của LĐQH: LĐQH p = (U,F) được gọi là

- *ở dạng chuẩn 1* (1NF) nếu mọi thuộc tính trong U đều không phải là thuộc tính phức hợp,
- *ở dạng chuẩn 2* (2NF) nếu p ở 1NF và mọi thuộc tính không khoá đều phụ thuộc đầy đủ vào mọi khoá,

- *ở dạng chuẩn 3* (3NF) nếu p ở 1NF và mọi thuộc tính không khoá đều phụ thuộc trực tiếp vào mọi khoá,
- *ở dạng chuẩn 3 mạnh* hoặc *Boyce-Codd* (3SNF, BCNF) nếu p ở 1NF và mọi PTH không tầm thường X→Y đều cho ta X là một siêu khóa.

4.b) Sơ đồ tương quan giữa các dạng chuẩn:

$$BCNF \Rightarrow 3NF \Rightarrow 2NF$$

4.c) Xác định dạng chuẩn cao nhất của LĐQH *h* sau:

$h = (U, F); U = ABCD, F = \{ CD \rightarrow B, A \rightarrow C, B \rightarrow ACD \}$

Giải thích vì sao?

Lược đồ *h* có 3 khóa: $K_1 = B$, vì $B^+ = U$;

$K_2 = CD$, vì $(CD)^+ = U, C^+ = C \neq U, D^+ = D \neq U.$

$K_3 = AD$, vì $(AD)^+ = U, A^+ = AC \neq U, D^+ = D \neq U.$

Tập thuộc tính khóa là $U_K = ABCD = U$, vậy *h* không có thuộc tính không khóa.

Ta có PTH $A \rightarrow C$ mà *A* không phải là siêu khóa nên *h* không thể ở dạng chuẩn BCNF.

Do *h* không có thuộc tính không khỏa nên không có phụ thuộc bắc cầu của thuộc tính không khóa vào các khóa. Vậy *h ở dạng chuẩn 3NF*.

Câu 5.

Giả thiết:

$p = (U, F); U = ABCDEHG;$

$F = \{ DE \rightarrow G, H \rightarrow C, E \rightarrow A, CG \rightarrow H, DG \rightarrow EA, D \rightarrow B \}$..

a. Tìm tập *M* là giao của toàn bộ các khoá của *p*. Cho biết *p* có đúng 1 khoá hay không?

$M = U - ABCGEH = D$. $M^+ = D^+ = BD \neq U$ nên p có hơn 1 khóa.

b. Tìm 1 khoá của *p*: *K = DEH, vì* $(DEH)^+ = DEHGCAB$,

$(DE)^+ = DEGAB \neq U$, $(DH)^+ = DHCB \neq U$, $(EH)^+ = EHAC \neq U$.

c. Tập *BCE* có phải là khoá của *p* không? Vì sao?

Không, vì BCE không chứa D là giao các khóa.

d. Hãy thêm hoặc bớt 1 phụ thuộc hàm cho *F* để LĐQH có đúng 1 khoá. *Thêm chẳng hạn PTH* $D \to ABCEGH$. *Khi đó D vẫn là giao các khóa và* $D^+ = U$ *nên p có đúng 1 khóa.*

BÀI GIẢI

ĐỀ THI TUYỂN NGHIÊN CỨU SINH ĐẠI HỌC BÁCH KHOA HÀ NỘI NĂM 2002

Bài 1.

1a) Định nghĩa lược đồ quan hệ:

LĐQH là một cặp $p = (U,F)$ trong đó U là tập hữu hạn các thuộc tính, F là tập PTH trên U.

1b) Định nghĩa khoá của lược đồ quan hệ: Cho LĐQH $p = (U,F)$. Khóa của LĐQH là tập con thuộc tính K của U thỏa

(i) $K^+ = U$

(ii) $\forall A \in K$: $(K-\{A\})^+ \neq U$.

1c) Thuật toán tìm một khoá của lược đồ quan hệ:

Tư tưởng: Xuất phát từ một siêu khóa K, loại dần các thuộc tính trong K, bảo toàn bất biến $K^+ = U$.

Algorithm Key

```
Format: Key(U,F)
Input:          - Tập thuộc tính U
                - Tập PTH F
Output: - Khóa K ⊆ U thỏa
```

```
    •   K+=U
    •   ∀ A ∈ K: (K-{A})+ ≠ U
Method
    K:=U;
    for each attribute A in U do
        if (K-{A})+ = U then
            K := K-{A}
        endif;
    endfor;
    return K;
end Key;
```

Bài 2. Giả thiết:

$p = (U,F)$; $U = ABCDEH$

$F = \{ BC \rightarrow E,\ D \rightarrow A,\ C \rightarrow A,\ AE \rightarrow D,\ BE \rightarrow CH \}$.

a) Tìm một khoá K của lược đồ p:

LĐ p có khóa $K = BC$ vì

$(BC)^+ = BCEADH = U$, $B^+ = B \neq U$, $C^+ = CA \neq U$.

b) Ngoài khoá K, lược đồ p còn khoá nào khác không? Vì sao? Giao của các khóa: $M = U - EADCH = B$; $M^+ = B^+ = B \neq U$. Vậy LĐ có hơn 1 khóa.

c) Tập BCH có phải là khoá của p không? Vì sao? BCH không phải là khóa của p vì BCH chứa thực sự khóa $K = BC$ tìm được ở câu a.

d) Tập BD có phải là khoá của p không? Vì sao?

Vì $(BD)^+ = BDA \neq U$ nên BD không phải là khóa của p.

e) Tính $Z = (X^+ \cup Y)^+ \cap K^+ - (X \cup Y)$ với $X = AB$, $Y = D$, K là một siêu khoá của p.

Vì K là siêu khóa nên $K^+ = U$, mặt khác $(X^+ \cup Y)^+ = (X \cup Y)^+$, do đó $Z = (X \cup Y)^+ - (X \cup Y) = (ABD)^+ - ABD = ABD\text{-}ABD = \varnothing$.

f) Hãy thêm cho F một phụ thuộc hàm để p có đúng một khoá. Giải thích cách làm:

Vì giao các khóa là B nên ta có thể thêm PTH $B \rightarrow ACDEH$. Khi đó giao các khóa vẫn là B và ta có $B^+ = U$ nên LĐ có đúng 1 khóa.

Bài 3. Một lược đồ quan hệ ở dạng chuẩn 3 và có một khoá duy nhất có thể ở dạng chuẩn Boyce-Codd không? Vì sao? *Có.*

Chứng minh: Giả sử LĐ p ở 3NF, có một khóa duy nhất K nhưng không ở BCNF. Vậy phải có một PTH không tầm thường $X \rightarrow A$, $A \not\subseteq X$ sao cho X không phải là siêu khóa. Ta có $K \rightarrow X$ vì K là khóa, $X\ ! \rightarrow K$ vì X không phải là siêu khóa. Nếu A không phải là thuộc tính khóa thì phụ thuộc bắc cầu $K \rightarrow X$, $X\ ! \rightarrow K$, $X \rightarrow A$, $A \not\subseteq X$ chứng tỏ LĐ đã cho không ở 3NF. Vậy A phải là thuộc tính khóa, tức là A nằm trong khóa duy nhất K. Xét tập $M = (K\text{-}A)X$. Vì $M \supseteq X$ nên theo tiên đề phản xạ Armstrong $M \rightarrow X$, mặt khác $X \rightarrow A$ theo giả thiết, nên theo tiên đề bắc cầu Armstrong ta suy ra $M \rightarrow A$. Khi đó M^+ chứa A và do đó chứa K. Tức là M là siêu khóa không chứa A. Từ siêu khóa này ta có thể tìm được một khóa $K' \neq K$ (vì K chứa A mà K' không chứa A). Hai khóa này mâu thuẫn với điều kiện khóa K là duy nhất. Vậy giả thiết về sự tồn tại PTH không tầm thường $X \rightarrow A$, $A \not\subseteq X$ sao cho X không phải là siêu khóa là sai. LĐ phải ở dạng BCNF, đpcm.

BÀI GIẢI

ĐỀ THI TUYỂN CAO HỌC NGÀNH GIS NĂM 2003
ĐẠI HỌC BÁCH KHOA TP HỒ CHÍ MINH
Môn: Cơ sở dữ liệu, đề số 1

Thời gian 120 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, XY = X∪Y. Các từ viết tắt: LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1.

a) Định nghĩa phép trừ hai quan hệ:

Phép trừ (theo lý thuyết tập hợp hoặc *lấy phần riêng*) hai quan hệ tương thích *R*(*U*) và *S*(*U*), ký hiệu *R*-*S*, cho ta quan hệ chứa các bộ của quan hệ *R* không có trong quan hệ *S*,

$$P(U) = R\text{-}S = \{\, t \mid t \in R,\ t \notin S \,\}$$

b) Thuật toán xác định phép trừ hai quan hệ

```
Algorithm Substraction
Format: P = R-S
Input:         - Quan hệ R(U)
               - Quan hệ S(U)
```

```
Output: - Quan hệ R-S={t|t∈R,t∉S}
Method
      Create(P,Attr(R));
        for each tuple u in R
              with u not_in S do
                    add u to P;
        endfor;
        return P;
end  Substraction;
```

Câu 2.

a) *Định nghĩa LĐQH: Lược đồ quan hệ (LĐQH) là một cặp p = (U,F), trong đó U là tập hữu hạn các thuộc tính, F là tập các phụ thuộc hàm trên U.*

b) *Định nghĩa bao đóng của tập thuộc tính trong LĐQH: Cho LĐQH p = (U,F), trong đó U là tập thuộc tính, F là tập PTH trên U. Bao đóng của tập thuộc tính X, ký hiệu X^+ là tập thuộc tính*

$$X^+ = \{A \in U \mid X \rightarrow A \in F^+\}$$

c) Thuật toán tìm bao đóng của tập thuộc tính trong LĐQH:

Ý tưởng: *Cho tập PTH F trên tập thuộc tính U và một tập con các thuộc tính X trong U. Để xác định bao đóng của tập thuộc tính X, X^+ ta xuất phát từ tập X và bổ sung dần cho X các thuộc tính thuộc vế phải R của các PTH $L \rightarrow R \in F$ thỏa điều kiện $L \subseteq X$. Thuật toán sẽ dừng khi không thể bổ sung thêm thuộc tính nào cho X.*

```
Algorithm Closure
Format: C = X⁺
Input:          - LĐQH p =(U,F)
                - Tập thuộc tính X ⊆ U
Output: - Y = X⁺ = {A∈U | X→A∈F⁺}
Method
        Y:=X;
     repeat
            Z:=Y;
            for each FD L→R in F do
                 if L ⊆ Y then
                        Y:=Y ∪ R;
                 endif;
            endfor;
        until Y=Z;
        return Y;
  end Closure;
```

Câu 3.

Giả thiết:

LĐQH $p = (U, F)$, $U = ABCDE$,

$F = \{ DE \to A,\ B \to C,\ E \to AD \}$.

a) Tìm một khoá của lược đồ p. Lược đồ p có khoá $K=BE$, vì $(BE)^+=U$ và $B^+ = BC \neq U$; $E^+ = ADE \neq U$.

b) Tập BCE có phải là khoá của p không? Vì sao?

Không, vì BCE chứa BE, mà BE là khoá.

c) Tập AD có phải là khoá của p không? Vì sao?

Không, vì $(AD)^+ = AD \neq U$.

d) Lược đồ p còn khoá nào nữa không? Vì sao?

Ta tìm tập *M* là giao của toàn bộ các khoá của *p*:

$$M = U - \bigcup_{L \to R \in F} (R - L).$$

M = ABCDE - ACD = BE = K do đó LĐQH p chỉ có một khoá duy nhất.

e) Tính $Z = (X^+Y)^+ \cap (K^+ - Y)$ biết X = DE, Y = AD, K là một siêu khoá của p.

Nhận xét: Vì K là siêu khoá của p nên $K^+ = U$. Mặt khác, theo công thức $(X^+Y)^+ = (XY)^+$ nên ta có

$Z = (XY)^+ \cap (U - Y) = (DEA)^+ \cap BCE = DEA \cap BCE = E$.

f) Có thể thêm vào F một PTH để thu được lược đồ có đúng 2 khoá không. Giải thích cách làm?

Thêm, chẳng hạn PTH $C \to B$. Khi đó ngoài khoá BE lược đồ có thêm khoá CE, vì $C \to B$. Dễ thấy có thể tách LĐQH p thành 2 LĐQH thành phần độc lập nhau là p_1=(ADE, {DE→A, E→AD}) với khoá K_1= E và p_2 =(BC, {B→C}) với khoá K_2 = B. Khi thêm PTH $C \to B$ cho LĐQH p_2 ta có thêm khoá K_3 = C. Hợp nhất 2 lược đồ này lại ta sẽ được đúng 2 khoá là

K_1K_2 = EB và K_1K_3 = EC.

BÀI GIẢI

ĐỀ THI TUYỂN CAO HỌC NGÀNH GIS NĂM 2003
ĐẠI HỌC BÁCH KHOA TP HỒ CHÍ MINH
Môn: Cơ sở dữ liệu, đề số 2

Thời gian 120 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, XY = X∪Y. Các từ viết tắt: LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1.

a) Định nghĩa phép toán kết nối tự nhiên hai quan hệ: Phép kết nối (tự nhiên) hai quan hệ *R*(*U*) và *S*(*V*), ký hiệu *R*S*, cho ta quan hệ chứa các bộ được dán từ các bộ *u* của quan hệ *R* với mỗi bộ *v* của quan hệ *S* sao cho các trị trên miền thuộc tính chung (nếu có) của hai bộ này giống nhau.

$$P(UV) = R*S = \{ u*v \mid u \in R, v \in S, u.M = v.M, M = U \cap V \}$$

b) Thuật toán xác định phép toán kết nối tự nhiên hai quan hệ:


```
Algorithm Join
Format: P = R*S
Input:          - Quan hệ R(U)
                - Quan hệ S(V)
Output: - Quan hệ
```

```
R*S={u*v|u∈R,v∈S,u.M=v.M,M=U∩V}
Method
        X = Attr(R) ∪ Attr(S);
        M = Attr(R) ∩ Attr(S);
    Create(P,X);
        for each tuple u in R do
              for each tuple v in S do
              if u.M = v.M then
                          add u*v to P ;
                    endif;
              endfor;
        endfor;
        return P;
end Join;
```

Câu 2.

a) *Định nghĩa LĐQH: Lược đồ quan hệ (LĐQH) là một cặp p = (U,F), trong đó U là tập hữu hạn các thuộc tính, F là tập các phụ thuộc hàm trên U.*

b) *Định nghĩa khoá của lược đồ quan hệ: Cho LĐQH p = (U,F). Khóa của LĐQH là tập con thuộc tính K của U thỏa*

(i) $K^+ = U$

(ii) $\forall A \in K: (K-\{A\})^+ \neq U$.

c) *Thuật toán tìm một khoá của lược đồ quan hệ:*

Tư tưởng: *Xuất phát từ một siêu khóa K, loại dần các thuộc tính trong K, bảo toàn bất biến $K^+ = U$.*

```
Algorithm Key
        Format: Key(U,F)
        Input:        - Tập thuọc tính U
                        Tạp PTH F
        Output: - Khóa K ⊆ U thỏa
            •   K+ = U
            •   ∀ A ∈ K: (K-{A})+ ≠ U
        Method
                K:=U;
                for each attribute A in U do
                        if (K-{A})+ = U then
                            K := K-{A}
                        endif;
                endfor;
                return K;
            end Key;
```

Câu 3.

Giả thiết:

LĐQH $p = (U, F)$; $U = ABCDE$;

$F = \{ EA \to B,\ C \to D,\ A \to BE \}$.

a) Tìm một khoá của lược đồ p:

Ta tìm tập *M* là giao của toàn bộ các khoá của *p:*

$$M = U - \bigcup_{L \to R \in F} (R - L)$$

M = ABCDE - BDE = AC, vì $(AC)^+$=U nên p có một khoá duy nhất là M = AC.

b) Tập CDA có phải là khoá của p không? Vì sao?

Không, vì CDA chứa khoá AC.

c) Tập BE có phải là khoá của p không? Vì sao?

Không, vì p chỉ có một khoá duy nhất là AC.

d) Lược đồ p còn khoá nào nữa không? Vì sao?

LĐQH p chỉ có một khoá duy nhất M=AC theo lý do trong câu a).

e) Tính $Z = (X^+ Y)^+ \cap K^+$ biết X = AE, Y = BE, K là một khoá của p.

Nhận xét: Vì K là khoá của p nên K^+ = U. Mặt khác, theo công thức $(X^+Y)^+ = (XY)^+$ nên ta có

$Z = (XY)^+ \cap U = (XY)^+ = (ABE)^+ = ABE$.

f) Có thể thêm vào F một PTH để thu được lược đồ có đúng 2 khoá không. Giải thích cách làm?

Thêm, chẳng hạn PTH $D \to C$ Khi đó ngoài khoá AC lược đồ có thêm khoá AD, vì $D \to C$. Dễ thấy có thể tách LĐQH p thành 2 LĐQH thành phần độc lập nhau là p_1=(ABE, {EA→B, A→BE}) với khoá K_1= A và p_2 =(CD, {C→D}) với khoá K_2 = C. Khi thêm PTH $D \to C$ cho LĐQH p_2 ta có thêm khoá K_3 = D. Hợp nhất 2 lược đồ này lại ta sẽ được đúng 2 khoá là

K_1K_2 = AC và K_1K_3 = AD.

BÀI GIẢI

ĐỀ THI TUYỂN CAO HỌC NGÀNH GIS NĂM 2003
ĐẠI HỌC BÁCH KHOA TP HỒ CHÍ MINH
Môn: Cơ sở dữ liệu, đề số 3

Thời gian 120 phút, không tham khảo tài liệu

Trong đề thi, tập các thuộc tính được viết dưới dạng xâu ký tự, ký hiệu XY được hiểu là phép hợp 2 tập X và Y, XY = X∪Y. Các từ viết tắt: LĐQH: lược đồ quan hệ, PTH: phụ thuộc hàm.

Câu 1.

a) Định nghĩa phép toán lấy giao hai quan hệ:

Phép giao (theo lý thuyết tập hợp hoặc *lấy phần chung*) hai quan hệ tương thích *R*(*U*) và *S*(*U*), ký hiệu *R*&*S*, cho ta quan hệ chứa các bộ xuất hiện đồng thời trong cả hai quan hệ thành phần,

$$P(U) = R\&S = \{\, t \mid t \in R,\ t \in S \,\}$$

b) Thuật toán giao hai quan hệ:


```
Algorithm Intersection
Format: P = R&S
Input:          - Quan hệ R(U)
                - Quan hệ S(U)
Output: - Quan hệ R&S={t|t∈R, t∈S}
```

```
Method
    Create(P,Attr(R));
      for each tuple u in R
            with u in S do
                  add u to P;
      endfor;
      return P;
end Intersection;
```

Câu 2.

a) Định nghĩa bao đóng của tập thuộc tính trong LĐQH: Cho LĐQH p = (U,F), trong đó U là tập thuộc tính, F là tập PTH trên U. Bao đóng của tập thuộc tính X, ký hiệu X^+ là tập thuộc tính

$$X^+ = \{ A \in U \mid X \to A \in F^+ \}$$

c) Thuật toán tìm bao đóng của tập thuộc tính trong LĐQH:

Ý tưởng: *Cho tập PTH F trên tập thuộc tính U và một tập con các thuộc tính X trong U. Để xác định bao đóng của tập thuộc tính X, X^+ ta xuất phát từ tập X và bổ sung dần cho X các thuộc tính thuộc vế phải R của các PTH $L \to R \in F$ thỏa điều kiện $L \subseteq X$. Thuật toán sẽ dừng khi không thể bổ sung thêm thuộc tính nào cho X.*

```
Algorithm Closure
Format: C = X+
Input:        - LĐQH p =(U,F)
              - Tập thuộc tính X ⊆ U
```

```
Output: - Y = X+ = {A ∈ U | X→ A ∈ F+}
Method
        Y:=X;
      repeat
            Z:=Y;
            for each FD L→R in F do
                  if L ⊆ Y then
                        Y:=Y ∪ R;
                  endif;
            endfor;
        until Y=Z;
        return Y;
end Closure;
```

Câu 3.

Giả thiết:

LĐQH p = (U, F); U = ABCDE;

$F = \{ AB \to C,\ D \to E,\ B \to CA \}$

a) Tìm một khoá của lược

Ta tìm tập *M* là giao của to ộ các khoá của *p:*

$$M = U - \bigcup_{L \to R \in F} (R - L)$$

M = ABCDE - CAE = BD, vì (=U nên p có một khoá

M = BD.

168

01.07.20

xuất

XB do

ng và nộp

b) Tập DEB có phải là khoá của p không? Vì sao?

Không, vì DEB chứa thực sự khoá M = BD.

c) Tập CA có phải là khoá của p không? Vì sao?

Không vì p chỉ có một khoá duy nhất M = BD.

d) Lược đồ p còn khoá nào nữa không? Vì sao?

Không, với lý do trên.

e) Tính $Z = (X^+Y)^+ \cap K^+$ biết $X = AB$, $Y = CA$, K là một siêu khoá của p.

Nhận xét: Vì K là khoá của p nên $K^+ = U$. Mặt khác, theo công thức $(X^+Y)^+ = (XY)^+$ nên ta có

$$Z = (XY)^+ \cap U = (XY)^+ = (ABC)^+ = ABC.$$

f) Có thể thêm vào F một PTH để thu được lược đồ có đúng 2 khoá không. Giải thích cách làm?

Thêm, chẳng hạn PTH $E \rightarrow D$ Khi đó ngoài khoá BD lược đồ có thêm khoá BE, vì $E \rightarrow D$. Dễ thấy có thể tách LĐQH p thành 2 LĐQH thành phần độc lập nhau là $p_1=(ABC,\{AB \rightarrow C, B \rightarrow CA\})$ với khoá $K_1 = B$ và $p_2 = (DE, \{D \rightarrow E\})$ với khoá $K_2 = D$. Khi thêm PTH $E \rightarrow D$ cho LĐQH p_2 ta có thêm khoá $K_3 = E$. Hợp nhất 2 lược đồ này lại ta sẽ được đúng 2 khoá là

$K_1K_2 = BD$ và $K_1K_3 = BE$.

01.07.2003

BÀI TẬP CƠ SỞ DỮ LIỆU

Chịu trách nhiệm xuất bản:
CÁT VĂN THÀNH

In 1000 cuốn khổ 14,5 × 20,5 cm tại Xưởng in Nhà xuất bản Thống kê. Giấy phép xuất bản số: 155-205/XB-QLXB do Cục xuất bản cấp ngày 03 tháng 03 năm 2003. In xong và nộp lưu chiểu tháng 10 năm 2003.