

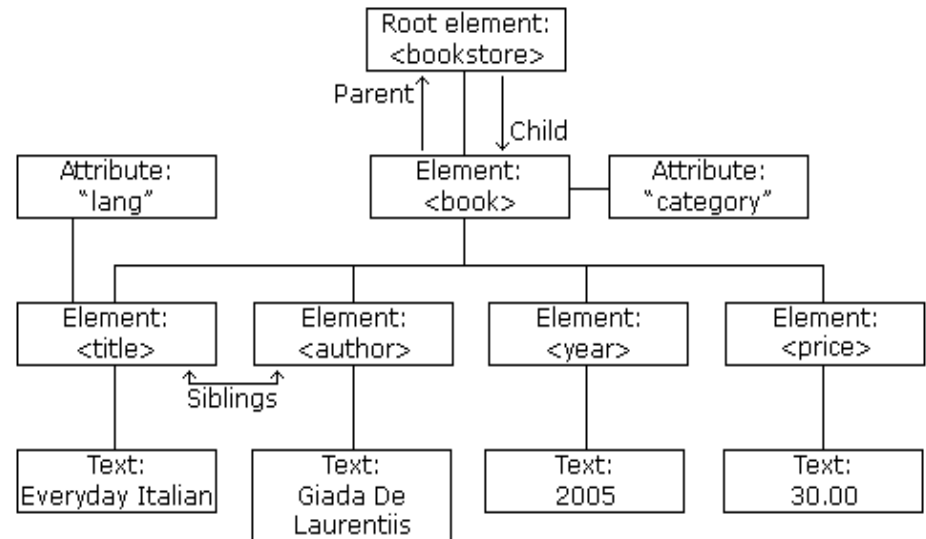
Tutorial – XML

CSCC43 - Introduction to Databases

Note: some examples are taken from <http://www.w3schools.com/XML/>

Tree Structure of XML

```
<bookstore>
  <book category="COOKING">
    <title lang="en">Everyday Italian</title>
    <author>Giada De Laurentiis</author>
    <year>2005</year>
    <price>30.00</price>
  </book>
  <book category="CHILDREN">
    <title lang="en">Harry Potter</title>
    <author>J K. Rowling</author>
    <year>2005</year>
    <price>29.99</price>
  </book>
  <book category="WEB">
    <title lang="en">Learning XML</title>
    <author>Erik T. Ray</author>
    <year>2003</year>
    <price>39.95</price>
  </book>
</bookstore>
```



Exercise – Draw XML Tree

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<endangered_species>
  <animal>
    <!--only one name element is permitted -->
    <name language="English">Tiger</name>
    <threats>
      <threat>poachers</threat>
      <threat>habitat destruction</threat>
      <threat>trade in tiger bones for traditional Chinese medicine (TCM)</threat>
    </threats>
    <!--the weight and length elements are _not_ optional -->
    <weight>500 pounds</weight>
    <length>3 yards from nose to tail</length>
    <source sectionid="120" newspaperid="21"/>
    <picture filename="tiger.jpg" x="200" y="197"/>
    <!--only one subspecies element is permitted -->
    <subspecies>
      <name language="English">Amur or Siberian</name>
      <name language="Latin">P.t. altaica</name>
      <region>Far East Russia</region>
      <population year="1999">445</population>
    </subspecies>
  </animal>
</endangered_species>
```

Well-Formed XML

- All XML documents must have a root element.
- All XML elements must have a closing tag
- XML tags are case sensitive
- XML elements must be properly nested
- XML attribute values must be quoted
 - using `"..."` or `'...'`
- Reserved characters cannot be used directly
 - using entity reference instead

Exercise – Find errors

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- sample XML document for home sale listing ->
<Listing date = 01/01/2007>
  <Home id='h1'>
    <Address>
      <Street> 20 Happy Street </street>
      <City> Toronto </City>
    </Address>
    <Bedrooms> 3 </Bedroom>
    <Price> 250000 </Price>
    <Room type="kitchen & dinning" length=10>
  </Home>
  <Home>
    <Address>
      <Street><Num>15</Num>Sunny Avenue
      <City></Street> Toronto </City>
    </Address>
    <Bedrooms>2</Bedrooms> <Size> Lot size > 20000 </Size>
  </Listing>
<Listing/ date = '01/02/2007'>
  <Home id = 'h3'><Address/><Price>300000</Price></Home>
</Listing>
```

Exercise – Find errors (Sol)

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- sample XML document for home sale listing -> # invalid comment
<Listing date = 01/01/2007> # attribute values must be quoted
  <Home id='h1'>
    <Address>
      <Street> 20 Happy Street </street> # tags are case sensitive
      <City> Toronto </City>
    </Address>
    <Bedrooms> 3 </Bedroom> # start and end tag do not match
    <Price> 250000 </Price>
    <Room type="kitchen & dinning" length=10> # attribute values must be quoted; missing <Room> end tag
  </Home> # need entity reference & for &;
  <Home> # possibly missing id (not a well-formed error though)
    <Address>
      <Street><Num>15</Num>Sunny Avenue
      <City></Street> Toronto </City> # tags not properly nested
    </Address>
    <Bedrooms>2</Bedrooms> <Size> Lot size > 20000 </Size> # need to use entity reference &gt; for >
  </Listing> # missing close tag </Home>
  <Listing/ date = '01/02/2007'> # wrong empty element
    <Home id = 'h3'><Address/><Price>300000</Price></Home>
  </Listing>

# missing root tag <Listings> ... </Listings>
```

Exercise – XML

A company database needs to store information about employees, including their

- **social insurance numbers**
- **status** (full time vs. part time),
- **salary** (monthly, hourly),
- **sex, date of birth,**
- **phone numbers** (home, work and mobile),
- **home address** (*street, province, city, postal code*).

Create an XML document instance according this description which contain data for at least two employees.