Chapter 16 – Part 2

The World Wide Web

"On the Internet, nobody knows you're a dog."

The New Yorker, Peter Steiner, July 5, 1993
Layers of a Computing System

Communication

Application

Operating System

Programming

Hardware

Information
Chapter Goals

• Compare and contrast the Internet and the World Wide Web

• Describe general Web processing

• Write basic HTML documents

• Describe several specific HTML tags and their purposes
Chapter Goals

• Describe the processing of Java applets and Java server pages
• Compare and contrast HTML and XML
• Define basic XML documents and their corresponding DTDs
• Explain how XML documents are viewed
• Hands-On HTML Coding
Figure 16.2 A browser retrieving a Web page
Structure of URIs

http://www.cs.wisc.edu/~dbbook/index.html

- URI has three parts:
  - **Naming schema** (http)
  - **Name of the host computer** (www.cs.wisc.edu)
  - **Name of the resource** (~dbbook/index.html)

- URLs are a subset of URIs
Hypertext Transfer Protocol

◆ What is a communication protocol?
  ▪ Set of standards that defines the structure of messages
  ▪ Examples: TCP, IP, HTTP

◆ What happens if you click on
www.cs.wisc.edu/~dbbook/index.html?

  ▪ Client (web browser) sends HTTP request to server
  ▪ Server receives request and replies
  ▪ Client receives reply; makes new requests
HTML

- Web pages are created (or built) using a language called the Hypertext Markup Language, or HTML.

- The term markup language comes from the fact that the primary elements of the language take the form of tags that we insert into a document to annotate the information stored there.
HTML Fundamentals

- <h1>An important heading</h1>
- <h2>A slightly less important heading</h2>
- <p>This is the first paragraph.</p> <p>This is the second paragraph.</p>
- This is a really <em>interesting</em> topic!
HTML Fundamentals

An important heading

A slightly less important heading

This is the first paragraph.

This is the second paragraph.

This is a really interesting topic!
Trees (cont’d)

- Hierarchical structures are called *trees*
- Binary trees
  - Each node has no more than two children
  - The beginning of the tree is a unique starting node called the *root*
  - The node to the left of a node, if it exists, is called its *left child*
  - The node to the right of a node, if it exists, is its *right child*
  - If a node in the tree has no children, it is called a *leaf node*

Figure 9.16  A binary tree
Linked Implementation

- **Linked implementation**  An implementation based on the concept of a *node*

- A node is made up of two pieces of information
  - the item that the user wants in the list, and
  - a pointer to the next node in the list
Figure 9.4 Anatomy of a linked list

User's data

Pointer to next item in the list

End of list

List implementation

Hofstra University – Overview of Computer Science, CSC005
Linked Implementation

Figure 9.5 An unsorted linked list
Images and Links

• Many tags can contain attributes that indicate additional details about the information or how the enclosed information should be displayed
  – An image can be incorporated into a web page using the IMG element, which takes an attribute that identifies the image file to display
  – `<IMG SRC = "myPicture.gif">`
Images and Links (cont.)

• A link is specified using the element A, which stands for anchor.

• The tag includes an attribute called **HREF** that specifies the URL of the destination document.

For example

```
<A HREF = "http://duke.csc.villanova.edu/docs/" >
Documentation Central!</A>
```
Interactive Web Pages

• When HTML was first developed, there was no way to interact with the information and pictures presented in a web page

• As users have clamoured for a more dynamic web, new technologies were developed to accommodate these requests

• Many of the new ideas were offshoots of the newly developed Java programming language
Java Applets

- **Java applet** A program that is designed to be embedded into an HTML document and transferred over the Web to someone who wants to run the program.

  An applet is embedded into an HTML document using the APPLET tag:

  ```html
  <APPLET code="MyApplet.class" width=250 height=150 ></APPLET>
  ```
Java Applets

• A browser has a **built-in interpreter** that executes the applet, allowing the user to interact with it.
  – Consider the difficulties inherent in this situation
  – How can we execute a program that was written on one type of computer on possibly many other types of computers?
Java Applets

• Java programs are compiled into **Bytecode**, a low-level representation of a program that is not the machine code for any particular type of CPU

• Java applets are restricted as to what they can do
  – The Java language has a carefully constructed **security** model
  – An applet, for instance, cannot access any local files or change any system settings
Java Server Pages

• A Java Server Page, or JSP, is a web page that has **JSP scriptlets** embedded in them

• **Scriptlet**  A small piece of executable code intertwined among regular HTML content

• Also called a **Servelet**!
Java Server Pages

• A JSP scriptlet is encased in special tags beginning with `<%` and ending with `%>

• Imagine JSP scriptlets as having the expressive power of a full programming language

```html
<H3>
<%
out.println("hello there");
%>
</H3>
```
Java Server Pages

• Note that JSPs are executed on the server side where the web page resides.

• By the time it arrives at your computer, all active processing has taken place, producing a static (though dynamically created) web page.

• JSPs are particularly good for coordinating the interaction between a web page and an underlying database.
XML

• HTML has a **predefined set of tags** and each tag has its own meaning

• There is nothing about HTML tags that describes the true content of a document

• The **Extensible Markup Language**, or **XML**, allows the creator of a document to describe its contents by **defining** his or her own set of tags
XML

• **Metalanguage** A language for talking about, or defining, other languages

• XML is a metalanguage
XML

• Like HTML, an XML document is made up of tagged data

```xml
<?xml version="1.0" ?>
<!DOCTYPE books SYSTEM "books.dtd">
<books>
  <book>
    <title>The Hobbit</title>
    <authors>
      <author>J.R.R. Tolkien</author>
    </authors>
    <publisher>Ballantine</publisher>
    <pages>287</pages>
    <isbn>0-345-27257-9</isbn>
    <price currency="USD">7.95</price>
  </book>
  <book>
    <title>A Beginner’s Guide to Bass Fishing</title>
    <authors>
      <author>J. T. Angler</author>
      <author>Ross G. Clearwater</author>
    </authors>
    <publisher>Quantas Publishing</publisher>
    <pages>750</pages>
    <isbn>0-781-40211-7</isbn>
    <price currency="USD">24.00</price>
  </book>
</books>
```

Figure 16.5 An XLML document containing data about books
XML

- **Document Type Definition (DTD)** A specification of the organization of the document

- The structure of a particular XML document is described by its corresponding DTD document

```xml
<!ELEMENT books (book*) >
<!ELEMENT book (title, authors, publisher, pages, isbn, price)>
<!ELEMENT authors (author+)>
<!ELEMENT title (#PCDATA)>
<!ELEMENT author (#PCDATA)>
<!ELEMENT publisher (#PCDATA)>
<!ELEMENT pages (#PCDATA)>
<!ELEMENT isbn (#PCDATA)>
<!ELEMENT price (#PCDATA)>
<!ATTLIST price currency CDATA #REQUIRED>
```

*Figure 16.6 The DTD document corresponding to the XML books document*
XML

- XML represents a standard format for organizing data without tying it to any particular type of output

- **Extensible Stylesheet Language (or XSL)** A language for defining transformations from XML documents to other output formats

Figure 16.7
An XML document can be transformed into many output formats
AJAX

- Asynchronous JavaScript and XML
- The latest craze in interactive pages
  - http://ajaxphp.packtpub.com/ajax/quickstart
  - http://labs.google.com/suggest
  - Gmail
  - http://maps.google.com
The Basics of HTML

- http://www.w3.org/MarkUp/Guide/ - Getting started with HTML
- http://www.w3.org/MarkUp/Guide/Advanced.html - More advanced features
- http://www.w3.org/MarkUp/Guide/Style.html - Adding a touch of style

- http://www.w3.org/People/Raggett/tidy/ - HTMLTidy
- http://validator.w3.org/ - Ensures proper HTML code
Other Useful Websites

- http://www.sitepoint.com/ - lots of helpful information and excellent books
- http://www.w3.org/XML/ - everything you ever wanted to know about XML
- http://www.w3.org/Style/XSL/ - definitive source for XSL
Welcome to Google Page Creator

Create your own web pages, quickly and easily.

Google Page Creator is a free online tool that makes it easy for anyone to create and publish useful, attractive web pages in just minutes.

- **No technical knowledge required.**
  Build high-quality web pages without having to learn HTML or use complex software.

- **What you see is what you’ll get.**
  Edit your pages right in your browser, seeing exactly how your finished product will look every step along the way.

- **Don’t worry about hosting.**
  Your web pages will live on your own site at http://yourgmailusername.googlepages.com

Google Page Creator is a Google Labs project, and is still in an early testing phase. If you’re interested in taking it for a test drive, login with your Gmail account to begin making pages. If you don’t have a Gmail account and you have a mobile phone, you can sign up here. We invite you to let us know what you think by sending us your feedback and suggestions.
Previous Homework

- Get a gmail account...
- ...if you don't want to use your mobile (or you don't have one), send an email to: papacosta@gmail.com
- When you get an account, send me a message! papacosta@gmail.com
- This becomes our 1\textsuperscript{st} Assignment!
Assignment One

- Create A **Simple** Home Page – 6/10 pts
- Create A **Complex** Home Page – 8/10 pts
  - Links, images, color fonts, multi-column layouts
- Create A **Multilevel Site** – 10/10 pts
  - Tree structure
  - Multiple pages

- **Due Next Wednesday**
  - No lateness
Homework

• Read Chapter Sixteen (Again!)

...be careful on the web!