

# Les Contextes XPath

par [Mathieu Lemoine](#)

Date de publication :

Dernière mise à jour : 09/10/2006

Une notion généralement peu connue, et pourtant essentielle, du langage XPath est la notion de contexte. Je vais essayer de l'expliquer dans ce tutoriel.

- I - Qu'est-ce qu'un contexte ?
- II - Les opérations possibles sur les contextes
  - II-A - Le ré-axage ou déplacement
  - II-B - La précision ou restriction
  - II-C - L'alternative ou union
  - II-D - L'isolement ou regroupement
- III - Précisions sur l'isolement
  - III-A - Elle n'a aucun effet si elle est la dernière effectuée.
  - III-B - Aucun effet si il n'y a pas une alternative avant ou une précision sur la position après.
  - III-C - Doit être en position de générer un contexte premier.
- IV - Divers
  - IV-A - Remerciements
  - IV-B - Téléchargement

## I - Qu'est-ce qu'un contexte ?

Un chemin XPath complet est défini par la succession de chemin XPath atomiques présentés sous cette forme :

```
/axe::test[prédicat]
```

- L'axe indique *où* chercher dans le document
- Le *test* indique *quoi* chercher dans l'axe
- Le *prédicat* pose des *conditions* sur les éléments trouvés

Chaque chemin XPath atomique définit un contexte, c'est à dire un ensemble de caractéristiques que les noeuds sélectionnés par le chemin auront forcément. De même, pour un document XML donné, on peut associer chaque contexte à un ensemble de noeuds... Le mot contexte est souvent utilisé de manière ambiguë, pour décrire et l'ensemble de balise, et l'ensemble de caractéristiques... Afin de bien faire la séparation, nous parleront de contexte abstrait dans le cas général, et de contexte concret quand c'est appliqué à un document.

- Le contexte concret regroupe les balises sélectionnées (niveau 0)
- Le contexte abstrait décrit les balises sélectionnables (niveau 1)

Ainsi, nous parlerons par exemple d'un contexte (abstrait) ciblant un noeud qui est un *attribut* [axe], dont le nom est *href* [test] et dont *la longueur est supérieure à 50 caractères* [prédicat].

Ceci correspondrait au chemin XPath atomique :

```
/attribute::href[string-length(.) > 50]
```

La succession des différents chemins XPath atomiques pour construire le chemin final, construit en même temps une succession de contexte représentant le chemin parcouru peu à peu, chacun étant relié au précédent par une suite d'opérations que l'on peut définir précisément.

Le premier contexte, correspondant au premier chemin XPath atomique (qui est basé sur la racine, ou sur l'ensemble de tous les noeuds sélectionnables), échappe à cette règle. Pour le distinguer des autres, on le qualifiera ainsi de contexte **premier**.

## II - Les opérations possibles sur les contextes

Ces Opérations sont au nombre de quatre, chacune correspond à une partie d'un chemin XPath atomique, ou à un élément particulier dans les chemins XPath (tels que les parenthèses ou la barre verticale).

J'ai à chaque fois utilisé deux noms pour les désigner, un imageant l'effet de l'opération pour un contexte abstrait (caractéristiques) et un le faisant pour un contexte concret (ensemble de noeuds).

- 1 Le *ré-axage* ou *déplacement* (on dira aussi éventuellement *extension*)
- 2 La *précision*, ou *restriction*
- 3 L'*alternance*, ou *union*
- 4 L'*isolement*, ou *regroupement*

### II-A - Le ré-axage ou déplacement

Cette opération est représentée par l'axe XPath. Elle permet de changer les axes de portées pour les caractéristiques.

L'effet produit sur les contextes concrets est un déplacement de l'ensemble de sélection.

*Quelques fois, le test peut également représenter un ré-axage, par exemple, `attribute()` ou `comment()`, qui permettent de s'axer sur les attributs d'un noeud ou les commentaires contenus dans ce noeud.*

*Les axes `descendant-or-self` et `ancestor-or-self` sont particuliers car les contextes (concrets) modifiés contiennent les noeuds du contexte concret avant opération.*

*Dans leur cas, l'appellation `extension` semble donc plus appropriée que `déplacement`.*

*L'axe `self` est spécial car il représente l'opération d'identité : Il ne modifie pas le contexte.*

Exemples :

AAA/child::\*

```
<root>
  <AAA>
    <BBB>
      <CCC/>
    </BBB>
  <BBB>
    <CCC/>
  </BBB>
</AAA>
<AAA>
  <BBB/>
</AAA>
<AAA>
  <CCC/>
</AAA>
</root>
```

attribute::\*

```
<root>
  <AAA status="good">
    <BBB>
      <CCC status="bad"/>
    </BBB>
  </AAA>
```

attribute::\*

```

<AAA>
  <BBB status="good" />
</AAA>
<AAA>
  <CCC type="empty" />
</AAA>
</root>

```

BBB/following-sibling::\*

```

<root>
  <AAA>
    <BBB/>
    <DDD>
      <CCC/>
    </DDD>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>
    <CCC/>
    <BBB/>
  </AAA>
  <AAA>
    <CCC/>
  </AAA>
</root>

```

## II-B - La précision ou restriction

Cette opération est représentée par le prédicat. Elle permet d'affiner, de préciser les caractéristiques demandées.

L'effet produit sur les contextes concrets est une restriction des noeuds sélectionnés, un "filtrage".

*Quelques fois, le test peut également représenter une précision, par exemple, la précision directe du nom du noeud.*

Exemples :

AAA/BBB[position()=1]

```

<root>
  <AAA>
    <BBB>
      <CCC/>
    </BBB>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>
    <BBB/>
  </AAA>
  <AAA>
    <CCC/>
  </AAA>
</root>

```

CCC/ancestor::\*[name()='AAA']

```

<root>
  <AAA>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>

```

```
CCC/ancestor::*[name()='AAA']
```

```
<BBB/>
</AAA>
<AAA>
  <CCC/>
</AAA>
</root>
```

```
*[@status='good']
```

```
<root>
  <AAA status="good">
    <BBB>
      <CCC status="bad" />
    </BBB>
  </AAA>
  <AAA>
    <BBB status="good" />
  </AAA>
  <AAA>
    <CCC status="good" />
  </AAA>
</root>
```

## II-C - L'alternative ou union

Cette opération est effectuée par l'utilisation de la barre verticale (|). Elle permet d'indiquer un contexte abstrait satisfaisant alternatif.

Les contextes concrets produits sont l'union des contextes concrets correspondant à chacun des deux contextes abstraits.

Exemple :

```
AAA[position()=1]|CCC
```

```
<root>
  <AAA>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>
    <BBB/>
  </AAA>
  <AAA>
    <CCC/>
    <DDD/>
  </AAA>
</root>
```

## II-D - L'isolement ou regroupement

Cette opération est effectuée par l'utilisation des parenthèses ('(...)'). Elle permet d'isoler les caractéristiques.

Il n'y a pas à proprement parler d'effet produit sur les contextes concrets. En fait, cette opération permet d'indiquer un nouveau contexte **premier** c'est à dire un nouvel ensemble de noeuds sélectionnables.

Exemples :

```
(AAA[position()=1]|BBB)/child:.*
```

```
<root>
  <AAA>
```

(AAA[position()=1]|BBB)/child::\*

```

    <DDD>
      <CCC/>
    </DDD>
  </AAA>
  <AAA>
    <BBB/>
  </AAA>
  <AAA>
    <BBB>
      <CCC/>
      <DDD/>
    </BBB>
  </AAA>
</root>

```

(AAA/BBB)[position()=1]

```

<root>
  <AAA>
    <BBB>
      <CCC/>
    </BBB>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>
    <BBB/>
  </AAA>
  <AAA>
    <CCC/>
  </AAA>
</root>

```

(BBB/following-sibling::\*)

```

<root>
  <AAA>
    <BBB/>
    <DDD>
      <CCC/>
    </DDD>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>
    <CCC/>
    <BBB/>
  </AAA>
  <AAA>
    <CCC/>
  </AAA>
</root>

```

### III - Précisions sur l'isolement

Cette opération a plusieurs particularités :

#### III-A - Elle n'a aucun effet si elle est la dernière effectuée.

*cf : II-D : Exemple 3*

#### III-B - Aucun effet si il n'y a pas une alternative avant ou une précision sur la position après.

```
(CCC/ancestor::*)[name()='AAA']
```

```
<root>
  <AAA>
    <BBB>
      <CCC/>
    </BBB>
  </AAA>
  <AAA>
    <BBB/>
  </AAA>
  <AAA>
    <CCC/>
  </AAA>
</root>
```

#### III-C - Doit être en position de générer un contexte premier.

Une parenthèse ouvrante **doit** être le premier caractère d'un chemin XPath **ou** n'être précédée **que** par des parenthèses ouvrantes. Sinon générer un contexte premier qui serait soumis lui même à un contexte préalable n'aurait pas de sens...

```
AAA/(descendant-or-self::BBB|ancestor::root)
```

```
INVALIDE !!!
```



## IV - Divers

*Sera complété au fur et à mesure...*

### IV-A - Remerciements

Merci à [GrandFather](#), [Erwy](#) et [Yogui](#) pour m'avoir conseiller pendant la rédaction de cet article.

### IV-B - Téléchargement

- [Version PDF](#)
- [Version HTML zippée](#)