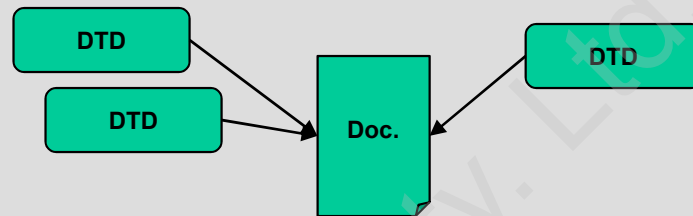


More XML Syntax

ml

- Attempt to make the definition of elements/attributes globally unique
 - ‘new’ to XML
 - *promulgated Jan. 14, 1999*
 - simplify the process of combining portions of different DTDs: lets you write an XML document that uses two or more sets of XML tags in modular fashion



- a simple idea
 - *give things a (hopefully!) globally unique name*
 - then `<law:bill>` and `<currency:bill>` can define different bills without conflict...
 - *similar to Java's package idea*

- Each element/attribute is associated with/grouped into a space which is named by a URI

- example:

- *BOOKS and COMMENT are in the default namespace*
- *BOOK lives in the 'bk' namespace*
- *the currency attribute is defined in, and obtained from, the 'money' namespace*

```
<?xml version="1.0" ?>
<BOOKS>
  <bk:BOOK xmlns:bk="urn:BookLovers.org:BookInfo"
            xmlns:money="urn:Finance:Money">
    <bk:TITLE>A Suitable Boy</bk:TITLE>
    <bk:PRICE money:currency="US Dollar">22.95</bk:PRICE>
    <COMMENT xmlns="">A damn good read!</COMMENT>
  </bk:BOOK>
</BOOKS>
```

- note the use of the colon

- *separates prefix and name*
- *shouldn't use it in a normal name*

- Namespaces are really simply about name munging
 - Creating an expanded name from a shortcut version

```
<bob:stuff xmlns:bob="www.transentia.com.au/bob/stuffns">  
...  
</bob:stuff>
```



```
<www.transentia.com.au/bob/stuffns:stuff>  
...  
</www.transentia.com.au/bob/stuffns:stuff>
```

From <http://www.xml.com/pub/1999/01/namespaces.html>:

“What Do Namespace Names Point At?”

One of the confusing things about all this is that namespace names are URLs; it's easy to assume that since they're Web addresses, they must be the address of something. They're not; these are URLs, but the namespace draft doesn't care what (if anything) they point at. Think about the example of the XML.com programmer looking for book titles; that works fine without the namespace name pointing at anything.

The reason that the W3C decided to use URLs as namespace names is that they contain domain names (e.g. www.xml.com), which work globally across the Internet.”

“...It has nothing special to do with multiple inheritance, formatting, versioning, brevity, or anything besides making sure that names are globally unique and globally defined.”

- Can be applied to:
 - element
 - Attribute
- Apply to all children
 - Unless overridden

```
<h:html xmlns:xdc="http://www.xml.com/books"
        xmlns:h="http://www.w3.org/HTML/1998/html4">
  <h:head><h:title>Book Review</h:title></h:head>
  <h:body>
    <xdc:bookreview>
      <xdc:title h:style="font-family: sans-serif;">
        XML: A Primer</xdc:title>
      <h:table>
        <h:tr align="center">
          <h:td>Author</h:td><h:td>Price</h:td>
          <h:td>Pages</h:td><h:td>Date</h:td></h:tr>
        <h:tr align="left">
          <h:td><xdc:author>Simon St. Laurent</xdc:author></h:td>
          <h:td><xdc:price>31.98</xdc:price></h:td>
          <h:td><xdc:pages>352</xdc:pages></h:td>
          <h:td><xdc:date>1998/01</xdc:date></h:td>
        </h:tr>
      </h:table>
    </xdc:bookreview>
  </h:body>
</h:html>
```

```
<!-- same example with default namespace -->
<html xmlns="http://www.w3.org/HTML/1998/html4"
      xmlns:xdc="http://www.xml.com/books">
  <head><title>Book Review</title></head>
  <body>
    <xdc:bookreview>
      <xdc:title>XML: A Primer</xdc:title>
      <table>
        ...
      </table>
    </xdc:bookreview>
  </body>
</html>
```

- Potentially problematic...

- DTDs must be written in such a fashion that a processor that knows nothing about namespaces can still parse and validate the document

“Namespaces in an XML is an official W3C recommendation. The W3C considers it complete, aside from possible minor errors and elucidations. Nonetheless, of all the XML specifications from the W3C, this one is the most controversial. Many people feel very strongly that this standard contains fundamental flaws. The main objection argues that namespaces are, in practice, incompatible with DTDs and validation. ...”

*“...you cannot use the same DTD for both documents with namespaces and documents without, even if they use essentially the same vocabulary.
In fact, you can't even use the same DTD for documents that use the same tag sets and namespaces, but different prefixes, because DTDs are tied to the actual prefixes rather than the URIs of the namespaces.”*

```
<!DOCTYPE doc [  
<!ELEMENT doc (x)>  
<!ELEMENT x EMPTY>  
<!ATTLIST x xmlns CDATA #FIXED "some.uri">  
<doc><x/></doc>
```

```
<!DOCTYPE doc [  
<!ELEMENT doc (x)>  
<!ELEMENT x EMPTY>  
<doc><x xmlns="some.uri"/></doc>
```

- Schemas (will) provide better facilities

- Designed to encode the logical structure of XML documents
 - “Any XML document is part of a set of XML documents that are logically equivalent within an application context, but which vary in physical representation based on syntactic changes permitted by XML”
 - introduces the notion of equivalence between XML documents which can be tested at the syntactic level
 - two XML documents whose Canonical-XML form is identical will be considered equivalent for the purposes of many applications

Hallelujah, I'm a bum!

```
<?xml version="1.0" ?>
<!-- file: canon.xml -->
<!DOCTYPE d [
  <!ENTITY lsb '['>
  <!ENTITY rsb ']'>
  <!ENTITY bum SYSTEM "hall.xml">
]>
<d>&lsb; &bum; &rsb;</d>
```

canonicalise

```
<?xml version="1.0" ?>
<!-- produced canonical version -->
<d>[Hallelujah, I'm a bum!]</d>
```

C:\Documents and Settings\Bob\Desktop\canon.xml - Micro

Address C:\Documents and Settings\Bob\Desktop\canon.xml

```
<?xml version="1.0" ?>
<!-- file: canon.xml -->
<!DOCTYPE d (View Source for full doctype...)>
<d>[Hallelujah, I'm a bum! ]</d>
```