1. XML开头都需要一个声明 <?和?>表明这是一个处理指令

<?xml version=”1.0” encoding=”UTF-8” ?>

1. <book id=”1.0”>

</book>

等同于：

<book>

<id>1.0</id>

</book>

1. XML格式能够表达层次结构（树），并且重复的元素不会被曲解
2. XML与HTML区别
   1. XML是大小写敏感的
   2. XML不可省略结束标签
   3. XML只有单个标签时要**以/结尾**，使解释器同时认其为结束标签<hr/>（注意区别<a></a>）
   4. XML属性值必须用引号括起来，而HTML引号可有可无
   5. XML属性必须有属性值，HTML属性可以没有属性值
      1. HTML :<intput type=”radio” name=”language” value=”java” checked>其中checked属性没有属性值是允许的
3. XML文档结构

示例：

<?xml version="1.0" encoding="UTF-8"?> //文档头

<!DOCTYPE …….> //文档类型定义**DTD(文档类型定义)**，用以确保文档正确的机制，但不是必须的

<bookstore> //这是根root元素

<book id="1">

<name>冰与火之歌</name>

<author>乔治马丁</author>

<year>2014</year>

<price>89</price>

</book>

<book id="2">

<name>安徒生童话</name>

<year>2004</year>

<price>77</price>

<language>English</language>

</book>

</bookstore>

1. 在设计XML文档的时候，最好使元素要么只包含子元素，要么只包含文本，尽量不要混合式内容(mixed content)，可以简化解析过程(如getFirstChild())
   * 1. <font> //**不推荐的混合式**
        1. Hell
        2. <size>36</size>

</font>

1. XML元素可以包含属性：
   * 1. <size unit=”pt”>36</size>

常用法则：属性只应该用来修改值的解释，而不是用来指定值

如果值这样表示<font name=”Aero” size=”36”/> ，那么给size增加单位就必须添加到属性值中去：< font name=”Aero” size=”36 pt”/>，这样不够清真！耦合了，还是应该如下：

1. <font>
   * 1. <name>Areo</name>
     2. <size unit=”pt”>36</size>

</font>

这样增加或修改单位时就在size标签里指定属性就好

1. 元素(Element)和文本(Text)是XML文档的主要支撑元素，还有其他的一些标记如下：
   1. **字符引用**(character reference): &#十进制值、&#x十六进制值
      1. &#233、&#xE9表示同一个字符
   2. **实体引用**(entity reference): &name 同样为了显示这些特殊符号而不触发命令
      1. &lt、&gt、&amp、&quot、&apos
      2. 分别表示字符：小于、大于、&、引号、省略号
      3. 还可以在DTD中定义其他的实体引用
   3. **CDATA部分**(CDATA Section): 是用来表示字符数据的一种特殊形式，可以**囊括那些含有<,>,&之类字符的字符串，而不会触发对应命令**，用 <![CDATA[ 和 ]]> 限定界限，
      1. 注意CDATA中不能包含]]>字符串
      2. 这一特性常被用来当作将遗留数据偷偷纳入XML文档的一个后门
   4. **处理指令**(processing instruction)是给那些处理XML文档的应用程序使用的指令，**用 <? 和 ?> 来限定界限**，例如：
      1. <?xml-stylesheet href=”mystyle.css” type=”text/css”?>
      2. 而**XML文档开头也是处理指令**<?xml version=”1.0” encoding=”UTF-8”?>

* 1. **注释**(comment)**用 <!— 和 --> 限定界限**
     1. <!—This is a comment. --> 注释内容不能含有字符串 “—“