PXML Master MIAGE M1 2011-2012

Cours 6: Transformer des données avec XSLT

## Modèle de données de XPath 2.0

Le résultat d'une requête XPath est une séquence de nœuds ou de valeurs.

Qui dit séquence dit

- ordonné
- doublons possibles

## XPath 2.0

## Séquences

- Construction d'une séquence : Un moyen de construire une séquence est d'utiliser la virgule comme opérateur de concaténation. L'opérateur to permet de définir un intervalle.
  - (10, 1, 2, 3, 4) séquence de 5 entiers
  - (10, (1, 2), (), (3, 4)) même séquence que précédemment
  - (10, 1 to 4) toujours la même séquence

#### 2.Expressions de chemin

- //bouteilles/bouteille/(codebarre,volume) séquence avec tous les fils codebarre et volume des éléments bouteille.
- (21 to 29)[5] l'entier 25.
- //bouteille[position()==(1 to 2)] le premier et le second élément bouteille du document.

# Séquences

- 3. Opérateurs sur les séquences :
- union.
- intersect.
- except.
- 4. Comparaison de deux séquences : quantification existentielle
- (1,2) = (2,3) vaut vrai (Donc l'égalité n'est pas transitive! (2,3) = (3,4) vaut vrai, pourtant (1,2) = (3,4) vaut faux.)

#### Instruction conditionnelle if

```
"if" "(" Expr ")" "then" Expr
```

"else" Expr

#### Exemple:

if (count(//bouteille)>2) then //codebarre else //volume

renvoie la séquence des éléments **codebarre** des bouteilles dont le volume est supérieur à 50cl.

### Plein de fonctions!

Voir la fiche synthétique

http://www.mulberrytech.com/quickref/functions.pdf

accessible via le portail dans l'onglet document et la norme W3C.

# Expression quantifiée

C'est une expression booléenne dont la syntaxe est :

```
("some" | "every") "$"VarName
"in" Expr ("," "$" VarName "in" Expr)*
"satisfies" Expr
```

#### Exemple:

```
some x in (1, 2, 3), y in (2, 3, 4) satisfies x + y = 4 every x in (1, 2, 3), y in (2, 3, 4) satisfies x + y = 4 La première expression vaut vrai, la seconde vaut faux.
```

#### La boucle for

```
"for" "$" VarName "in" Expr ("," "$" VarName "in" Expr)*
"return" Expr
```

#### Exemple:

```
for $b in //bouteille , $e in id($b/attribute::ref_eau)
return $e/nom eau
```

renvoie la séquence (avec des doublons) des noms d'eaux mises en bouteilles. Ici, la boucle **for** concatène tous les résultats sans enlever les doublons. On a donc autant d'items dans la séquence résultat que d'éléments bouteille.

## Motivations

- 1. Motivation à l'origine : associer un style à un document XML
  - XSL = XML Stylesheet Language
  - Les CSS que l'on utilise pour XHTML ne permettent pas d'afficher les valeurs des attributs, de transformer la structure du document, ni de créer de nouvelles données.
  - De plus CSS pour XML est un peu lourd car il n'y a pas de style par défaut comme en XHTML
  - XSL = XSL-FO (pour l'aspect formatage) + XSL-T (pour l'aspect transformation)
- 2. Transformation d'un document XML en un autre document XML
  - Transformation d'un arbre source en un arbre cible
  - Une transformation est donnée par un ensemble de règles
  - XSLT 1.0 utilise XPath 1.0
  - XSLT 2.0 utilise XPath 2.0

## **XSLT**

### Définir une feuille XSLT

- XSLT est un dialecte XML, défini par un schéma d'espace de noms http://www.w3.org/1999/XSL/Transform
- version 1.0 qui date de novembre 99 (et version 2.0 de janvier 2007).
- La racine d'un document XSLT est un élément xsl:stylesheet, ou bien xsl:transform (synonyme).

```
<xsl:stylesheet version="1.0"
   xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
   xmlns="http://www.w3.org/TR/xhtml1/strict">
...
</xsl:stylesheet>
```

## Appliquer une feuille XSLT à un document

 Attacher la feuille de style au document, comme on le fait pour une feuille CSS.

```
<?xml-stylesheet type="text/xsl" href="transfo1.xsl" ?>
```

2. Utiliser un programme qui applique la transformation au document pour produire un autre document

## Règles de transformation

- Une règle est définie avec un élément xsl:template.
- On v trouve :
- ✓un critère de sélection de nœuds dans le document source (attribut match). C'est une requête XPath restreinte aux axes verticaux descendants (mais qui peut contenir des prédicats).
- √ parfois des paramètres
- ✓ Un constructeur de la séquence résultat. Il est évalué pour produire une séquence d'items qui sont écrits dans l'arbre résultat.
- L'élément xsl:apply-templates permet d'appliquer explicitement une règle sur une séquence de nœuds définie par une sélection XPath.
- L'attribut match du template est nécessaire, sauf si le template a un attribut name. Dans ce cas, l'appel au template se fait par son nom :

```
<xsl:call-template name="..."/>
```

### Préciser le format de sortie

Pour préciser un format de sortie autre que XML qui est le format de sortie par défaut, on utilise l'élément xsl:output (c'est une instruction de traitement) en donnant une valeur à son attribut method parmi les valeurs html, xml ou text. Cet élément dispose aussi d'un attribut indent qu'on peut positionner à yes ou no

```
<xsl:stylesheet version="1.0"
    xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
    xmlns="http://www.w3.org/TR/xhtml1/strict">
    <xsl:ouput method="html" indent="yes"/>
    ...
</xsl:stylesheet>
```

# Exemple

## Application des règles

- L'attribut select de l'élément apply-templates permet de définir la séquence de nœuds sur lesquels il faut appliquer une règle. Sa valeur est une requête XPath, qu'on évalue à partir du nœud courant. Le résultat de cette requête XPath doit être une séquence de nœuds.
- En l'absence de select, l'instruction apply-templates traite tous les nœuds enfants du nœud courant, y compris les nœuds textes mais pas les attributs.
- L'idée est de traiter les nœuds de façon descendante, mais si le chemin XPath valeur de l'attribut select permet de remonter dans le document, il est possible de créer une transformation qui ne s'arrête pas (boucle).

## Constructeur de séquence

Séquence de nœuds frères dans la feuille de style (donc en particulier dans un *template*).

Chacun d'eux est, au choix,

1.une instruction XSLT

2.un littéral élément

3.un nœud texte

# Exemple de transformation qui boucle

En sortie: <a><b><a>.....

# Constructeur de séquence

- Le constructeur de séquence est évalué pour chaque item de la séquence s des nœuds à traiter (dans un template, s est la séquence renvoyée par le select, ou bien la séquence des nœuds fils)
- Quand un constructeur de séquence est évalué, le processeur conserve donc une trace (appelée focus) de la séquence s des items en cours de traitement :
- ✓ L'item contexte (souvent nœud contexte),
- √ la position de l'item contexte dans la séquence s.
- √la taille de la séquence s.

#### Résolution des conflits

- Il est possible que plusieurs règles puissent s'appliquer sur l'item courant.
- Pour deux règles qui sont applicables :
- 1.On compare les attributs priority des règles, s'ils existent
- 2. Sinon, on calcule une priorité en fonction de la "sélectivité" du pattern de la rèale.
- 3.Si malgré tout, il subsiste plusieurs règles, alors le processeur peut déclencher une erreur, ou bien choisir la dernière règle dans l'ordre du document XSLT.

### Les modes

```
<xsl:template match="/">
  <h2>Marques</h2>
  <xsl:apply-templates select="voiture" mode="marque"/>
  <h2>Modèles</h2>
  <xsl:apply-templates select="voiture" mode="modele"/>
  </xsl:template>

<xsl:template match="voiture" mode="mk">
    ...
  </xsl:template>

<xsl:template match="voiture" mode="modele">
    ...
  </xsl:template>
```

#### Les modes

- Les modes permettent pour une même règle de traiter un même nœud du document source, en produisant des résultats différents.
- Grâce à l'attribut mode de template et apply-template, on peut définir pour quel mode une règle s'applique
- L'instruction apply-templates dispose aussi d'un attribut mode, qui permet de dire dans quel mode on veut appliquer une règle.
- Il existe un mode par défaut, qui n'a pas de nom, utilisé quand aucun nom de mode n'est donné explicitement

# Règles prédéfinies

- S'il n'existe pas de règle qui s'applique sur un nœud sélectionné par un apply-templates, on évalue une règle prédéfinie (quelque soit le mode)
- Pour un nœud élément, on traite les nœuds fils :

```
<xsl:template match="*|/">
  <xsl:apply-templates/>
</xsl:template>
```

 Pour un nœud texte ou attribut, on construit un nœud texte qui contient la valeur textuelle du nœud contexte.

```
<xsl:template match="text()|@*">
  <xsl:value-of select="."/>
</xsl:template>
```

# Règles prédéfinies (2)

La feuille suivante (vide !) permet d'extraire tout le texte (pas les attributs) d'un document :

```
<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet version="1.0"
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"/>
```

# Copie récursive (deep copy)

- Instruction xsl:copy-of avec un attribut select obligatoire.
- Les items de la séquence résultat du select sont traités de la manière suivante :
- 1.Si l'item est un nœud élément ou un nœud document, alors on ajoute au résultat un nouveau nœud du même type et de même contenu (attributs, texte, sous-éléments ...) que l'item source
- 2.Si l'item est un nœud d'un autre type (nœud attribut, un nœud texte, ...), alors la copie est la même qu'avec l'instruction xsl:copy
- 3.Si l'item est de valeur atomique, sa valeur est ajoutée à la séquence résultat.

# Recopie d'un nœud (shallow copy)

- L'élément xsl:copy permet de copier l'item contexte.
- Si l'item contexte est une valeur atomique, l'instruction xsl:copy retourne cette valeur, et ne tient pas compte du constructeur de séquence.
- Si l'item contexte est un nœud élément ou un nœud document, l'instruction xsl:copy retourne un nœud de même type que le nœud contexte, et qui contient le résultat de l'évaluation du constructeur de séquence contenu dans l'élément xsl:copy. Donc, les attributs et le contenu du nœud contexte n'est pas recopié.
- Si l'item contexte est un autre type de nœud (nœud attribut, un nœud texte, ...), l'instruction xsl:copy retourne un nœud de même type que le nœud contexte, et qui contient la même valeur texte que le nœud contexte.

## Exemple: transformation identité

```
<xsl:template match="/">
  <xsl:copy-of select="."/>
</xsl:template>
```

#### équivalent à :

```
<xsl:template match="@*|node()">
  <xsl:copy>
     <xsl:apply-templates select="@*|node()"/>
  </xsl:copy>
</xsl:template>
```

#### Créer des nœuds

1.Création d'éléments

2.Création d'attributs

3. Création de nœuds texte

## Création d'éléments

#### est transformé en

```
<?xml version="1.0" encoding="UTF-8"?>
<renault xmlns="http://www.w3.org/1999/xhtml">
    <megane>2005</megane>
</renault>
```

## Création d'éléments

- On peut utiliser un littéral élément, en écrivant directement les balises que l'on veut en sortie
- On peut utiliser un élément xsl:element, ça permet de construire un élément dont le nom (ou le contenu) est calculé dynamiquement et communiqué dans l'attribut name sous la forme d'une expression entre accolades.

## Création d'attributs

- On peut écrire l'attribut "en dur" dans un littéral élément
- On peut aussi utiliser des accolades autour de l'expression définissant la valeur de l'attribut; {exp} représente la valeur de l'évaluation de l'expression exp. Par exemple:

```
<xsl:template match="photograph">
  <img src="/images/{href}" width="{size/@width}"/>
  </xsl:template>
```

évalué sur :

```
<photograph>
  <href>headquarters.jpg</href>
  <size width="300"/>
</photograph>
```

a pour résultat :

<img src="/images/headquarters.jpg" width="300"/>

#### Création d'attributs

• On peut aussi utiliser un élément xsl:attribute pour définir un nœud attribut. Par exemple :

## Création d'un nœud texte

Autre exemple :

```
<xsl:template match="/">
  <les-titres>
     <xsl:value-of select=".//titre"/>
  </les-titres>
</xsl:template/>
```

## Création d'un nœud texte

- On peut écrire directement du texte dans le constructeur de séquence
- On peut utiliser l'élément xsl:text

```
<xsl:text>blabla</xsl:text>
```

• On peut générer un nœud texte dont le contenu n'est pas statique, grâce à l'élément xsl:value-of et son attribut select

```
<xsl:value-of select="./ville"/>
```

## Structures de contrôle

1.Répétition: for each

2.Conditionnelles: if et choose

3.Tri:sort

# Répétition

- L'instruction xsl:for-each traite chaque item d'une séquence d'items et pour chacun évalue le constructeur de séquence
- Le résultat d'un xsl:for-each est la séquence concaténation des séquences obtenues pour chaque item.

# Transformation équivalente(1)

# Exemple

```
<xsl:template match="/">
  <html xmlns="http://www.w3.org/1999/xhtml">
     <head><title>Clients</title></head>
     <body>
        <xsl:for-each select="clients/client">
                <t.h>
                     <xsl:apply-templates select="nom"/>
                   <xsl:for-each select="cmde">
                     <xsl:apply-templates/>
                   </xsl:for-each>
                </xsl:for-each>
          </body>
  </html>
  </xsl:template>
```

# Transformation équivalente(2)

#### Instructions conditionnelles

- Instruction xsl:if avec un attribut test.
- √Le test est une condition XPath.
- √Si le test est évalué à vrai, alors le constructeur de séquence à l'intérieur de l'instruction if est évalué ; sinon, la séquence vide est retournée.
- Instruction xsl:choose

```
<xsl:choose>
  <xsl:when test = "...">...</xsl:when>
    ...
  <xsl:when test = "...">...</xsl:when>
    <xsl:otherwise>...</xsl:otherwise>
</xsl:choose>
```

## Variables et paramètres

### Tri des nœuds avant traitement

Par défaut les nœuds sont traités dans l'ordre du document. On peut les traiter dans un ordre différent à l'aide de l'élément xsl:sort qui ne peut être fils que d'un élément xsl:for-each ou d'un élément xsl:apply-template.

### Définition de variable

- Élément xsl:variable, avec un attribut name obligatoire, et des attributs select et as optionnels.
- La valeur de la variable est
- √ soit la valeur de l'évaluation de l'expression du select,
- √ soit la valeur de l'évaluation du constructeur de séquence de l'élément xsl:variable.
- Si l'attribut select est présent, le constructeur de séquence doit être vide.
- L'attribut as donne le type de la variable.
- S'il est absent, alors la variable prend le type de sa valeur. S'il est présent alors la valeur de la variable est convertie en une valeur de ce type. Une variable sans attribut as et sans attribut select dont le contenu est non vide, est évaluée en un nœud document qui a pour fils l'évaluation du constructeur de séquence (cf exemples)

## Exemples

```
<xsl:variable name="i" as="xs:integer*" select="1 to 3"/>
a pour valeur la séquence d'entiers (1 2 3)
```

<xsl:variable name="i" as="xs:integer" select="@size"/> a
pour valeur l'entier valeur de l'attribut size

<xsl:variable name="doc"><c/></xsl:variable> a pour valeur un
nœud document qui a pour fils un élément vide c

# Règles paramétrées

- L'élément xsl:param est utilisé pour définir un paramètre d'un template.
- La définition d'un paramètre ressemble fort à la définition d'une variable.
   L'attribut select ou le constructeur de séquence servant à donner une valeur par défaut.
- On a aussi un attribut required, faux par défaut, qui indique si le paramètre est obligatoire.
- Quand un paramètre est obligatoire, il ne peut pas avoir de valeur par défaut.
- Dans un apply-templates ou un call-templates, on donne une valeur au paramètre grâce à l'élément with-param.

## Utilisation des variables

## Exemple (on reprend l'exemple des clients)

```
<xsl:template match="client">
<!-- le parametre indique le nombre minimum de commandes
obligatoires-->
    <xsl:param name="mini" required="yes"/>
<!-- la variable qui contient le nombre de commandes du
client courant -->
    <xsl:variable name="nb-cmdes" select="count(cmde)"/>
    <xsl:if test="$mini &lt;= $nb-cmdes">
       <xsl:apply-templates select="nom"/>
          <xsl:apply-templates select="cmde"/>
       </t.r>
    </xsl:if>
 </xsl:template>
 <xsl:template match="cmde">
    <xsl:apply-templates/>
 </xsl:template>
```

## Exemple (2)

On ne veut conserver que les clients qui ont au moins 2 commandes

## exemple

```
<?xml version="1.0" encoding="UTF-8" ?>
<xsl:stylesheet version="2.0"</pre>
 xmlns:xsl="http://www.w3.org/1999/XSL/Transform" >
  <xsl:ouput method="html" indent="yes"/>
  <xsl:template match="/">
    <html>
      <head>
        <title>biblio</title>
      </head>
      <body>
        <h1>Les Livres</h1>
        <xsl:apply-templates select="//book"/>
        <h1>Les Auteurs</h1>
        <xsl:apply-templates select="//author"/>
      </body>
    </html>
  </xsl:template>
</xsl:stylesheet>
```

## Spécificités de XSLT 2.0

- Recommandation du W3C en 2007
- Objectifs
- √accroître la simplicité d'utilisation
- √gestion de plusieurs documents de sortie
- √ possibilité de créer des fonctions
- √utilisation de XML Schema
- √simplifier le regroupement des contenus en sortie
- √ compatibilité ascendante

## Feuille de style simplifiée

```
Conditions
```

```
√ feuille XSLT ne contenant qu'une règle pour la racine (<xs1:template match="/">).
```

- √ feuille ne contenant pas d'élément xsl:apply-template
- Application
- √combiner le contenu des éléments xsl:stylesheet et xsl:template dans la balise ouvrante du document résultat.

```
√supprimer les éléments xsl:stylesheet, xsl:ouput et xsl:template
```

# Feuille de style simplifiée

## Production de xhtml

```
<?xml version="1.0"?>
<xsl:stylesheet
   xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
   version="2.0"
   xmlns="http://www.w3.org/1999/xhtml">

<xsl:template match="/">
   <xsl:template match="/">
   <xsl:result-document method="xhtml">
        <html><head><title>
...
        <xsl:for-each select="./nom[@langue='Francais']">
              <xsl:value-of select="."/>
        </ssl:for-each>
...
        </body>
        </html>
    </xsl:result-document>
...
```

## Production de xhtml

```
<?xml version="1.0"?>
<xsl:stylesheet
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  version="2.0"
  xmlns="http://www.w3.org/1999/xhtml">
<xsl:output method="xhtml"/>

<xsl:template match="/">
  <html><head>
  ...
```

## Production de plusieurs documents en sortie

#### **Fonctions**

```
<?xml version="1.0"?>
<xsl:stylesheet</pre>
 xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
 version="2.0"
 xmlns:mfn="http://www.lifl.fr/mfn"
 exclude-result-prefixes="mfn">
<xsl:output method="html"/>
<xsl:template match="/">
  <body>
    <xsl:value-of select="mfn:bonjour(//nom)"/>
</xsl:template>
<xsl:function name="mfn:bonjour">
  <xsl:param name="salutation"/>
 Bonjour, <xsl:value-of select="$salutation"/>
</xsl:function>
</xsl:stylesheet>
```

## Validation du résultat

```
<?xml version="1.0"?>
<xsl:stylesheet
  xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
  xmlns:ici="http://www.ici.fr"
  version="2.0">
  <xsl:output method="xml" indent="yes"/>
  <xsl:import-schema
    namespace="http://www.ici.fr"
    schema-location="schema.xsd"/>
  <xsl:template match="/">
    <ici:clubs xsl:validation="strict">
...
```

# groupement des résultats

```
<+r>
   ID journal
   ID section
   <xsl:for-each-group select="//source" group-by="@idjournal">
    <+d>>
       <xsl:value-of select="@idjournal"/>
      <t.d>
      <xsl:value-of
       select="current-group()/@idsection"/>
    </xsl:for-each-group>
```

## Conclusion

- Langage de transformation d'arbres
- Langage fonctionnel, récursif
- Utilisation de XPath
- Gestion des espaces de noms
- Modularité et importation de feuilles
- Définition de fonctions
- Paramétrage de la feuille de style (pas seulement des règles)
- Génération de plusieurs arbres résultats
- Génération de résultats intermédiaires

• ...