

TD5 – XML, DTD, XML Schema

Exercice 1

Etant donné le document ci-dessous,

1. Trouver les 6 erreurs qui font de ce document un document qui n'est pas bien formé.
2. Donner l'arbre correspondant.

```
<?xml encoding="utf-8" version="1.0"?>
<cours jour = "Jeudi" salle="B300">
  <sujet>XML</sujet>
  <Horaire>8h00</horaire>
</cours>
<td jour=Mardi>
  <sujet>XML</sujet>
  <horaire>15h30</horaire>
<tp jour="Mercredi">
  <sujet>XML</sujet>
  <horaire>13h30</horaire>
</TP>
```

Exercice 2

Les éléments XML doivent suivre ces règles de nommage (cf. http://www.w3schools.com/xml/xml_elements.asp)

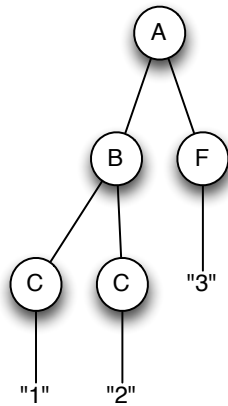
- *Element names are case-sensitive*
- *Element names must start with a letter or underscore*
- *Element names cannot start with the letters xml (or XML, or Xml, etc)*
- *Element names can contain letters, digits, hyphens, underscores, and periods*
- *Element names cannot contain spaces*

Parmi les balises suivantes, indiquer lesquelles sont des balises correctes :

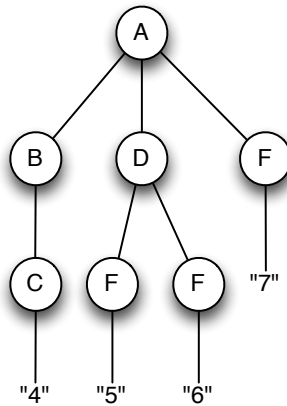
<123>	<a-b-c>
<__>	<dd;ee>
<_a>	<ASDF>
<_:-->	<1qwerty>
<a:2>	<a2:eee>
<.aaa>	<aa __>
<a.b.c>	<xmltest>
<_aaa-a:_->	

Exercice 3

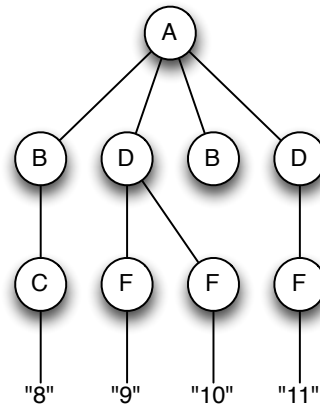
On considère les arbres XML suivants a_1 , a_2 , a_3 correspondant respectivement à des fichiers `a1.xml`, `a2.xml`, `a3.xml` :



a_1



a_2



a_3

On considère aussi les DTD suivants :

d1.dtd

```
<!ELEMENT A ((B|D)+,F)>
<!ELEMENT B (C*)>
<!ELEMENT C (#PCDATA)>
<!ELEMENT D (F*)>
<!ELEMENT F (#PCDATA)>
```

d2.dtd

```
<!ELEMENT A ((B,D)+,F?)>
<!ELEMENT B (C*)>
<!ELEMENT C (#PCDATA)>
<!ELEMENT D (F*)>
<!ELEMENT F (#PCDATA)>
```

d3.dtd

```
<!ELEMENT A (B+,D+,F*)>
<!ELEMENT B (C*)>
<!ELEMENT C (#PCDATA)>
<!ELEMENT D (F*)>
<!ELEMENT F (#PCDATA)>
```

1. Pour chacun de ces DTD, dire s'il valide ou non chacun des documents `a1.xml`, `a2.xml` et `a3.xml`.
2. Donner la forme sérialisée de chacun des documents `a1.xml`, `a2.xml` et `a3.xml`.

Exercice 4

Considérons le DTD suivant :

```
<!ELEMENT movies (movie*)>
<!ELEMENT movie (title, year, runtime, directors, actor*, summary)>
<!ATTLIST movie id ID #REQUIRED>
<!ELEMENT title (#PCDATA)>
<!ELEMENT year (#PCDATA)>
<!ELEMENT runtime (#PCDATA)>
<!ELEMENT directors (director*)>
<!ELEMENT director (name)>
<!ELEMENT actor (name, role)>
<!ELEMENT name (firstname, lastname)>
<!ELEMENT firstname (#PCDATA)>
<!ELEMENT lastname (#PCDATA)>
<!ELEMENT role (#PCDATA)>
<!ELEMENT summary (#PCDATA)>
```

1. Écrire un fichier XML valide par rapport à ce DTD.
2. Pour l'élément `name` comment changer la définition pour autoriser le nom et le prénom à apparaître dans n'importe quel ordre (mais exactement un nom et un prénom) ?
3. Même question pour `movie` (on veut autoriser les éléments à apparaître dans n'importe quel ordre, mais avec les même cardinalités que dans le DTD initial).

Exercice 5

Considérons le XML Schema suivant :

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema">
  <xs:element name="personnes">
    A COMPLETER
  </xs:element>
  <xs:element name="personne">
    A COMPLETER
  </xs:element>
  <xs:complexType name="type-evenement">
    A COMPLETER
  </xs:complexType>
</xs:schema>
```

Donner les définitions manquantes :

1. **personnes** : contient un nombre arbitraire d'éléments de type **personne**.
2. **personne** : contient (dans cet ordre) le nom, le prénom, la date et le lieu de naissance, et optionnellement la date et le lieu de décès. On retrouve également un attribut **id** obligatoire.
3. **type-evenement** : contient une date et un lieu (c'est le type à utiliser pour les naissances et les décès).

Par exemple, le document suivant doit être valide pour le XML schema résultat.

```
<?xml version="1.0" encoding="UTF-8"?>
<personnes>
  <personne id="p-034">
    <nom>Martin</nom>
    <prenom>Jules</prenom>
    <naissance>
      <date>1918-04-12</date>
      <lieu>Angers</lieu>
    </naissance>
    <deces>
      <date>1973-08-22</date>
      <lieu>Paris</lieu>
    </deces>
  </personne>
  <personne id="p-098">
    <nom>Martin</nom>
    <prenom>Henri</prenom>
    <naissance>
      <date>1931-08-14</date>
      <lieu>Paris</lieu>
    </naissance>
  </personne>
</personnes>
```