Trường Đại học Công nghệ thông tin, ĐHQG-HCM Ôn thi cao học CNTT năm 2009



Phụ thuộc hàm và các dạng chuẩn (Functional Dependency and Normal Forms)

> Giảng viên: PGS.TS. Đỗ Phúc Khoa Hệ thống thông tin



Nội dung

- Phu thuôc hàm
- Hê tiên đề Armstrong
- Bao đóng của tập thuộc tính
- Bài toán thành viên
- Phủ tối tiểu
- Khóa và thuật toán tìm khóa
- Các dang chuẩn
- Chuẩn hóa lược đô quan hệ
- Các thuật toán kiểm tra phân rã nối không mất tin
- Thuật toán phân rã LĐQH đạt DC3

2



Phụ thuộc hàm Functional Dependencies (1)

- PTH (FDs) được dùng để đo mức độ hoàn thiện của thiết kế các quan hệ
- PTH và khóa được dùng để xác định dạng chuẩn của quan hệ
- PTH là các ràng buộc (constraints) được suy từ ý nghĩa và các liên hệ giũa các thuộc tính dữ liệu



Định nghĩa phụ thuộc hàm Functional Dependencies (2)

- X -> Y đúng nếu bất kỳ khi nào hai bô (tuple) có cùng giá trị X phải có cùng giá trị Y
- Với bất kỳ hai bộ t1 và t2 trong thể hiện quan hệ r(R): Nếu t1[X]=t2[X] thì t1[Y]=t2[Y]
- X -> Y trong R xác định ràng buộc trên tất cả thể hiện r(R)
- Ký hiệu X -> Y (đọc là "X xác định duy nhất Y")
- PTH được suy từ các ràng buộc trong thế giới thực trên các thuộc tính.

4



Ví dụ về phụ thuộc hàm (1)

- Mã nhân viên xác định tên nhân viên SSN -> ENAME
- Mã đề án xác định tên đề án và địa điểm PNUMBER -> {PNAME, PLOCATION}
- Mã nhân viên và mã đề án xác định giờ làm việc trong tuần của nhân viên cho đề án {SSN, PNUMBER} -> HOURS



Ví dụ về ràng buộc PTH (2)

- PTH là một tính chất của các thuộc tính trong lược đồ R
- Ràng buộc phải thỏa trên tất cả các thể hiện của quan hệ r(R)
- Nếu K là khóa của R thì K xác định phụ thuộc với tất cả các thuộc tính của R (vì chúng ta không bao giờ có hai bộ phân biệt mà t1[K]=t2[K])



Bài toán tìm tất cả PTH khả dĩ

Cho thể hiện của quan hệ, tìm tất cả các PTH khả dĩ:

R (A	В	C)
1	2	4
1	2	4
2	5	7

3 5

Các PTH: A->B, A->C, B->C,AB->C,.... Thuật toán?



Tính F+, bao đóng của tập các PTH F

- Khi thiết kế CSDL quan hệ,chúng ta bắt đầu bằng cách xem xét tập các PTH khả dĩ.
- Khảo sát được tất cả các PTH là điều quan trọng, do vậy làm thế nào để có tất cả PTH.
- Bao đóng (closure) của tâp PTH F là tâp tất cả PTH có thể suy diễn logic từ F. Ta ký hiệu bao đóng của F là F+
- Ta tính F+ bằng cách áp dùng hệ tiên đề Armstrong



Các luật suy diễn cho PTH (1)

Cho tập PTH F, ta có thể suy ra thêm các PTH khác

Hệ tiên đề Armstrong:

- IR1. (Luật phản xạ) Nếu Y ⊆ X, thì X -> Y Vd: ABC → BC
- IR2. (Luật tăng trưởng) nếu X -> Y thì XZ -> YZ Vd: nếu C → D thì ABC → ABD
- IR3. (Luât bắc cầu) Nếu X -> Y và Y -> Z thì X -> Z Vd: Nếu AB \rightarrow CD và CD \rightarrow EF thì AB \rightarrow EF
- Hệ tiên đề Armstrong là đúng và đủ



Các luật suy diễn khác của PTH (2)

- Có thể suy thêm các luật hữu ích khác từ hệ tiên đề
- (Luật hợp) Nếu X -> Y và X -> Z thì X -> YZ Vd: Nếu AB → CD và AB → EF thì AB → CDEF
- (Luât tách) Nếu X -> YZ thì X -> Y và X -> Z Vd: Nếu AB \rightarrow CDEF thì AB \rightarrow CD và AB \rightarrow EF $var AB \rightarrow C var AB \rightarrow D var AB \rightarrow E var AB \rightarrow F$
- (Luật bắc cầu giả) Nếu X -> Y và WY -> Z thì WX->Z Vd: Nếu AB \rightarrow EF và DEF \rightarrow G thì ABD \rightarrow G



Bài tập:

- Cho quan hệ R = (A, B, C, G, H, I) và tập PTH: $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$
- Dùng hệ tiên đề Armstrong và các luật mở rộng của nó để tính
- (Luật phản xạ) Nếu $Y \subseteq X$, thì $X \rightarrow Y$
- (Luật tăng trưởng) Nếu X -> Y thì XZ -> YZ
- (Luật bắc cầu) Nếu X -> Y và Y -> Z thì X -> Z (Luật hợp) Nếu X -> Y và X -> Z thì X -> YZ
- (Luật tách) Nếu X -> YZ thì X -> Y và X -> Z
- (Luật bắc cầu giả) Nếu X -> Y và WY -> Z thì WX -> Z

Ap dụng hệ tiên đề Armstrong

Bài tập: Hãy dùng hệ tiên đề Armstrong để chứng minh:

Nếu X -> Y và U -> V thì XU -> YV

- 1. X -> Y (gt)
- 2. XU-> YU, (tăng trưởng U vào (1))
- 3. U -> V (gt)
- 4. YU -> YV (tăng trưởng Y vào (3)
- 5. XU -> YV (bắc cầu (2) và (4))



Bao đóng của tập thuộc tính The Closure of Attribute Sets



Bao đóng của tập thuộc tính

- F+ có thể có kích thước lớn, tìm F+ mất nhiều công sức
- Nếu chúng ta có thể xác định duy nhất tất cả thuộc tính trong R bằng một tập con thuộc tính X thì X là siêu khóa của R
- Bao đóng của X trên F (được ký hiệu là X+) là tập con của tập thuộc tính được xác định duy nhất bởi X qua các PTH trong F

14



Thuật toán tính X+ (bao đóng của X trên F)

```
repeat
oldX+=X
// khởi tạo
repeat
oldX+=X+
for each FD Y \rightarrow Z in F do
// kiểm tra từng PTH
if Y is a subset of X+, then // nếu vế trái trong X+
X+=X+UZ
// thêm vế phải vào X+
until X+== oldX+
// lặp trong khi X+ thay đổi
```

15



Ví dụ tính bao đóng của tập thuộc tính

R=(A,B,C,D,E,H)

- $F \{ AB \rightarrow C, BC \rightarrow AD, D \rightarrow E, CE \rightarrow B \}$
- How to compute closure of {A,B}, i.e., {A,B}+?
- 1. Start with X={A,B}
- 2. Add C to X due to AB \rightarrow C; X={A,B,C}
- 3. Add A,D to X due to BC \rightarrow AD; X={A,B,C,D}
- 4. Add E to X due to D \rightarrow E; X={A,B,C,D,E}
- 5. No more attributes can be added to X
- 6. $\{A,B\}^+ = \{A,B,C,D,E\}$

16



Ví dụ tính bao đóng X+

Cho quan hệ R = (A, B, C, G, H, I) và tập PTH $F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\},$ Tính (AG)+?

Ta muốn kiểm tra AG là siêu khóa của R.

- Dùng thuật toán tính bao đóng (AG)+.
- Nếu (AG)+ chứa tất cả thuộc tính của R thì AG là siêu khóa của R.

17

Ví dụ tiếp theo

Cho R = (A, B, C, G, H, I) và F = $\{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$, tính(AG)+

```
X+ = X

repeat

oldX+ = X+

for each FD Y \rightarrow Z in F do

if Y is a subset of X+, then

X+ = X+ UZ

until X+ == oldX+
```



- Cho F = { AB -> E, BE -> I, E -> C, CI -> D }
 Chứng minh AB -> CD suy được từ F.
- Simple Proof:
- (AB)+ = ABEICD which includes CD.
- Proof using Armstrong's Axioms:
- 1. AB --> E, Given
- 2. E --> C, Given
- 3. AB --> C, Transitivity on (1) and (2)
- 4. AB --> BE, Augment (1) by B
- 5. BE --> I, Given

19



Tiếp theo

- 6. AB --> I, Transitivity on (4) and (5)
- 7. ABC --> CI, Augment (6) by C
- 8. AB --> ABC, Augment (3) by AB
- 9. AB --> CI, Transitivity on (7) and (8)
- 10. CI --> D, Given
- 11. AB --> D, Transitivity on (9) and (10)
- 12. ABC --> CD, Augment (11) by C
- 13. AB --> CD, Transitivity on (8) and (12)

20



Tìm khóa Find Keys



Định nghĩa khóa và các thuộc tính tham gia vào khóa (1)

- Siêu khóa (superkey) của một lược đồ quan hệ R = {A1, A2,, An} là tập thuộc tính S ⊆ R thỏa tính chất không có hai bộ t1 và t2 trong một trạng thái hợp lệ r của R mà t1[S] = t2[S]
- Khóa (key) K là siêu khóa với tính chất bổ sung là khi xóa thuộc tính nào khỏi K sẽ khiến K không còn là siêu khóa.

22



Định nghĩa khóa và các thuộc tính tham gia vào khóa (2)

- Nếu lược đồ quan hệ có nhiều hơn một khóa, mỗi khóa sẽ được gọi là khóa dự tuyển (candidate key). Một trong các khóa dự tuyển được lựa chọn làm khóa chính (primary key), các khóa còn lại làm khóa phụ (secondary keys).
- Thuộc tính khóa là thuộc tính nằm trong một khóa dự tuyển.
- Thuộc tính không khóa không phải là thuộc tính khóa



Khóa của quan hệ

Định nghĩa: Cho quan hệ r(R), tập $K \subset R$ được gọi là khóa của quan hệ r nếu:K+=R nếu bớt một phần tử khỏi K thì bao đóng của nó sẽ khác R. Như thế tập $K \subset R$ nếu K+=R và $(K \setminus A)+ \neq R$, $\forall A \subset R$.

Một quan hệ có thể có nhiều khóa.



Ví dụ về khóa

Cho R = { A, B, C, D, E, G } và
F= {AB->C, D->EG , BE -> C , BC -> D ,
CG ->BD, ACD ->B, CE -> AG}
Ta sẽ thấy các tập thuộc tính:
K1 = { A, B} , K2 = {B,E} , K3={C,G} ,
K4={C,E} , K5 = {C,D}, K6={B,C} đều
là khóa

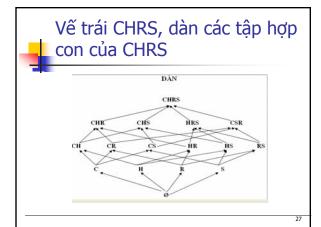


Tìm khóa (1)

- Gọi LHS là tập các thuộc tính nằm ở vế trái của các PTH. Đây là tập các ứng viên của siêu khóa.
- Dùng bao đóng của tập thuộc tính và tính chất của siêu khóa để kiểm tra ứng viên có có phải là siêu khóa hay không? Nếu không phải thì bỏ qua ứng viên đó.

26







Tìm khóa (1)

- Nếu ứng viên là siêu khóa thì tiếp tục kiểm tra các tập con của ứng viên này.
- Lặp lại cho đến khi gặp siêu khóa nhỏ nhất
- Siêu khóa nhỏ nhất là khóa dự tuyển (candidate keys)

28



Ví dụ tìm khóa (1)

 Cho LĐQH r(R) với R = {A, B, C, D, E, G, H,I} và tập PTH F F={ AB -> CDEFGH,

C -> BEI,

G -> H }

Tìm tất cả các khóa dự tuyển của LĐQH trên và chỉ định khóa chính?



Ví dụ tìm khóa (2)

- Bước 1: Tìm siêu khóa
 - Tập LHS: ABCG
- Bước 2: Kiểm tra các tập con để tìm siêu khóa ; tập con nhỏ nhất để tìm khó dư tuyển
 - Kiểm tra ABCABC+ = ABCDEGHI = R



Ví dụ tìm khóa (3)

- Kiểm tra ACG
- ACG+= ABCDEGHI =R
- Kiểm tra BCG

BCG+= BCGEHI ≠ R ? KHÔNG

Can ABG

ABG+ = ABCDEGHI = R

VẬy ABC, ABG, ACG là các siêu khóa

21



Ví dụ tìm khóa (4)

Lặp lại bước 2 cho ABC:

- Xét AB
- AB+ = ABCDEGHI = R
- Xét AC

AC+ = ABCDEGHI = R

• Xét BC determine A? No

 $BC+ = BCEI \neq R ? KHÔNG$

Vây AB và AC là siêu khóa

32



Ví dụ tìm khóa (5)

- Lặp lại bước 2 cho ACG:
 - Xét AC

AC+ = ABCDEGHI = R

- Xét AG
- $AG+ = AGH \neq R$? **KHÔNG**
- Xét CG

 $CG+ = CGBEHI \neq R ? KHÔNG$

Vậy AC là siêu khóa

1

Ví dụ tìm khóa (6)

- Lặp lại bước 2 cho ABG:
 - Xét AB

AB+=ABCDEGHI=R

Xét AG

 $AG+ = AGH \neq R$? **KHÔNG**

Xét BG

BG+= BGH ≠ R ? KHÔNG

Vây AB là siêu khóa

34



Ví dụ tìm khóa (7)

- Tiếp tục lặp bước 2 cho AB:
 - Xét A

 $A+=A \neq R$

- Xét B
- $B+=B \neq R$
- Vậy AB là siêu khóa nhỏ nhất (minimal superkey)



Ví dụ tìm khóa (8)

- Tiếp tục lặp bước 2 cho AC:
 - Xét A

 $A+=A \neq R$

- Xét C
- $C+=CBEI \neq R$
- Vậy AC là siêu khóa nhỏ nhất
- Tóm lại ta có AB và AC là các khóa dự tuyển, chọn một trong hai làm khóa chính



Bài tập tìm khóa:

- cho R = { A,B,C,D,E,G,H,I} F= { AC \rightarrow B, BI \rightarrow ACD, ABC \rightarrow D , H \rightarrow I , ACE \rightarrow BCG , CG \rightarrow AE } Tìm khóa ? (1)
- Bước 1: Gán K = R = {A,B,C,D,E,G,H,I}
- Bước 2: Lần lượt loại các thuộc tính của K

37



Tìm khóa? (1)

- Loại phần tử A: ta có {B,C,D,E,G,H,I}+ = R vì pth CG → AE khiến A thuộc về {B,C,D,E,G,H,I}+ nên K = {B,C,D,E,G,H,I}.
- Loại phân tử B, ta có {C,D,E,G,H,I}+ = R vì pth CG → AE khiến A thuộc về {C,D,E,G,H,I}+ và pth AC → B nên K {C,D,E,G,H,I}.

38



Tìm khóa ? (2)

- Loại phần tử C, ta có {D,E,G,H,I}+ ≠ R nên K vẫn là {C, D,E,G,H,I}
- Loại phần tử D, ta có: {C, E,G,H,I}+ = R vì pth CG → AE khiến A thuộc về {C, E,G,H,I}+ và pth AC → B nên K ={C,E,G,H,I}.
- Loại phần tử E, ta có: {C, G,H,I}+ = R vì pth CG → AE, AC → B, ABC→ D nên K ={C,G,H,I}.



Tìm khóa ? (3)

- Loại phân tử G, ta có: {C, H,I}+ ≠ R nên K vẫn là {C, G,H,I}.
- Loại phần tử H, ta có: {C, G,I}+ ≠ R nên K vẫn là {C, G,H,I}.
- Loại phần tử I, ta có: {C,G,H}+ = R vì CG → AE, AC → B, ABC→ D nên K={C,G,H}.
- Vậy K={ C,G,H} là một khóa của r (R)

40



Bài tập 1

Cho LĐQH r(R) với R=ABCD và tập PTH F ={AB->C, C->D, D->A}

- Tìm các khóa dư tuyển ?
- Xét các tập con của ABCD
- Khóa AB,DB
- AB+=ABCD=R và DB=DBAC=R



Bài tập 2

- Cho LĐQH r(R) với R=ABCD và tập PTH F ={AB->C, B->D, D->B}
- Tìm các khóa dự tuyển ?
- Xét các tâp con của ABD
- Khóa AB,AD
- AB+=ABCD=R và AD=DBAC=R



- Cho tập PTH
- F={XY->W, Y->Z, WZ->P, WP->QR, Q->X}
- CM XY->P có thể suy được từ F
- (1) Ta có XY-> W (cho)
- (2) Ta có WZ->P (cho)
- (3)XYZ->P (bắc cầu giả (1) và (2)
- (4) Y->Z (cho)
- (5) XY->P (bắc cầu giả (4) và (3)
- Có thể dùng bao đóng XY+=XYWZPQR đọcm

.11



Tương đương giữa các tập PTH



Bài toán thành viên

Procedure Member

Vào: F và PTH X → Y

 $\mathbf{Ra}:$ Đúng nếu F \mid = X \rightarrow Y và Sai nếu nguợc lại

Cách thức:

Member(F, $X \rightarrow Y$)

Begin

If $Y \subset Closure(X,F)$ Then return (True)

Else Return (False)

End



Tương đương của tập các PTH

- Hai tập PTH F và G là tương đương nếu:
 - Mọi PTH của F đều có thể suy được từ G và
 - Mọi PTH của G đều có thể suy được từ F
 - Do vậy F và G là tương đương nếu F + = G +
- Định nghĩa: F phủ G nếu mọi PTH của G đều suy được từ F (G+ ⊆ F+)
- F và G là tương đương nếu F phủ G và G phủ
 F

46



Thuật toán DERIVES kiểm tra F |= G

- Vào: hai tập PTH F và G
- Ra: Đúng nếu F |= G và sai nếu ngược lại.
- Cách thức:
- DERIVES (F,G)Begin
- V:= true:
- For each X → Y ∈ G do
- $V := V \text{ AND member}(F, X \rightarrow Y)$
- Return (V);
- Dựa trên hàm DERIVES để xây dựng hàm tương đương của hai tập PTH.



Thuật toán EQUIVALENCE kiếm tra F tương đương G

- Vào: hai tấp PTH F và G
- Ra: Đúng nếu F tương đương G, sai nếu ngược lại
- Cách thức

EQUIVALENCE(F, G)

Beain

V := Derives (F,G) AND Derives(G,F) Return (V)

End.



Bài tập kiểm tra tương đương

- Kiểm tra 2 tập PTH tương đương?
- Giải thích lý do

F = { AB -> C, B -> C, A -> D } G = { A -> B, B -> C, C -> D }

Giải: Hai tập PTH F và G là không tương đương vì

- PTH A -> B ∈ G nhưng không thể suy được từ F (Bao đóng A+ trên F là AD và không chứa B).
- PTH C -> D ∈ G cũng không suy được từ F (Bao đóng C+ trên F là C and và cũng không chứa D)

49



Phủ của tập PTH, tập PTH tương đương

- Cho F và G là tập các PTH, ta nói F tương đương với G và ký hiệu F ≡ G nếu F+= G+.
- Nếu F và G là tương đương thì ta nói F phủ G hay G phủ F.

50



Tương đương giữa 2 tập PTH

 <u>Bài tập</u>: Cho quan hệ Q(ABCDE) với:
 F = {A → BC, A → D CD → E }
 Và
 G = { A → BCE A → ABD CD → E

51



Tương đương giữa 2 tập PTH

- Ta cần CM: $F \models G \Leftrightarrow F \models G$
- Xét PTH A → E ∈ G suy được từ F nhờ vào các luật dẫn.Trong F,ta có:
- {A → C; A → D} | {A → CD; CD | E} | A → E (bắc cầu)
- <u>Kết luận</u>: F ⊨ G
 - Ta nhận thấy $F \subseteq G$, do đó hiển nhiên $G \models F$
- Kết luân: F ≡ G

52



PTH đầy đủ

- Cho quan hệ r(U), F và X,Y ⊂ U,
- PTH f: X → Y∈ F là đầy đủ với X nếu không tồn tai X′ ⊂ X sao cho F|=X′→ Y.



Phủ tối tiểu của PTH Minimal Covers of FDs (1)

Tập PTH là tối tiểu nếu thỏa các điều kiện sau:

- (1) Mọi PTH của F đều có một thuộc tính duy nhất ở vế phải (RHS).
- (2) Không thể loại bất kỳ PTH nào khỏi F mà tập các PTH còn lại vẫn tương đương với F.
- (3) Không thể thay thế PTH X -> A của F với PTH Y -> A khi Y ⊂ X mà vẫn có tập PTH tương đương với F (tất cả các PTH đều là PTH đầy đủ)



Phủ tối tiểu của PTH

- Mọi tập PTH đều tương đương với một phủ tối tiểu
- Có thể tương đương với nhiều phủ tiểu
- Có thuật toán để tính phủ tối tiểu

55



56



Ví dụ tìm phủ tối thiểu

- Cho tập thuộc tính R = {PCHART} và tập phụ thuộc hàm F như sau:
- $F = \{ P \rightarrow CHART$ $CH \rightarrow PART$ $C \rightarrow T$ $A \rightarrow R$ }



Các bước của thuật toán (1)

- Bước 1: G = F.
- Bước 2: $G = \{ P \rightarrow C; P \rightarrow H; P \rightarrow A; P \rightarrow R; P \rightarrow T; CH \rightarrow P; CH \rightarrow A; CH \rightarrow R; CH \rightarrow T; C \rightarrow T; A \rightarrow R \} (11 PTH)$
- Buớc 3: Kiểm tra PTH đầy đủ. Xóa lần lượt các thuộc tính trong vế trái của các PTH mà vế trái có nhiều thuộc tính, ví dụ CH → P; CH → A; CH → R; CH → T (Vế trái có 2 thuộc tính).

58



Các bước của thuật toán (2)

- 3.a. (Xóa thuộc tính C từ CH → P), chứng tỏ có thể suy H → P từ G. (sai)
- 3.b. (Xóa thuộc tính H from CH → P) chứng tỏ có thể suy C→P có thể được suy ra từ G. (sai)
- 3.c.(Xóa thuộc tính C từ CH → A) chứng tỏ có thể suy H→A từ G. (sai)
- 3.d.(Xóa thuộc tính h từ CH → A) chứng tỏ có thể suy C→A từ G. (sai)

4

Các bước của thuật toán (3)

- 3.e.(Xóa thuộc tính C từ CH → R) chứng tỏ có thể suy H→R từ G. (sai)
- 3.f.(Xóa thuộc tính H từ CH → R) chứng tỏ có thể suy C→R từ G. (sai)
- 3.g(Xóa thuộc tính C từ CH → T) chứng tỏ có thể suy H→T từ G. (sai)
- 3.h. (Xóa thuộc tính H từ CH → T) chứng tỏ có thể suy C→T từ G. (đúng)



Các bước của thuật toán (4)

- Do vậy, vào cuối bước 3, ta có :
- G= { P \rightarrow C; P \rightarrow H; P \rightarrow A; P \rightarrow R; P \rightarrow T; CH \rightarrow P; CH \rightarrow A; CH \rightarrow R; C \rightarrow T; A \rightarrow R } (10 PTH)
- Loai CH → T

Các bước của thuật toán (5)

Bước 4: Kiểm tra có thể loại bỏ các PTH trong G . Mỗi lần loại bỏ sẽ phát sinh tập các PTH mới G'.
4.a. Có thể loại P→C ? (không)
4.b. Có thể loại P→ H? (không)
4.c. Có thể loại P→A ? (không)
4.d. Có thể loại P→R ? (được vi P→A và A→R)
4.e. Có thể loại P→ T? (được vì P→C và C→T)
4.f. Có thể loại CH → P? (không)
4.g. Có thể loại CH → A? (được, vì CH→ P và P→A)
4.h. Có thể loại CH→ R? (được vì CH→ P and P→R)



Các bước của thuật toán (6)

4.i. Có thể loại C→ T ? (không)
 4.j. Có thể loại A→ R ? (không)

Do vậy vào cuối bước 4, ta có: $G = \{P \rightarrow C; P \rightarrow H; P \rightarrow A;$ $CH \rightarrow P; C \rightarrow T; A \rightarrow R \}$ Kết quả: Phủ tối thiểu của F là : $G = \{P \rightarrow C; P \rightarrow H; P \rightarrow A;$ $CH \rightarrow P; C \rightarrow T; A \rightarrow R \}$

63



Bài tập tìm phủ tối tiểu

- Cho R = ABCDE và
- F = { A -> C, BD -> E, B -> D, B -> E, C -> AD }.
- 1) Tìm phủ tối thiểu của PTT(F) ? Đáp
- PTT(F)= { A-> C, B -> D, B-> E, C-> A, C-> D }
- 2) Tìm tất cả các khóa của F?
- Đáp: AB và BC

64

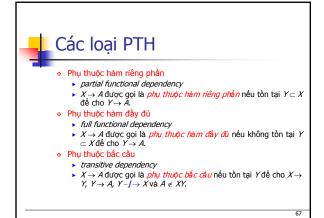


Các dạng chuẩn Normal Forms



Các dạng chuẩn

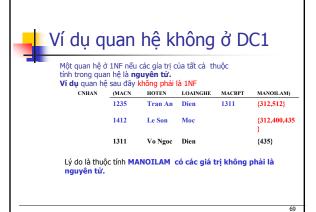
- Dạng chuẩn 1 (DC1): First Normal Form (1NF)
- Dạng chuẩn 2 (DC2): Second Normal Form (2NF)
- Dạng chuẩn 3 (DC3): Third Normal Form (3NF)
- Để chuẩn hóa 1NF-> 2NF-> 3NF





- · Định nghĩa
 - Quan hệ R ở dạng chuẩn 1 (1NF First Normal Form) nếu mọi thuộc tính của R đều chứa các giá trị nguyên tố (atomic value), giá trị này không là một danh sách các giá trị hoặc các giá trị phức hợp (composite value).
- \diamond Các thuộc tính của quan hệ R
 - ▶ Không là thuộc tính đa tri (*multivalued attribute*).
 - ▶ Không là thuộc tính phức hợp (composite attribute).

68





Dạng chuẩn 2

- · Định nghĩa
 - Quan hệ R ở dạng chuẩn 2 (2NF Second Normal Form) nếu R ở dạng chuẩn 1 và mọi thuộc tính không khóa đều phụ thuộc hàm đầy đủ vào mọi khóa của R.

70



Ví dụ về quan hệ ở DC2

Ví du: Cho LĐQH R = {A,B,C,D,E,G} và F = { A \rightarrow BC, C \rightarrow DE, E \rightarrow G }

- Ta thấy A là khóa vì A+ = R (tập thuộc tính của quan hệ).
- Các thuộc tính không khóa là {B,C,D,E,G}.
- Do khóa chỉ có một thuộc tính nên quan hệ R ở 2NF.



Dạng chuẩn 3 Third Normal Form (2)

- Lược đồ quan hệ R ở 3NF nếu nó ở 2NF và không có thuộc tính không khóa nào trong A trong R là phụ thuộc bắc cầu vào khóa
- Có thể phân rã quan hệ thành các quan hệ ở 3NF qua tiến trình chuẩn hóa 3NF



Dạng chuẩn 3(2)

- Lược đồ quan hệ R ở 3NF nếu nó ở 2NF và không có thuộc tính không khóa A nào trong R là phụ thuộc bắc cầu vào bất kỳ khóa nào của R
- Định nghĩa: Lược đồ quan hệ R ở 3NF nếu có PTH X -> A thỏa trên R thì:
 - X là siêu khóa của R, hay
 - A là thuộc tính khóa của R



Ví dụ về quan hệ không ở DC3

- Xét quan hệ CNHAN nhu sau: CNHAN(MACN LOAINGHE HESOTHUONG)
- Khóa của quan hệ là MACN
- Ta thấy có các pth trong quan hệ: MACN → LOAINGHE MACN → HESOTHUONG $\textbf{LOAINGHE} \rightarrow \textbf{HESOTHUONG}$
- Pth bắt cầu: MACN -> LOAINGHE và **LOAINGHE -> HESOTHUONG**
- Thuộc tính **không khóa** HESOTHUONG **phụ thuộc bắc cầu vào thuộc tính khóa** MACN, do đó quan hệ CNHAN **không phải là 3NF.**



Bài tập 1

- Xét LĐQH r(R) với R=ABCDE và tâp PTH $F = \{AB \rightarrow CE, E \rightarrow AB, C \rightarrow D\}$
- Dạng chuẩn cao nhất của quan hệ này là gì?
- 1) Tìm khóa: AB, E

AB+=ABCDE và E+=ABCDE



Bài tập 1 (Xét DC cao nhất)

- Thuộc tính không khóa {C,D}
 2) Xác định dạng chuẩn 2
- Các thuộc tính không khóa phải phụ thuộc đầy đủ vào khóa. Đúng nên ở DC2
- 3) Xác định dạng chuẩn 3
- Các thuộc tính không khóa phải không được phụ thuộc bắc cầu vào khóa
- Hoặc nếu có PTH X->A thì X phải là siêu khóa hoặc A là thuộc tính
- Xét PTH C->D ta có C không phải siêu khóa, D cũng không phải thuộc tính khóa

Ta có AB->C và C-> D, vậy D phụ thuộc bắc cầu vào khóa AB



Bài tập 2(Xét DC cao nhất)

- Cho LĐQH r(R) với R=ABCD
- F= {A->C, D->B, C->ABD}
- Xác định dang chuẩn cao nhất
- Xác định khóa: A, C
- A+=ACBD
- C+ACBD
- Do tất cả các PTH đều có vế trái chỉ chứa 1 thuốc tính nên DC2



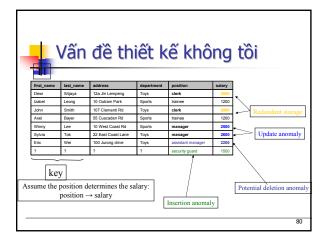
Bài tập 2(Xét DC cao nhất)

- Các thuôc tính khóa A,C
- Các thuộc tính không khóa B,D
- PTH D->B có vế trái D không phải là siêu khóa.
- Ta có C->D và D->B phụ thuộc bắc cầu: không ở DC3
- Dang chuẩn cao nhất là DC2



Thiết kế CSDL quan hệ

- CSDLQH yêu cầu tìm tập các lược đồ quan hê "tốt"
- Các PTH được dùng để lọc sơ đồ ER (phân rã quan hệ phổ quát)
- Thiết kế tồi có thể đưa đến sai sót







Chuẩn hóa

- Chuẩn hóa là tiến trình phân rã lược đồ quan hệ R thành các into **mảnh** (vd bảng nhỏ hơn) R₁, R₂,..., R_n sao cho:
 - Phân rã không mất tin (Lossless decomposition): Các mảnh chứa cùng thông tin như bản gốc. Ngược lại sẽ bị mất tin.

82



Chuẩn hóa

- Bảo toàn phụ thuộc (Dependency preservation): Cần bảo toàn các phụ thuộc trong từng quan hệ R_i
- Dạng tốt (Good form): Các mảnh R_i không nên bị dư (bảng bị dư nếu có PTH mà vế trái LHS không phải là khóa).

Phân rã kết không mất tin

Phân rã R không mất tin thành 2 quan hệ R1 và R2 nếu có thể khôi phục lại quan hệ R từ R1 và R2:

R= R1 ⋈ R2

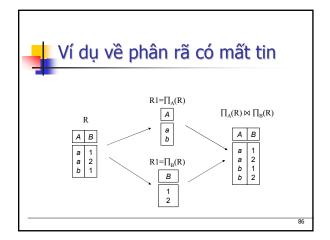
 SELECT first_name, last_name, address, department, T2.position, salary
 FROM T2, T3
 WHERE T2.position = T3.position



Phân rã kết không mất tin

- Phân rã R thành R₁ và R₂ là không mất tin (lossless) nếu và chỉ nếu ít nhất một trong các PTH sau có mặt trong F+:
 - ${\color{red} \bullet} \ R_1 \cap R_2 \to R_1$
 - $R_1 \cap R_2 \rightarrow R_2$
- Nói cách khác, tập thuộc tính chung của R₁ và R₂ phải là khóa dự tuyển cho R₁ hay R₂.

85





Phân rã bảo toàn phụ thuộc

- Phân rã lược đồ quan hệ R với tập PTH
 F thành tập các bảng (mảnh) R_i với PTH
 F_i
- F_i là tập các phụ thuộc trong F+ (bao đóng của F) chỉ bao hàm các thuộc tính có trong R_i.
- Phân rã bảo toàn phụ thuộc nếu và chỉ nếu (∪_i F_i)+ = F+

87

VÍ dụ về phân rã không bảo toàn phụ thuộc R=(A,B,C),F=(A)→(B), (B)→(C), (A)→(C)}. Key: A Co F(H,(B)→(C), với vẻ tráikhông phái khóa (có thể bị dư trong R.) Gran fách thánh 2 bảng R1(A,B), R2(A,C) (chuẩn hóa)

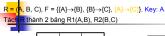
A	В	C
1	2	3
2	2	3
3	2	3
4	2	4

В	Α	C
2	1	3
2	2	3
2	3	3
2	4	4

Phân rã là không mất tin vì thuộc tính chung là A (khóa của R1 và R2) Phân rã là không bảo toàn phụ thuộc vì F1={{A} \rightarrow {B}}, F2={{A} \rightarrow {C}} và (F1 \cup F2)+ \pm F+ PTH {B} \rightarrow {C}}. Đã bị mất

88

Ví dụ về phân rã bảo toàn phụ thuộc



A	В	C
1	2	3
2	2	3
3	2	3
4	2	4



Phân rã là không mất tin vì thuộc tính chung là B (khóa của R2) Phân rã bảo toàn phụ thuộc vì F1={{A} \rightarrow {B}}, F2={{B}} \rightarrow {C}} và (F1 \cup F2)+=F+

•

Thuật toán kiểm tra phép phân rã không mất tin

- Vào: Lược đồ quan hệ R={A1, A2, . . . , An}, tập các pth F và phép tách
 - ρ(R1,R2, . . . , Rk)
- Ra: Kết luận phép tách ρ không mất tin.
- Các bước của thuật toán:
- Thiết lập một bảng với n cột(thuộc tính) và k dòng (quan hệ), cột thứ j ứng với thuộc tính A_j, dòng thứ i ứng với lược đồ R_i.
- Tại dòng i và cột j , ta điền ký hiệu a_j nếu thuộc tinh $A_i \in R_i$. Ngược lại ta điền ký hiệu b_{ii} .



Ví dụ: dùng thuật toán kiểm tra phép phân rã không mất tin

- Cho LĐQH r(R) với R={A,B,C,D}
- Và tập PTH F
- F={A->B, AC->D}
- Phép phân rã ρ(AB,ACD)
- Kiểm tra phép phân rã ρ không mất tin



Tạo bảng

	А	В	С	D
AB	a1	a2	b13	b14
ACD	a1	b22	a3	a4

PTH: A->B, AC->D

92



Làm bằng dùng PTH A->B

	Α	В	С	D
AB	a1	a2	b13	b14
ACD	a1	b22/a2	a3	a4

Dùng PTH A->B để làm bằng phần tử b22 thành a2

Vì bảng có dòng 2 chứa toàn ai nên phân rã trên là không mất tin



Kiểm tra phép phân rã mất tin

- Cho LĐQH r(R) với R={A,B,C,D,E}
- Và tập PTH F
- F={A->C, B->C,C->D,DE->C,CE->A}
- Phép phân rã ρ(AD,AB,BE,CDE)
- Kiểm tra phép phân rã ρ là mất tin

94



Phân rã bảo toàn phụ thuộc

- Một phân rã ρ=(R1,R2, . . . , Rk) của lược đồ quan hệ R trên tập PTH F bảo toàn phụ thuộc nếu có thể suy ra được F từ các hình chiếu của F trên Ri.
- Hình chiếu của F trên một tập các thuộc tính Z, ký hiệu là Π_Z(F) là tập các phụ thuộc X → Y thuộc F+ sao cho XY ⊂ Z (chú ý X → Y có thể không thuộc F và chỉ thuộc F+).
- Ta nói phân rã $\stackrel{.}{\rho}$ bảo toàn tập PTH F nếu hợp tất cả các PTH trong $\Pi_{Ri}(F)$ với i=1,2,...,k khẳng định logic tất cả PTH trong F.



Thuật toán kiểm tra bảo toàn phu thuộc

- Vào: Phân rã ρ=(R1,R2, . . ., Rk) và tập nth F
- Ra: Đúng nếu ρ bảo toàn phụ thuộc và sai nếu ngược lại
- Gọi G là hợp của tất cả Π_{Ri}(F) với i=1,...,k
- Dùng thuật toán EQUIVALENCE để xem xét G có tương đương với F hay không. Nếu có trả về đúng, ngược lại trả về sai.



Thuật toán phân rã lược đồ thành DC3 không mất tin, bảo toàn phu thuộc

- Bước 1: Loại khỏi R tất cả các thuộc tính không có mặt trong về trái và về phải của tất cả các PTH trong F
- Bước 2: Nếu có PTH X->Y mà XY=R thì kết quả chính là R
- Bước 3: Với từng PTH X->A của F, tạo lược đồ XA.
- Bước 4: Nếu có các PTH X->A₁, X->A₂,...,X->A_n thì tạo lược đô XA₁A₂...A_n thay cho tùng lược đô XA₁.
- Bước 5: Nếu tất cả các phần tử của khóa K không xuất hiện ở về trái trong bất kỳ quan hệ nào được tạo ở bước 4 thì tạo quan hệ mới có thuộc tính là khóa K

07



Ví du

- Cho LĐQH r (R) với R=XYZWQ và tập PTH F
- F={X->Y, XZ->W, YW->Q}
- Tìm một khóa dự tuyển: XZ
- XZ+=XZWYQ
- Bước 1:Không có thuộc tính nào thỏa nên không thực hiên bước 1.
- Bước 2:Không có PTH nào thỏa nên không thực hiện bước 2.
- Bước 3:Xét PTH X->Y, ta có quan hệ R(XY)

98



Ví dụ

- Bước 3:Xét PTH XZ->W, ta có quan hệ R(XZW)
- Bước 3:Xét PTH YW->Q, ta có quan hệ R(YWQ)
- Bước 4: Không có quan hệ nào có chung vế trải của các PTH nên không gộp
- Bước 5: Các thuộc tính khóa XZ xuất hiện ở vế trái của PTH XZ->W, do vậy KHÔNG tạo quan hệ mới
- Kết quả: R1(XY), R2(XZW) và R3(YWQ) đạt DC3 không mất tin và bảo toàn phụ thuộc.

Bài tập 1

- Cho LĐQH r(R) với R={A,B,C,D,E} và tập PTH F={A→BC}. Một phân rã R1(A,B,C) and R2(A,D,E)
- Phân rã này không mất tin?
 Đúng vì thuộc tính chung A là thuộc tính khóa của R1
- Phân rã này có bảo toàn phụ thuộc? Đúng vì PTH:A→BC được duy trì trong R1
- Phân rã R1(A,B,C) và R2(C,D,E) không mất tin?

Không vì C (thuộc tính chung) không phải là khóa của bất kỳ bảng nào.

100



Bài tập 2

- Cho LĐQH r(R) với R= {A, B, C, D, E} và tập PTH F = {A->BC, CD->E, B->D, E->A}. Xét phân rã R₁= (A, B, C) R₂= (A, D, E)
- Phân rã không mất tin?
 Đúng vì thuộc tính chung A là khóa của R1.
- Phân rã bào toàn phụ thuộc? Không: ta mất các PTH CD->E và B->D

7

Bài tập 3:

- Cho LĐQH r(R) với R={A, B, C, D} và tâp PTH F={AB->CD, B ->C}.
- R ở DC2?
 - Khóa: AB vì AB+ = ABCD=R
 - B →C vi phạm DC2 vì B là tập con của khóa và C không phải là thuộc tính khóa.



Cho LĐQH r(R) với $R=\{A, B, C, D\}$ and $F=\{AB->CD, C->D\}$.

R ở dạng chuẩn 2?

Khóa: AB ta có AB+= ABCD
Ta có C->D, C không phải là tập con của khóa vậy R ở dạng chuẩn 2

103



Bài tấp 5

- Cho LĐQH r(R) với R={A, B, C, D, E} và tập PTH
- F={A->B, BC->E, ED->A}

Tìm các khóa của R

ACD, BCD, CDE

Phân rã R đạt 3NF. R đã ở DC3 vì tất cả các thuộc tính đều là thuộc tính khóa.

104



Bài tập 6

- Cho quan hệ
- Sale (Customer, Store, Product, Price)
- Và ràng buộc:
- Khách mua hàng chỉ mua hàng trong 1 cửa hàng duy nhất. Trong một của hàng các sản phẩm có một đơn giá duy nhất ứng với từng sản phẩm của cửa hàng.
- Mô tả trên ứng với các PTH sau:
- {Customer} → {Store}
- $\{Store, Product\} \rightarrow \{Price\}$
- Khóa dư tuyển:
- {Customer, Product}

4

Bài tập 6 (tt)

- Quan hệ Sale có ở dạng chuẩn 3?
 Không vì cả 2 PTH đều vi phạm định nghĩa DC3.
- Phân rã Sale thành DC3
 R1(Customer,Store), R2(Store,Product,Price)
- Phân rã bảo toàn phụ thuộc?
 Đúng vì từng PTH đều nằm trong từng quan hê

106

Bài tập 7

Phân rã (<u>Customer,</u>Store), (<u>Store,Product</u>,Price) là mất tin vì thuộc tính chung Store không phải là khóa của bất kỳ bảng con nào

KEY: Customer, Product

Custome	Store	Produc	Price
E1	sl	5 1	prl
c1	sl	p2	pr2
c2	sl	pl	pr1

Customer	Store		Store	Product	Price
cl	sl		sl	p1	prl
c2	sl		sl	p2	pr2
		•			

Kết 2 bảng con (trên thuộc tính chung **store**) không khôi phục trở lại bảng gốc . Kết quả kết tạo ra 4 mấu tin trong khi bảng gốc chỉ có 3 mấu tin. Vấn đề phát sinh vì không có bảng con nào chứa khóa dự tuyển **Lời giải:** thêm bảng con (Customer, Product) trong phân rã. (thuật toán phân rã đạt ĐC3)

Customer	Product
c1	pl
c1	p2
c2	pl

Bài tâp 8

Cho LĐQH r(R) với R ={A, B, C, D, E, F, G}, F={AB ->CD, C->EF, G->A, G->F, CE->F}

- Phân rã thành các quan hệ ở DC3
- Tính phủ tối tiểu F_C={AB->CD, C->EF, G->AF}
- Các khóa dư tuyển : GB
- Phân rã đat DC3

 $R1={ABCD}$, $R2={CEF}$, $R3={GAF}$, $R4={GB}$



- Cho LĐQH r(R) với R = {A,B,C,D}.
- Phân rã R thành các quan hệ ở DC3 theo từng PTH sau: {B}→{C}, {D}→{A}
- Khóa dự tuyển: BD
- Phân rã: (B,C), (D,A), (B,D)

109



- Cho LĐQH r(R) với R = {A,B,C,D}. Phân rã R thành các quan hệ ở DC3 theo từng PTH sau{A,B,C}→{D}, {D}→{A} Khóa dự tuyển: ABC, BCD
- Phân rã R(A,B,C,D) đã ở DC3

 Quan hệ r (R) có dư thừa?
- Có vì vế trái của PTH {D}→{A} có thuộc tính D không phải là khóa

Phân rã R(A,B,C,D) theo cách bảo toànn phụ thuộc và không dư thừa?

Không có {A,B,C}→{D} chứa 4 thuộc tính. Nếu phân rã R sẽ mất PTH này.

111



Bài tập 10

xác định dạng chuẩn cao nhất cỉa các LĐQH sau:

- · R1(A, C, B, D, E), F1={A→B, C→D} Khóa là ACE
- DC1 (cả 2 PTH đều vi phạm DC2)
- R2(A, B, C, F), F2={AB→C, C→F}, Khóa là
 2NF (C→F vi phạm DC3)
- R3(A, B, C), F3={AB→C, C→B}
 Khóa là AB, AC, Tất cả thuộc tính của quan hệ đều là thuộc tính khóa, nên ở DC3.

111

Bài tập 11: Phân rã không mất tin (LJD)

 $cho LDQH r(R) với R = {A,B,C,D,E}.$

 $F = \{A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A \}.$ Các phân rã sau là không mất tin ?

- 1) $R1 = \{A,B,C\}, R2 = \{A,D,E\}.$
- 2) $R1 = \{A,B,C\},$
- $R2 = \{C, D, E\}.$
- 1) Vì R1 ∩ R2 = A và A là khóa của R1, phân rã là không mất tin.
- 2) Do R1 ∩ R2 = C và C không phải là khóa của R1 R2,nên phân ra này không phải là phân ra không mất tin.

112

Bài tập 12: Phân rã bảo toàn PTH



 $F = \{A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A\}.$

 $R1 = \{A,B,C\},$

 $R2 = \{A,D,E\}.$

Phân rã trên có bảo toàn phụ thuộc?

Không vì mất CD \rightarrow E và B \rightarrow D.

113

Bài tập 13

Cho LĐQH r(R) với R =(A, B, C, D). $T = \{C \rightarrow D, C \rightarrow A, B \rightarrow C\}.$

- 1) Xác định các khóa dư tuyển.
- 2) Xác định dang chuẩn cao nhất.
- 3) Phân rã R thành các quan hệ ở DC3.

Bài tập 13 (1)

$$R = (A, B, C, D).$$

 $- \{C \rightarrow D, C \rightarrow A, B \rightarrow C\}.$

1) Xác định các khóa dự tuyển.

$$B^+ = B$$
 $(B\rightarrow B)$
 $= BC$ $(B\rightarrow C)$
 $= BCD$ $(C\rightarrow D)$
 $= ABCD$ $(C\rightarrow A)$
Khóa dự tuyển là B.

B là khóa dự tuyển DUY NHẤT

Bài tập 13 (2)

$$R = (A, B, C, D).$$

 $- \{C \rightarrow D, C \rightarrow A, B \rightarrow C\}.$

2) Xác định dạng chuẩn cao nhất.

Khóa là B

R không ở DC3 vì:

PTH C→D vi phạm,

C→D không hiển nhiên ({0} ⊄ {C}).

C không phải là siêu khóa.

D không phải là thành phàn của khóa bắt kỳ.

C→A gây ra vi phạm

B→C không gây vi phạm

116

Bài tập 13 (3)

$$R = (A, B, C, D).$$

$$T = \{C \rightarrow D, C \rightarrow A, B \rightarrow C\}.$$

3) Phân rã R thành các quan hệ ở DC3.

Phủ chính tắ (canonical cover) là $F_c = \{C \rightarrow DA, B \rightarrow C\}.$

Đối với từng PTH trong F_c tạo bảng $R_1 = \{C, D, A\}, R_2 = \{B, C\}.$

 $R_1 = \{C, D, A\}, R_2 = \{B, C\}.$

Bảng R₂ chứa khóa dự tuyển cho R -kết thúc.

1

Bài tập 14

R = (A, B, C, D) F = {AB \rightarrow C, AB \rightarrow D, C \rightarrow A, D \rightarrow B} Đúng. Xác định tất cả các khóa dự tuyển :

AB, BC, CD, AD

Kiểm tra tất cả PTH thỏa DC3

118



Bài tập 15

R = (A, B, C, D) $F = \{ABC \rightarrow D, D \rightarrow A\}$

- 1. Xác định tất cả khóa dự tuyển của R ABC, BCD
- Xác định dạng chuẩn cao nhất
 Dạng chuẩn 3 vì tất cả các thuộc tính đều là thuộc tính khóa.

4

Tài liêu tham khảo

- David Maier, Theory of Relational Database, Computer Science Press, 1983
- Đỗ Phúc, Nguyễn Đăng Tỵ: Cơ sở dữ liệu , NXB ĐHQG-HCM, 2004
- Lê Tiến Vượng: Nhập môn cơ sở dữ liệu quan hệ, NXb Thống kê, 2003
- Nguyễn Xuân Huy, Lê Hoài Bắc, Bài tập CSDL, NXB Thống Kê, 2003
- Nguyễn Văn Tâm, Nguyễn Hữu Hạnh, Cơ sở dữ liệu quan hệ: Lý thuyết và thực hành. NXB Thống kê, 2002