

## Le langage XPath

### Objectif

L'objectif de ce TP est de mettre en pratique les constructions proposées par le langage XPath. Nous utilisons pour cela un outil en ligne de commande qui extrait les fragments d'un document qui vérifient un chemin de localisation XPath. **Attention**, une expression XPath sélectionne un ensemble de noeuds de l'arbre du document, cet outils affiche la représentation XML de tout les **sous-arbres** dont ces noeuds sont les racines.

### Pré requis

Parcourir et garder la recommandation XPath (<http://www.w3.org/TR/xpath>) ouverte.

Outils utilisés pour l'application des chemins XPath :

- Parser xml : Xerces (<http://xml.apache.org/xerces2-j/index.html>)
- Processeur XSLT : Xalan (<http://xml.apache.org/xalan-j/index.html>)

La commande pour extraire les fragments correspondants à une expression est la suivante :

`/usr/local/bin/XPath <XMLFileName> ' <XPathExpression> '` (Regarder le script !)

A noter : sous Unix, l'expression XPath doit être mise entre ' ' (ou entre " ") pour éviter une interprétation de la ligne de commande.

A partir du document décrivant une liste de livres du TP sur l'édition de documents XML, proposer une ou plusieurs expressions XPath qui répondent aux questions suivantes (certaines questions n'ont pas de réponses dans le document, dans ce cas vous ajouterez des informations dans celui-ci) :

### 1 Expressions XPath

1. En utilisant unique les axes child, descendant, et attribute.
  - Le document complet, les livres, les catégories, les catégories des livres
2. En utilisant les positions dans les prédicats.
  - Le titre du premier livre, le titre du dernier livre, les catégories du second et de l'avant-dernier livre
3. En utilisant des expressions de chemin dans les prédicats.
  - Les catégories ayant une note
  - Les livres ayant au moins deux auteurs.
4. Le code de la catégorie " Science "
5. Le titre des livres de la catégorie " C "
6. Le titre des livres dont la quantité est égale à 0
7. Les livres dont le titre contient le mot " Animal "
8. Les livres dont le nom de l'auteur a six lettres (y-a-t-il un pb ?)
9. Les livres dont l'auteur est " Bonner " ou la catégorie " F "

- Tous les éléments
- Le texte sans balisage
- Les livres qui sont avant le premier livre de la catégorie X
- Le livre qui suit immédiatement le dernier livre de l'auteur " bonner "
- Afficher les instructions de traitements (de quoi s'agit-il ?)
- Afficher un livre sur deux
- Le (ou les) livre(s) du milieu de la liste
- La description de la catégorie du premier livre

## 2 XPath, Id, Idref et documents multiples

Que font les fonctions `id()` et `document()` ? Que doit contenir le document ? Proposer des exemples de requêtes utilisant ces fonctions.

Proposer un ensemble d'expressions XPath qui permettent de sélectionner des fragments des documents décrivant des programmes de conférences du TP sur l'édition de documents XML.

## 3 Pour aller plus loin

En utilisant, les documents décrivant des conférences que vous avez structurés lors du TP 1, proposer un ensemble de requêtes XPath permettant de sélectionner et d'extraire les informations qui vous semblent importantes dans ce document. Votre balisage est-il selon vous suffisant ?