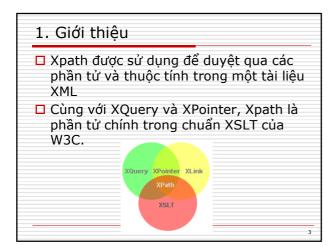
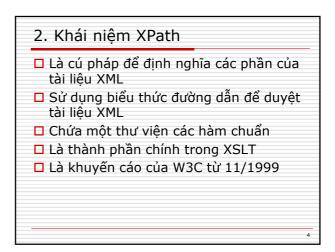
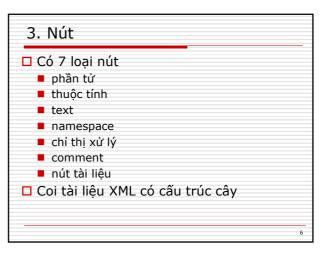
# XPath Nguyễn Hồng Phương Email: phuong.nguyenhong@hust.edu.vn Site: http://is.hut.edu.vn/~phuongnh Bộ môn Hệ thống thông tin Viện Công nghệ thông tin và Truyền thông Đại học Bách Khoa Hà Nội

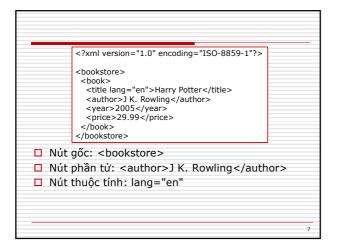
## Nội dung 1. Giới thiệu 2. Khái niệm XPath 3. Nút 4. Cú pháp 5. Một số ví dụ





2. Khái niệm Xpath (tiếp)
 Biểu thức đường dẫn
 Sử dụng biểu thức đường dẫn để chọn lựa các nút/tập nút trong tài liệu XML
 Khá giống với hệ thống file máy tính truyền thống
 Hàm chuẩn
 Trên 100 hàm chuẩn được tích hợp sẵn: các hàm giá trị chuỗi, giá trị số, so sánh date time, Boolean,...





☐ Giá trị nguyên tố: nút không có nút cha và nút con

■ Ví dụ: J K. Rowling, "en"

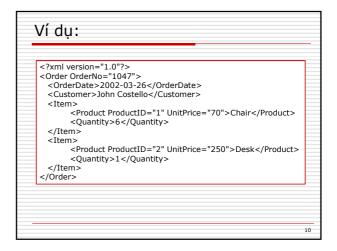
☐ Item: là giá trị nguyên tố hoặc nút

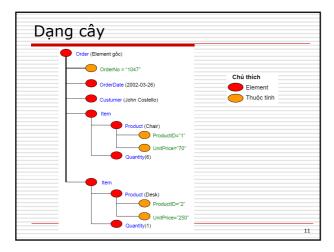
☐ Quan hệ giữa các nút:

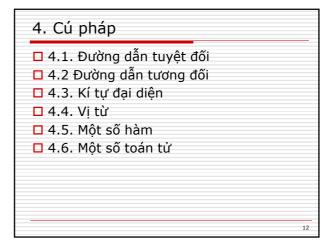
■ Cha: mỗi phần tử và thuộc tính đều có nút cha (trừ nút gốc)

■ Con: mỗi nút có thể có 0, 1 hoặc nhiều nút con

Nút anh em: có cùng nút cha
 Tổ tiên: nút cha, nút cha của cha,...
 Hậu duệ: Nút con, nút con của con,...







### 4.1. Đường dẫn tuyệt đối

- □ Nếu đường dẫn XPath bắt đầu bởi dấu / thì có nghĩa đây là một đường dẫn tuyệt đối bắt đầu từ phần tử gốc.
- ☐ Ví du: muốn chon nút Order ta viết như sau
  - Cú pháp đầy đủ: /child::Order
  - Cú pháp tắt: /Order
- ☐ Đi ra nhánh con Custumer bằng XPath như
  - Cú pháp đầy đủ:

/child::Order/child::Custumer

Cú pháp tắt: /Order/Custumer

13

- Muốn đi đến thuộc tính của nút, cần phải chỉ rõ từ khóa Attribute trong cú pháp đầy đủ hoặc @ trong cú pháp tắt.
- ☐ Để lấy thuộc tính OrderNo của nút Order ta dùng cú pháp Xpath:
  - Cú pháp đầy đủ:

/child::Order/Attribute::OrderNo

Cú pháp tắt: /Order/@OrderNo

### 4.2. Đường dẫn tương đối

- ☐ Khi muốn trích một phần tử nào đó mà chỉ biết tên của phần tử này, không biết là phần tử này nằm ở vị trí nào thì có thể dùng đường dẫn tương đối.
- □ Dùng dấu // để chỉ cho trình phân tích biết đây là đường dẫn tương đối.
- ☐ Ví du: để trích các phần tử có tên là Product
  - Cú pháp đầy đủ: //child::Product
  - Cú pháp tắt: //Product

4.3. Ký tự đại diện

- □ Bất kỳ nút nào: ký tự đại diện \*
- dể lấy tất cả các phần tử con của phần tử Order, ta viết như sau:
  - ☐ Cú pháp đầy đủ: /child::Order/child::\*
  - ☐ Cú pháp tắt: /Order/\*
- □ Bất kỳ nút thuộc tính nào: @\*
- □ Bất kỳ nút của bất kỳ kiểu nào: node()
- □ Ví du

/bookstore/\* Chọn tất cả nút con của phần tử bookstore Chọn tất cả phần tử trong tài liệu //title[@\*] Chon tất cả phần tử title có bất kỳ thuộc tính

4.4. Vị từ

- ☐ Được sử dụng để tìm một nút xác định hoặc một nút chứa một giá trị xác định
- □ Được đặt trong cặp ngoặc vuông []
- □ Ví du:
  - chọn phần tử book đầu tiên là con của phần tử bookstore: /bookstore/book[1]
  - chọn phần tử book cuối cùng là con của phần tử bookstore: /bookstore/book[last()]
  - chọn 2 phần tử book đầu tiên là con của phần tử bookstore:

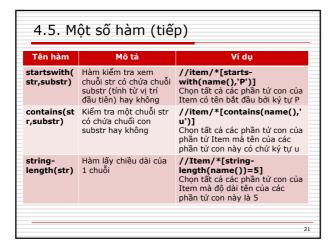
/bookstore/book[position()<3]

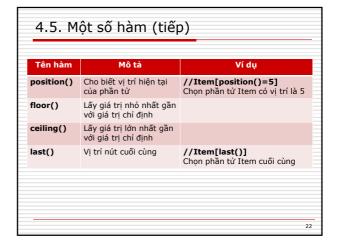
Vị từ: ví dụ

- lấy tất cả phần tử title có thuộc tính tên là lang: //title[@lang]
- lấy tất cả phần tử title có thuộc tính lang nhận giá trị 'eng': //title[@lang='eng']
- lấy ra phần tử book là con của bookstore có phần tử price với giá trị lớn hơn 35: /bookstore/book[price>35.00]
- /bookstore/book[price>35.00]/title



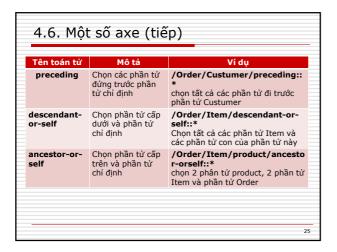
4.5. N	1ột số hàm	1
Tên hàm	Mô tả	Ví dụ
count()	Hàm lấy tổng số nút con của một phần tử nào đó	//Item[count(*)=2] Chọn tất cả các phần tử Item có số phần tử con là 2
name()	Lấy tên của phần tử	/Order/*[name()='Item'] Chọn tất cả các phần tử con của Order có tên là Item
not()	Hàm phủ định	//Item/*[not(@*)] Chọn tất cả các phần tử con của Item không chứa thuộc tính nào
	Hàm bỏ khoảng trắng	//Item/*[normalizespace(@ ProductID)='abc'] Chọn tất cả các phần tử con của Item có thuộc tính ProductID=abc (không phân biệt khoảng trắng)
		20

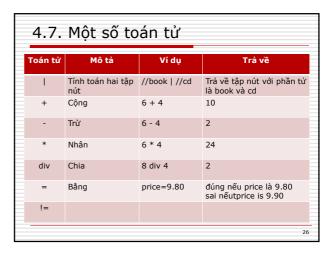


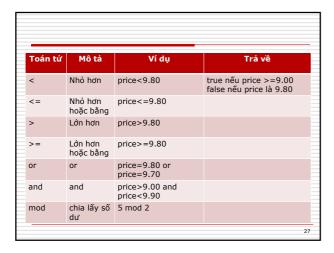


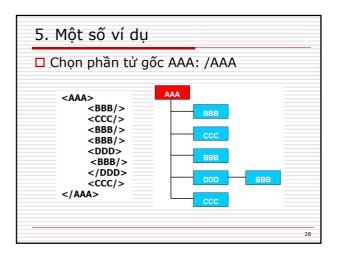
4.6. Một số axe				
Tên	Mô tả	Ví dụ		
ı	Toán tử hoặc dùng để chọn ra một lần nhiều phần tử có điều kiện khác nhau	//Item/*[starts-with(name(),'U')   startswith(name(),'Q') ] Chọn tất cả các phần tử là con của Item có có tên bắt đầu bởi ký tự P hoặc Q		
descendant	Chọn phần tử con của phần tử chỉ định	/Order/Item/Product/descendan t::* Chọn tất cả các phần tử là con của /Order/Item/Product		
ancestor	Chọn phần tử cấp trên	/Order/Item/Product/ancestor:: * chọn 2 phần tử Item và phần tử Order		
		23		

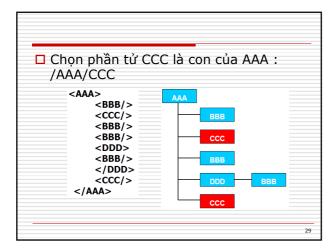
Tên	Mô tả	Ví dụ
following- sibling	Chọn phần tử cùng cấp kế tiếp	/Order/OrderDate/following- sibling::* chọn các phần tử Custumer và hai phần tử Item theo sau và cùng cấp với phần tử OrderDate
preceding- sibling	Chọn phần tử cùng cấp trước đó	/Order/Custumer/preceding- sibling::* chọn phần tử OrderDate
following	Chọn phần tử theo sau phần tử chỉ định	/Order/OrderDate/following::* chọn phần tử Custumer và 2 phần tử Item và các phần tử con của Item

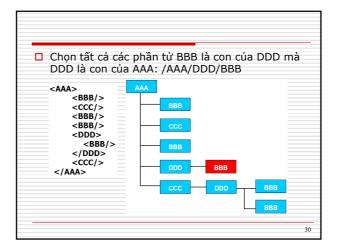


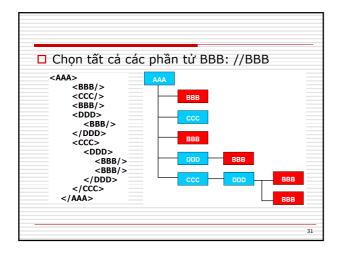


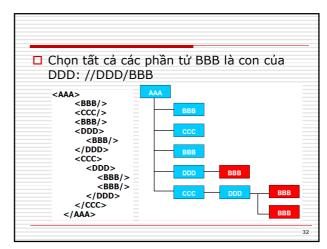


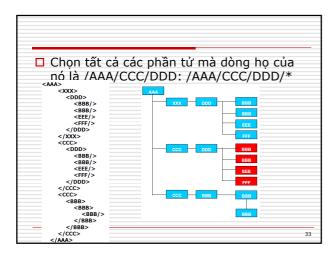


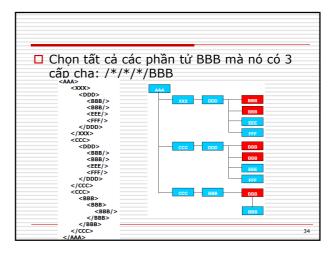


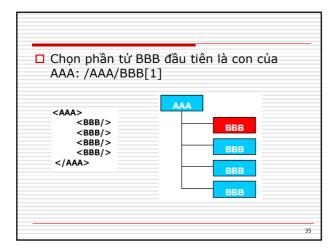


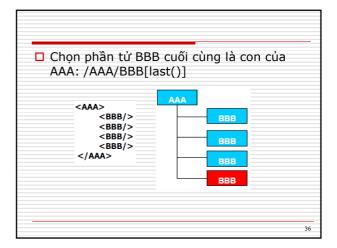


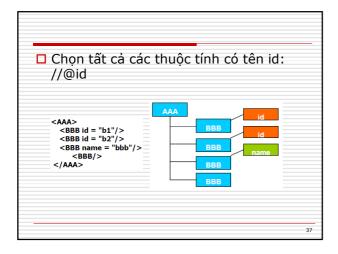


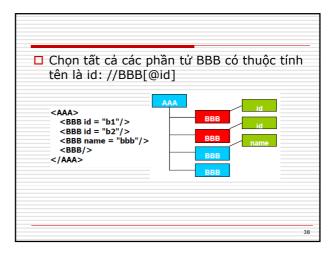


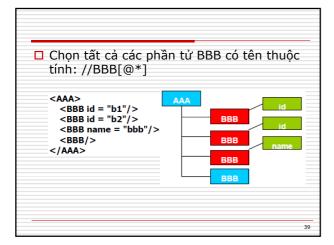






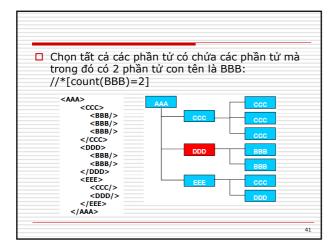


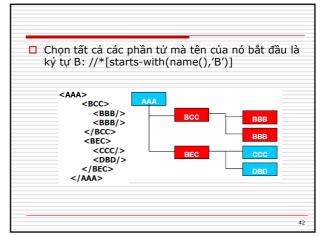


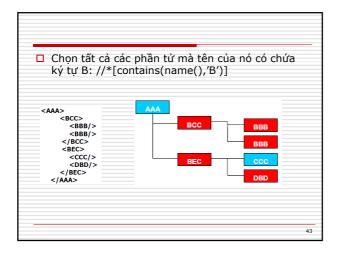


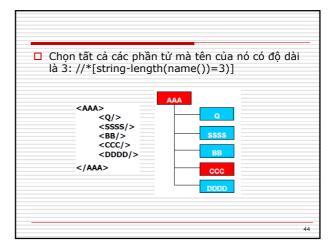
Chọn tất cả các phần tử BBB có tên thuộc tính là bbb, không phân biệt khoảng trắng:
//BBB[normalize-space(@name)='bbb']

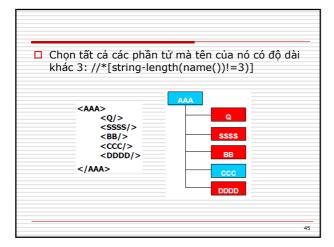
<AAA>
<BBB id = "b1"/>
<BBB name="bbb"/>
<BBB name="bbb"/>
<ABB name="bbb"/>
<ABB name="bbb"/>
<BBB name="bbb"/>
<ABB name="bbb"/>
<BBB name="bbb"/>
<ABB name="b

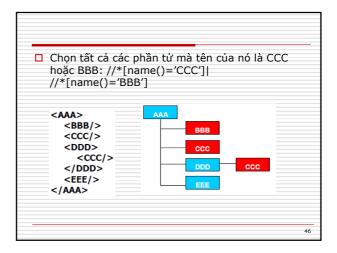


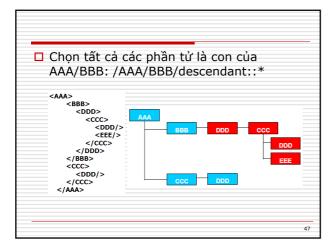


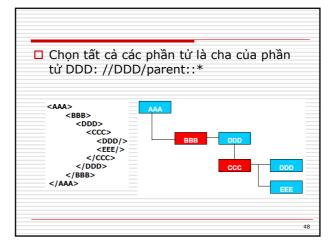


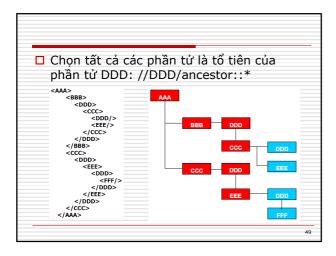


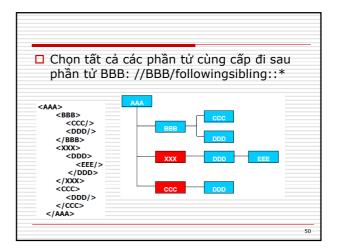


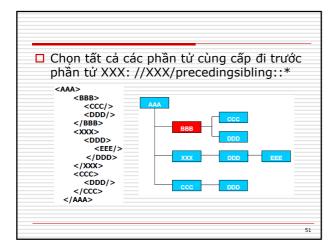


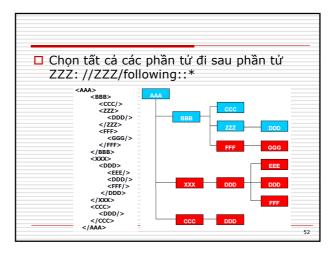


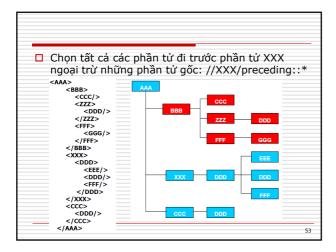


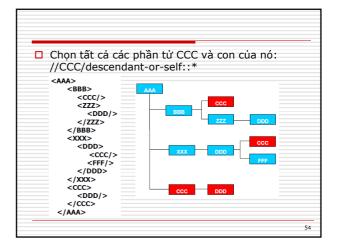


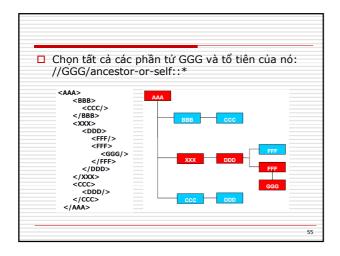


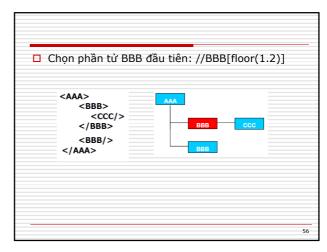


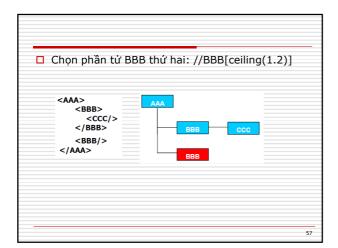


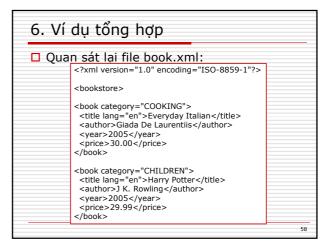


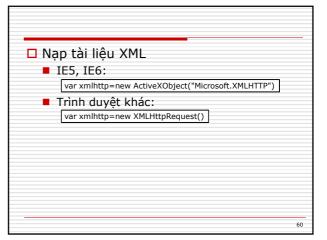


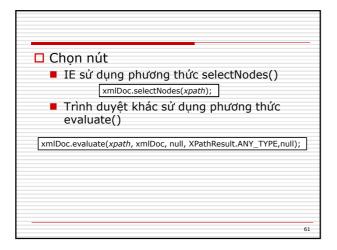












```
Chọn tất cả các title

/bookstore/book/title

/!DOCTYPE html>
<html>
<html>
<body>
<script>
function loadXMLDoc(dname){
if (window.XMLHttpRequest){
    xhttp=new XMLHttpRequest();
    }else{
    xhttp=new ActiveXObject("Microsoft.XMLHTTP");
    }
    xhttp.send("");
    return xhttp.responseXML;
}

xml=loadXMLDoc("books.xml");
    path="/bookstore/book/title"
```

```
Chọn tất cả các title (tiếp)

while (result)
{
    document.write(result.childNodes[0].nodeValue);
    document.write("<br>");
    result=nodes.iterateNext();
    }
}
</script>
</body>
</html>

L Kết quả:

Everyday Italian
Harry Potter
XQuery Kick Start
Learning XML
```

```
Chọn title của book đầu tiên

/bookstore/book[1]/title

/IDOCTYPE html>
<html>
<html>
<body>
<script>
function loadXMLDoc(dname){
if (window.XMLHttpRequest) {
    xhttp=new XMLHttpRequest();
    }
else {
    xhttp=new ActiveXObject("Microsoft.XMLHTTP");
    }
    xhttp.open("GET",dname,false);
    xhttp.send("");
    return xhttp.responseXML;
}
```

```
Chọn title của book đầu tiên (tiếp)

Chú ý: IE5 trở về sau cho rằng, phần tử đầu tiên bắt đầu từ 0

Giải quyết vấn đề này:

[xml.setProperty("SelectionLanguage","XPath");
xml.selectNodes("/bookstore/book[1]/title");

// code for IE
if (window.ActiveXObject){
xml.setProperty("SelectionLanguage","XPath");
var nodes=xml.selectNodes(path);
for (i=0;i<nodes.length;i++)
{
document.write(nodes[i].childNodes[0].nodeValue);
document.write("<br/>br>");
}
// code for Mozilla, Firefox, Opera, etc.
.....
```

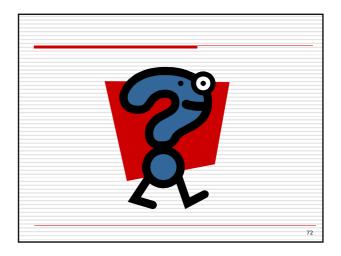
```
Chọn text của tất cả nút price

// bookstore/book/price/text()

| Jooctype html>
| html>
| html>
| sbody>
| script>
| function loadXMLDoc(dname) |
| if (window.XMLHttpRequest) |
| knttp=new XMLHttpRequest(); |
| else |
| xhttp=new ActiveXObject("Microsoft.XMLHTTP"); |
| xhttp.open("GET",dname,false); |
| xhttp.send(""); |
| return xhttp.responseXML; | 69
```

```
Chọn text của tất cả nút price (tiếp)

| xml=loadXMLDoc("books.xml");
| path="/bookstore/book/price/text()"
| // code for IE
| if (window.ActiveXObject)
| {
| var nodes=xml.selectNodes(path);
| for (i=0;i<nodes.length;i++)
| {
| document.write(nodes[i].nodeValue);
| document.write("<br>");
| }
| // code for Mozilla, Firefox, Opera, etc.
```



## Lời hay ý đẹp

"Một vết chém của con dao có thể chữa khỏi, nhưng một vết chém của lưỡi thì khó lòng chữa được" Ngạn ngữ Tây Ban Nha

73