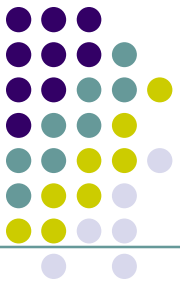


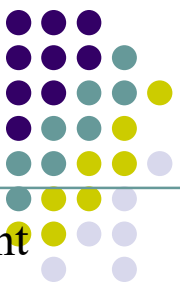


Université Gaston Berger de Saint-Louis
Année académique 2016-2017



XPATH

Syntaxe de définition des parties d'un document XML



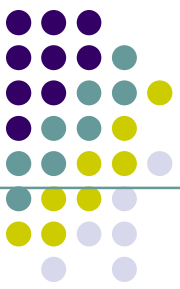
Références

1. XML Programming Success in a Day: Beginner's Guide to Fast, Easy, and Efficient Learning of XML Programming. Sam Key, 2015.
2. XML Programming: The Ultimate Guide to Fast, Easy, and Efficient Learning of XML Programming. Christopher Right, 2015
3. Beginning XML. Joe Fawcett and Danny Ayers, 2012.
4. XML, Cours et exercices. Modélisation, Schémas et DTD, design patterns, XSLT, DOM, Relax NG, XPath, SOAP, XQuery, XSL-FO, SVG, eXist. *Alexandre Brilliant*, Édition : Eyrolles 2^eédition, 2010
5. Schémas XML, Jean-Jacques Thomasson, Edition Eyrolles 2002
6. World Wide Web Consortium (W3C) : w3.org 00336241511074
7. Tutoriel XSLT de World Wide Web School : w3schools.com/xsl
8. xsl.developpez.com
9. Youtube.com



Sommaire

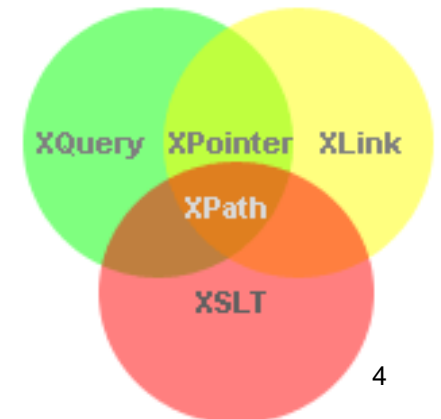
- Introduction
- Les nœuds
- La syntaxe
- Les axes
- Les opérateurs
- Exemple pratique
- Exercice



Introduction

Qu'est-ce que XPATH?

- XPath signifie XML Path Language
- XPath est une syntaxe pour définir les parties d'un document XML
- Il utilise des expressions de chemin pour naviguer dans un doc XML
- C'est un langage de sélection, pas un langage de requête comme SQL
- Il ne permet pas de transformer le résultat final, tâche dédiée à des outils comme XSLT ou XQuery.
- Il contient une bibliothèque de fonctions standards
- C'est un élément majeur dans XSLT

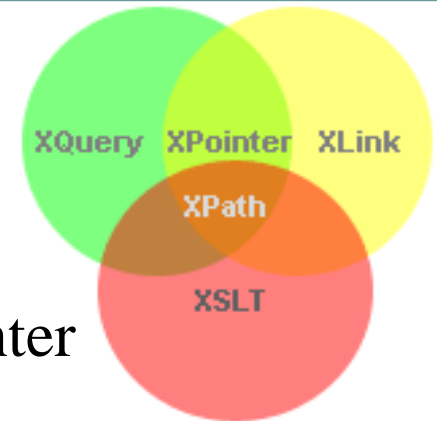




Introduction

Qu'est-ce que XPATH?

- C'est un élément majeur dans XSLT
- XPath a été conçu pour être utilisé par XSLT, XPointer et d'autres outils d'analyse de XML.
- Aujourd'hui, les expressions XPath peuvent également être utilisées en JavaScript, Java, XML Schema, PHP, Python, C et C ++, et beaucoup d'autres langues.
- Il est une recommandation du W3C depuis novembre 1999.
- Sa dernière version XPath 3.0 est devenu une recommandation du W3C depuis le 8 avril 2014.





Introduction

Expressions de chemin de XPath

- XPath utilise des expressions de chemin pour sélectionner des nœuds ou ensembles de nœuds dans un document XML.
- Ces expressions de chemin ressemblent beaucoup à celles qu'on utilise lorsqu'on travaille dans un système de fichier traditionnel d'un ordinateur.

Fonctions XPath standard

- XPath comprend plus de 200 fonctions intégrées.
- Il y a des fonctions de chaînes de caractères, de valeurs numériques, de comparaison de date et heure, de manipulation de nœuds et de QName, de manipulation de séquence, de valeurs booléennes, etc.



Les nœuds XPath

Terminologie:

Nœud:

- Dans XPath, il y a sept types de nœuds: les élément, les attributs, du texte, les espaces de noms, les instructions de traitement, les commentaires et les nœuds du document.
- Les documents XML sont traités comme des arbres de nœuds.
- L'élément le plus haut de l'arbre est appelé l'élément racine.

Exemple:

```
<?xml version="1.0" encoding="UTF-8"?>
<librairie>
  <livre categorie="Histoire">
    <titre>Long walk to freedom</titre>
    <auteur>Nelson Mandela</auteur>
    <annee>1995</annee>
  </livre>
</librairie>
```

Diagram illustrating the XML structure and node types:

- Nœud élément racine**: Points to the root element `<librairie>`.
- Nœud attribut**: Points to the attribute `categorie="Histoire"` of the `<livre>` element.
- Nœud élément**: Points to the element `<annee>1995</annee>`.



Les nœuds XPath

Terminologie:

Valeurs atomique:

Une valeur atomique est un nœud sans enfant ou parent.

Exemple : 1955, "Histoire"

Item:

Un item est une valeur atomique ou un nœud.



Les nœuds XPath

Relation entre nœuds

Parent: Chaque élément et attribut a un parent.

Dans l'exemple suivant livre est le parent de titre, auteur et annee.

Enfant :

Un nœud élément peut avoir zéro, un ou plusieurs enfants.

Dans l'exemple suivant; titre, auteur et annee sont des enfants de livre:

```
<?xml version="1.0" encoding="UTF-8"?>
<librairie>
  <livre categorie="Histoire">
    <titre>Long walk to freedom</titre>
    <auteur>Nelson Mandela</auteur>
    <annee>1995</annee>
  </livre>
</librairie>
```



Les nœuds XPath

Relation entre nœuds

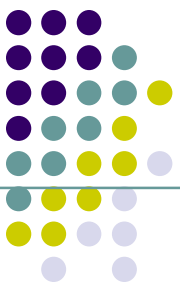
Frères : Ce sont des nœuds ayant le même parent.

Dans l'exemple suivant, titre, auteur et annee sont tous frères.

Ancêtres : Le parent d'un nœud, le parent de parent, etc.

Dans l'exemple suivant, les ancêtres de titre sont livre et librairie:

```
<?xml version="1.0" encoding="UTF-8"?>
<librairie>
  <livre categorie="Histoire">
    <titre>Long walk to freedom</titre>
    <auteur>Nelson Mandela</auteur>
    <annee>1995</annee>
  </livre>
</librairie>
```



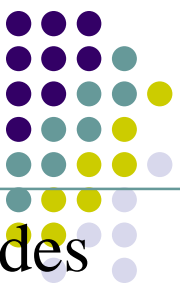
Les nœuds XPath

Relation entre nœuds

Descendant: C'est l'enfant d'un nœud, l'enfant d'un enfant, etc.

Dans l'exemple suivant, les descendants de librairie sont livre, titre, auteur et annee.

```
<?xml version="1.0" encoding="UTF-8"?>
<librairie>
  <livre categorie="Histoire">
    <titre>Long walk to freedom</titre>
    <auteur>Nelson Mandela</auteur>
    <annee>1995</annee>
  </livre>
</librairie>
```



Syntaxes XPath

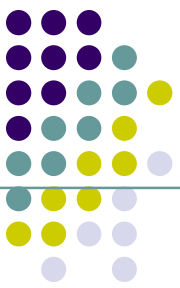
- XPath utilise des expressions de chemin pour sélectionner des nœuds ou ensembles de nœuds dans un document XML.
- Un nœud est choisi en suivant un chemin ou des étapes.

Exemple: Soit le document XML suivant:

```
<?xml version="1.0" encoding="UTF-8"?>
<librairie>

<livre>
  <titre language="fr">Algo et programmation</titre>
  <prix>3000</prix>
</livre>

<livre>
  <titre lang="en">Learning XML</titre>
  <prix>3995</prix>
</livre>
</librairie>
```



Syntaxes XPath

Sélection de nœuds:

XPath utilise des expressions de chemin pour sélectionner des nœuds dans un document XML. Le nœud est choisi en suivant un chemin ou des étapes. Voici les expressions de chemin les plus usuelles :

Expression	Description
<i>nodename</i>	Sélectionne tous les nœuds ayant le nom " <i>nodename</i> "
<i>/</i>	Sélectionne à partir du nœud racine
<i>//</i>	Sélectionne, à partir du nœud courant, les nœuds correspondant à la sélection, peu importe où ils sont
<i>.</i>	Sélectionne le nœud courant
<i>..</i>	Sélectionne les parents du nœud courant
<i>@</i>	Sélectionne les attributs



Syntaxes XPath

Sélection de nœuds:

Exemples:

Expression de chemin	Résultat
librairie	Sélectionne tous les nœuds de "librairie"
/librairie	Sélectionne l'élément racine librairie. Remarque : Si le chemin commence par une barre oblique (/), il représente toujours un chemin absolu à un élément!
librairie/livre	Sélectionne tous les éléments livre enfants de librairie
//livre	Sélectionne tous les éléments livre, peu importe où ils se trouvent dans le document
librairie//livre	Sélectionne tous les éléments livre descendant de l'élément librairie, peu importe où ils sont sous l'élément librairie
//@langue	Sélectionne tous les attributs nommés langue



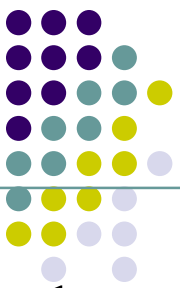
Syntaxes XPath

Prédicats: Un prédicat est utilisé pour trouver un nœud spécifique ou un nœud ayant une valeur spécifique. Les prédicats sont toujours incorporés entre crochets.

Exemple:

Expression de chemin	Résultat
<code>/librairie/livre[1]</code>	Sélectionne le premier élément livre de l'élément librairie
<code>/librairie/livre[last()]</code>	Sélectionne le dernier élément livre de l'élément librairie
<code>/librairie/livre[last()-1]</code>	Sélectionne l'avant dernier élément livre de l'élément librairie
<code>/librairie/livre[position()<3]</code>	Sélectionne les deux premiers éléments Livre de librairie
<code>//titre[@lang]</code>	Sélectionne tous les éléments titre ayant un attribut lang
<code>//titre[@lang='en']</code>	Sélectionne les élt titre ayant un attribut "lang" avec "en" comme valeur
<code>/librairie/livre[prix>35.00]</code>	Sélectionne les élt livre de l'élt librairie dont le prix > 35.00
<code>/librairie/livre[prix>35.00]/titre</code>	Sélectionne tous les éléments titre des éléments livre de l'élément librairie ayant un élément prix avec une valeur supérieure à 35.00

Note: Dans IE 5,6,7,8,9 le 1e nœud est [0], mais pour W3C, c'est [1]. Pour résoudre ce problème ds IE, mettre SelectionLanguage ds XPath:
Ds JavaScript: xml.setProperty("SelectionLanguage","XPath");



Syntaxes XPath

Sélection de nœuds inconnus

Les méta-caractères peuvent être utilisés pour sélectionner des nœuds XML inconnus.

Méta-caractère	Description
*	Correspond à tout nœud d'élément
@*	Correspond à tout nœud d'attribut
node()	Correspond à tout nœud de toute nature

Exemple:

Expression	Résultat
<code>/librairie/*</code>	Sélectionne tous les nœuds élément enfant de librairie
<code>//*</code>	Sélectionne tous les éléments du document
<code>//titre[@*]</code>	Sélectionne tous les éléments titre ayant au moins un attribut de toute nature



Syntaxes XPath

Sélection de plusieurs chemins

En utilisant le opérateur | dans une expression XPath, on peut sélectionner plusieurs chemins.

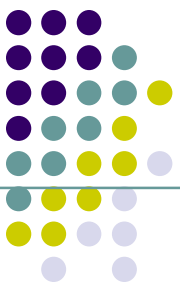
Expression de chemin	Résultat
//livre/titre //livre/prix	Sélectionne tous les éléments titre ET tous les éléments prix de tous les éléments livre
//titre //prix	Sélectionne tous les éléments titre ET les éléments prix du document
//librairie/livre/titre //prix	Sélectionne tous les éléments titre de l'élément livre de l'élément librairie ET tous les éléments prix du document



Axes XPath

Un axe définit un ensemble de nœuds par rapport au nœud courant.

Nom axe	Résultat
ancestor	Sélectionne les ancêtres (parents, grands-parents, etc.) du nœud courant
ancestor-or-self	Sélectionne tous les ancêtres (parents, grands-parents, etc.) du nœud courant et le nœud courant lui-même
attribute	Sélectionne tous les attributs du nœud courant
child	Sélectionne tous les enfants du nœud courant
descendant	Sélectionne les descendants (enfants, petits-enfants, etc.) du nœud courant
descendant-or-self	Sélectionne tous les descendants (enfants, petits-enfants, etc.) du nœud courant et le nœud courant lui-même
following	Sélectionne tout ds le document après la balise de fermeture du nœud courant
following-sibling	Sélectionne tous les frères et sœurs après le nœud courant
namespace	Sélectionne tous les nœuds d'espace de noms du nœud courant
parent	Sélectionne le parent du nœud courant
preceding	Sélectionne tous les nœuds apparaissant avant le nœud actuel dans le document, à l'exception des ancêtres, des nœuds attribut et de namespaces
preceding-sibling	Sélectionne tous les frères et sœurs avant le nœud courant
self	Sélectionne le nœud courant



Axes XPath

Chemin d'accès

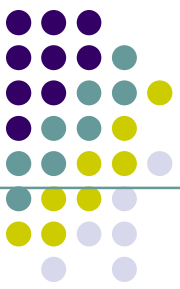
- Un chemin d'accès peut être absolu ou relatif.
- Un chemin d'accès absolu commence par une barre oblique (/) contrairement à un chemin d'accès relatif.
- Dans les deux cas, le chemin d'accès consiste en une ou plusieurs étapes, chacune séparée par une barre oblique:

Syntaxes: Chemin d'accès absolu: `/etape/etape/...`
 Chemin d'accès relatif: `etape/etape/...`

Chaque **étape** est évaluée par rapport à l'ensemble des nœuds de l'élément courant. Elle consiste à :

- un axe qui définit la relation d'arbre entre les nœuds choisis et le nœud courant
- un nœud-test qui identifie un nœud avec un axe
- zéro ou plusieurs prédicats (pour affiner l'ensemble de nœuds sélectionnés)

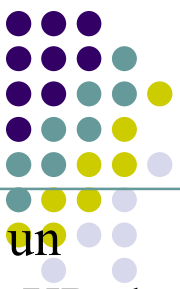
Syntaxe d'une étape: `nomaxe::nœud-test[predicat]`



Axes XPath

Exemple d'étape

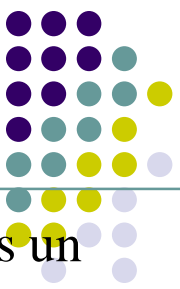
Etape	Résultat
child::livre	Sélectionne tous les nœuds livre qui sont enfants du nœud courant
attribute::langue	Sélectionne l'attribut langue du nœud courant
child::*	Sélectionne tous les éléments enfants du nœud courant
attribute::*	Sélectionne tous les attributs du nœud courant
child::text()	Sélectionne tous les nœuds texte enfants du nœud courant
child::node()	Sélectionne tous les enfants du nœud courant
descendant::livre	Sélectionne tous les livre descendant du nœud courant
ancestor::livre	Sélectionne tous les livre ancêtres du nœud courant
ancestor-or-self::livre	Sélectionne tous les livre ancêtres du nœud courant - ainsi que l'élément courant s'il est un nœud livre
child::* / child::prix	Sélectionne tous les prix petits-fils du nœud courant



Les opérateurs XPath

Un opérateur XPath renvoie un ensemble de nœuds, une chaîne, un booléen ou un nombre. Voici une liste des opérateurs pouvant être utilisés dans une expression XPath:

Opérateur	Description	Exemple
	Calcule deux ensembles de nœuds	//livre //cd
+	Addition	6 + 4
-	Soustraction	6 - 4
*	Multiplication	6 * 4
div	Division	8 div 4
=	Egalité	prix=9.80
!=	Différence	prix!=9.80
<	Inférieur strict	prix<9.80
<=	Inférieur ou égal	prix<=9.80
>	Supérieur strict	prix>9.80
>=	Supérieur ou égal	prix>=9.80
or	Ou	prix=9.80 or prix=9.70
and	Et	prix>9.00 and prix<9.90
mod	Modulo (Reste de la division)	5 mod 2



Exemple pratique

Considérons le code suivant permettant d'appliquer des expressions XPath dans un document book.xml.

```
<!DOCTYPE html>
<html><body>
<script>
function loadXMLDoc(dname){
    if (window.XMLHttpRequest){
        xhttp=new XMLHttpRequest();
    }
    else{
        xhttp=new ActiveXObject("Microsoft.XMLHTTP");
    }
    xhttp.open("GET", dname, false);
    try { xhttp.responseType="msxml-document" } catch(err) { } // Helping IE
    xhttp.send("");
    return xhttp;
}
...
```



Exemple pratique

Considérons le code suivant permettant d'appliquer des expressions XPath dans un document book.xml.

...

```
var x=loadXMLDoc("books.xml");
```

```
var xml=x.responseXML;
```

```
path="/bookstore/book/title";
```

```
// code for IE
```

```
if (window.ActiveXObject || xhttp.responseType=="msxml-document"){
```

```
    xml.setProperty("SelectionLanguage","XPath");
```

```
    nodes=xml.selectNodes(path);
```

```
    for (i=0;i<nodes.length;i++){
```

```
        document.write(nodes[i].childNodes[0].nodeValue);
```

```
        document.write("<br>");
```

```
    }
```

```
}
```

...



Exemple pratique

Considérons le code suivant permettant d'appliquer des expressions XPath dans un document book.xml.

```
...  
// code for Chrome, Firefox, Opera, etc.  
else  
if (document.implementation && document.implementation.createDocument){  
    var nodes=xml.evaluate(path, xml, null, XPathResult.ANY_TYPE, null);  
    var result=nodes.iterateNext();  
  
    while (result){  
        document.write(result.childNodes[0].nodeValue);  
        document.write("<br>");  
        result=nodes.iterateNext();  
    }  
}  
</script>  
</body>  
</html>
```




Exercice

1. Adapter l'exemple ci-dessous à votre document XML
2. Modifier le document obtenu pour que les requêtes proviennent d'un champ de texte d'un formulaire
3. Donner les expressions XPath permettant d'avoir:
 - a. le prénom du premier étudiant
 - b. tous les noms
 - c. les nœuds age dont $\text{age} > 25$
 - d. les nœuds etudiant ayant un $\text{age} > 35$
4. Appliquer chaque expression du 3. à la page obtenue dans 2.



Merci de votre attention