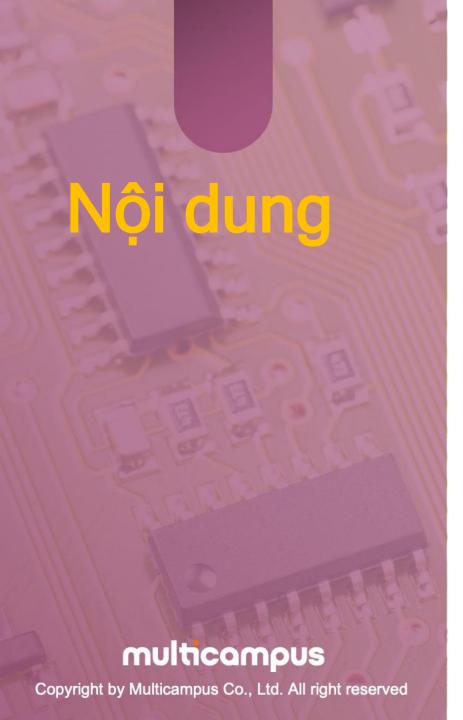


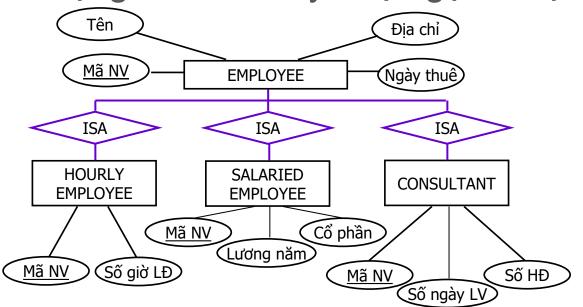
multicampus



- 1. Biết cách mô tả cấu trúc mô hình thực thể quan hệ mở rộng (EER) và biểu diễn chúng trong ER
- 2. Biết cách thể hiện nhóm thực thể có quan hệ đặc biệt
- 3. Cách lựa chọn một khóa chính tốt và các đặc điểm của khóa
- 4. Áp dụng các giải pháp linh hoạt cho các trường hợp mô hình hóa dữ liệu đặc biệt

Mô hình thực thể quan hệ mở rộng - EER

- Cách gọi khác là mô hình thực thể nâng cao (mở rộng)
- Thêm nhiều cấu trúc ngữ nghĩa hơn
 - Thực thể cha Entity Supertype
 - Thực thể con Entity Subtype
 - Cụm thực thể Entity Clustering
- Lược đồ mà sử dụng mô hình này được gọi là lược đồ EER EERD



Thực thể cha và Thực thể con

- Thực thể cha Entity Supertype:
 - Kiểu thực thể có chứa nhiều đặc điểm chung của các thực thể con

Hourly Employee

Hourly_Rate

(subtype)

- Chứa các đặc tính phổ biến
- VD: Nhân viên(Mã NV, Họ tên, Ngày sinh)
- Thực thể con Entity Subtype:
 - Chứa đặc tính duy nhất của mỗi thực thể con
 - VD:
 - Nhân viên theo giờ(Mã NV, số giờ lao động)
 - Nhân viên theo năm(Mã NV, Lương năm, Cổ đông)
 - Nhân viên tư vấn(Mã NV, Ngày làm việc, Số hợp đồng)

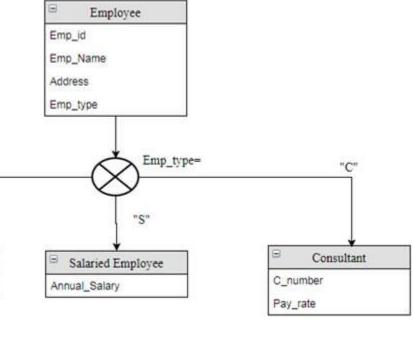


FIGURE 5.1 NULLS CREATED BY UNIQUE ATTRIBUTES

Database name: Ch05_AirCo

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_LICENSE	EMP_RATINGS	EMP_MED_TYPE	EMP_HIRE_DATE
100	Kolmycz	Xavier	T				15-Mar-88
101	Lewis	Marcos		ATP	SEL/MEL/Instr/CFII	1	25-Apr-89
102	Vandam	Jean					20-Dec-93
103	Jones	Victoria	R				28-Aug-03
104	Lange	Edith		ATP	SEL/MEL/Instr	1	20-Oct-97
105	Williams	Gabriel	U	COM	SEL/MEL/Instr/CFI	2	08-Nov-97
106	Duzak	Mario		COM	SEL/MEL/Instr	2	05-Jan-04
107	Diante	Venite	L				02-Jul-97
108	Wiesenbach	Joni					18-Nov-95
109	Travis	Brett	T	COM	SEL/MEL/SES/Instr/CFII	1	14-Apr-01
110	Genkazi	Stan					01-Dec-03

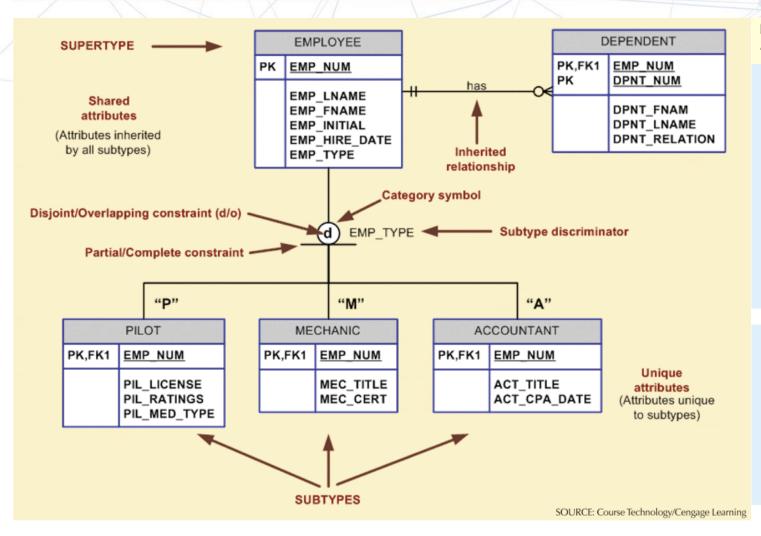
- Nhân viên có chung thuộc tính: EMP_NUM, EMP_LNAME, EMP FNAME, EMP HIRE DATE.
- Phi công có giới hạn giờ bay, kiểm tra chuyến bay và đào tạo định kỳ,...
- Nếu lưu tất cả nhân viên khác vào cùng bảng trên thì sẽ có nhiều giá trị trống vì các nhân viên khác không có thuộc tính trên

Sự phân cấp thực thể

- Mô tả sự sắp xếp của thực thể cha ở mức cao và thực thể con ở mức độ thấp
- Các mối quan hệ này được mô tả dưới dạng ISA
- Subtype chỉ tồn tại trong ngữ cảnh của Supertype
- Mỗi subtype có thuộc tính riêng biệt liên quan trực tiếp
- Có thể có nhiều cấp độ của mối quan hệ Supertype/Subtype

Thực thể phân cấp

multicampus



The EMPLOYEE-PILOT supertype-subtype relationship 5.3

Table name: EMPLOYEE

EMP_NUM	EMP_LNAME	EMP_FNAME	EMP_INITIAL	EMP_HIRE_DATE	EMP_TYPE
100	Kolmycz	Xavier	T	15-Mar-88	
101	Lewis	Marcos		25-Apr-89	P
102	Vandam	Jean		20-Dec-93	A
103	Jones	Victoria	R	28-Aug-03	
104	Lange	Edith		20-Oct-97	P
105	v∕villiams	Gabriel	U	08-Nov-97	P
106	Duzak	Mario		05-Jan-04	P
107	Diante	Venite	L	02-Jul-97	M
108	√lesenbach	Joni		18-Nov-95	M
109	Travis	Brett	T	14-Apr-01	P
110	Genkazi	Stan		01-Dec-03	A

Table name: PILOT

EMP_NUM	PIL_LICENSE	PIL_RATINGS	PIL_MED_TYPE
101	ATP	SEL/MEL/Instr/CFII	1
104	ATP	SELMEL/Instr	1
105	COM	SEL/MEL/Instr/CFI	2
106	COM	SELMELInstr	2
109	COM	SEL/MEL/SES/Instr/CFII	1

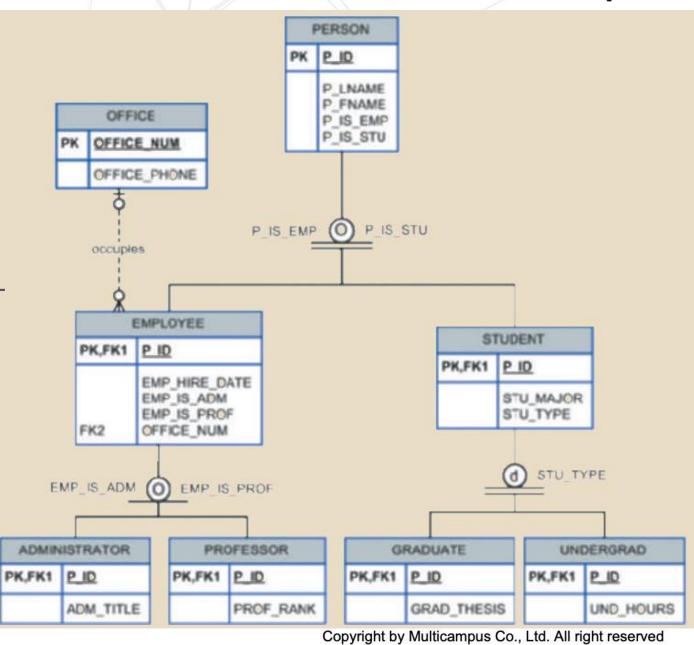
Thực thể thừa kế

- Cho phép thực thể con kế thừa thuộc tính và mối quan hệ của thực thể cha
- Tất cả thực thể con kế thừa thuộc tính khóa chính của thực thể cha
- Khi triển khai, thực thể cha và thực thể con có mối quan hệ 1:1
- Thực thể con kế thừa toàn bộ mối quan hệ mà thực thể cha tham gia
- Thực thể con ở mức thấp kế thừa toàn bộ thuộc tính và quan hệ của thực thể con ở mức cao hơn

Thực thể thừa kế

multicampus

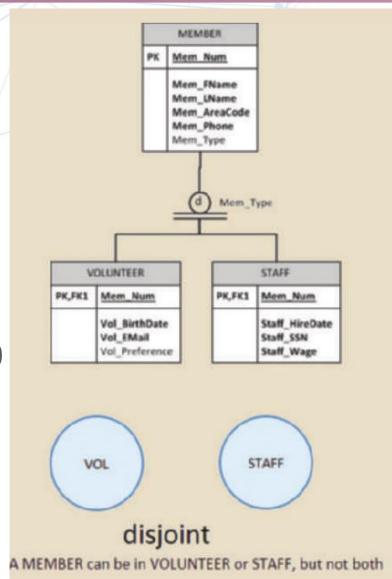
- EMPLOYEE, STUDET k
 é th
 ùa PERSON
 - P_IS_EMP, P_IS_STU để cho biết là đối tượng nào
 - Néu P_IS_EMP là "Y", P_IS_STU là "N"
 - Và ngược lại
- Tương tự, ADMINISTRATOR và PROFESSOR
- GRADUATE, UNDERGRAD kế thừa STUDENT
 - STU_TYPE xác định STUDENT là kiểu sinh viên nào

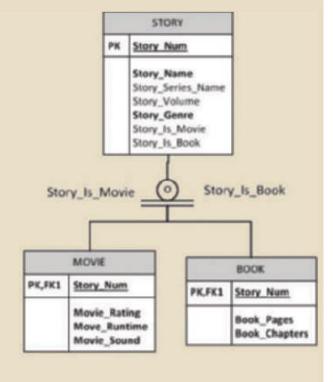


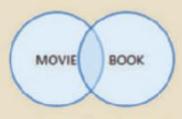
Ràng buộc rời rạc hay chồng chéo

multicampus

- MEMBER chỉ có thể là
 - VOLUNTEER hoặc STAFF
 - Không thể là cả 2 được
- STORY có thể là
 - MOVIE hoặc BOOK hoặc cả hai
- M1, Name1, V
- M2, Name2, S
 - Volum: M1, 1/1/1990
 - Stafff: M2, 1/6/2021, 1000
- S1,Sname1, Y,N
- S2,sname2, N, Y
- STORY: S3,sname3, Y,Y
 - MOVIE: s3, rating3
 - BOOK: s3, title, 30







overlapping

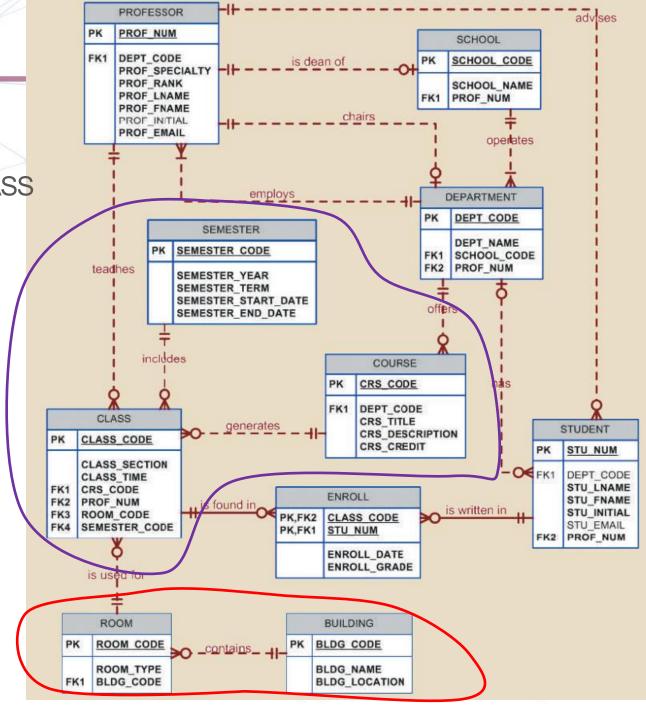
A STORY can be in MOVIE, BOOK, or both

Cụm thực thể - Entity Clustering

- Khi triển khai mô hình ER trên thực tế có thể có rất nhiều thực thể, không thể trình hết trên màn hình
 - Nên nhóm các thực thể và mối quan hệ của nó thành 1 cụm
- Cụm thực thể là thực thể không có thực ảo
- Nó đại diện cho nhiều thực thể và mối quan hệ, nhằm đơn giản mô hình ER
- Ví dụ:
 - Cụm OFFERING đại diện cho 3 thực thể: SEMESTER, COURSE, CLASS
 - Cụm LOCATION đại diện cho 2 thực thể: ROOM, BUILDING

Cụm thực thể

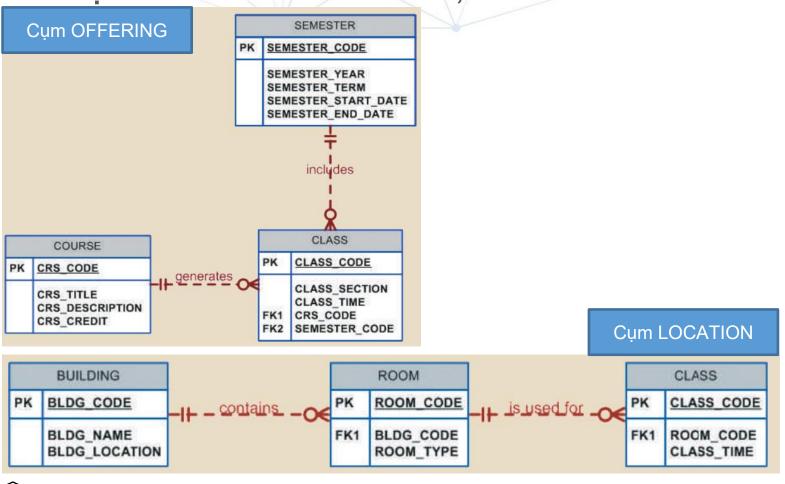
- Từ ERD nhóm:
 - Cum OFFERING: SEMESTER, COURSE, CLASS
 - Cum LOCATION: ROOM, BUILDING
- Kết quả như hình bên

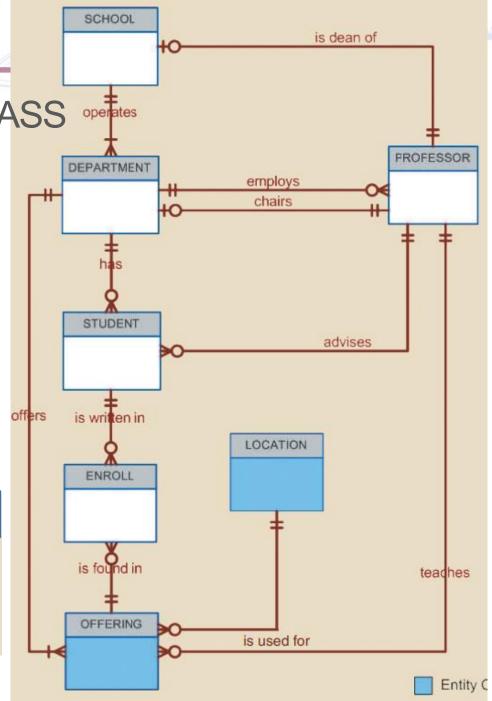


Cụm thực thể

Cum OFFERING: SEMESTER, COURSE, CLASS

Cum LOCATION: ROOM, BUILDING





Toàn vẹn thực thể: lựa chọn khóa chính

- · Khóa chính là tính chất quan trọng nhất của thực thể
 - Là thuộc tính đơn hoặc thuộc tính kết hợp
- Khóa chính và khóa ngoại kết hợp với nhau để triển khai các mối quan hệ
- Lựa chọn khóa chính ảnh hưởng trực tiếp đến hiệu quả thực thi cơ sở dữ liệu

Khóa tự nhiên và Khóa chính

- Là khái niệm về việc xác định duy nhất thường gặp trong thế giới thực
 - VD: Số hóa đơn, số thẻ tín dụng, số thẻ căn cước công dân,...
- Khóa tự nhiên hoặc mã định danh tự nhiên là mã trong thế giới thực để nhận dạng duy nhất 1 đối tượng
- Quen thuộc với người dùng cuối
- Nói chung mô hình dữ liệu thường lấy mã định danh tự nhiên làm khóa chính của thực thể
- Thay vào đó, có thể sử dụng khóa chính tổng hợp hoặc khóa thay thế

Hướng dẫn chọn khóa chính

- Khóa là 1 hoặc 1 tập thuộc tính xác định duy nhất 1 thực thể
- Một số câu hỏi cần quan tâm:
 - Có thể lấy 12 thuộc tính làm khóa được không?
 - Tại sao EMP_NUM được chọn là khóa chính mà không phải sự kết hợp của EMP_LNAME, EMP_ FNAME, EMP_INITIAL và EMP_DOB?
 - Một thuộc tính kiểu text kích thước 256 byte là khóa thì sao?
- Không có câu trả lời duy nhất cho các câu hỏi trên
- Cần quan tâm đến chức năng của khóa là gì?
 - Xác định duy nhất 1 thực thể hoặc hàng trong bảng
 - Đảm bảo tính toàn vẹn của thực thể, không phải để mô tả thực thể
 - Khóa chính và khóa ngoại dùng để kết nối giữa các thực thể
 - Ẩn với người dùng cuối

Đặc điểm PK	Giải thích
Giá trị duy nhất	Đảm bảo xác định duy nhất từng thực thể, PK là giá trị duy nhất, không chứa giá trị NULL
Không thông minh	PK không nên nên có ý nghĩa ngữ nghĩa, ngoài việc dùng để xác định thực thể hơn là dùng để mô tả. VD: 1 mã sinh viên: 650973 hơn là Smith, Martha
Không thay đổi theo thời gian	Nếu PK có ý nghĩa thì có thể phải cập nhật lại. VD lấy tên làm PK, vậy khi đổi tên thì phải đổi lại? PK thay đổi thì FK phải thay đổi theo. Hơn nữa việc thay đổi PK có nghĩa là cố tình thay đổi danh tính của thực thể. Vậy nên PK là vĩnh viễn không bao giờ thay đổi
Tốt nhất là một thuộc tính	Nên đặt PK là 1 thuộc tính nhưng không bắt buộc. PK là 1 thuộc tính sẽ đơn giản trong việc triển khai FK. Nếu PK có nhiều thuộc tính dẫn đến FK cũng nhiều thuộc tính và trở nên cồng kềnh vì FK tham chiếu đến PK
Tốt nhất là số	Giá trị duy nhất sẽ được quản lý tốt hơn nếu ở dạng số vì DBMS có thể triển khai dạng số tự động tăng
Tuân thủ bảo mật	PK không nên chọn bao gồm các thuộc tính có thể được coi là rủi ro bảo mật. VD dùng số an sinh xã hội làm khóa chính trong bảng EMPLOYEE

Khi nào nên chọn khóa tổng hợp

- PK là khóa tổng hợp sẽ hữu ích trong trường hợp sau:
 - Là số định danh trong thực thể kết hợp, được tách ra trong mối quan hệ N:M.
 - Là số định danh của thực thể yếu, trong đó thực thể yếu có mối quan hệ xác định chặt chẽ với thực thể cha
- Ví dụ:
 - STUDENT và CLASS có mối quan hệ N:M qua thực thể kết hợp ENROLL
 - Đảm bảo 1 sinh viên không thể đăng ký nhiều hơn 1 trong cùng 1 lớp
 - EMPLOYEE và DEPENDENT là mối quan hệ phụ thuộc tồn tại và
 PK của DEPENDENT là khóa tổng hợp chứa khóa của thực thể cha.

Khi nào nên chọn khóa tổng hợp

multicampus

FIGURE 5.6

The M:N relationship between STUDENT and CLASS

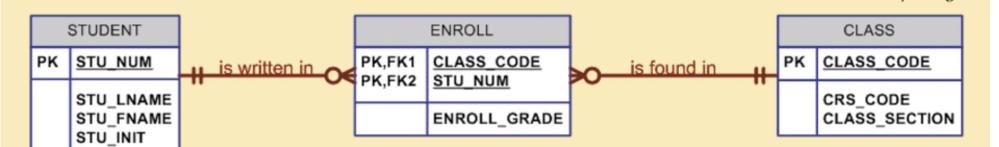


Table name: STUDENT (first four fields)

STU_NUM	STU_LNAME	STU_FNAME	STU_INIT
321452	Bowser	William	С
324257	Smithson	Anne	K
324258	Brewer	Juliette	
324269	Oblonski	Walter	Н
324273	Smith	John	D
324274	Katinga	Raphael	P
324291	Robertson	Gerald	T
324299	Smith	John	В

Table name: ENROLL

CLASS_CODE	STU_NUM	ENROLL_GRADE
10014	321452	С
10014	324257	В
10018	321452	A
10018	324257	В
10021	321452	С
10021	324257	С

Table name: CLASS (first three fields)

CLASS_CODE	CRS_CODE	CLASS_SECTION
10012	ACCT-211	1
10013	ACCT-211	2
10014	ACCT-211	3
10015	ACCT-212	1
10016	ACCT-212	2
10017	CIS-220	1
10018	CIS-220	2
10019	CIS-220	3
10020	CIS-420	1
10021	QM-261	1
10022	QM-261	2
10023	QM-362	1
10024	QM-362	2
10025	MATH-243	1
COLUDOR	o = 1	1 10 1

Database name: Ch05_Tinycollege

SOURCE: Course Technology/Cengage Learning

- Lý do chọn khóa thay thế:
 - PK không tồn tại trong thế giới thực hoặc khóa tự nhiên
 - Khóa có nội dung
 - Số CCCD: 031080xxxxxx; 001078xxxxxx
 - Khóa quá dài và cồng kềnh
- Trong trường hợp này sẽ tạo khóa thay thế
 - Là khóa được tạo bởi nhà thiết kế cơ sở dữ liệu để đơn giản việc xác định các thực thể
 - Sử dụng các ràng buộc duy nhất và không rỗng

Ví dụ cho bảng sau

DATE	TIME_START	TIME_END	ROOM	EVENT_NAME	PARTY_OF
6/17/2018	11:00 a.m.	2:00 p.m.	Allure	Burton Wedding	60
6/17/2018	11:00 a.m.	2:00 p.m.	Bonanza	Adams Office	12
6/17/2018	3:00 p.m.	5:30 p.m.	Allure	Smith Family	15
6/17/2018	3:30 p.m.	5:30 p.m.	Bonanza	Adams Office	12
6/18/2018	1:00 p.m.	3:00 p.m.	Bonanza	Boy Scouts	33
6/18/2018	11:00 a.m.	2:00 p.m.	Allure	March of Dimes	25
6/18/2018	11:00 a.m.	12:30 p.m.	Bonanza	Smith Family	12

Lược đổ:

- EVENT (DATE, TIME_START, TIME_END, ROOM, EVENT_NAME, PARTY_OF)
 - PK: (DATE, TIME_START, ROOM) or (DATE, TIME_END, ROOM)
- RESOURCE (RSC_ID, RSC_DESCRIPTION, RSC_TYPE, RSC_QTY, RSC_PRICE)
- EVENT và RESOURCE có mối quan hệ N:M nên tách thành 1:M
- EVNTRSC (DATE, TIME_START, ROOM, RSC_ID, QTY_USED)
- → Cồng kềnh nên lựa chọn khóa thay thế

Tình huống thiết kế: thiết kế linh hoạt

- Thiết kế và mô hình hóa đòi hỏi các kỹ năng có được qua kinh nghiệm thực tế
- Kinh nghiệm có được thông qua thực hành (tế) và lặp lại
- Các trường hợp thiết kế đặc biệt cần quan tâm:
 - Thiết kế linh hoạt
 - Xác định đúng khóa chính
 - Ví trí của khóa ngoại

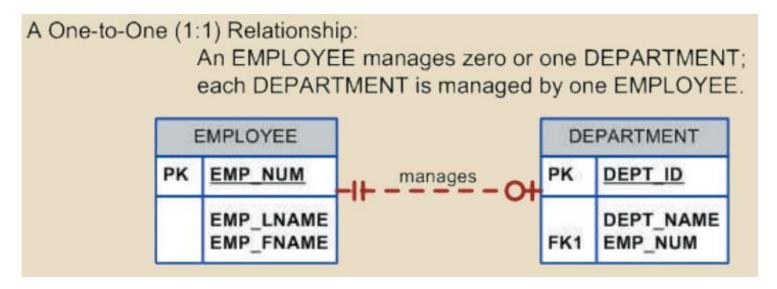
Tình huống 1: quan hệ 1:1

- FK làm việc với PK để thực hiện đúng mối quan hệ
 - Quy tắc: đặt PK của bên "một" (thực thể cha) ở bên "nhiều" (thực thể phụ thuộc) làm khóa ngoại
- Ví dụ, EMPLOYEE và DEPARTMENT là quan hệ 1:1
 - Có 2 trường hợp để chọn:
 - Đặt FK vào cả 2 thực thể: (không khuyến khích vì trùng lặp)
 - Đặt EMP NUM làm khóa ngoại cho DEPARTMENT
 - Đặt DEPT_ID làm khóa ngoại cho EMPLOYEE
 - Đặt FK vào 1 trong 2 thực thể
 - Nhưng PK nào nên đặt làm khóa ngoại ? Thể hiện trong bảng sau

TH	Ràng buộc quan hệ ER	Hành động
I	Một bên là bắt buộc và bên kia là lựa chọn	Đặt PK bên bắt buộc làm FK bên lựa chọn và đặt FK bắt buộc
II	Cả 2 bên là lựa chọn	Chọn FK có ít giá trị NULL nhất, hoặc đặt FK vào thực thể mà vai trò của mối quan hệ được thực thi
Ш	Cả 2 bên là bắt buộc	Xem xét như TH2 hoặc xem xét lại mô hình

Tình huống 1: quan hệ 1:1

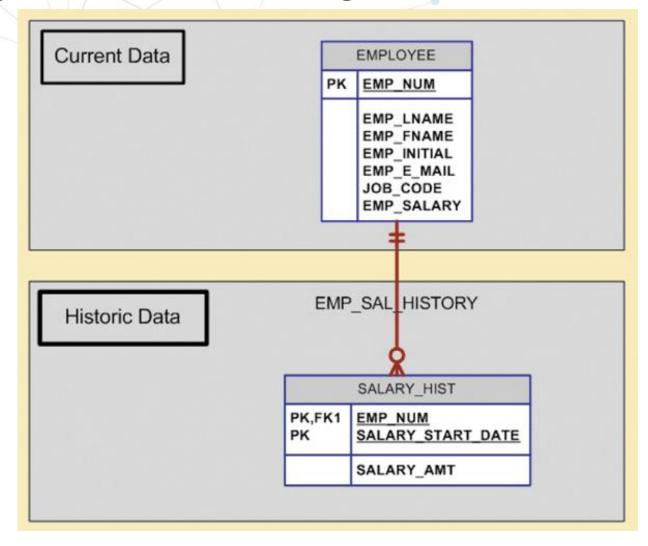
- Ví dụ:
 - EMPLOYEE quản lý DEPARTMENT và
 - EMPLOYEE là bắt buôc với DEPARTMENT
 - Vậy đặt EMP_NUM làm FK trong bảng DEPARTMENT



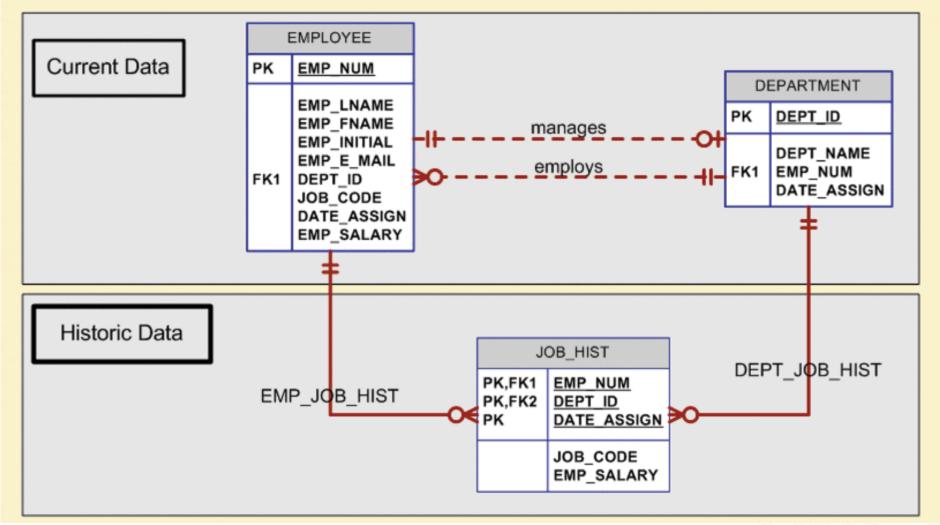
- Các câu hỏi thường đặt ra:
 - Làm thế nào để so sánh lợi nhuận của công ty năm nay với các năm trước?
 - Xu hướng bán hàng của sản phẩm XYZ là gì?
- Do vậy phải cần cả dữ liệu hiện tại và dữ liệu lịch sử
- Tùy vào từng dữ liệu mà có biến thể theo thời gian:
 - Ngày sinh, Số thẻ CCCD,... không biến đổi theo thời gian
 - Điểm TB, số dư tài khoản,...có biến đổi theo thời gian
- Thông thường, những giá trị thuộc tính hiện tại được thay thế bằng giá trị mới mà không liên quan đến giá trị trước
- Dữ liệu biến thể theo thời gian:
 - Giá trị thay đổi theo thời gian
 - Phải giữ lịch sử thay đổi dữ liệu

- Dữ liệu biến thể theo thời gian:
 - Giá trị thay đổi theo thời gian
 - Phải giữ lịch sử thay đổi dữ liệu
- Giữ lịch sử của dữ liệu theo thời gian đồng nghĩa với việc phải có thuộc tính đa trị trong thực thể
- Phải tạo ra một thực thể mới có mối quan hệ 1:M với thực thể ban đầu
- Thực thể mới chứa giá trị mới và ngày thay đổi
- Ví dụ, muốn theo dõi lịch sử tiền lương của nhân viên thì
 - Thuộc tính EMP_SALARY sẽ là thuộc tính đa trị.
 - Vì vậy mỗi nhân viên sẽ có 1 hoặc nhiều bản ghi trong SALARY_HIST(MaNV, Ngay, Tienluong)

Ví dụ, duy trì lịch sử tiền lương



Ví dụ, duy trì lịch sử công việc

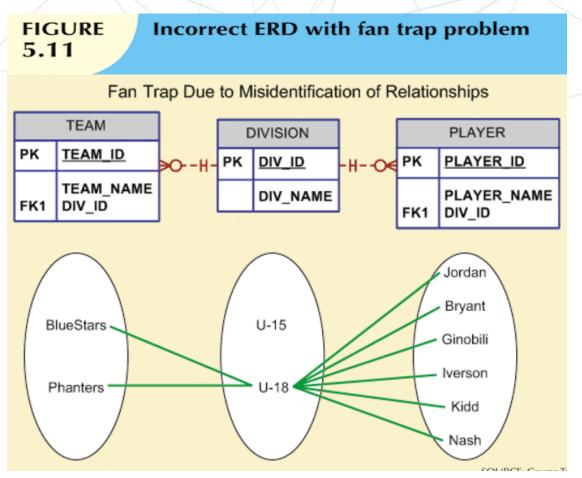


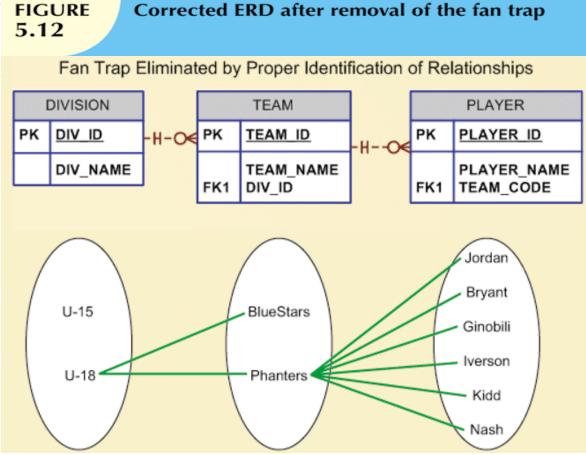
Tình huống 3: Bẫy thiết kế

- Bẫy thiết kế xảy ra khi mối quan hệ được xác định không đúng hoặc không đầy đủ
 - Được trình bày theo cách không phù hợp với thế giới thực
- Bẫy thiết kế phổ biến nhất được gọi là bẫy quạt (Fan Trap)
- Bẫy quạt xảy ra khi một thực thể có hai mối quan hệ 1: M với các thực thể khác
 - Tạo ra sự liên kết giữa các thực thể khác không được thể hiện trong mô hình
- Ví dụ, liên đoàn bóng đã JCB có nhiều bộ phận, mỗi bộ phận có nhiều người chơi, mỗi bộ phận có nhiều đội
 - Đúng ra là: mỗi bộ phận có nhiều đội, mỗi đội có nhiều người chơi

Tình huống 3: Bẫy thiết kế

multicampus



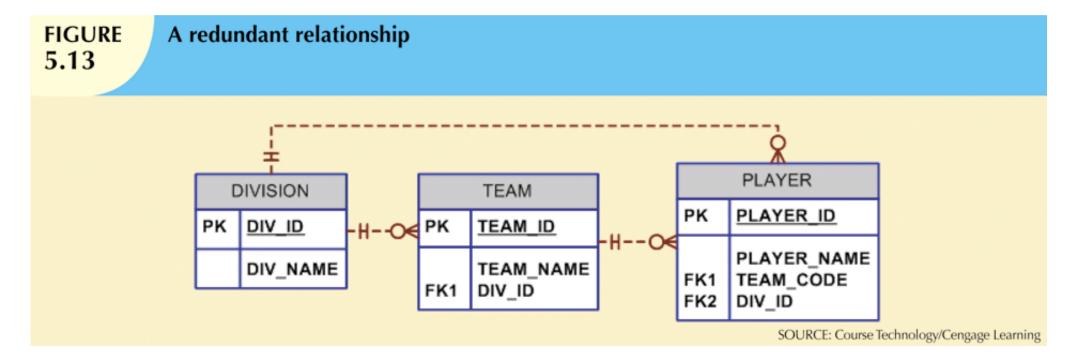


Tình huống 4: mối quan hệ dư thừa

- Dư thừa (Dự phòng) thường tốt trong môi trường máy tính
 - · Có nhiều bản sao ở nhiều nơi, khi có sự cố có thể khôi phục
- Nhưng dư thừa sẽ không tốt trong môi trường cơ sở dữ liệu
 - Dư thừa gây ra những bất thường về dữ liệu: chèn, sửa, xóa
- Xảy ra khi có nhiều đường dẫn quan hệ giữa các thực thể có liên quan
- Vấn đề chính là dư thừa mối quan hệ vẫn đảm bảo tính nhất quán trên mô hình
- Điều cần chú ý là thiết kế sử dụng mối quan hệ dư thừa như cách để đơn giản việc thiết kế
 - 1 bộ phận quản lý nhiều đội bóng, một đội bóng có nhiều người chơi
 - Lại có thêm 1 mối quan hệ giữa bộ phận quản lý nhiều người chơi

Tình huống 4: mối quan hệ dư thừa

- Ví dụ về dư thừa mối quan hệ
 - 1 bộ phận quản lý nhiều đội bóng, một đội bóng có nhiều người chơi
 - Lại có thêm 1 mối quan hệ giữa bộ phận quản lý nhiều người chơi



- Mô hình thực thể mở rộng (EER) bổ sung ngữ nghĩa cho mô hình ER
 - Thêm ngữ nghĩa qua thực thể cha, thực thể con và cụm thực thể
 - Thực thể cha là thực thể chung có quan hệ với 1 hoặc nhiều thực thể con
- Phân cấp thực thể:
 - Mô tả sự sắp xếp và mối quan hệ giữa thực thể cha và thực thể con
- Thực thể kế thừa:
 - Thực thể con kế thừa thuộc tính và mối quan hệ của thực thể cha

- Cụm thực thể là thực thể ảo:
 - Đại diện cho nhiều thực thể và mối quan hệ trong ERD
 - Kết hợp nhiều thực thể và mối quan hệ có liên quan thành một đối tượng thực thể
- Khóa tự nhiên là khóa tồn tại trong thế giới thực
 - Đôi khi sẽ là khóa chính tốt
- Đặc điểm của khóa chính:
 - Phải có giá trị duy nhất
 - Nên không thông minh (chứa các giá trị quan trọng)
 - Không thay đổi theo thời gian
 - Tốt nhất là số và là 1 thuộc tính đơn lẻ

- Khóa tổng hợp rất hữu ích trong biểu diễn:
 - Mối quan hệ N:M
 - Mối quan hệ thực thể yếu khỏe
- Khóa thay thế sẽ hữu ích khi không có khóa tự nhiên phù hợp làm khóa chính
- Trong mối quan hệ 1:1, hãy đặt PK của thực thể bắt buộc:
 - Như FK của thực thể lựa chọn
 - Là FK trong thực thể có ít giá trị NULL
 - Là nơi FK đóng vai trò

- Dữ liệu theo thời gian:
 - Dữ liệu có giá trị thay đổi theo thời gian
 - Yêu cầu giữ lịch sử thay đổi
- Để duy trì lịch sử của dữ liệu biến đổi theo thời gian:
 - Tạo thực thể mới chứa giá trị mới, ngày thay đổi, dữ liệu liên quan đến thời gian khác
 - Thực thể duy trì mối quan hệ 1:M với thực thể mà lịch sử được duy trì

- Bẫy lỗi thiết kế:
 - 1 thực thể có 2 mối quan hệ 1:M với 2 thực thể khác
 - Liên kết giữa các thực thể khác nhau không được thể hiện trong mô hình
- Mối quan hệ dư thừa xảy ra khi có nhiều đường dẫn mối quan hệ giữa các thực thể có liên quan
 - Vấn đề cần lưu ý là nó vẫn nhất quán trên toàn mô hình

Từ khóa

multicampus

completeness constraint

design trap

disjoint subtype

EER diagram (EERD)

entity cluster

entity subtype

entity supertype

extended entity relationship model (EERM)

fan trap

generalization

inheritance

natural key (natural identifier)

nonoverlapping subtype

overlapping subtype

partial completeness

specialization

specialization hierarchy

subtype discriminator

surrogate key

time-variant data

total completeness

Bài tập

- Học viên tạo tệp:
 - STT_HoVaTen_Chapter5_Advanced Data Modeling.docx
- Copy các câu hỏi trong tệp
 - Chapter5_Advanced Data Modeling.docx
 - Trả lời các câu hỏi và làm bài tập phía dưới câu hỏi
- Nộp bài lên hệ thống

THANKYOU

multicampus

Copyright by Multicampus Co., Ltd. All right reserved