**CHƯƠNG I : TỔNG QUAN VỀ HỆ CƠ SỞ DỮ LIỆU**

**1. Hệ thống quản lý dữ liệu bằng fil**e

* Đặc điểm: ( Lưu ý )
* Dữ liệu được lưu trữ trên các file riêng biệt
* Chương trình, định nghĩa và quản lý DL của chính nó

***Nhược điểm ( Lưu ý )***

* Dư thừa dữ liệu
* Sự mâu thuẫn dữ liệu
* Sự cô lập dữ liệu
* Ko có vấn đề bảo mật
* Ko điều khiển truy xuất cạnh tranh

***Tổ chức DL để khai thác hiệu quả là 1 sự cần thiết, trước đây người ta quản lý dữ liệu bằng file nhưng dở vì …… ( 5 nhược điểm ) ⇒ Học thuộc***

**2. Hệ quản trị cơ sở dữ liệu**

* ***Cơ sở dữ liệu*** *là một tập hợp dữ liệu tác nghiệp, có cấu trúc, lớn, được lưu trữ và mang tính chất dùng chung* ***( Học thuộc )***
* ***Hệ quản trị cơ sở dữ liệu*** *là một tập hợp các chương trình, cho phép người dùng:*
* *Tạo ra cơ sở dữ liệu*
* *Duy trì hoạt động của cơ sở dữ liệu*
* ***Điều Khiển tất cả truy xuất đến CSDL******( học thuộc )***
* **Hệ cơ sở dữ liệu** ( cơ sở dữ liệu ) là chương trình, **hệ quản trị cơ sở dữ liệu** là phần mềm
* **Hệ CSDL** nghĩa là gồm CSDL và Hệ quản trị CSDL (Oracle)

**3. Các kiểu mô hình dữ liệu**

3.1 Kiến trúc cơ sở dữ liệu

* Kiến trúc cơ sở dữ liệu có **3 mức/tầng /lớp**
* **Mức ngoài** ( mức người dùng ) <=====> Lấy dữ liệu ra

# Ở mức này còn có các **view** ( Khung nhìn ). View là khái niệm cho phép người dùng quan sát dữ liệu theo nhiều cách khác nhau trong khi dữ liệu được lưu trữ trong mức vật lý theo một cách duy nhất.

* **Mức khái niệm** ( mức logic → thiết kế ) <=====> Cách lưu trữ VD: Queue, mảng…
* **Mức trong** ( lưu trữ vật lý )<=====> Lưu trữ ntn
* **Mức trong (Internal level):**

Thiết kế cấu trúc bảng dữ liệu, bao gồm các trường, kiểu dữ liệu và các ràng buộc, ví dụ như ràng buộc khóa chính và ràng buộc duy nhất.

Lập kế hoạch để tối ưu hóa cơ sở dữ liệu, bao gồm việc sắp xếp và phân bổ dữ liệu để tối đa hóa hiệu suất truy xuất dữ liệu.

* **Mức khái niệm (Conceptual level):**

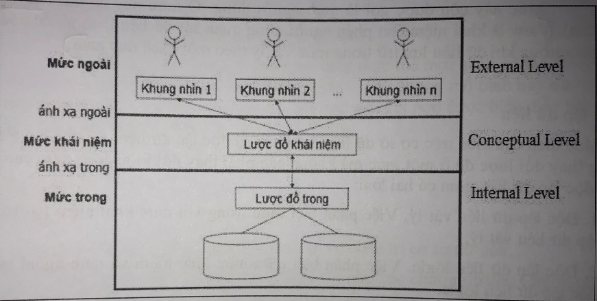
Thiết kế mô hình dữ liệu cho toàn bộ hệ thống cơ sở dữ liệu, bao gồm các thực thể, quan hệ và thuộc tính của chúng.

Xác định các ràng buộc và quy tắc kinh doanh để đảm bảo rằng dữ liệu được lưu trữ và quản lý đúng cách.

* **Mức ngoài (External level):**

Thiết kế giao diện người dùng để cho phép người dùng truy cập và thực hiện các thao tác trên cơ sở dữ liệu, bao gồm các form, báo cáo và truy vấn.

Xác định các cấu trúc và kiểu dữ liệu được truy xuất từ một ứng dụng hoặc hệ thống bên ngoài.



3.2 Độc lập dữ liệu

* Độc lập dữ liệu vật lý: phân biệt mức trong và mức khái niệm
* Độc lập dữ liệu logic: phân biệt mức khái niệm và mức ngoài

**3.3 Các kiểu mô hình dữ liệu**

* Gồm 3 nhóm:
* Mô hình dữ liệu dựa trên đối tượng
* Mô hình dữ liệu dựa trên mẫu tin
* Mô hình dữ liệu dựa trên vật lý
* Các mô hình thường được sử dụng là **(Học Kỹ)** :
* ***Mô hình thực thể mối quan hệ ERM (Rất quan trọng)***
* ***Mô hình dữ liệu quan hệ ( Mấu chốt ) (Học Kĩ)***

*#* Mô hình dữ liệu quan hệ mô tả dữ liệu bằng nhiều bảng (table), mỗi bảng là một tập hợp các bộ (type) hay còn gọi là mẫu tin ( record ). Mẫu tin thể hiện trên một dòng gồm nhiều giá trị được thể hiện trên nhiều cột hay còn gọi là thuộc tính

* Mô hình dữ liệu phân cấp
* Mô hình dữ liệu mạng

4. Hệ quản trị cơ sở dữ liệu

* **Hệ quản trị cơ sở dữ liệu** là một **tập hợp các chương trình**. Chúng cho phép người dùng tạo ra cơ sở dữ liệu, duy trì hoạt động của cơ sở dữ liệu và điều khiển tất cả truy xuất cạnh tranh đến cơ sở dữ liệu. **( Nhắc lại )**

**4.3 Ngôn ngữ cơ sở dữ liệu**

1. Ngôn ngữ định nghĩa dữ liệu DDL (lệnh create tạo bảng)
2. Ngôn ngữ thao tác dữ liệu DML ( nhớ 3 thao tác: Insert, Update, Delete
3. Ngôn ngữ truy vấn có cấu trúc SQL(lệnh select đồ đó)

# SQL: dùng để truy vấn dữ liệu cần thiết

**CHƯƠNG II: ĐẠI SỐ QUAN HỆ**

1. Cơ sở dữ liệu quan hệ

1.1 Tập hợp **thuộc tính** và **miền giá trị**

* Một giá trị được gọi là **nguyên tố** hoặc là giá trị đơn, khi không được phép phân chia ra nữa trong ngữ cảnh đang xét
* Tập hợp các thuộc tính ký hiệu là Ω = A1A2…..An
* Tập hợp các miền giá trị ký hiệu là Δ = D1D2….Dn

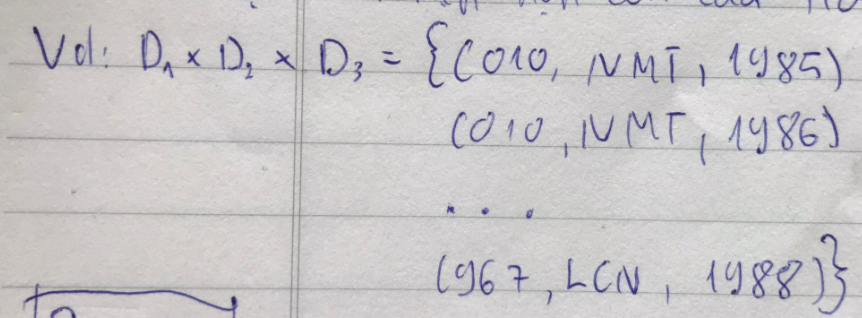
**#** Miền Di lấy giá trị tại thuộc tính Ai được ký hiệu là: **dom(Ai) = Di (i = 1 -> n)**

**vd:** D1 = dom(mssv) = (12214124,23423423,32525)

1.2 Quan hệ

* Quan hệ là một tập hợp con của tích Descartes D1 x D2 x ….Dn.
* Quan hệ là một tập hợp các bộ t có dạng:

=> R = { t = (d1,d2,....dn) / di ∈ Di , i = 1,2….,n }



1.3 Table (quan hệ có thể biểu diễn bằng bảng)

* **Cột (thuộc tính)**
* **Dòng (Bộ)(Thể hiện: thể hiện ra dòng đó có các thuộc tính gì)**
* **Cơ sở dữ liệu quan hệ là cơ sở dữ liệu mà trong đó dữ liệu được tổ chức dưới dạng bảng là hình ảnh trực quan của quan hệ** (**học thuộc**)

1.4 Bậc và lực lượng của 1 quan hệ

* Bặc: số cột
* LL của 1 quan hệ: số dòng

**1.5 Lược đồ quan hệ và thể hiện của quan hệ (Chú ý)**

* Nói đến lược đồ quan hệ là nói đến tất cả các quan hệ sinh ra từ đó mà không phải riêng cho một quan hệ nào
* Lược đồ dùng để mô tả tất cả các quan hệ
* **Thể hiện**: Thể hiện của lược đồ quan hệ là một quan hệ cụ thể được sinh ra từ lược đồ quan hệ đó m
* Một quan hệ cụ thể được thể hiện bởi các bộ/dòng
* **Lược đồ cơ sở dữ liệu**: là tập hợp tất cả các lược đồ quan hệ và các mối kết hợp có trong cơ sở dữ liệu đó
* **Thể hiện của cơ sở dữ liệu**: là nội dung thật sự của cơ sở dữ liệu tại một thời điểm nào đó
* Lược đồ quan hệ (Tổng quát)(Cột)(Cố định -> tượng đối)
* Lược đồ mô tả tổng quát các quan hệ bằng cách mô tả các cột
* Thể hiện (Cụ thể)(Dòng)(Thay đổi)
* Mang tính cụ thể (dòng/ bộ)

1.6 Khóa

* Khóa chính
* Khóa ngoại: không phải khóa của quan hệ đó nhưng là khóa của quan hệ khác. Thường được dùng để tạo sự thông thương dữ liệu

**CHƯƠNG III: PHỤ THUỘC HÀM VÀ KHÓA**

1.Ràng buộc toán vẹn

***Khắc sâu chữ ĐIỀU KIỆN***

1.1 Ví dụ ràng buộc toàn vẹn

* Ví dụ có 1 cơ sở dữ liệu quản lý sinh viên thì các ràng buộc toàn vẹn cũng như các điều kiện trên cơ sở đó như là : mỗi sinh viên có 1 mã sv riêng biệt, mỗi môn học có 1 mã riêng biệt

1.2 Tính hợp lý trên cơ sở dữ liệu

* Ràng buộc để đảm bảo tính hợp lý của các CSDL
* Trong 1 CSDL, luôn tồn tại nhiều ràng buộc: giữa các thuộc tính, giữa các bộ, giữa các quan hệ
* Các ràng buộc về cơ bản là các điều kiện được kiểm tra khi thực hiện các thao tác **Insert, Update, Delete**
* Mục đích xây dựng ràng buộc là đảm bảo tính độc lập và tính toàn vẹn

1.3 Biểu diễn ràng buộc toàn vẹn

* **Phụ thuộc hàm là công cụ biểu diễn ràng buộc toàn vẹn (Học Thuộc)**

1.4 Phạm vi của ràng buộc toàn vẹn (coi trong sách)

1.5 Phân loại ràng buộc toàn vẹn

* Phân loại ràng buộc theo **ý nghĩa** hoặc **phạm vi** (Nhiu đây là đủ )

1.6 Bảng tầm ảnh hưởng tổng hợp

* Khi thực hiện thao tác xen, sửa, xóa trên phạm vi của một ràng buộc toàn vẹn C thì phải kiểm tra thao tác đó có vi phạm C không

Bảng tầm ảnh hưởng của 1 ràng buộc toàn vẹn C:

| C | Thêm | Sửa | Xóa |
| --- | --- | --- | --- |
| R1 | +/-/ | +/-/\* | +/- |
| R2 | +/- | +/-/\* | +/- |
| …… | … | … | … |
| Rn | +/- | +/-/\* | +/- |

Dấu + cho biết cần phải kiểm tra C (*kiểm tra rồi mới làm*)

Dấu - cho biết không cần kiểm tra C (*cứ làm*)

Dấu \* cho biết không được thao tác đối với giá trị khóa (*cứ làm nhưng trừ khóa*)

2. PHỤ THUỘC HÀM

2.1 Phụ thuộc hàm

***Định nghĩa:(Học thuộc)***

1. ***X → Y: được gọi là với mỗi giá trị X xác định duy nhất 1 giá trị Y hay Y phụ thuộc X***
2. ***Với mỗi quan hệ R sinh ra từ lược đồ quan hệ s(Ω), nếu 2 bộ bất kỳ có cùng giá trị trên X thì chúng cũng có cùng giá trị trên Y***

***∀ R ∈ (Ω) ∀ t1,t2 ∈ R, nếu t1[X] = t2[X] thì t1[Y] = t2[Y]***

2.3 Phụ thuộc hàm được suy diễn từ tập hợp phụ thuộc hàm f

* Suy diễn: từ những lập luận đã cho suy diễn nên những điều kiện khác
* Định nghĩa:
* (Như ý 2 ở trên)
* Khi quan hệ R sinh ra từ lược đồ quan hệ s = (Ω,f) thỏa mọi phụ thuộc hàm trong f thì gọi là R thỏa f
* Về suy diễn :

X → Y ∉ f , X →Y được suy diễn từ f ⇔ ∀ R ∈ s(Ω,f)

nếu R thỏa f thì R thỏa X→Y

2.4 Hệ tiên đề Armstrong về các phụ thuộc hàm (Làm bài tập, xem thêm trong sách)

* **Hệ tiên đề Armstrong:**
* Tính chất phản xạ: Nếu **Y ⊆ X** thì **X —>Y**
* Tính chất tăng cường: nếu **X —> Y** thì **XZ —> YZ**
* Tính chất bắc cầu: Nếu **X —> Y** và **Y—>Z** thì **X—>Z**
* Định lý: Hệ tiên đề Armstrong {A1, A2, A3} là đúng đắn và đầy đủ
* Suy diễn bằng tiên đề Armstrong (Bài Tập)

3 Bao đóng(Làm bài tập)

* Định nghĩa: Bao đóng của f, ký hiệu f+, là tất cả phụ thuộc hàm được suy diễn từ f
* Suy diễn = hệ tiên đề Armstrong
* Được định nghĩa theo cách **đệ quy**
* Mệnh đề: Phụ thuộc hàm X→Y thuộc vào bao đóng f+ khi và chỉ khi X→Y được suy diễn từ f (theo hệ tiên đề Armstrong)

**(X) + ⊇ X**

3.2 Bao đóng của tập hợp các thuộc tính(Phần bài tập xem trong sách)

*\*Thuật toán tìm bao đóng của tập hợp các thuộc tính*

*\*Mệnh đề xác định 1 PTH có được suy diễn từ tập hợp PTH f hay không*

4 Phủ tối tiểu(Phần bài tập xem trong sách)

*\*Phụ thuộc hàm thừa*

* **Định nghĩa**: Phụ thuộc hàm X→Y ∊ f được gọi là phụ thuộc hàm thừa trong f khi và chỉ khi nó được suy diễn logic từ f \ X→Y. Tức là khi đó f tương đương với

f \ X→ Y, do đó f+ = (f \ x→Y)+

*\*Thuật toán tìm PTH thừa*

*\*Thuộc tính thừa ở vế trái của 1 PTH*

* **Định nghĩa:** Ai thừa ⇔ thay AiAj→R bằng Aj→R thì bao đóng cũng không thay đổi (*Nếu vế trái có 1 thuộc tính thì ta KO xét*)

*\*Thuật toán tìm thuộc tính thừa ở vế trái*

*\*Phủ tối thiểu của tập hợp PTH*

* **Định nghĩa:** Tập hợp phụ thuộc hàm f được gọi là tối thiểu nếu f thỏa mãn các điều kiện sau:
* Vế phải của mỗi PTH chứa duy nhất 1 thuộc tính, gọi là thuộc tính đơn
* Không tồn tại PTH thừa
* Không tồn tại PTH mà vế trái của nó có thuộc tính thừa

*\*Các bước tìm phủ tối thiểu*

5 Khóa trên lược đồ quan hệ

Trên lược đồ quan hệ luôn tồn tại một thuộc tính hoặc một tập hợp các thuộc tính mà giá trị của chúng cho phép **xác định duy nhất tất cả giá trị của các thuộc tính còn lại.**

5.1 Các định nghĩa

* **Khóa (Key):**
* Tập hợp con K ⊆ Ω được gọi là khóa của lược đồ quan hệ s khi và chỉ khi thỏa các điều kiện sau (2 điều kiện):

1. Mọi thuộc tính đều phụ thuộc hàm K, tức K → Ω hay (K)+ = Ω
2. Không tồn tại tập hợp con thực sự X ⊂ K mà X cũng xác định

Ω (X ↛ Ω) (↛ là dời cái dấu / dô giữa cái thân mũi tên)

* Các thuộc tính tham gia vào các khóa gọi là **thuộc tính khóa** , các thuộc tính còn lại gọi là **thuộc tính không khóa**
* **Siêu Khóa (Super Key)**
* Nếu K ⊆ Ω thỏa điều kiện đầu ở định nghĩa khóa bên trên mà không thỏa điều kiện 2 thì K được gọi là siêu khóa
* **Khóa chính và khóa ngoại**(Giống ở trên định nghĩa trước đó r)
* Khóa ngoại thường được dùng để *kết nối dữ liệu*

5.2 Mệnh đề

1. Đk 1 và 2 khẳng định các thuộc tính không khóa phụ thuộc hàm đầy đủ vào khóa.
2. ***K là khóa ⇔ (K)+ = Ω theo nghĩa tối tiểu*** *(Điều này rất tiện dụng cho việc một tập hợp các thuộc tính có phải là khóa hay không).*
3. Giá trị của các thành phần của khóa không thể là giá trị null.
4. Với 2 bộ bắc kỳ của quan hệ nếu chúng có giá trị giống nhau trên các thuộc tính khóa thì cũng có giá trị giống nhau trên các thuộc tính không khóa và khi chỉ giống nhau trên một số thuộc tính khóa thì không thể giống nhau trên các thuộc tính không khóa.

5.6 Thuật toán tìm tất cả các khóa (Coi trong sách làm bài tập)

**CHƯƠNG IV: CHUẨN HÓA LƯỢC ĐỒ QUAN HỆ**

1 Khái niệm về chuẩn hóa cơ sở dữ liệu

1.1 Mức độ chuẩn hóa

* **Ông EF.Codd** đưa ra khái niệm chuẩn hóa lần đầu tiên.
* **Ông R.Fagin** đưa ra khái niệm dạng chuẩn thứ 4.
* Có nhiều cách tách/phân hoạch

1.2 Quá trình chuẩn hóa

* Chuẩn hóa là việc ***PHÂN HOẠCH***
* Quá trình chuẩn hóa cơ sở dữ liệu là việc **tách / phân rã** một lược đồ quan hệ thành 1 nhóm các lược đồ quan hệ, gọi là các lược đồ quan hệ chiếu, sao cho **khi *KẾT NỐI TỰ NHIÊN* chúng không làm tổn thất thông tin và bảo toàn các phụ thuộc hàm**

2 Phép tách không tổn thất thông tin

2.1 Phép tách

* Xét lược đồ quan hệ s = (Ω,f), trong đó Ω =A1, A2, ….., An là tập hợp các thuộc tính, và f là tập hợp các phụ thuộc hàm
* Định nghĩa:
* Phép tách φ[Ω1,Ω2,....,Ωn] tách lược đồ quan hệ s = (Ω,f) thành các lược đồ quan hệ si = (Ωi , fi), i = 1,2,.......,p được xác định như sau:

1. Ωi ⊂ Ω, i = 1,2…..,p
2. Ω1 U Ω2 U ……. U Ωn = Ω
3. fi = {X → Y ∈ f+ /XY ⊆ Ωi }, i = 1,2,....,p

* Ký hiệu:
* Ký hiệu ***si = (Ωi , fi) = πΩi(s)*** được gọi là các **lược đồ chiếu** của

s = (Ω,f) trên tập các thuộc tính

* Ký hiệu ***fi = πΩi(f)*** được gọi là tập hợp các **phụ thuộc hàm** chiếu của f trên tập hợp thuộc tính Ωi
* Khi R là một quan hệ được sinh ra từ lược đồ quan hệ s = (Ω,f) thì ký hiệu ***Ri = πΩi(R)*** được gọi là các **quan hệ tham chiếu** của R trên Ωi

2.2 Phép tách tự nhiên

* Phép tách φ[Ω1,Ω2,....,Ωn] trên lược đồ quan hệ s = (Ω,f) được gọi là ***phép tách tự nhiên*** khi và chỉ khi kết quả của phép kết nối tự nhiên \* các lược đồ chiếu

s1 \* s2\* ... \* sp là một lược đồ quan hệ trên tập hợp các thuộc tính Ω

***R1 \* R2 \*... \* Rn ⊇ R***

2.3 phép tách không tổn thất thông tin

* Phép tách φ[Ω1,Ω2,....,Ωn] trên lược đồ quan hệ s = (Ω,f) được gọi là ***phép tách không tổn thất thông tin*** ⇔ với mỗi quan hệ R được sinh ra từ lược đồ quan hệ

s = (Ω,f): quan hệ kết quả của phép kết nối tự nhiên \* của các quan hệ chiếu Ri bằng chính R

***R1 \* R2 \*... \* Rn = R***

2.4 Thuật toán kiểm trả một phép tách có tổn thất thông tin hay không (coi trong sách đi)

2.5 Định lý

* Phép tách φ[Ω1,Ω2,....,Ωn] trên lược đồ quan hệ s = (Ω,f) là không tổn thất thông tin nếu:
* Hoặc Ω1 U Ω2 → Ω1\Ω2
* Hoặc Ω1 U Ω2 → Ω2\Ω1

3 Phép tách bảo toàn phụ thuộc hàm

3.1 Phụ thuộc hàm chiếu

* Phép tách φ[Ω1,Ω2,....,Ωn] trên lược đồ quan hệ s = (Ω,f), trong đó

Ω =A1, A2, ….., An là tập hợp các thuộc tính, f là tập hợp các phụ thuộc hàm

* Xét Φ ⊆ Ω là một tập hợp con các thuộc tính
* Tập hợp các phụ thuộc hàm chiếu của f trên Φ được ký hiệu và định nghĩa như sau  
   πΦ (f) = {L → R є f+ / LR ⊆ Φ}

4 Các dạng chuẩn 1NF, 2NF, 3NF, BCNF

4.1 Dạng chuẩn 1-1NF (First Normal Form)

Định nghĩa:

* Lược đồ quan hệ s = (Ω,f) được gọi là thỏa dạng chuẩn 1NF khi và chỉ khi mọi thuộc tính của nó đề chỉ chứa các **giá trị nguyên tố**
* Thông thường để đưa một lược đồ không là 1NF về 1NF người ta thường làm **phẳng và tách** thành các lược đồ chiếu

4.2 Dạng chuẩn 2-2NF (Second Normal Form)

* Định nghĩa:
* Lược đồ quan hệ s được gọi là thỏa dạng 2NF khi và chỉ khi là dạng chuẩn 1NF và mọi **thuộc tính không khóa phụ thuộc hàm đầy đủ vào khóa**

‘’ Y phụ thuộc hàm đầy đủ vào X khi và chỉ khi Y phụ thuộc hàm vào X và Y không phụ thuộc vào bất kỳ tập hợp con thực sự nào của X ”

* Để biết quan hệ R có thỏa dạng chuẩn 2NF hay không người ta phải xác định **tất cả khóa của R**
* Chỉ có các quan hệ có 2 thuộc tính trở lên thì mới có thể không là dạng chuẩn 2NF

(coi ví dụ trong sách)

4.3 Dạng chuẩn 3-3NF (Third Normal Form)

* Định nghĩa:
* Lược đồ quan hệ s thỏa dạng chuẩn 3NF khi và chỉ khi dạng chuẩn 2NF và mọi thuộc tính **không khóa không phụ thuộc hàm bắc cầu** vào khóa thông qua thuộc tính không khóa

Xem thuật toán trong sách

4.4 Dạng chuẩn BCNF(Boyce Codd Normal Form)

* Định nghĩa:
* Lược đồ quan hệ s được gọi là dạng chuẩn BCNF khi và chỉ khi với mỗi phụ thuộc hàm X → Y ∊ f+ thì:

1. Hoặc đó là phụ thuộc hàm tầm thường (tức là Y ⊆ X)
2. Hoặc là vế trái X chính là siêu khóa (tức là (X) = Ω)

Xem thêm ví dụ trong sách

**HẾT.**