



FPT POLYTECHNIC



Bài 5:

CHU N HÓA C S D LI U

- Giới thiệu ngôn ngữ SQL
- Giới thiệu Oracle SQL
- Tìm hiểu các nhóm ngôn ngữ của SQL

- Tìm hiểu các bộ phận khi cấp nhu cầu
- Giải thích số lượng thiết bị chuẩn hóa dữ liệu
- Tìm hiểu về khái niệm:
 - Khóa cửa quanh
 - Phân thu c hàm
- Tìm hiểu quy trình chuẩn hóa CSDL
 - Tìm hiểu 3 độ chuẩn: 1NF, 2NF, 3NF
 - Trình tự chuẩn hóa theo 3 độ chuẩn
- Tìm hiểu khái niệm khung chuẩn

- Ví dụ về CSDL là Bảng hóa đơn (INVOICE) cho khách hàng như sau:

Acme Industries INVOICE					
Customer Number: 1454837			Terms: Net 30		
Customer: W. Coyote			Ship Via: USPS		
General Delivery Falling Rocks, AZ 84211 (599) 555-9345			Order Date: 12/01/2008		
Product No.	Description	Quantity	Unit Price	Extended	Amount
SPR-2290	Super Strength Springs	2	24.00	\$48.00	
STR-67	Foot Straps, leather	2	2.50	\$5.00	
HLM-45	Deluxe Crash Helmet	1	67.88	\$67.88	
SFR-1	Rocket, solid fuel	1	128,200.40	\$128,200.40	
ELT-1	Emergency Location Transmitter	1	79.88	**FREE GIFT**	
TOTAL ORDER AMOUNT:			\$128,321.28		

■ Bút tháp khi thêm dữ liệu:

- Không thể thêm một khách hàng vào CSDL nếu khách hàng không mua một mặt hàng nào.

■ Bút tháp khi cập nhật dữ liệu:

- Muốn thay đổi cách của khách hàng, phải thay đổi cách trên tất cả các hóa đơn của khách hàng ở bởi vì cách của khách hàng luôn được thay đổi trong tất cả các hóa đơn của khách hàng.

■ Bút tháp khi xóa dữ liệu:

- Nếu xóa hóa đơn cuối cùng của một khách hàng thì tất cả dữ liệu về khách hàng đó.

- M c ích c a chu n hóa CSDL là nhóm các thu c tính vào các quan h nh m gi m thi u d th ad li u, lo i b các b t th ng khi c p nh t CSDL.
-> C n có các b c chu n hóa d li u t m t CSDL ch a chu n hóa sang chu n hóa.
- D ng ch a chu n hóa (unnormalized form - UNF): quan h ch a chu n hóa là quan h ch a các b d li u b l p l i giá tr .

- Bằng hóa đơn (INVOICE) trên bìa đính kèm sau.
- Bằng này không có số

INVOICE

Customer Number	Customer	Terms	Ship Via	Order Date	Product No.	Description	Quantity	Unit Price	Extended Amount
1454837	W.Coyote	Net 30	USPS	12/01/2008	SPR-2290	Super Strength Springs	2	24.00	\$48.00
1454837	W.Coyote	Net 30	USPS	12/01/2008	SRT-67	Foot Straps, leather	2	2.50	\$5.00
1454837	W.Coyote	Net 30	USPS	12/01/2008	HLM-45	Deluxe Crash Helmet	1	67.88	\$67.88
1454837	W.Coyote	Net 30	USPS	12/01/2008	SFR-1	Rocket, Solid fuel	1	128,200.4	\$128,200.4
1454837	W.Coyote	Net 30	USPS	12/01/2008	ELT-1	Emergency Location Transmitter	1	79.88	\$79.88

- Bảng này có chia thành hai bảng sau:

Customer Number	Customer	Terms	Ship Via	Order Date
1454837	W.Coyote	Net 30	USPS	12/01/2008

Customer Number	Product No.	Description	Quantity	Unit Price	Extended Amount
1454837	SPR-2290	Super Strength Springs	2	24.00	\$48.00
1454837	SRT-67	Foot Straps, leather	2	2.50	\$5.00
1454837	HLM-45	Deluxe Crash Helmet	1	67.88	\$67.88
1454837	SFR-1	Rocket, Solid fuel	1	128,200.4	\$128,200.4
1454837	ELT-1	Emergency Location Transmitter	1	79.88	\$79.88

- Khóa (Primary Key) là một thuộc tính (một nhóm) thuần túy có ý nghĩa quan hệ cho phép xác định mỗi biến khác nhau trong quan hệ.
- Ví dụ :
 - Trong quan hệ **SINHVIEN**: MaSV, HoTenSV, NgaySinh, DiaChi thì thuộc tính khóa là **MaSV**.

- Cho hai thuộc tính A và B
- B có giá trị là phương thức hàm vào A nếu có không quá một giá trị của thuộc tính B liên kết với một giá trị của thuộc tính A.

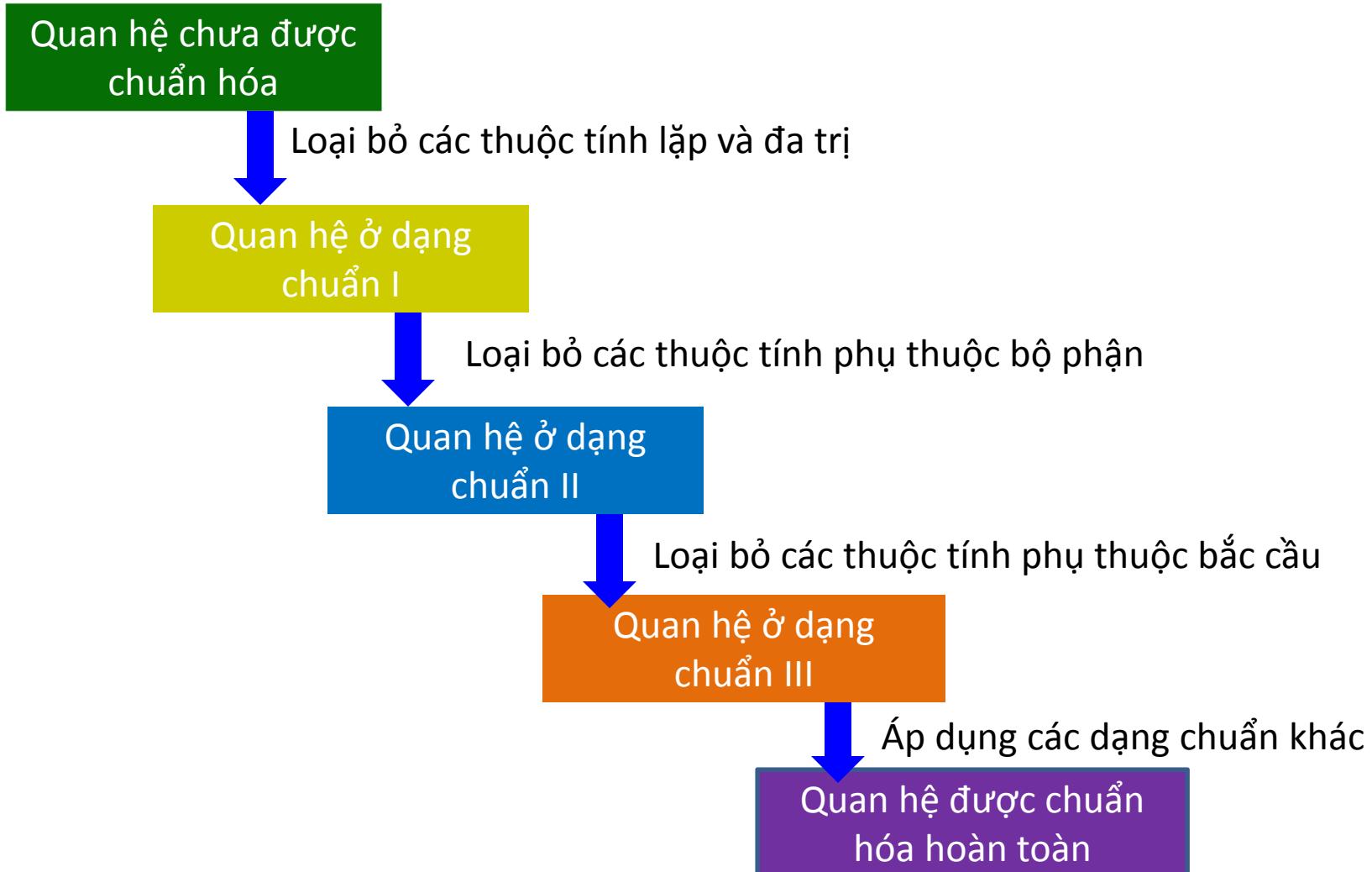
■ Định nghĩa cách khác:

- Cho r là một quan hệ định nghĩa trên tập quan hệ R .
- X và Y là hai tập con (khác nhau) các thuộc tính của R . Ta nói X xác định hàm Y , ký hiệu: $X \rightarrow Y$ là một **phụ thu c hàm** định nghĩa trên R nếu: $\forall t_1, t_2 \in r(R): t_1(X) = t_2(X) \Rightarrow t_1(Y) = t_2(Y)$

■ Ví dụ :

- Quan hệ **SINHVIEN**: **MASV**, **HOTENSV**, **NAMSINH** có phụ thu c **MASV**, **HOTENSV** \rightarrow **NAMSINH**

Quy trình chuẩn hóa dữ liệu



- Chu n hóa là quá trình phân rã l c quan h d a trên m t t p ph thu c hàm nh m m các l c quan h tho mān 2 tính ch t:
 - Trùng l p d li u ít nh t
 - Gi m thi u kh n ng gāy ra b t th ng khi c p nh t d li u.
- Các d ng chu n:
 - D ng chu n 1 (1NF)
 - D ng chu n 2 (2NF)
 - D ng chu n 3 (3NF)

Ví d 1: B ng d ng ch a chu n

- Xét b ng d li u NV nh sau:

ĐƠN VỊ	Mã số ĐV	Tên ĐV	Mã số NQL	Địa điểm
	5	Nghiên cứu	NV002	Nam Định, Hà Nội, Bắc Ninh
	4	Hành chính	NV014	Hà Nội
	1	Lãnh đạo	NV061	Hà Nội

- Ta thấy trong a i m c a n v 5 ch a thu c tính a tr (nhiều hơn 1 giá trị) l u 3 a danh khác nhau: Nam Định, Hà Nội, Bắc Ninh. Vì vậy, b ng d li u này ch a t d ng chu n hóa.

Ví d 2: B ng d ng ch a chu n

■ Xét b ng d li u NHÂN VIÊN_D ÁN:

NHÂN VIÊN_D ÁN	Mã số DA	Tên DA	Tên nhân viên	Số giờ
	1	DA01	Vân	15
			Nam	20
	2	DA02	Nam	10
			Thanh	12
			Băng	28
	3	DA03	Thanh	20
			Vân	12

- Ta th y nhóm thu c tính {Tên nhân viên, s gi } b l p l i n h i u l n trong các d án khác nhau. Vì v y, b ng d li u này ch a t d ng chu n h óa.

Ví d 3: B ng d ng ch a chu n

■ Xét b ng d li u Invoice

INVOICE

Customer Number	Customer	Terms	Ship Via	Order Date	Product No.	Description	Quantity	Unit Price	Extended Amount
1454837	W.Coyote	Net 30	USPS	12/01/2008	SPR-2290	Super Strength Springs	2	24.00	\$48.00
1454837	W.Coyote	Net 30	USPS	12/01/2008	SRT-67	Foot Straps, leather	2	2.50	\$5.00
1454837	W.Coyote	Net 30	USPS	12/01/2008	HLM-45	Deluxe Crash Helmet	1	67.88	\$67.88
1454837	W.Coyote	Net 30	USPS	12/01/2008	SFR-1	Rocket, Solid fuel	1	128,200.4	\$128,200.4
1454837	W.Coyote	Net 30	USPS	12/01/2008	ELT-1	Emergency Location Transmitter	1	79.88	\$79.88

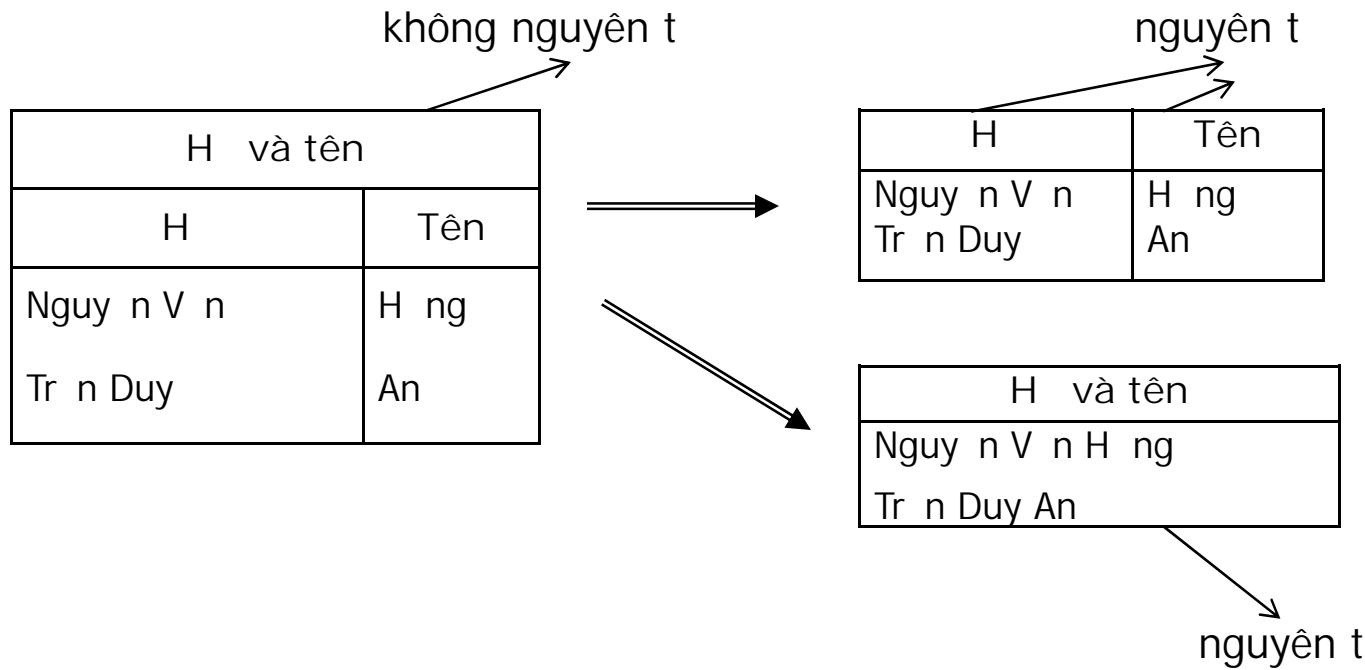
ây là 1 hóa n c a 1 kh ách hàng. N u thêm 1 hóa n kh ác th i các thu c t ính {Product No, Description, Quantity, Unit Price, Extended Mount} s l p l i trong hóa n m i. Vì v y, b ng d li u n ày ch a t d ng ch a n h óa.

- Khái niệm: Là cơ quan hạch giá là 1NF nếu và chỉ nut t có các thuộc tính của R tho mãn các điều kiện sau:
 - Là nguyên tử
 - Giá trị của các thuộc tính trên các bộ là nhất quán, không chia nhóm lẻ
 - Không có một thuộc tính nào có giá trị mà có thể tính toán từ một số thuộc tính khác

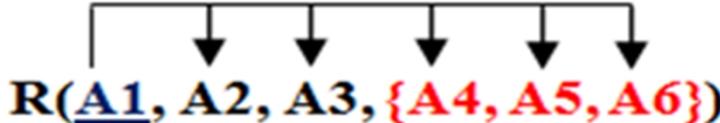
■ Ví dụ 1: Mô hình quan hệ trong 1NF

<u>MSKH</u>	TÊNKH	TP	PVC	<u>MSMH</u>	TÊNMH	G	SL
S1	An	HCM	01	P1	Táo	650	300
S1	An	HCM	01	P2	Chanh	500	200
S1	An	HCM	01	P3	Táo	450	400
S2	Hòa	HN	02	P1	Chanh	650	100
S2	Hoà	HN	02	P3	Táo	450	300
S3	Thanh	NT	03	P2	Chanh	500	200
S4	Trang	NT	03	P2	Chanh	500	210

■ Ví d 2: Quan h sau c ch ứ n h óa v d ạng ch ứ n 1



- Nguyên tắc chung: Loại bỏ thuộc tính lặp hoặc các tr
- Các bước thực hiện:
 - Tách nhóm thuộc tính lặp / a tr sang một tập riêng.
 - Khóa_candidate là Khóa candidate và khóa nhóm lặp
 - Bộ còn lại là bộ gồm có khóa và các thuộc tính còn lại.
- Ví dụ: cho quan hệ

$$R(\underline{A1}, A2, A3, \{A4, A5, A6\})$$


A1 là khóa chính của quan hệ, {A4, A5, A5} là thuộc tính lặp. Vì vậy, quan hệ R sẽ

$$R1(\underline{A1}, A4, A5, A6)$$


cách thành 2 quan hệ R1, R2

$$R2(\underline{A1}, A2, A3)$$


Ví d 1: a b ng d li u v d ng chu n 1NF

- Cho b ng N V

ĐƠN VỊ	Mã số ĐV	Tên DV	Mã số NQL	Địa điểm
	5	Nghiên cứu	NV002	Nam Định, Hà Nội, Bắc Ninh
	4	Hành chính	NV014	Hà Nội
	1	Lãnh đạo	NV061	Hà Nội

- Ta có quan h sau R(Mã số DV, Tên DV, Mã số NQL, { a i m})
- Thu c tính a tr là { a i m}. Vì v y, quan h này c tách thành:

R1(Mã số DV, a i m)

R2(Mã số DV, Tên DV, Mã số NQL)

Bảng ĐƠN VỊ - ĐỊA ĐIỂM

Mã số DV	Địa điểm
5	Nam định
5	Hà nội
5	Bắc Ninh
4	Hà nội
1	Hà nội

Bảng ĐƠN VỊ

Mã số DV	Tên DV	Mã số NQL
5	Nghiên cứu	NV002
4	Hành chính	NV014
1	Lãnh đạo	NV061

Ví d 2: a b ng d li u v d ng chu n 1NF

- Cho b ng NHÂN VIÊN_D ÁN

NHÂN VIÊN_D ÁN	MãsôDA	TênDA	Tênnhânviên	Sôgiờ
	1	DA01	Vân	15
			Nam	20
	2	DA02	Nam	10
			Thanh	12
			Băng	28
			Thanh	20
	3	DA03	Vân	12

- Quan h này là R(Mãs DA, TênDA, {Tênnhânviên, S gi })
- Thu c tính l p là {Tênnhânviên, S gi }. Vì v y, quan h này c tách thành:

R1(Mãs V, Tênnhânviên, S gi)

R2(Mãs V, TênDA)

NHÂN VIÊN_D ÁN

MãsôDA	Tênnhânviên	Sôgiờ
1	Vân	15
1	Nam	20
2	Nam	10
2	Thanh	12
2	Băng	28
3	Thanh	20
3	Vân	12

D ÁN

MãsôDA	TênDA
1	DA01
2	DA02
3	DA03

- Xét b ng INVOICE c vi t thành quan h sau (thêm Invoice Number làm thu c tính khóa cho b ng Invoice vì m i hóa n có S h oá n là Khóa)
R(Invoice Number (PK), Customer Number, Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone, Terms, Ship Via, Order Date, {Product Number, Product Description, Quantity, Unit Price, Extended Amount, Total Order Amount })
- Thu c tính l p c bôi và t trong d u { } quan h c tách
R1-INVOICE (Invoice Number (PK), Product Number, Product Description, Quantity, Unit Price, Extended Amount, Total Order Amount)
R2-INVOICE LINE ITEM (Invoice Number (PK), Customer Number, Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone, Terms, Ship Via, Order Date)

- Kết quả, bảng được tách thành hai bảng sau

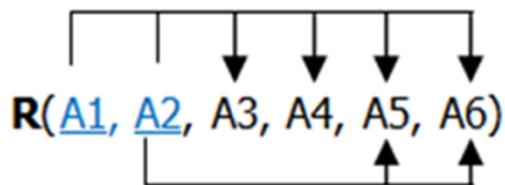
Customer Number	Customer	Terms	Ship Via	Order Date
1454837	W.Coyote	Net 30	USPS	12/01/2008

Product No.	Customer Number	Product No.	Description	Quantity	Unit Price	Extended Amount
SPR-2290	1454837	SPR-2290	Super Strength Springs	2	24.00	\$48.00
SRT-67	1454837	SRT-67	Foot Straps, leather	2	2.50	\$5.00
HLM-45	1454837	HLM-45	Deluxe Crash Helmet	1	67.88	\$67.88
SFR-1	1454837	SFR-1	Rocket, Solid fuel	1	128,200.4	\$128,200.4
ELT-1	1454837	ELT-1	Emergency Location Transmitter	1	79.88	\$79.88

- Quan hệ R là dạng chuẩn 2 nếu là 1NF và tất cả các thuộc tính không khóa đều thuộc tham mưu vào khóa.
- Thuộc tính không khóa: là thuộc tính không tham gia vào bất kỳ khóa nào.

- Quy tắc chun hóa t 1NF - 2NF:

- Bắc 1: Loại bỏ các thuộc tính không khoá phu thu c vào m t b phn khoá chính và tách thành ra m t b ng riêng, khoá chính c a b ng là b phn khoá mà chúng phu thu c vào.
- Bắc 2: Các thuộc tính còn l i l p thành m t quan h , khoá chính c a quan h này là khoá chính ban đầu.



Khóa chính là bộ $\{\underline{A_1}, \underline{A_2}\}$ (tức là từ $\{\underline{A_1}, \underline{A_2}\} \rightarrow \{A_3, A_4, A_5, A_6\}$)

Có phụ thuộc hàm phụ thuộc vào một phần của khóa $\{\underline{A_2}\} \rightarrow \{A_5, A_6\}$

Quan hệ được tách thành hai quan hệ sau

$R1(\underline{A_2}, A_5, A_6)$ - Tách các thuộc tính phụ thuộc vào 1 phần của khóa

$R2(\underline{A_1}, \underline{A_2}, A_3, A_4)$ - Bảng ban đầu còn lại

Ví d 1: Chu n hóa t 1NF-2NF

- Ví d : Cho quan h R(MSKH, TÊNKH, TP, PVC, MSMH, TÊNMH, G, SL) và t p ph thu c hàm F c nh ngh a trên R:

- Tập phụ thuộc hàm:

$$\begin{aligned}
 F = \{ & \underline{\text{MSKH}} \rightarrow \text{TÊNKH}, \text{TP} \\
 & \underline{\text{MSMH}} \rightarrow \text{TÊNMH}, \text{ĐG} \\
 & \underline{\text{MSKH}}, \underline{\text{MSMH}} \rightarrow \text{SL} \\
 & \text{TP} \rightarrow \text{PVC} \}
 \end{aligned}$$

<u>MSKH</u>	TÊNKH	TP	PVC	<u>MSMH</u>	TÊNMH	ĐG	SL
S1	An	HCM	01	P1	Táo	650	300
S1	An	HCM	01	P2	Chanh	500	200
S1	An	HCM	01	P3	Cam	450	400
S2	Hoa	HN	02	P1	Táo	650	100
S2	Hoa	HN	02	P3	Cam	450	300
S3	Thanh	NT	03	P2	Chanh	500	200
S4	Trang	NT	03	P2	Chanh	500	210

- Quan h này c tách thành các quan h sau
 R1(MSKH, TÊNKH, TP, PVC)
 R2(MSKH, MSMH, TÊNMH, G, SL)
- R2 ch a ph thu c hàm {MSMH} → {TÊNMH, G} nên c tách thành
 R21(MSMH, TÊNMH, G)
 R22(MSKH, MSMH, SL)

Ví d 1: Chu n hóa t 1NF-2NF (ti p)

■ K t qu c a quá trình chu n hóa

R ₁		ĐH	
MSKH	TÊNKH	TP	PVC
S1	An	HCM	01
S2	Hoà	HN	02
S3	Thanh	NT	03
S4	Trang	NT	03

$F_{R_1} = \{MSKH \rightarrow TÊNKH, TP$

$TP \rightarrow PVC\}$

MH		
MSMH	TÊNMH	ĐG
P1	Táo	650
P2	Cam	500
P3	Chanh	450

$F_{MH} = \{MSMH \rightarrow TÊNMH, ĐG\}$

MSKH	MSMH	SL
S1	P1	300
S1	P2	200
S1	P3	400
S2	P1	100
S2	P3	300
S3	P2	200
S4	P2	210

$F_{ĐH} = \{MSKH, MSMH \rightarrow SL\}$

Ví d 2: Chu n hóa t 1NF-2NF

NHÂN VIÊN_DỰ ÁN(Mã số NV, Mã số DA, Số giờ, Họ tên NV, Tên DA, Địa điểm DA)

với các phụ thuộc hàm:

$$\{\text{Mã số NV}, \text{Mã số DA}\} \rightarrow \{\text{Số giờ}\}$$

$$\{\text{Mã số NV}\} \rightarrow \{\text{Họ tên NV}\}$$

$$\{\text{Mã số DA}\} \rightarrow \{\text{Tên DA, Địa điểm DA}\}$$

<u>Mã số NV</u>	<u>Mã số DA</u>	Số giờ	Họ tên NV	Tên DA	Địa điểm DA
NV01	DA01	10	Nguyễn Lan	Dự án 1	Hà Nội
NV02	DA02	20	Lê Trung	Dự án 2	Hà Nam
NV01	DA02	15	Nguyễn Lan	Dự án 2	Hà Nam
NV02	DA01	30	Lê Trung	Dự án 1	Hà Nội
NV03	DA01	12	Đinh Trung	Dự án 1	Hà Nội
NV03	DA03	15	Đinh Trung	Dự án 3	Hải Phòng
NV02	DA03	26	Lê Trung	Dự án 3	Hải Phòng
NV04	DA02	16	Hoài Nam	Dự án 2	Hà Nam

- Quan hệ này được tách thành các quan hệ sau

R1 (Mã số NV, Họ tên NV)

R2 (Mã số NV, Mã số DA, Tên DA, Địa điểm DA, Số giờ)

Trong quan hệ R2 chứa phụ thuộc hàm $\{\text{Mã số DA}\} \rightarrow \{\text{Tên DA, Địa điểm DA}\}$ tách thành

R21(Mã số DA, Tên DA, Địa điểm DA)

R22(Mã số NV, Mã số DA, Số giờ)

■ K t qu c a quá trình chu n hóa

NHÂN VIÊN

Mã số NV	Họ tên NV
NV01	Nguyễn Lan
NV02	Lê Trung
NV03	Đinh Trung
NV04	Hoài Nam

DỰ ÁN

Mã số DA	Tên DA	Địa điểm DA
DA01	Dự án 1	Hà Nội
DA02	Dự án 2	Hà Nam
DA03	Dự án 3	Hải Phòng

NHÂN VIÊN_DỰ ÁN

Mã số NV	Mã số DA	Số giờ
NV01	DA01	10
NV02	DA02	20
NV01	DA02	15
NV02	DA01	30
NV03	DA01	12
NV03	DA03	15
NV02	DA03	26
NV04	DA02	16

Xét Quan h INVOICE LINE ITEM:

(Invoice Number (PK), Product Number (PK), Product Description,
Quantity, Unit Price, Extended Amount)

- Ta th y có ph thu c hàm vào m t ph n c a khóa nh sau
{Product Number} -> {Product Description, Unit Price}
- Quan h này c tách thành 2 quan h sau:
PRODUCT (Product Number, Product Description, Unit Price)
INVOICE LINE ITEM (Invoice Number (PK), Product Number (PK),
Quantity, ExtAmount)

- Cho quan hệ R , F là tập phâ thu c hàm nh nghia trên R .
- R là dạng chuẩn 3 nếu tho mãn các điều kiện sau:
 - R là dạng chuẩn 2NF
 - Mọi thuộc tính không khóa không phải thuộc bộ chung vào khóa (tức là tất cả các thuộc tính phải được suy ra trực tiếp từ khóa).

- Ví dụ : Quan hệ $R(\underline{\text{MSKH}}, \text{TÊNKH}, \text{TP}, \text{PVC})$, $F_R = \{\underline{\text{MSKH}} \rightarrow \{\text{TÊNKH}, \text{TP}, \text{PVC}\}$

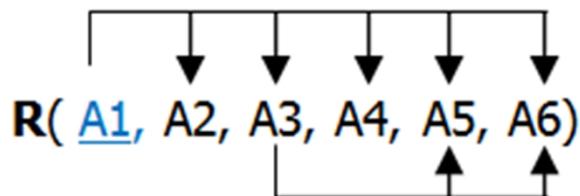
MSKH là khóa chính

$\{\text{TP}\} \rightarrow \text{PVC}$

- Nhận xét chuẩn 3NF vì thuộc tính PVC c suy ra thuộc tính TP (không phải là khóa).

■ Quy tắc chun hóa v 3NF:

- Bác 1: Loại bỏ các thuộc tính phụ thuộc bắc cầu ra khỏi quan hệ và tách chúng thành một quan hệ riêng có khoá chính là thuộc tính bắc cầu.
- Bác 2: Các thuộc tính còn lại lập thành một quan hệ có khoá chính là quan hệ ban đầu.



Khóa chính là $\{\underline{A1}\}$ (tức là từ $\{\underline{A1}\} \rightarrow \{A2, A3, A4, A5, A6\}$)

Có phụ thuộc hàm phụ thuộc bắc cầu $\{A3\} \rightarrow \{A5, A6\}$

Quan hệ được tách thành hai quan hệ sau

R1(A3, A5, A6)

- Tách các thuộc tính phụ thuộc bắc cầu

R2(A1, A2, A3, A4)

- Bảng ban đầu còn lại

Ví d 1: Chuy n sang chu n 3NF

- Ví d : Quan h R(MSKH, TÊNKH, TP, PVC) c tách thành 2 b ng
 R-VC (TP, PVC)
 R-KHÁCHHÀNG (MSKH, TÊNKH, TP)

KH			VC	
<u>MSKH</u>	TÊNKH	TP	<u>TP</u>	PVC
S1	An	HCM	HCM	01
S2	Hoà	HN	HN	02
S3	Bình	NT	NT	03
S4	Trang	NT		

$F_{KH} = \{MSKH \rightarrow TÊNKH, TP\}$

$F_{VC} = \{TP \rightarrow PVC\}$

MH			ĐH		
<u>MSMH</u>	TÊNMH	ĐG	<u>MSKH</u>	<u>MSMH</u>	SL
P1	Táo	650	S1	P1	300
P2	Cam	500	S1	P2	200
P3	Chanh	450	S1	P3	400
			S2	P1	100
			S2	P3	300
			S3	P2	200
			S4	P2	210

$F_{MH} = \{MSMH \rightarrow TÊNMH, ĐG\}$

$F_{DH} = \{MSKH, MSMH \rightarrow SL\}$

Ví dụ 2: Chuyển sang chuẩn 3NF

Cho quan hệ sau:

NHÂN VIÊN - ĐƠN VỊ(Mã số NV, Họ tên NV, Ngày sinh, Địa chỉ, Mã số ĐV, Tên ĐV, Mã số NQL)

Có chứa các phụ thuộc hàm dưới đây:

{Mã số NV} → {Họ tên NV, Ngày sinh, Địa chỉ, Mã số ĐV, Tên ĐV, Mã số NQL} Mã số NV là khóa chính của bảng

{Mã số ĐV} → {Tên ĐV, Mã số NQL}

Chứa phụ thuộc hàm bắc cầu vào thuộc tính không phải khóa

<u>Mã số NV</u>	Họ tên NV	Ngày sinh	Địa chỉ	Mã số ĐV	Tên ĐV	Mã số NQL
NV01	Tùng	1/1/1983	16 Thụy Khuê HN	DV01	Phòng tin học	NV02
NV02	Cúc	2/2/1980	10 Hàng Đào HN	DV01	Phòng tin học	NV02
NV03	Trúc	3/3/1983	Đà Nẵng	DV02	Hành chính	NV05
NV04	Mai	5/5/1982	Hà Giang	DV02	Hành chính	NV05
NV05	Sen	6/6/1986	Hải Phòng	DV02	Hành chính	NV05
NV06	Hồng	6/8/1984	Hà Nam	DV03	Kinh doanh	NV06

- Quan hệ này chia tách thành các quan hệ sau

ĐƠN VỊ (Mã số ĐV, Tên ĐV, Mã số NQL)

NHÂN VIÊN (Mã số NV, Họ tên NV, Ngày sinh, Địa chỉ, Mã số ĐV)

Ví dụ 2: Chuyển sang chuẩn 3NF

ĐƠN VỊ

<u>Mã số DV</u>	Tên DV	Mã số NQL
DV01	Phòng tin học	NV02
DV02	Hành chính	NV05
DV03	Kinh doanh	NV06

NHÂN VIÊN

<u>Mã số NV</u>	Họ tên NV	Ngày sinh	Địa chỉ	Mã số DV
NV01	Tùng	1/1/1983	16 Thụy Khuê HN	DV01
NV02	Cúc	2/2/1980	10 Hàng Đào HN	DV01
NV03	Trúc	3/3/1983	Đà Nẵng	DV02
NV04	Mai	5/5/1982	Hà Giang	DV02
NV05	Sen	6/6/1986	Hải Phòng	DV02
NV06	Hồng	6/8/1984	Hà Nam	DV03

- Xét quan h **INVOICE** à c tách nh sau:
 - **INVOICE**: Invoice Number (PK), Customer Number, Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone, Terms, Ship Via, Order Date, Total Order Amount
- Ta th y **INVOICE** có ch a thu c tính b c c u sau
{Customer Number} -> {Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone}
- Quan h này c tách thành hai b ng sau:
CUSTOMER (Customer Number, Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone)
INVOICE (Invoice Number (PK), Customer Number, Terms, Ship Via, Order Date, Total Order Amount)

- Quy trình

- Chu n h oá quan h INVOICE v d ng chu n 1NF
- Ti p t c chu n h oá v d ng 2NF
- Chu n h oá v d ng 3NF

- Sau khi chuy n sang d ng chu n 1NF, CSDL s g m hai quan h nh sau:
 - Quan h INVOICE: Invoice Number (PK), Customer Number, Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone, Terms, Ship Via, Order Date, Total Order Amount
 - Quan h INVOICE LINE ITEM: Invoice Number (PK), Product Number (PK), Product Description, Quantity, Unit Price, Extended Amount

■ Ví d 2: Quan h INVOICE c phân rã v d ng 2NF
nh sau:

- **INVOICE**: Invoice Number (PK), Customer Number, Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone, Terms, Ship Via, Order Date, Total Order Amount
- **INVOICE LINE ITEM**: Invoice Number (PK), Product Number (PK), Quantity, Sale Unit Price, Extended Amount
- **PRODUCT**: Product Number (PK), Product Description, List Unit Price

- Quan h INVOICE c phân rã v d ng 3NF nh sau:
 - INVOICE: Invoice Number (PK), Customer Number, Terms, Ship Via, Order Date
 - INVOICE LINE ITEM: Invoice Number (PK), Product Number (PK), Quantity, Sale Unit Price
 - PRODUCT: Product Number (PK), Product Description, List Unit Price
 - CUSTOMER : Customer Number (PK), Customer Name, Customer Address, Customer City, Customer State, Customer Zip Code, Customer Phone

Cho lược đồ quan hệ R = {A, B, C,D,E,F,G,H,I,J} có khóa chính là A,B

Với tập các phụ thuộc hàm :

A, B → C,D,E,F,G,H,I,J

A → E,F,G,H,I,J

F → I, J

Hãy chuẩn hóa quan hệ R về dạng chuẩn 3NF

áp án: R1(E,I,J), R2(A,E,F,G,H), R3(A,B,C,D)

Bài tập vở nhà:

Chuẩn hóa CSDL UTLA Academic Tracking và Computer Books Company (giáo trình) sau ôn bài giải cho giảng viên.

- Quá trình chuẩn hóa làm tăng thêm số lượng các bảng và các liên kết.
- Nếu sử dụng CSDL có thể gặp vấn đề hữu dụng mà không thể giải quyết bằng cách sử dụng các phép toán có (chuyển đổi thành các câu lệnh lưu học nâng cấp phần cứng) -> Cần phải khách hàng.
- Các bước khách hàng:
 - Kết hợp lại các quan hệ đã chia thành một số quy tắc chuẩn hóa
 - Lưu trữ dữ liệu thô trong các bảng
 - Lưu trữ dữ liệu tóm tắt trong các bảng

- Các thao tác trên dữ liệu (thêm, sửa, xóa) trên CSDL chưa chuẩn hóa có thể gây ra các bất thường trên CSDL.
- Chuẩn hóa CSDL để đảm bảo dữ liệu không bị thay đổi và loại bỏ các bất thường trên CSDL
- 3 bước chuẩn hóa thông thường:
 - Chuẩn hóa với 1NF: Loại bỏ dữ liệu lặp
 - Chuẩn hóa với 2NF: Loại bỏ các phần thuộc c hàm bị phân
 - Chuẩn hóa với 3NF: Loại bỏ các phần thuộc c hàm bị cung cấp

XIN C M N!