

Chương 4: Thiết kế dữ liệu lưu trữ

Môn học: Phân tích và thiết kế phần mềm

Nội dung

- Lưu trữ dữ liệu bằng CSDL quan hệ
 - Một số kỹ thuật để ánh xạ sơ đồ lớp sang sơ đồ logic (tổ chức lưu trữ dữ liệu bằng CSDL quan hệ)
- Lưu trữ dữ liệu bằng XML
 - Giới thiệu sơ lược về XML
 - Cách lưu trữ dữ liệu
- So sánh giữa cách sử dụng lưu trữ bằng XML, lưu trữ bằng CSDL quan hệ, và kết hợp cả 2 cách.

LƯU TRỮ DỮ LIỆU BẰNG CSDL QUAN HỆ

Quy tắc #1

- Thông thường, mỗi lớp đối tượng đơn giản được ánh xạ thành **một bảng**

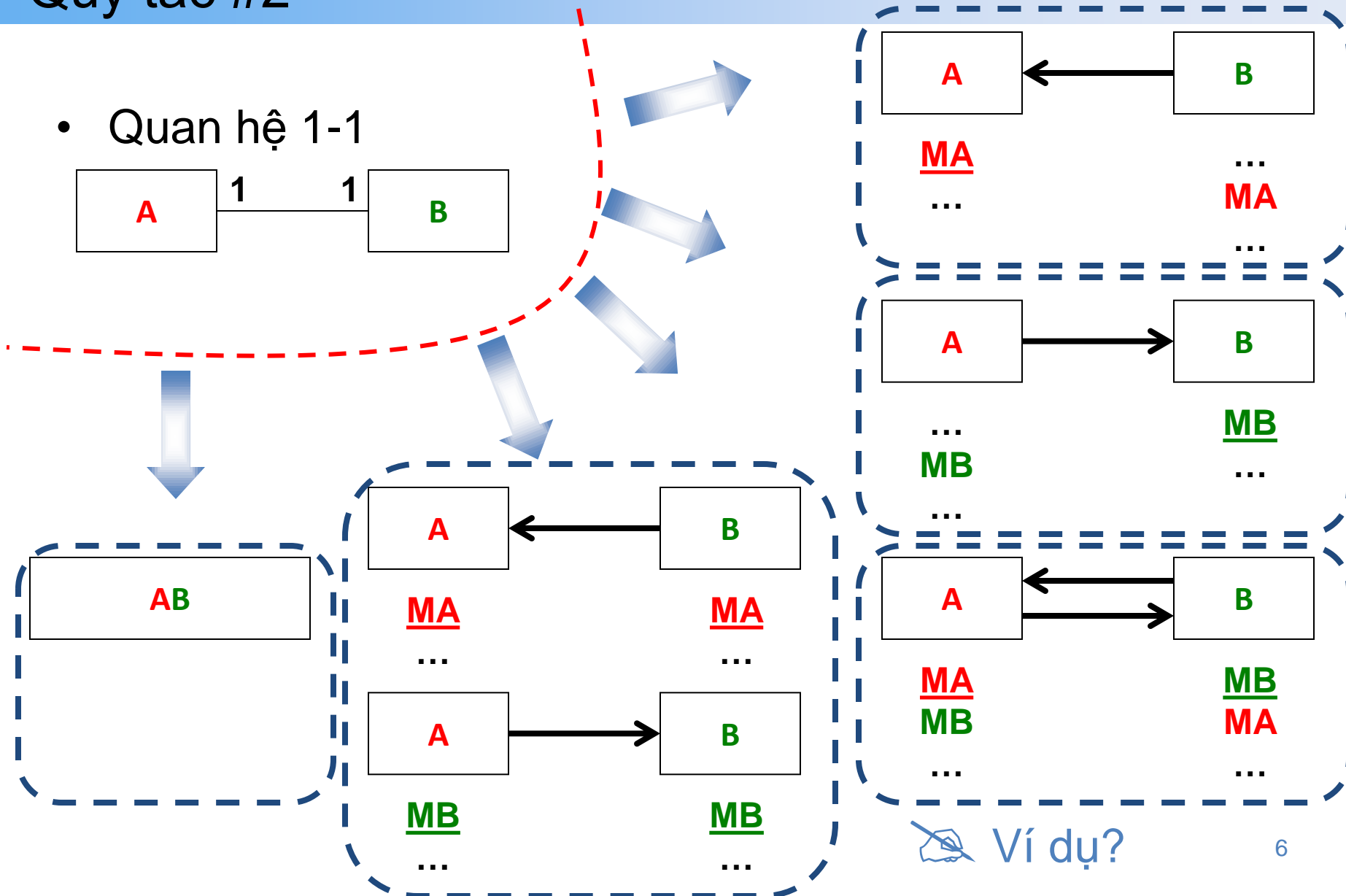
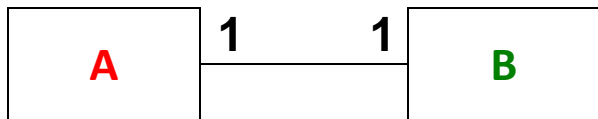
 Ví dụ?

Quy tắc #1 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các yêu cầu sau:
 - **Bài toán 1:** Viết phần mềm quản lý học sinh bao gồm các thông tin: Họ tên, ngày sinh, địa chỉ, điểm trung bình.
 - **Bài toán 2:** Viết phần mềm quản lý sách với các thông tin tên sách, số lượng, ngày nhập.
 - **Bài toán 3:** Viết phần mềm quản lý xe buýt với các thông tin tên chuyến, thời gian bắt đầu, thời gian kết thúc, thời gian giãn cách giữa hai chuyến.

Quy tắc #2

- Quan hệ 1-1



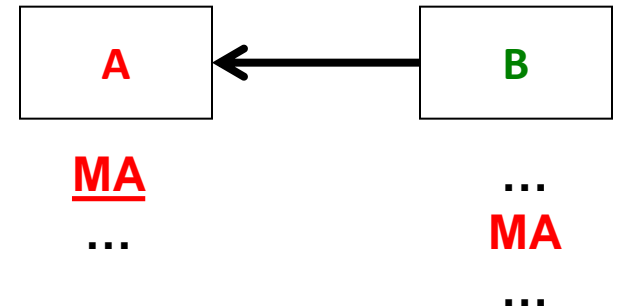
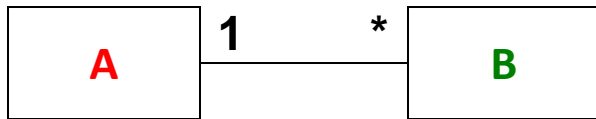
 Ví dụ?

Quy tắc #2 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho tất cả ***trường hợp có thể*** với các đặc tả sau:
 - **Bài toán 1**: Viết phần mềm quản lý các cá nhân và chứng minh nhân dân của họ. Các thông tin của cá nhân đó bao gồm (tên, địa chỉ, điện thoại). Các thông tin về chứng minh nhân dân gồm có (Số CMND, ngày cấp, nơi cấp).
 - **Bài toán 2**: Viết phần mềm cho các hành khách trên hãng máy bay gồm các thông tin họ tên, địa chỉ, hộ chiếu. Chi tiết thông tin hộ chiếu bao gồm (mã hộ chiếu, ngày cấp, nơi cấp, ngày hết hạn)

Quy tắc #3

- Quan hệ 1-n



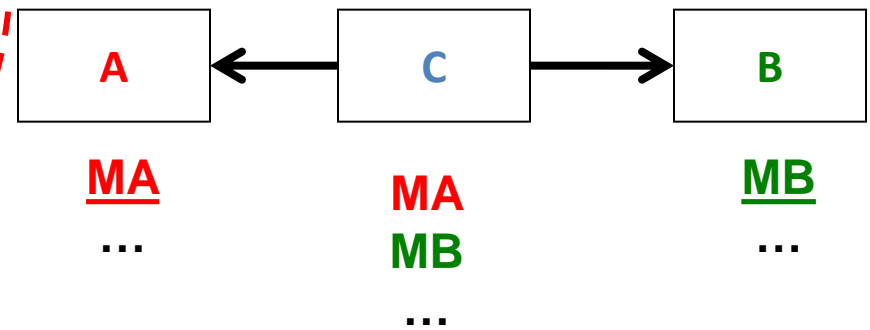
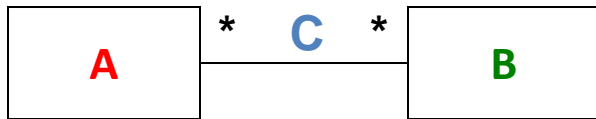
 Ví dụ?

Quy tắc #3 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - **Bài toán 1:** Viết phần mềm quản lý bộ môn gồm các thông tin tên bộ môn, giờ mở cửa, giờ đóng cửa. Bộ môn sẽ bao gồm các giáo viên gồm các thông tin tên, thâm niên, địa chỉ
 - **Bài toán 2:** Viết phần mềm quản lý sổ tiết kiệm có kỳ hạn của các chủ thẻ. Thông tin về sổ tiết kiệm bao gồm : số tiền gửi, ngày mở, ngày tắt toán.

Quy tắc #4

- Quan hệ m-n



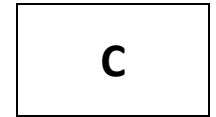
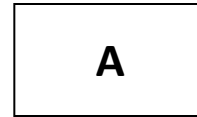
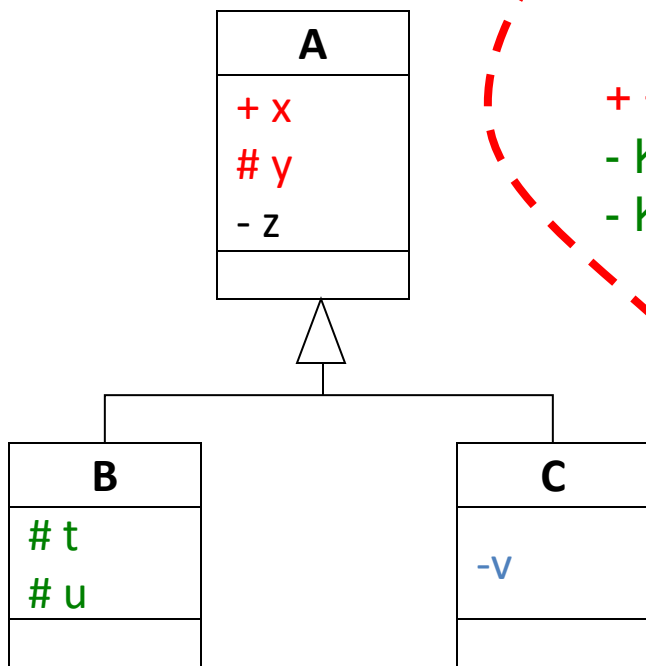
 Ví dụ?

Quy tắc #4 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - **Bài toán 1:** Viết phần mềm quản lý số lượng và đơn giá của các sản phẩm trong giỏ hàng. Sản phẩm gồm các thông tin bao gồm tên, mô tả, hình ảnh minh họa.
 - **Bài toán 2:** Viết phần mềm quản lý các ga trung gian của các tuyến tàu. Tuyến tàu bao gồm các thông tin tên tuyến, thời gian bắt đầu, thời gian kết thúc, thời gian giãn cách. Ga gồm có thông tin về tên ga, mô tả vị trí, bản đồ.

Quy tắc #5

- Quan hệ kế thừa



$A(\underline{MA}, x, y, z)$ $B(\underline{MB}, x, y, t, u)$ $C(\underline{MC}, x, y, v)$

+ Đơn giản

- Không thấy mối liên hệ giữa các loại đối tượng A, B, C
- Khó thống kê tổng quát

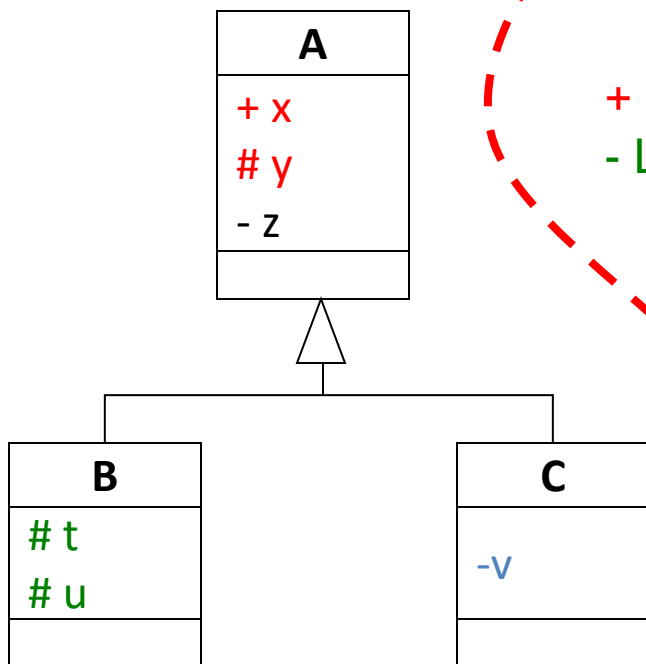
 Ví dụ?

Quy tắc #5 – Ví dụ minh họa

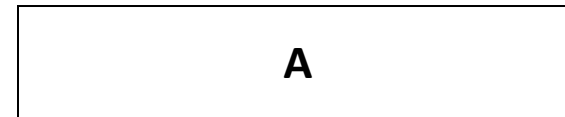
- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - **Bài toán 1:** Quản lý việc tính lương cho 3 loại nhân viên sau:
 - **Đối với nhân viên văn phòng**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số ngày làm việc} * 200.000 + \text{Trợ Cấp}$
 - **Đối với nhân viên sản xuất**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số Sản Phẩm} * 5.000$
 - **Đối với nhân viên quản lý**
 - $\text{Lương} = \text{Lương Cơ Bản} * \text{Hệ số chức vụ} + \text{Thưởng}$

Quy tắc #5

- Quan hệ kế thừa



- + Có được cái nhìn tổng quát về các đối tượng
- Lãng phí không gian lưu trữ



$A(\underline{M}, \text{Loại}, x, y, z, t, u, v)$

Loại	x	y	z	t	u	v
A						
B						
C						

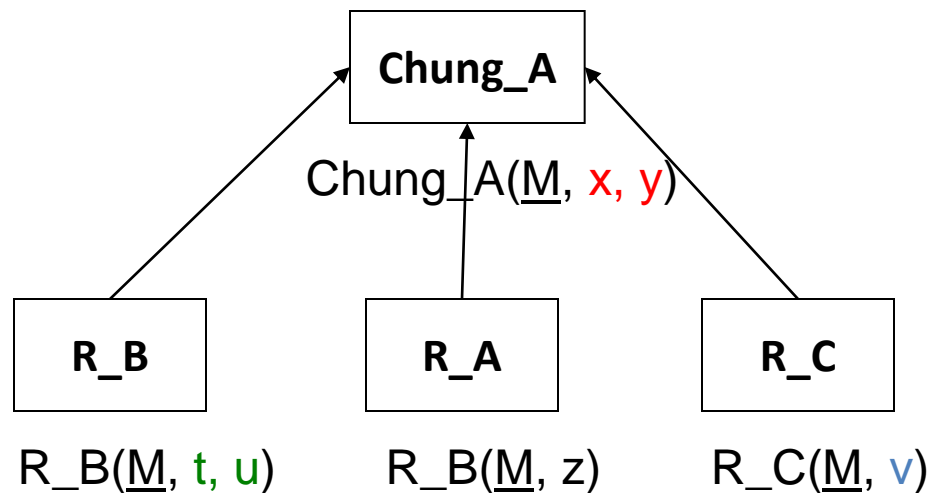
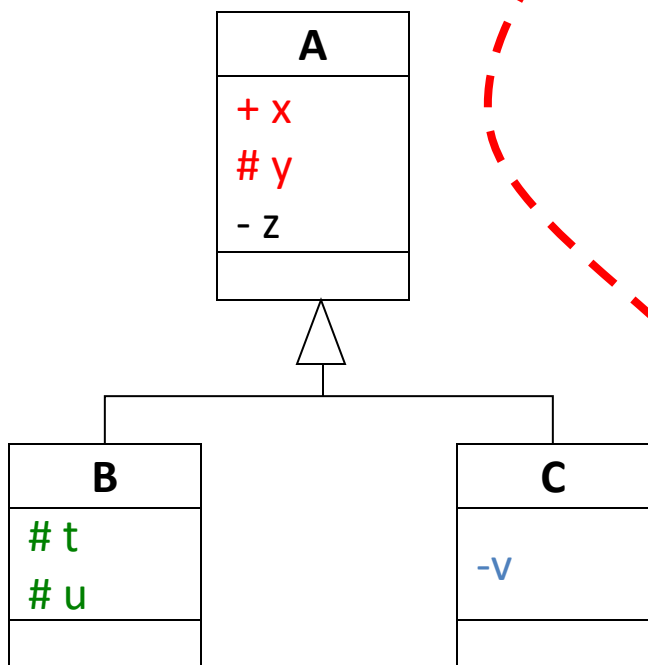
 Ví dụ?

Quy tắc #5 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - **Bài toán 1:** Quản lý việc tính lương cho 3 loại nhân viên sau:
 - **Đối với nhân viên văn phòng**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số ngày làm việc} * 200.000 + \text{Trợ Cấp}$
 - **Đối với nhân viên sản xuất**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số Sản Phẩm} * 5.000$
 - **Đối với nhân viên quản lý**
 - $\text{Lương} = \text{Lương Cơ Bản} * \text{Hệ số chức vụ} + \text{Thưởng}$

Quy tắc #5

- Quan hệ kế thừa



- + Tiết kiệm không gian lưu trữ
- + Cho phép có cái nhìn tổng quát
- Tổ chức khá phức tạp

 Ví dụ?

Quy tắc #5 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - **Bài toán 1:** Quản lý việc tính lương cho 3 loại nhân viên sau:
 - **Đối với nhân viên văn phòng**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số ngày làm việc} * 200.000 + \text{Trợ Cấp}$
 - **Đối với nhân viên sản xuất**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số Sản Phẩm} * 5.000$
 - **Đối với nhân viên quản lý**
 - $\text{Lương} = \text{Lương Cơ Bản} * \text{Hệ số chức vụ} + \text{Thưởng}$

Quy tắc #6

- Lớp đối tượng có thuộc tính có **cấu trúc phức tạp**
- Tách thành bảng phụ** để lưu trữ thuộc tính có cấu trúc phức tạp đó

```
class A
```

```
{
```

```
...
```

```
  B
```

```
  x;
```

```
...
```

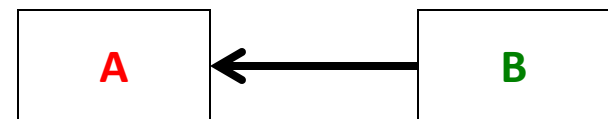
```
}
```

Cấu trúc của B gồm:

Thuộc tính 1

Thuộc tính 2

...



MA

...

...

MA

ThuocTinh1

ThuocTinh2

...



Ví dụ?

Quy tắc #6 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - Bài toán 1: Viết phần mềm tính chu vi của một hình tròn gồm tâm và bán kính.
 - Bài toán 2: Viết phần mềm tính diện tích của một tam giác

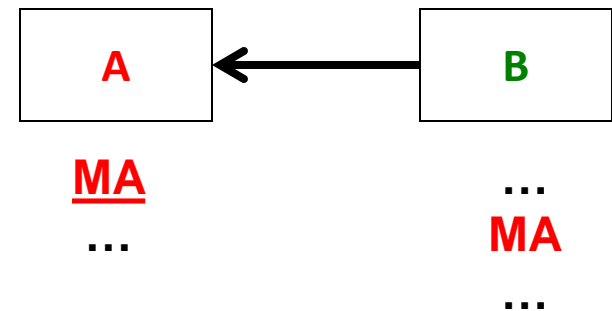
Thiết kế cơ sở dữ liệu (Bài 1)

Thiết kế cơ sở dữ liệu (Bài 2)

Quy tắc #7

- Lớp đối tượng có thuộc tính kiểu mảng
- Tách thành bảng chi tiết

```
class A
{
  ...
  B[] x;
  ...
}
```



 Ví dụ?

Quy tắc #7 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - Bài toán 1: Viết phần mềm tính phần giao của hai hình đa giác
 - Bài toán 2: Viết phần mềm tính giá trị của đa thức gồm các đơn thức.

Thiết kế Cơ sở dữ liệu (Bài 1)

Quy tắc #8

- Thuộc tính có giá trị rời rạc: Tách thành bảng danh mục

 Ví dụ?

Quy tắc #8 – Ví dụ minh họa

- Thiết kế cơ sở dữ liệu cho các đặc tả sau:
 - **Bài toán 1:** Viết phần mềm quản lý sách với các thông tin tên sách, ngôn ngữ, thể loại.
 - **Bài toán 2:** Viết phần mềm quản lý nhân viên với các thông tin tên, nơi sinh, địa chỉ, ngày sinh, phòng ban, chức vụ (giám đốc, trưởng phòng, nhân viên), tình trạng gia đình, số năm công tác,

Thiết kế Cơ sở dữ liệu (Bài 1)

Quy tắc #9

- Bảng tham số
 - Dạng 1

Khóa	ThamSố#1	ThamSố#2	...	ThamSố#N
...

Mỗi tham số tương ứng với một cột trong bảng tham số

Bảng tham số thường chỉ gồm 1 dòng (chứa giá trị các tham số hiện hành)

?

Cần bổ sung tham số mới?

Cần vô hiệu hóa tác dụng của một tham số?

Quy tắc #8 – Ví dụ minh họa

- Thiết kế bảng tham số và cho ví dụ cụ thể cho bài toán sau
 - **Bài toán 1:** Phần mềm quản lý sổ tiết kiệm với số tiền gửi tối đa là 500 000 000.
 - **Bài toán 2:** Phần mềm quản lý sổ nhà sách với số lượng sách mượn tối đa là 2 cuốn, thời gian mượn tối đa là 2 ngày. Nếu vượt quá hạn sẽ bị phạt 10 000 / 1 cuốn.

Thiết kế Bảng tham số (Bài 1)

Quy tắc #9

- Bảng tham số
 - Dạng 2

MãThamSố	TênThamSố	Kiểu	GiáTrị	TìnhTrạng
...

Được lưu
dạng chuỗi

Mỗi tham số tương ứng với một dòng trong bảng tham số

Giá trị hiện tại của tham số được lưu bằng dạng chuỗi

Mỗi tham số cần lưu trữ kiểu giá trị để phần mềm “hiểu” đúng nội dung giá trị hiện tại của tham số

?

Cần bổ sung tham số mới?

Cần vô hiệu hóa tác dụng của một tham số?

Quy tắc #8 – Ví dụ minh họa

- Thiết kế bảng tham số và cho ví dụ cụ thể cho bài toán sau
 - Bài toán 1: Phần mềm quản lý sổ tiết kiệm với số tiền gửi tối đa là 500 000 000.
 - Bài toán 2: Phần mềm quản lý sổ nhà sách với số lượng sách mượn tối đa là 2 cuốn, thời gian mượn tối đa là 2 ngày. Nếu vượt quá hạn sẽ bị phạt 10 000 / 1 cuốn.

Thiết kế Bảng tham số (Bài 1)

□

Bài toán tổng hợp

- **Thiết kế cơ sở dữ liệu cho bài toán sau:**
 - **Bài toán 1:** Công ty ABC là công ty sản xuất kinh doanh thú nhồi bông. Công ty có nhiều nhân viên làm việc trong ba bộ phận khác nhau: **bộ phận quản lý, bộ phận sản xuất, bộ phận văn phòng**. Việc tính lương cho nhân viên dựa vào các yếu tố sau:
 - **Đối với nhân viên văn phòng**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số ngày làm việc} * 100.000 + \text{Trợ Cấp}$
 - **Đối với nhân viên sản xuất**
 - $\text{Lương} = \text{Lương Cơ Bản} + \text{Số Sản Phẩm} * 2.000$
 - **Đối với nhân viên quản lý**
 - $\text{Lương} = \text{Lương Cơ Bản} * \text{Hệ số chức vụ} + \text{Thưởng}$

Thiết kế cơ sở dữ liệu

Bài toán tổng hợp

- **Thiết kế CSDL cho bài toán sau:**
- **Bài toán:** Một cái máy có nhiều chi tiết. Mỗi chi tiết máy có thể là chi tiết đơn hoặc là chi tiết phức.
 - **Chi tiết đơn:** là chi tiết không chứa bên trong nó chi tiết khác. Thông tin của chi tiết đơn bao gồm: *mã số chi tiết, giá tiền*.
 - **Chi tiết phức:** là chi tiết chứa bên trong nó nhiều chi tiết thành phần, mỗi một chi tiết thành phần này có thể là chi tiết đơn hoặc là chi tiết phức. Thông tin của chi tiết phức bao gồm: *mã số chi tiết, số lượng chi tiết thành phần, danh sách các chi tiết thành phần*. Giá tiền của chi tiết phức bằng tổng giá tiền của các chi tiết thành phần.

LƯU TRỮ DỮ LIỆU BẰNG XML

- Lược dịch từ slide chương 2 – XML
- trong môn Managing XML and Semistructured Data của GS. Dan Suciu,
- University of Washington
- Department of Computer Science & Engineering

Tài liệu gốc:

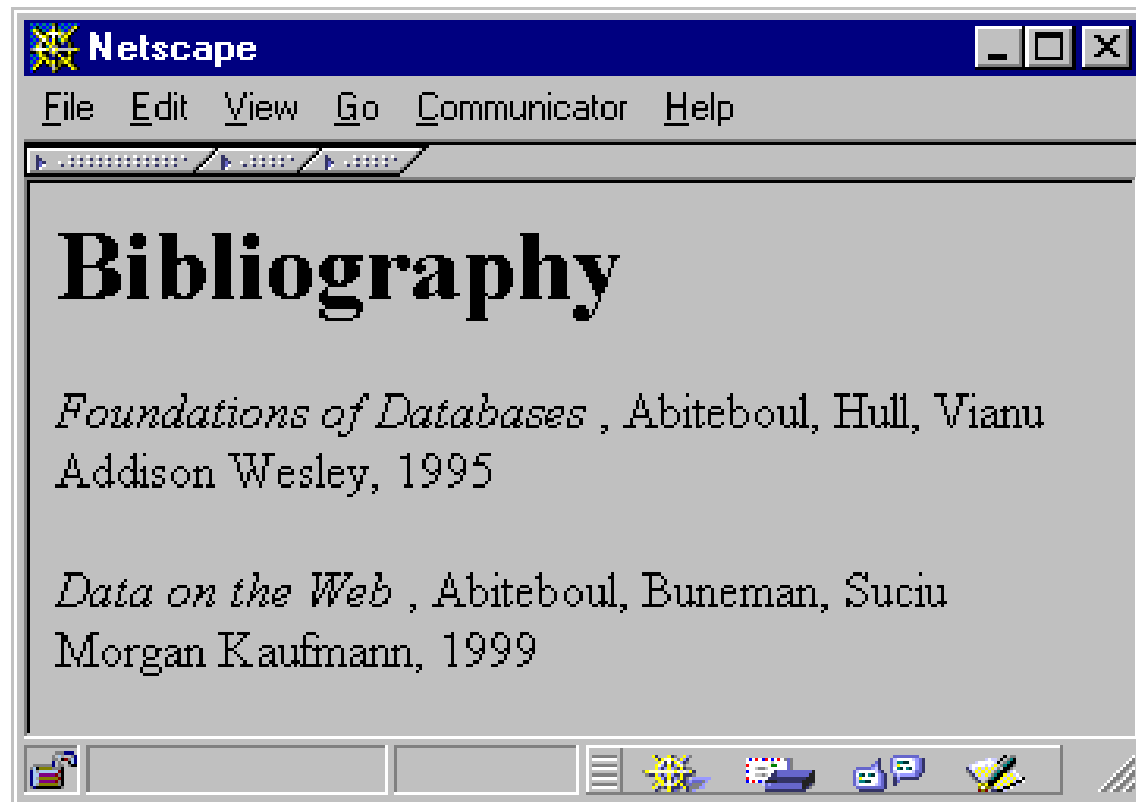
<http://www.cs.washington.edu/homes/suciu/COURSES/590DS/02xmlsyntax.ppt>

XML

- Chuẩn của W3C
- Ý tưởng ban đầu: văn bản có cấu trúc (SGML)
- <http://www.w3.org/TR/2000/REC-xml-20001006>
(version 2, 10/2000)

$$\text{HTML4.0} \in \text{XML} \subset \text{SGML}$$

HTML và XML



HTML mô tả “cách thể hiện”

HTML

`<h1> Bibliography </h1>`

`<p> <i> Foundations of Databases </i>`

Abiteboul, Hull, Vianu

`
` Addison Wesley, 1995

`<p> <i> Data on the Web </i>`

Abiteoul, Buneman, Suciu

`
` Morgan Kaufmann, 1999

XML

```
<bibliography>
  <book>
    <title> Foundations... </title>
    <author> Abiteboul </author>
    <author> Hull </author>
    <author> Vianu </author>
    <publisher> Addison Wesley </publisher>
    <year> 1995 </year>
  </book>
  ...
</bibliography>
```

XML mô tả nội dung

XML

<bibliography>

<book>

Start tag

<title> Foundations... </title>

<author> Abiteboul </author>

<author> Hull </author>

<author> Vianu </author>

<publisher> Addison Wesley </publisher>

<year> 1995 </year>

</book>

...

</bibliography>

End tag

XML

```
<bibliography>
  <book>
    <title> Foundations... </title>
    <author> Abiteboul </author>
    <author> Hull </author>
    <author> Vianu </author>
    <publisher> Addison Wesley </publisher>
    <year> 1995 </year>
  </book>
  ...
</bibliography>
```

Cơ bản về XML

- tag: **book**, **title**, **author**, ...
- start tag: **<book>**, end tag: **</book>**
- element: **<book>...<book>**, **<author>...</author>**
- Các element có thể lồng nhau
- Element rỗng: **<red></red>** hay **<red/>**
- Một tài liệu XML: có 1 *root element*

Tài liệu XML được gọi là “**well formed**” nếu các **tag bắt đầu** và **tag kết thúc** tương ứng xuất hiện **đầy đủ** và **đúng vị trí** (kể cả với các tag lồng nhau)

XML: Attribute

```
<book price = "55" currency = "USD">  
  <title> Foundations of Databases </title>  
  <author> Abiteboul </author>  
  ...  
  <year> 1995 </year>  
</book>
```

Attribute: cách khác để biểu diễn dữ liệu
(không dùng Element)

XML: Oid và Tham chiếu

```
<person id="o555"> <name> Jane </name> </person>  
  
<person id="o456"> <name> Mary </name>  
  <children idref="o123 o555"/>  
</person>  
  
<person id="o123" mother="o456"><name>John</name>  
</person>
```

OID và tham chiếu trong XML

Tham chiếu đến entity

- Cú pháp: &entityname;
- Ví dụ:
<element> this is less than < </element>
- Một số entity thông dụng:

<	<
>	>
&	&
'	'
"	“
&	Unicode char

Sử dụng Comment

<!-- Comment text... -->

XML Namespace

- <http://www.w3.org/TR/REC-xml-names> (1/99)
- $\text{name} ::= [\text{prefix:}] \text{localpart}$

```
<book xmlns:isbn="www.isbn-org.org/def">  
  <title> ... </title>  
  <number> 15 </number>  
  <isbn:number> .... </isbn:number>  
</book>
```

XML Namespace

- Cú pháp: `<number>` , `<isbn:number>`

`<tag xmlns:mystyle = "http://...">`

...

`<mystyle:title> ... </mystyle:title>`

`<mystyle:number> ...`

`</tag>`

Định nghĩa tại...

```
graph LR
    A["xmlns:mystyle = \"http://...\""] --> B(Định nghĩa tại...)
    C["mystyle:title"] --> B
    D["mystyle:number"] --> B
```

Mô hình truy vấn dữ liệu XML

- **Node** ::= DocNode |
ElemNode |
ValueNode |
AttrNode |
NSNode |
PINode |
CommentNode |
InfoltemNode |
RefNode

Mô hình truy vấn dữ liệu XML

Nút Element:

- `elemNode` : (`QNameValue`,
 `{AttrNode }`,
 `[ElemNode | ValueNode]`)
 → `ElemNode`
- `QNameValue` = “tên tag”
- `{...}` = “tập hợp gồm ...”
- `[...]` = “danh sách gồm ...”

Mô hình truy vấn dữ liệu XML

Ví dụ

```
<book price = "55"  
      currency = "USD">  
  <title> Foundations ... </title>  
  <author> Abiteboul </author>  
  <author> Hull </author>  
  <author> Vianu </author>  
  <year> 1995 </year>  
</book>
```

```
book1= elemNode(book,  
  {price2, currency3},  
  [title4,  
   author5,  
   author6,  
   author7,  
   year8])  
  
price2 = attrNode(...) /* next */  
currency3 = attrNode(...)  
title4 = elemNode(title, string9)  
...
```

Mô hình truy vấn dữ liệu XML

Nút Attribute:

- `attrNode : (QNameValue, ValueNode)`
 \rightarrow `AttrNode`

Mô hình truy vấn dữ liệu XML

Ví dụ

```
<book price = "55"  
      currency = "USD">  
  <title> Foundations ... </title>  
  <author> Abiteboul </author>  
  <author> Hull </author>  
  <author> Vianu </author>  
  <year> 1995 </year>  
</book>
```

```
price2 = attrNode(price,string10)  
string10 = valueNode(...) /* next */  
currency3 = attrNode(currency,  
                      string11)  
string11 = valueNode(...)
```

Mô hình truy vấn dữ liệu XML

Nút Value:

- `ValueNode` = `StringValue` |
 `BoolValue` |
 `FloatValue` ...
- `stringValue` : `string` → `StringValue`
- `boolValue` : `boolean` → `BoolValue`
- `floatValue` : `float` → `FloatValue`

Mô hình truy vấn dữ liệu XML

Ví dụ

```
<book price = "55"  
      currency = "USD">  
  <title> Foundations ... </title>  
  <author> Abiteboul </author>  
  <author> Hull </author>  
  <author> Vianu </author>  
  <year> 1995 </year>  
</book>
```

```
price2 = attrNode(price,string10)  
string10 = valueNode(stringValue("55"))  
currency3 = attrNode(currency, string11)  
string11 = valueNode(stringValue("USD"))  
  
title4 = elemNode(title, string9)  
string9 =  
valueNode(stringValue("Foundations..."))
```


Luyện tập

- Làm quen việc sử dụng XML để lưu trữ thông tin (thay cho cơ sở dữ liệu)
- Mở rộng:
 - Tìm hiểu cách khai báo XML Schema
 - Tìm hiểu về XQuery

SO SÁNH GIỮA CÁCH SỬ DỤNG LƯU TRỮ BẰNG CSDL VÀ BẰNG XML

So sánh

- Sử dụng CSDL
- Ưu điểm
- Khuyết điểm
- Trường hợp nên sử dụng



Thảo luận
trực tiếp
tại lớp