JavaScript Context

INFO/CSE 100, Autumn 2004

Fluency in Information Technology

http://www.cs.washington.edu/100

Readings and References

- References
  - » Wikipedia articles on HTML, XML, XHTML and scripting languages

    http://en.wikipedia.org/wiki/HTML

    http://en.wikipedia.org/wiki/XML

    http://en.wikipedia.org/wiki/XHTML

    http://en.wikipedia.org/wiki/Scripting_language
  - » IBM Home Page Reader 3.0

    http://www-3.ibm.com/able/solution_offerings/hpr.html

Language Layers

- Sure, JavaScript is fun. But where does it fit in the larger picture?
- *Markup language*
  - » information structure and content, hyperlinking
- *Lightweight scripting language*
  - » dynamic creation of HTML, response to events
- *Heavyweight programming language*
  - » active graphics creation, numerical computation

Markup Language - HTML

- HyperText Markup Language
  - » a language for describing the *content* and *presentation* of a web page
    - content: The meaning and structure of the web page
    - presentation: How the page is to be presented
  - » HTML pages have a basic *hierarchical structure* defined by the tags
    - <html>, <head>, <body>, and so on
  - » Basic HTML describes a *static* page
    - once parsed and rendered, the page doesn't change
    - hyperlinking was the big step forward in basic HTML

Content vs. Presentation

- Early versions of HTML mixed these two ideas
  - » For example: `<p align="left">`
  - » The `<p>` tag identifies the basic structure of the page content
  - » The align="left" attribute guides the presentation
- Newer versions are separating content and presentation
  - » All "presentation attributes" of HTML elements were *deprecated* in HTML 4.01.
  - » All "presentation attributes" of HTML elements are *not supported* in XHTML 1.0 Strict DTD.
- This separation is a very useful distinction to make
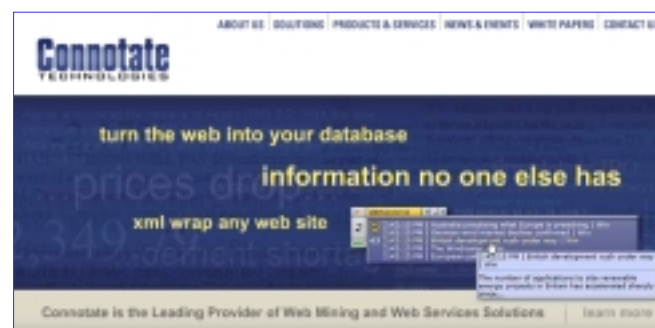 - » "what is the meaning?" vs "what is the format of the display?"

Why is this distinction useful?

- Once you have easy access to the structure and meaning, there are many useful ways to manage and present the information that is available
- Manage the information
  - » extract a travel map from a directions page
  - » filter an inventory list to select only certain attributes
- Present the information
  - » graphical browser on a full-size screen
  - » cell-phone browser display

Manage the information

- Information that is well structured can be
  - » filtered, reorganized, used as program input, presented in a variety of reports, ...
- This is what a database program is used for
  - » If we can separate the content from the presentation in web pages, we can use the web as a giant database
  - » Not a single database, but we can use it like one
- Web scraping, screen scraping
  - » extract the content and discard the presentation
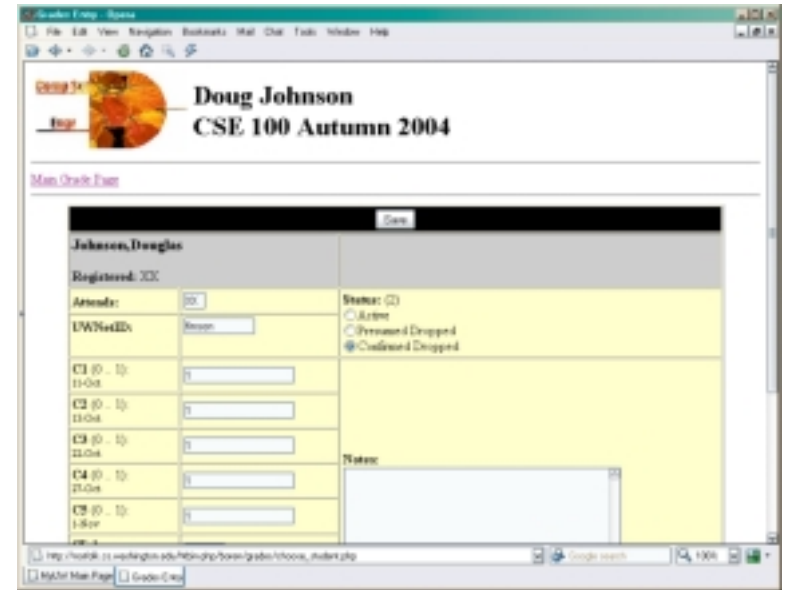
Web Scraping Example

New Brunswick, NJ, January 07, 2003 – Connotate Technologies, Inc., a leading provider of Web Mining solutions, today announced that Cinergy Marketing & Trading, L.P., a Houston affiliate of Cinergy Corp. has selected Connotate Technologies vTag Web Mining Server to **harvest energy market information from the web**. Using the server's automated monitoring, extraction, filtering, and delivery capabilities, Cinergy can now **access energy information from hundreds of Web sources**. ...
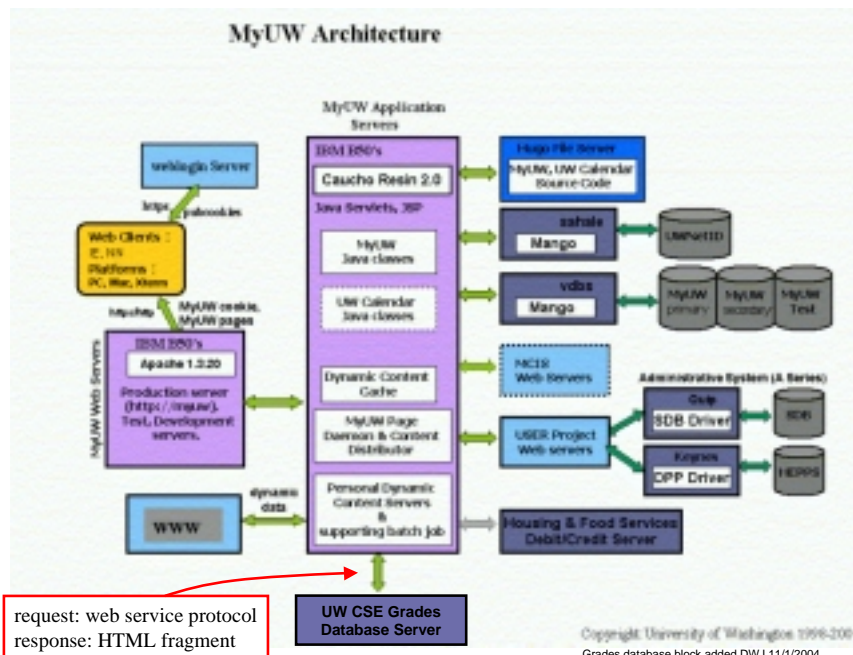
Structured Data Example

Structured Data Example

request: web service protocol
response: HTML fragment

Grades database block added DWJ 11/1/2004

Giving meaning to *all* web data

Semantic Web

The *Semantic Web* provides a common framework that allows *data* to be shared and reused across application, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. It is based on the Resource Description Framework (RDF), which integrates a variety of applications using XML for syntax and URIs for naming.

"The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation." -- Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American, May 2001

The Semantic Web is based on two fundamental concepts:
• The description of the meaning of the content in the Web
• The automatic manipulation of these meanings

Present the information

- Once you have a database with a defined structure, you can present the information it contains in many different ways
- Define a style for presenting the data
  - » Cascading Style Sheets associate style with structure
- Use a browser appropriate to the requirements
  - » desktop graphical: MSIE, Mozilla, Opera, Safari
  - » desktop text-only: Lynx, Links, WannaBe
  - » Audio: IBM Home Page Reader 3.0
  - » PDA and cell phone: Opera, WAP browsers, Pocket IE
- The information you want exists, but presenting it in a useful, timely, accessible fashion is a real challenge

Language Layers

- Sure, JavaScript is fun. But where does it fit in the larger picture?
- *Markup language*
  - » information structure and content, hyperlinking
- *Lightweight scripting language*
  - » dynamic creation of HTML, response to events
- *Heavyweight programming language*
  - » active graphics creation, numerical computation

Lightweight Scripting Language

- JavaScript is a *scripting language*
  - » has many features of larger languages but intended for more casual or rapid development
- Such languages are used to provide dynamic control of the content and display
  - » intended to be used by web page developers, system administrators, and others to quickly develop useful applications

Other Scripting Languages

- **PHP** is an HTML-embedded scripting language. Much of its syntax is borrowed from C, Java and Perl with a couple of unique PHP-specific features thrown in. The goal of the language is to allow web developers to write dynamically generated pages quickly.
- **Python** is an interpreted, interactive, object-oriented programming language. It is often compared to Tcl, Perl, Scheme or Java. Python is also usable as an extension language for applications that need a programmable interface.
- **Perl** is a language optimized for scanning arbitrary text files, extracting information from those text files, and printing reports based on that information. It's also a good language for many system management tasks. The language is intended to be practical (easy to use, efficient, complete) rather than beautiful (tiny, elegant, minimal).
- **VBScript**. Microsoft Visual Basic Scripting Edition brings active scripting to a wide variety of environments, including Web client scripting in Microsoft Internet Explorer and Web server scripting in Microsoft Internet Information Service.

Dynamic HTML

- Dynamic HTML is a technique of creating interactive web sites by using a combination of
  - » the static markup language HTML
  - » the style definition language Cascading Style Sheets
  - » a client-side scripting language
- A client-side scripting language provides
  - » Adaptivity … the page can be customized
  - » Interactivity … get user input and display results
  - » Applications … build a general purpose program

Language Layers

- Sure, JavaScript is fun. But where does it fit in the larger picture?
- *Markup language*
  - » information structure and content, hyperlinking
- *Lightweight scripting language*
  - » dynamic creation of HTML, response to events
- *Heavyweight programming language*
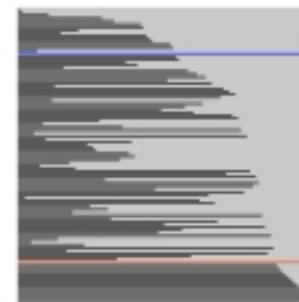 - » active graphics creation, numerical computation

Larger Programming Languages

- Web browsers are handy, but many applications are large and need specialized capabilities
- Large applications are generally written in full featured programming languages like
  - » Java, C++, C#, Fortran
  - » These languages come with extensive function libraries to support extended networking, graphics, data structures, etc.
- Many new languages straddle the line and are appropriate for both scripting and large-scale development

Recall the Sort Demo

- The code that actually does the sorting and graphics is written in Java.
- The programs are compiled and run as *applets*. They use the web page area for display, but are not working with HTML.

Bubble Sort   **Quick Sort**

http://java.sun.com/applets/jdk/1.0/demo/SortDemo/example1.html

Summary

- HTML forms the basic structure for web pages
  - » Modern trends are to separate
    - the structural description of the data
    - the presentation of the data to the user
  - » process the information according to the content
  - » display the information as appropriate to the user
- Programming languages make information management and display dynamic
  - » languages come in many flavors but they share many basic concepts and much of their basic syntax