

7.4 Data Storage: XML Files

- Extensible Markup Language
- Text files with a generic syntax
- One file that defines your specific XML syntax
 - DTD (Document Type Definition) file
 - or Schema
- One file to hold your actual data
- Advantages
 - There are already parsers for XML
 - It's a standard, there are lots of XML tools
 - DTD / Schema precisely defines your syntax
- Disadvantages
 - Not that easy to use an XML parser
 - Works with hierarchical data only

XML Example

```
<? xml version = "1.0" ?>
<!DOCTYPE addressbook SYSTEM "addressbook.dtd">
<addressbook>
  <person>
    <name> John C. Jack </name>
    <phone> 555-555-5555 </phone>
    <email> john@jack.com </email>
  </person>
  <person>
    <name> Fred Fredsson </name>
    <phone> 444-444-4444 </phone>
    <email> fred@nowhere.net </email>
  </person>
</addressbook>
```

XML Syntax

- Tags
 - opening and closing
 - `<name>`, `</name>`
- Elements
 - tags with data between them
 - `<name> John C. Jack </name>`
 - elements can contain elements: hierarchy
- Attributes
 - change the meaning of a tag
 - `<phone type="office">`
 - `<phone type="home">`
- Comments
 - `<!-- This is a comment -->`
- Version number

DTD

- Document Type Definition (DTD)
- Optional
 - Allows parsers to check XML files for validity
 - Sometimes replaced by schemas
 - use regular XML syntax
- DTD for our XML example

```
<? xml version="1.0" ?>
<!ELEMENT addressbook( person )* >
<!ELEMENT person( name, phone, email ) >
<!ELEMENT name( #PCDATA ) >
<!ELEMENT phone( #PCDATA ) >
<!ELEMENT email( #PCDATA ) >
```
- Syntax
 - * : repeated arbitrary times - #PCDATA: any string
 - | : either or

XML in the Java Library

- Part of the Java standard library since JDK 1.4
 - JAXP (Java API for XML Processing) for older versions
- org.xml.sax
 - Simple API for XML (SAX)
 - Reads in XML file sequentially
 - Generates a sequence of events
- org.w3c.dom
 - Document Object Model (DOM)
 - Reads in XML file as a whole
 - Generates a tree of objects

XML

- Standardized
 - standardized syntax
 - but no standardized semantics
- Human-readable
 - somewhat
- Used for all kinds of things
 - storage
 - data exchange
 - remote procedure calls
 - Web services
 - Web pages (XHTML)
- Hierarchical
 - may be a disadvantage
- Extensible
- Reusable Parsers

7.5 User Interfaces

- Interaction with a user
 - getting information to the user (output)
 - getting information from the user (input)
- Media / hardware
 - paper / printer
 - for interfacing with legacy bureaucracy
 - for signatures
 - for convenient reading
 - screen
 - color or black-and-white
 - resolution
 - fast enough for moving pictures?
 - audio
 - action sounds
 - speech
 - music and environmental sounds

Goals of User Interfaces

- Usable
 - how much effort to do a task?
 - example: often-used buttons easier to find
 - example: minimize need to switch between windows
- Learnable
 - how much effort to figure out how to do it?
- Immersive
 - you are thinking you are there (games)
- Adaptable
 - example: larger fonts for people with eye problems
- Internationalizable
 - translating into other languages / character sets
- Interactive
 - input and output happen concurrently