

# CS 186 Section Notes

## XML

XML in 3 points:

**1. XML is for structuring data**

**2. XML looks a bit like HTML**

**3. XML is text, but isn't meant to be read**

### XPath

- Think of XML as a tree (or directory) structure.
- XPath specifies *path expressions* that match XML data by navigating down (and occasionally up or across) the tree.

Basic constructs (very incomplete list):

/	root element, or separator between steps in path
*	matches any one element name
@X	matches attribute X of the current element
//	matches any descendant of the current element
[C]	evaluates condition C on the current element
contains(s1,s2)	returns TRUE if string s1 contains string s2
name()	returns tag of the current element
parent::	matches the parent of the current element
descendants::	matches any descendant of the current element
self::	matches the current element

### XPath Exercise:

bookstore.dtd

```
<!ELEMENT Bookstore (Book | Magazine)*>
<!ELEMENT Book (Title, Authors, Remark)*>
<!ATTLIST Book ISBN CDATA #REQUIRED
              Price CDATA #REQUIRED
              Edition CDATA #IMPLIED>
<!ELEMENT Magazine (Title)>
<!ATTLIST Magazine Month CDATA #REQUIRED
                  Year CDATA #REQUIRED>
```

```
<!ELEMENT Title (#PCDATA)>
<!ELEMENT Authors (Author+)>
<!ELEMENT Remark (#PCDATA)>
<!ELEMENT Author (First_Name, Last_Name)>
<!ELEMENT First_Name (#PCDATA)>
<!ELEMENT Last_Name (#PCDATA)>
```

1. Find all book titles in our sample XML doc
2. Find all book or magazine titles
3. Find all books costing less than \$70
4. Find all books whose title contains one of its author's last names
5. Find all magazines where there is a book of the same title

## XQuery

- Main construct: **FLWOR expressions**
  - FOR Set up iterator variables (and specify documents)
  - LET Set up other vars, usually for aggreg or common subqueries
  - WHERE Filtering condition
  - ORDER BY column
  - RETURN What to return
- XPath can be used within all clauses.
- XPath and XQuery Comparison
  - XPath is more of a node addressing language than of a query language
  - XQuery is much more complicated and (some say) convoluted than any other XML-related language
  - XPath is supposed to be the basis for languages like XPointer, XLink, and, of course, XQuery
  - The return type of XPath is a set of nodes, whereas XQuery returns an XML tree
  - XPath cannot be used as an XML transformation language

## **XQuery Exercise**

1. All titles of books costing  $< \$70$  where Hellerstein is an author
2. All author Last\_Name's of books related to databases
3. Average price of all database books
4. All database books priced above average over all books
5. Titles and prices of all books, sorted by price.