Cours 2

Le Langage XPath

Le langage XPath

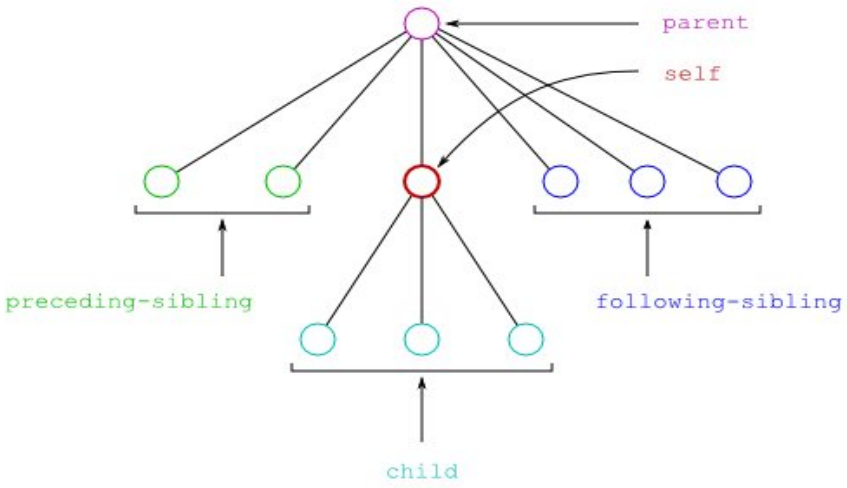
Exploitation de la structure du document

- Axes de parcours
- Test sur la valeur des éléments et des attributs
- Fonctions

Syntaxe générale

- 2 concepts
  - Noeud courant: c'est le noeud de départ qui peut être la racine ou tout autre noeud du document
  - Localisation: c'est le ou les noeuds recherchés
    - 3 éléments
      - Un axe
      - Un filtre
      - Un prédicat optionnel

Les axes



- Syntaxe d'une étape

NOM_DE_L_AXE::FILTRE[prédicat optionnel]

Localisation

- Axe: *la direction dans laquelle on se dirige à partir du noeud courant*

* self : *le noeud courant lui-même* ;
* child : *les enfants du noeud courant* ;
* descendant, descendant-or-self : *tous les descendants du noeud courant ;*
* parent : *le père du noeud courant* ;
* ancestor, ancestor-or-self : *les ancêtres du noeud courant* ;
* attribute : *les attributs du noeud courant* ;
* preceding, following : l*es noeuds, précédents ou suivants, du noeud courant, dans l'ordre de lecture du document* ;
* preceding-sibling, following-sibling : l*es frères, précédant ou suivant, le noeud courant* ;
* namespace : *les espaces de noms*.

Les filtres

- L'expression comment() sélectionne tous les noeuds commentaires
- L'expression text() sélectionne tous les noeuds ne contenant que du texte.
- L'expression node() sélectionne tous les noeuds fils
- *nom* = les éléments portant ce *nom* ;

Les prédicats

- Ils prennent la forme de tests que les noeuds sélectionnés devront vérifier.
- Ces tests peuvent impliquer des fonctions ou de nouveaux chemins XPath.

7

Des fonctions

- Ces fonctions peuvent apparaître dans des prédicats ou être utilisées directement dans un évaluateur d'expressions XPath.
- Il y a des fonctions sur les chaînes de caractères et qui vont porter sur les contenus textuels de noeuds :

concat : concaténe les chaînes de caractères passées en paramètres ;

contains, starts-with, ends-with : tests d'appartenance d'une chaîne dans une autre.

8

string-length(string **ch**) : retourne la longueur de la chaîne de caractères **ch**

substring("..",nb1,nb2) : extrait la chaîne de caractères à partir de l'indice nb1 sur une longueur nb2

/> xpath substring("email",1,3)

Object is a string : ema

9

Autres fonctions usuelles

Fonctions qui prennent en argument une requête XPath et vont porter sur des ensembles de noeuds :

count : le nombre de noeuds dans l'ensemble sélectionné par la requête ; / > xpath count(child::A) Object is a number : 3

name : le nom de l'élément courant. xpath name(self::node())

Object is a string : boiteAuxLettres

Fonctions sans paramètre mais liées au noeud courant :

position : le numéro du noeud courant dans la liste des noeuds considérés ;

last : le nombre de noeuds sélectionnés à l'étape courante.

position() !=last()

10

Chemins XPath

- Un chemin XPATH est constitué par un enchaînement éventuel de plusieurs étapes
- Une étape est constituée d'un axe ou d'un filtre ou d'un prédicat
- Exemple

```
descendant::mel[(position() mod 2) = 0]
```

- Recherche des noeuds descendants de la racine de nom mel, parmi ceux-ci prendre les noeuds en position paire.
- `descendant::mel[(position() mod 2) = 0]/child::expediteur`

- Si le chemin commence par un /, il s'agit d'un chemin absolu, c'est-à-dire prenant son origine à la racine du document et non pas au noeud courant.
- Il est possible de faire une disjonction de requêtes XPath avec l'opérateur | ; on obtient alors l'union des deux ensembles de noeuds correspondants.

Les notations abrégées

self::node() —> .

child::A —> A

parent::node() —> ..

attribute::x —> @x

/descendant-or-self::node()/child::A —> //A

descendant-or-self::node()/child::A —> .//A

child::A[position() = 2] —> A[2]

Exemple de requête

xmllint -shell boiteAuxLettres.xml

/ > xpath /*

Résultat

Object is a Node Set :
Set contains 1 nodes:
1 ELEMENT boiteAuxLettres

13

.//* → tous les noeuds descendants du noeud courant

Exemple: / > xpath //*
1 ELEMENT boiteAuxLettres
2 ELEMENT mel
3 ELEMENT expediteur
4 ELEMENT recepteurs
5 ELEMENT objet
6 ELEMENT contenu
ATTRIBUTE lang
TEXT
content=fr
.....
33 ELEMENT signature
ATTRIBUTE email
TEXT
content=duraton@iutv.univ-paris13.fr

@nom →attribute::*nom*

/ > xpath //@email
Object is a Node Set :
Set contains 4 nodes:
1 ATTRIBUTE email
TEXT
content=duraton@iutv.univ-paris13.fr
2 ATTRIBUTE email
TEXT
content=trucmuch@iutv.univ-paris13.fr
3 ATTRIBUTE email
TEXT
content=duraton@iutv.univ-paris13.fr
4 ATTRIBUTE email
TEXT
content=duraton@iutv.univ-paris13.fr

14

Tests de chemin XPath

- une solution parmi d'autres

xmllint --shell *fichier.xml*

- Exemple

xmllint --shell boiteAuxLettres.xml

Le « prompt » est />

/ > xpath descendant::mel[(position() mod 2) = 0]/expediteur/text()
/>exit (pour sortir)

Récapitulatif des opérateurs XPath

Opérateur	Action	Syntaxe	Exemples
,	Concaténation de listes	E1,E2	1,'Two',3.14,true()
for	Itération	for $i in E1 return E2	for $i in 1 to 5 return $i * $i
some	Quantification existentielle	some $i in E1 satisfies E2	
every	Quantification universelle	every $i in E1 satisfies E2	
if	Test	if (E1) then E2 else E3	if ($x > 0) then $x else 0
/	Enchaînement	E1/E2	
[ ]	Prédicat	E1[E2]	chapter[count(section) > 1]
and or not	Opérations logiques	E1 or E2	
to	Intervalle	E1 to E2	1 to 5
eq ne lt le gt ge	Comparaisons de valeurs atomiques	E1 eq E2	$x lt $y
= != < <= > >=	Comparaisons générales	E1 = E2	$x < $y
<< is >>	Comparaisons de nœuds	E1 is E2	
+ * - div idiv	Opérations arithmétiques	E1 + E2	$price * $qty
| union intersection except	Opérations sur les listes de nœuds	E1 | E2	/|*
instance of cast as castable as treat as	Changements de type	E1 instance of type	$x instance of xsd:string