

séance 1

Environnement XML et révisions Xpath

Parser, transformer, interroger son XML : Xpath - XSLT - Xquery

- Xpath est un langage pour parser/naviguer dans les documents XML :

```
TEI/text/body/lg/l
```

- XSLT est un langage de transformation pour les documents XML

```
<xsl:template match="TEI/text/body/lg/l">  
  <xsl:element name="l">  
    <xsl:value-of select="."/>  
  </xsl:element>  
</xsl:template>
```

- Xquery est un langage pour interroger les documents XML qui forment alors une sorte de base de données.

```
for $x in doc("books.xml")/bookstore/book
where $x/price>30
order by $x/title
return $x/title
```

Rappels Xpath : définition et syntaxe

Définition

« XSLT (extensible Stylesheet Language Transformations) est un langage de programmation fonctionnel utilisé pour spécifier comment un document XML est transformé en un autre document qui peut, mais qui n'est pas nécessairement, un autre document XML. Un processeur XSLT lit un arbre XML en entrée et une feuille de style XSL et produit un arbre résultat en sortie. »

Elliott Rusty Harold, W. Scott Means, Philippe Ensarguet [et al.], *XML en concentré*, Paris, O'Reilly, 2005, p. 519.

La syntaxe Xpath

Une expression de chemin correspond à une séquence d'étapes séparées par l'opérateur « / ». Sans indication particulière la relation se fait d'un élément parent vers un élément enfant.

Pour en savoir plus :

Michael R. Kay, XPath 2.0 programmer's reference, Indianapolis, IN, Wrox Press, 2004, p.215-216.

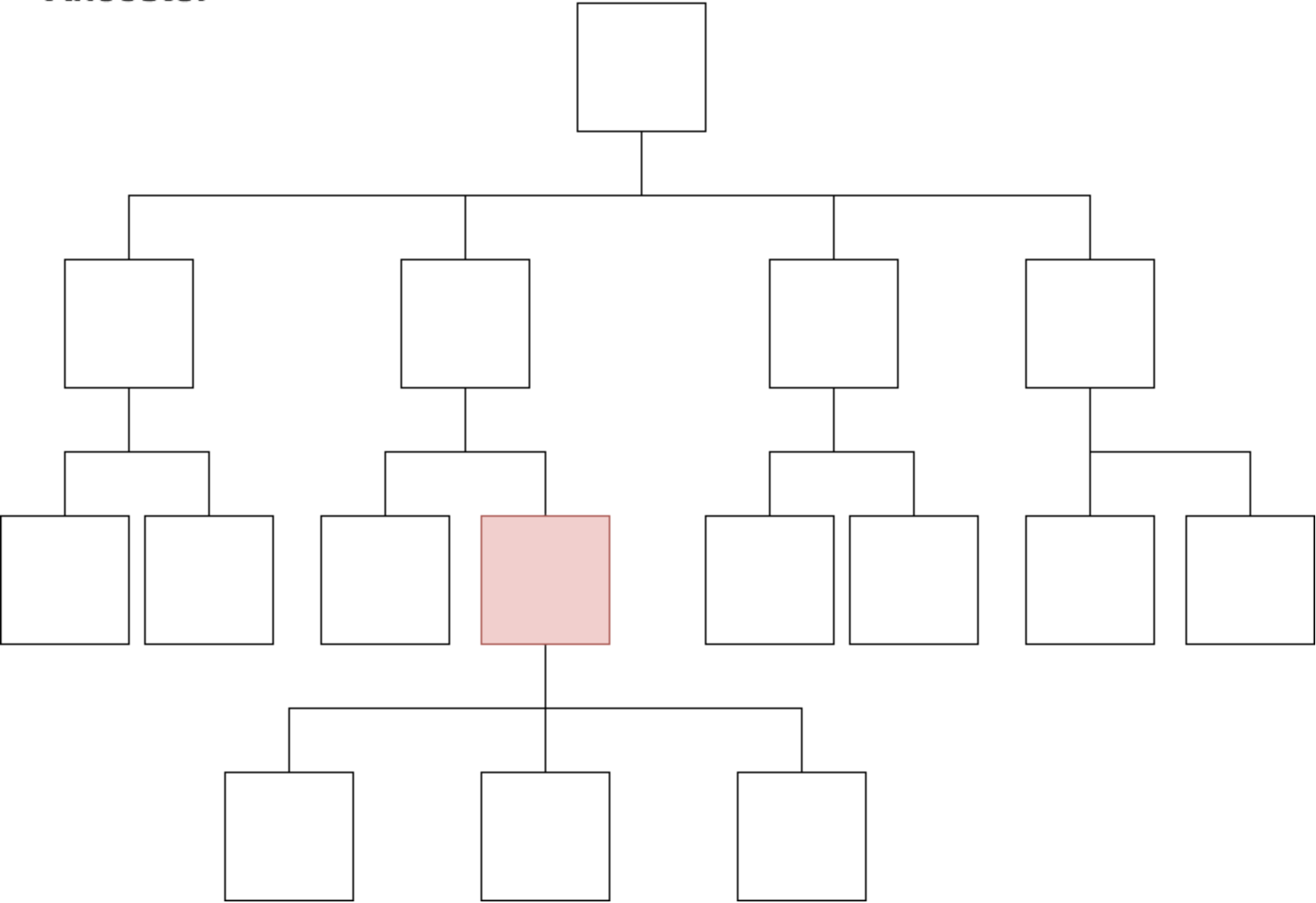
Exercice de repérage (1)

- Mon rêve familial
 - Donner le chemin de la racine TEI vers les vers d'un quatrain
 - Donner le chemin de la racine TEI vers le titre
- Le Misanthrope
 - Donner le chemin de la racine TEI vers les vers d'une scène
 - Donner le chemin de la racine TEI vers la déclaration d'un personnage
- Le Cid
 - Donner le chemin de la racine TEI vers la déclaration d'un nom de personnage dans la liste des personnages de la préface
 - Donner le chemin de la racine TEI vers les vers de la pièce

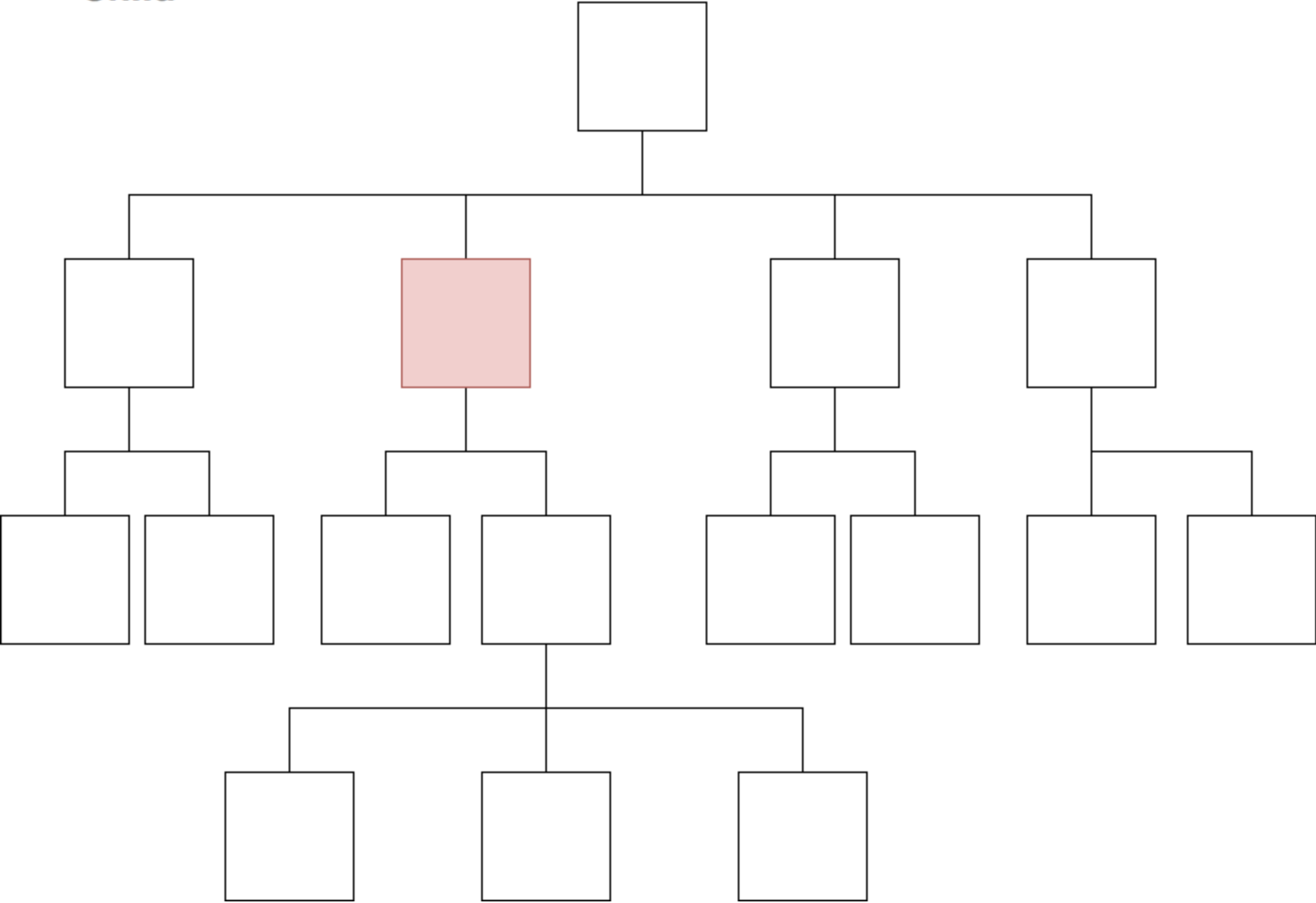
Xpath possède plusieurs axes de relation :

- ancestor : `ancestor::node()`
- ancestor-or-self : `ancestor-or-self::node()`
- parent : `parent::node()`
- descendant : `descendant::node()`
- descendant-or-self : `descendant-or-self::node()`
- child : `child::node()` ou `node()`
- following : `following::element`
- following-sibling : `following-sibling::node()`
- preceding : `preceding::node()`
- preceding-sibling : `preceding-sibling::node()`
- self : `self::node()` ou `.`
- attribute : `attribute::node()`

Ancestor



Child



Exercice de repérage (2)

- Mon rêve familial
 - Donner le chemin le plus court de la racine TEI vers les vers d'un quatrain
- Le Misanthrope
 - Donner le chemin le plus court depuis un vers vers le nom du personnage qui parle
 - Donner le chemin le plus court depuis les vers d'une scène vers le titre de la scène
- Le Cid
 - Donner le chemin le court de la racine TEI vers un des paragraphes de la préface

Pour affiner sa requête Xpath, on peut ajouter des **prédicats**.

Le prédicat est noté entre crochets droits après l'élément auquel il se rapporte.

On peut ainsi spécifier la position d'un élément, une relation d'un élément avec un autre élément, le nom ou la valeur de l'un de ses attributs.

Exemples :

- `lg[attribute::type='sizain'] ;`
- `lg[attribute::type='quatrain'][position()=1]` ou `lg[attribute::type='quatrain'][1] ;`
- `lg[attribute::type='quatrain']/l[last()] .`

Exercice de repérage (3)

- Mon rêve familial
 - Donner le chemin le plus rapide de la racine TEI vers le vers n°2 du premier quatrain;
- Le Misanthrope
 - Donner le chemin le plus rapide de la racine TEI vers le nom du locuteur de la troisième réplique de l'acte I, scène 1;
- Le Cid
 - Donner le chemin le plus rapide de la racine TEI vers le nom du deuxième personnage de la liste des personnages déclarée dans la préface.

Quelques abréviations

- descendant : `//`
- child : `node()`
- self : `.`
- attribute : `@`

Exercice

Écrire les chemins de l'exercice de repérage n°3 avec des abréviations

- Mon rêve familial
 - Donner le chemin le plus rapide de la racine TEI vers le vers n°2 du premier quatrain;
- Le Misanthrope
 - Donner le chemin le plus rapide de la racine TEI vers le nom du locuteur de la troisième réplique de l'acte I, scène 1;
- Le Cid
 - Donner le chemin le plus rapide de la racine TEI vers le nom du deuxième personnage de la liste des personnages déclarée dans la préface.

Pour aller plus loin : https://www.w3schools.com/xml/xpath_intro.asp