



TD/TP n° 1

XML & DTD

Master Ingénierie Informatique

U.E. « Bases de données avancées »

F. BERTRAND

Année universitaire 2021-2022

1 Environnement de travail

Pour manipuler les documents XML (ainsi que les DTD et les schémas RELAX NG) nous utiliserons l'éditeur *Exchanger XML* disponible dans le dossier « Applications » sous Windows. À noter que pour les personnes souhaitant installer un éditeur sur leur poste il est disponible à l'adresse suivante : <https://code.google.com/p/exchangerxml/>.

2 Notion de document bien formé

Le document XML ci-dessous n'est pas bien formé. Sans utiliser d'outil, indiquer la position (numéro de ligne) et la nature de chaque erreur de syntaxe :

```
1 <?xml version="1.0" ?>
2 <!-- Annuaire d'illustration -->
3 <annuaire>
4   <personne <!-- une 1ere personne --> >
5     <nom>Dupont</nom>
6     <prenom>Paul <!-- ou Polo --> </prenom>
7     <date format="ISO" format="fr-fr">1998-07-23
8     <telephone/>
9   </personne>
10  <personne> <!-- une 2nd personne -->
11    <nom>Alain</nom>
12    <prenom courant='Vrai'>Adrien</prenom>
13    <date format=ISO>1969-06-04</date>
14    <telephone>
15      <indicatif tel> 02 </indicatif tel>
16      <num#tel> 40 40 36 18 </num#tel>
17    </personne>
18  </telephone>
19 </annuaire>
20 <annuaire>
21   <personne>
22     <nom>DUPONT</nom>
23     <prenom>MARC</prenom>
24     <telephone/>
25   </personne>
26 </annuaire>
```

3 DTD et validité de documents XML

Indiquer pour quelles raisons les fichiers XML suivants sont valides ou non. Justifier votre réponse.

1. Ici la connaissance de la DTD n'est pas nécessaire...

```
1 <?xml version="1.0"?>
2 <!DOCTYPE tutorial SYSTEM "tutorial.dtd">
3
4 <text>This is an XML document</text>
```

2. Soit l'extrait de la DTD tutorial :

```
1 <!ELEMENT tutorial (XXX)>
2 <!ELEMENT XXX (AAA , BBB)>
3 <!ELEMENT AAA (#PCDATA)>
4 <!ELEMENT BBB (#PCDATA)>
```

et les fichiers

```
1 <?xml version="1.0"?>
2 <!DOCTYPE tutorial SYSTEM "tutorial.dtd">
3
4 <tutorial>
5 <XXX> <AAA>Start</AAA> <BBB>End</BBB> </XXX>
6 </tutorial>
```

```
1 <?xml version="1.0"?>
2 <!DOCTYPE tutorial SYSTEM "tutorial.dtd">
3
4 <tutorial>
5 <XXX> <AAA/> <BBB/> <BBB/> </XXX>
6 </tutorial>
```

3. Soit l'extrait de la DTD tutorial :

```
1 <!ELEMENT tutorial (XXX)>
2 <!ELEMENT XXX (AAA? , BBB+)>
3 <!ELEMENT AAA (CCC? , DDD*)>
4 <!ELEMENT BBB (CCC , DDD)>
5 <!ELEMENT CCC (#PCDATA)>
6 <!ELEMENT DDD (#PCDATA)>
```

et les fichiers

```
1 <?xml version="1.0"?>
2 <!DOCTYPE tutorial SYSTEM "tutorial.dtd">
3
4 <tutorial>
5 <XXX>
6   <AAA/> <BBB> <CCC/> <DDD/> </BBB>
7 </XXX>
8 </tutorial>
```

```
1 <?xml version="1.0"?>
2 <!DOCTYPE tutorial SYSTEM "tutorial.dtd">
3
4 <tutorial>
5 <XXX>
6   <AAA>
7     <CCC/><CCC/>
8     <DDD/><DDD/>
9   </AAA>
10  <BBB>
11    <CCC/><DDD/>
12  </BBB>
13 </XXX>
14 </tutorial>
```

4. Soit l'extrait de la DTD tutorial :

```
1 <!ELEMENT tutorial (XXX)>
2 <!ELEMENT XXX (AAA+ , BBB+)>
3 <!ELEMENT AAA (BBB | CCC )>
4 <!ELEMENT BBB (#PCDATA | CCC )*>
5 <!ELEMENT CCC (#PCDATA)>
```

et le fichier

```
1  <?xml version="1.0"?>
2  <!DOCTYPE tutorial SYSTEM "tutorial.dtd">
3
4  <tutorial>
5    <XXX>
6      <AAA>
7        <CCC>Un seul élément</CCC>
8      </AAA>
9      <AAA>
10       <BBB>
11         <CCC/>
12         <CCC/>
13         <CCC/>
14       </BBB>
15     </AAA>
16     <BBB/>
17     <BBB>
18       C'est <CCC/> une combinaison d'éléments <CCC> CCC </CCC>
19     </BBB>
20     <BBB>
21       Text only.
22     </BBB>
23   </XXX>
24 </tutorial>
```

Soit la DTD equipe.dtd :

```
1  <!ELEMENT equipe (responsable, (membre*)) >
2  <!ENTITY % identite "(nom,(prenom)+)" >
3  <!ELEMENT responsable %identite; >
4  <!ELEMENT membre %identite; >
5  <!ELEMENT titre (#PCDATA)>
6  <!ELEMENT nom (#PCDATA)>
7  <!ATTLIST nom titre (Dr | Pr) #IMPLIED >
8  <!ELEMENT prenom (#PCDATA)>
```

et les trois documents XML ci-dessous... Sont-ils valides vis-à-vis de cette DTD?... Sinon quelles sont les erreurs?...

```
1  <!DOCTYPE equipe SYSTEM "equipe.dtd">
2  <equipe>
3    <responsable>
4      <titre>Dr</titre>
5      <nom>Von Braun</nom>
6      <prenom>Werner</prenom>
7    </responsable>
8    <membre>
9      <nom>Lippisch</nom>
10     <prenom>Alexander</prenom>
11     <prenom>Martin</prenom>
12   </membre>
13   <membre>
14     <nom>Rudolph</nom>
15     <prenom>Arthur</prenom>
16   </membre>
17 </equipe>
```

```

1 <!DOCTYPE equipe SYSTEM "equipe.dtd">
2 <equipe>
3   <responsable>
4     <nom titre='Dr'>Von Braun</nom>
5     <prenom>Werner</prenom>
6   </responsable>
7   <membre>
8     <nom titre='Mr'>Lippisch</nom>
9     <prenom>Alexander</prenom>
10    <prenom>Martin</prenom>
11  </membre>
12  <membre>
13    <nom>Rudolph</nom>
14    <prenom>Arthur</prenom>
15  </membre>
16 </equipe>

```

```

1 <!DOCTYPE equipe SYSTEM "equipe.dtd">
2 <equipe>
3   <responsable>
4     <nom titre="Dr">Von Braun</nom>
5     <prenom>Werner</prenom>
6   </responsable>
7   <membre>
8     <nom>Lippisch</nom>
9   </membre>
10 </equipe>

```

Soit la DTD suivante :

```

1 <!-- ce qui suit permet d'utiliser tous les elements de HTML
2 -->
3
4 <!ENTITY % html PUBLIC
5   "-//W3C//DTD XHTML Basic 1.0//EN"
6   "http://www.w3.org/TR/xhtml1-basic/xhtml1-basic10.dtd">
7
8 %html;
9
10
11 <!-- l'element "article" doit contenir, en sequence, un element titre, un
12      element auteur, un element date, un element lieu et un element texte
13 -->
14
15 <!ELEMENT article (titre, auteur, date, lieu, texte)>
16 <!ELEMENT titre (#PCDATA)>
17 <!ELEMENT auteur (#PCDATA)>
18 <!ELEMENT date (#PCDATA)>
19 <!ELEMENT lieu (#PCDATA)>
20 <!-- l'element "texte" peut contenir du texte et des elements "grand" et
21      "image" en nombre indetermine
22 -->
23 <!ELEMENT texte (#PCDATA|grand|image)*>
24 <!ELEMENT grand (#PCDATA)>
25 <!-- l'element "image" doit contenir une balise HTML img et une legende
26      facultative
27 -->
28 <!ELEMENT image (html:img, (legende)?)>
29 <!ELEMENT legende (#PCDATA|grand)*>

```

Le document XML suivant est-il bien formé? Et est-il valide?...

```

1 <?xml version="1.0" encoding="ISO-8859-1"?>
2 <!DOCTYPE article SYSTEM "article.dtd">
3
4 <article xmlns:html="http://www.w3.org/1999/xhtml">
5   <date>14 juin 1972</date>
6   <auteur>David Durand</auteur>
7   <titre>Un journaliste accuse, un policier d'assaut</titre>
8   <lieu>Lyon</lieu>
9   <texte>
10     <grand>Un journaliste de la place accuse</grand>
11     les autorités ...
12     <image>
13       <html:img html:src="photo.gif" />
14       <legende>Legende de la photo</legende>
15     </image>
16     ...
17   </texte>
18 </article>

```

4 Fragments de DTD

- Quel fragment de DTD décrit correctement un numéro de téléphone avec un indicatif pays optionnel?
 1. <!ELEMENT numTelephone ((indPays)*,indRegion, numero)>
 2. <!ELEMENT numTelephone (indPays?,indRegion, numero)>
 3. <!ELEMENT numTelephone? (indPays,indRegion, numero)>
 4. <!ELEMENT numTelephone (indPays,(indRegion, numero)+)>
- De quelle manière indiquer, pour un élément A constitué de trois éléments B, C et D, que ces derniers peuvent apparaître dans n'importe quel ordre?
- De quelle manière indiquer qu'un élément A est constitué au minimum de 2 éléments B et au maximum de 4?

5 Création d'une DTD

On souhaite créer un document XML représentant une liste de commandes.

Une liste de commandes doit contenir une ou plusieurs commandes.

Une commande est constituée d'une ou plusieurs lignes de commandes (référéncant un seul produit) et peut posséder une remise. De plus, chaque commande possède un identifiant propre, l'identifiant du client qu'il l'a passée et son montant total qui s'il n'est pas fourni pourra être calculé par l'application traitant le fichier XML.

Une ligne de commande est constituée de l'identifiant de l'article concerné, de la quantité commandée et du prix unitaire de cet article.

6 Écriture d'une DTD à partir d'un document existant

À partir du document suivant, et sans utiliser le type ANY, écrire une DTD permettant de valider le document :

```

1 <?xml version='1.0' standalone="no" ?>
2 <!DOCTYPE musique SYSTEM "musique.dtd">
3 <musique>
4   <cd>
5     <titre>It's a wonderful life</titre>
6     <groupe>Sparklehorse</groupe>
7     <chanson>Eyepennies</chanson>
8     <chanson>Dog door</chanson>
9   </cd>

```

```

10 <cd>
11 <titre>BOF Le fabuleux destin...</titre>
12 <artiste>
13   <nom>Tiersen</nom>
14   <prenom>Yann</prenom>
15 </artiste>
16 <chanson>A quai</chanson>
17 <chanson duree='3:40'>Comptine d'un autre été</chanson>
18 </cd>
19 </musique>

```

7 Espaces de noms XML

Soit le document XML suivant :

```

1 <?xml version="1.1" ?>
2 <elementA xmlns:nsA="http://www.test.org">
3   <nsA:elementB xmlns="http://www.ns1.org" att1="val1">
4     <elementC xmlns="http://www.ns2.org" nsA:att2="val2">
5       <elementD>texte...</elementD>
6     </elementC>
7   </nsA:elementB>
8 </elementA>

```

Pour chaque élément et attribut présents dans ce document, donner son espace de noms (URI) auquel il appartient et son nom local.