CSP2103: Markup Languages

Lecture 1: Introduction to Markup Systems and Languages

Objectives

- Review the history of the Web, the Internet, and HTML
- Describe different HTML standards and specifications
- Learn about the basic syntax of HTML code
- Mark elements using two-sided and one-sided tags
- Insert an element attribute
- Create comments
- Describe block-level elements and inline elements
- Specify an element's appearance with inline styles

Objectives

- Create and format different types of lists
- Create boldfaced and italicized text
- Describe logical and physical elements
- Define empty elements
- Insert an inline image into a Web page
- Insert a horizontal line into a Web page
- Store meta information in a Web document
- Display special characters and symbols

The World Wide Web

- A network is a structure linking computers together for the purpose of sharing resources such as printers and files
- Users typically access a network through a computer called a host
- A computer that makes a service available to a network is called a server
- A computing device that requests and receives content is a client

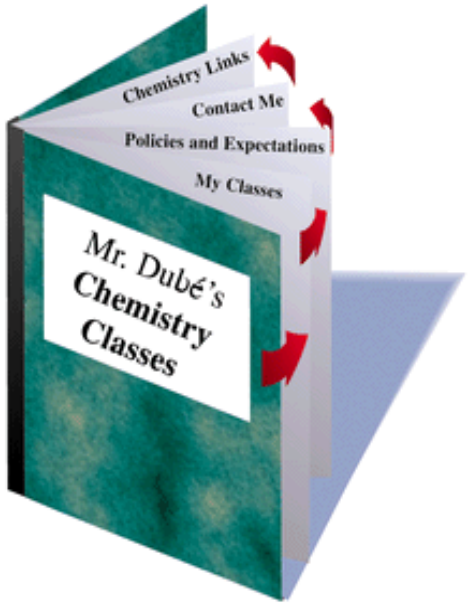
The World Wide Web

- A computer or other device that requests services from a server is called a client
- One of the most common network structures is the client-server network
- Typically, markup such as HTML, XML and XSLT take place on a client device (user end)
- Scripting systems such as PHP, ASP, Java run on the server (provider) end

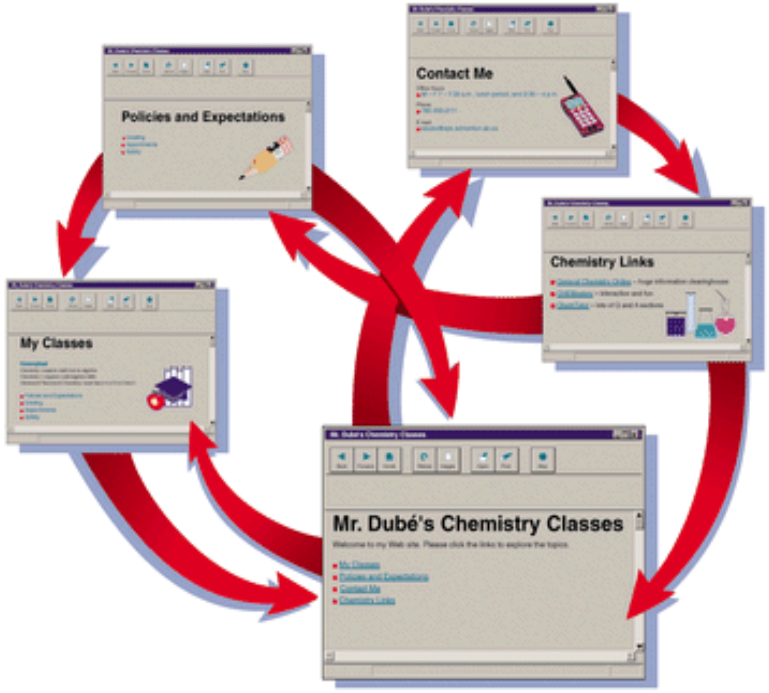
The Development of the Word Wide Web

- Sir Tim Berners-Lee and other researchers at the CERN nuclear research facility near Geneva, Switzerland laid the foundations for the World Wide Web, or the Web, in 1989
- They developed a system of interconnected hypertext documents that allowed their users to easily navigate from one topic to another via a graphical rather than purely text interface
- Hypertext is a method of organizing information that gives the reader control over the order in which the information is presented (ie non-linear)

Linear vs Hypertext

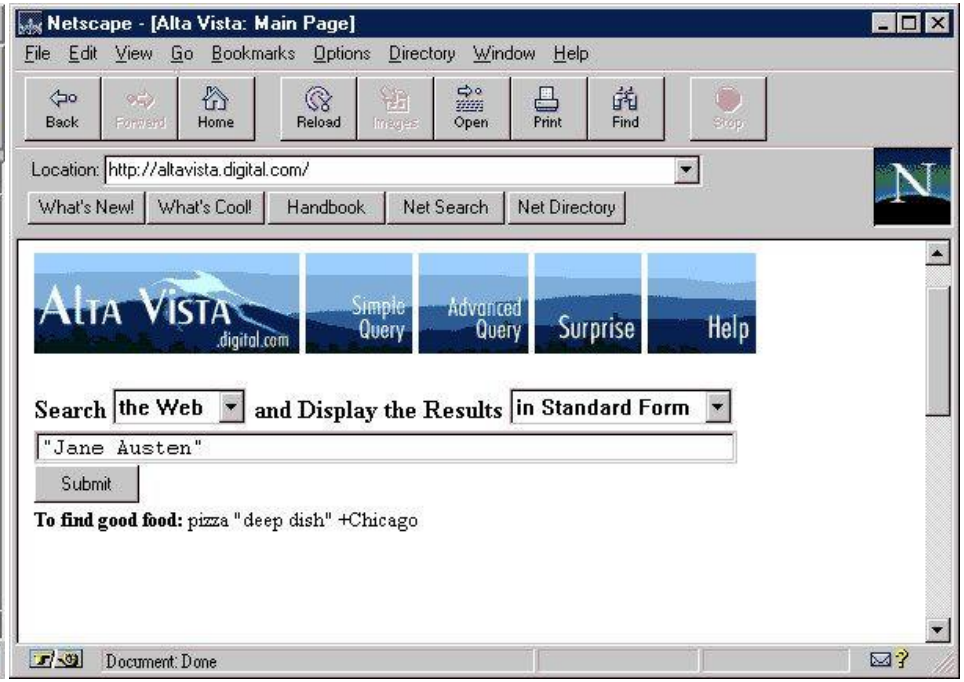


Reading a linear document



Reading a hypertext document

A Blast From the Past



http://www.techpaparazzi.com/how-web-browsers-looked-from-in-90s/

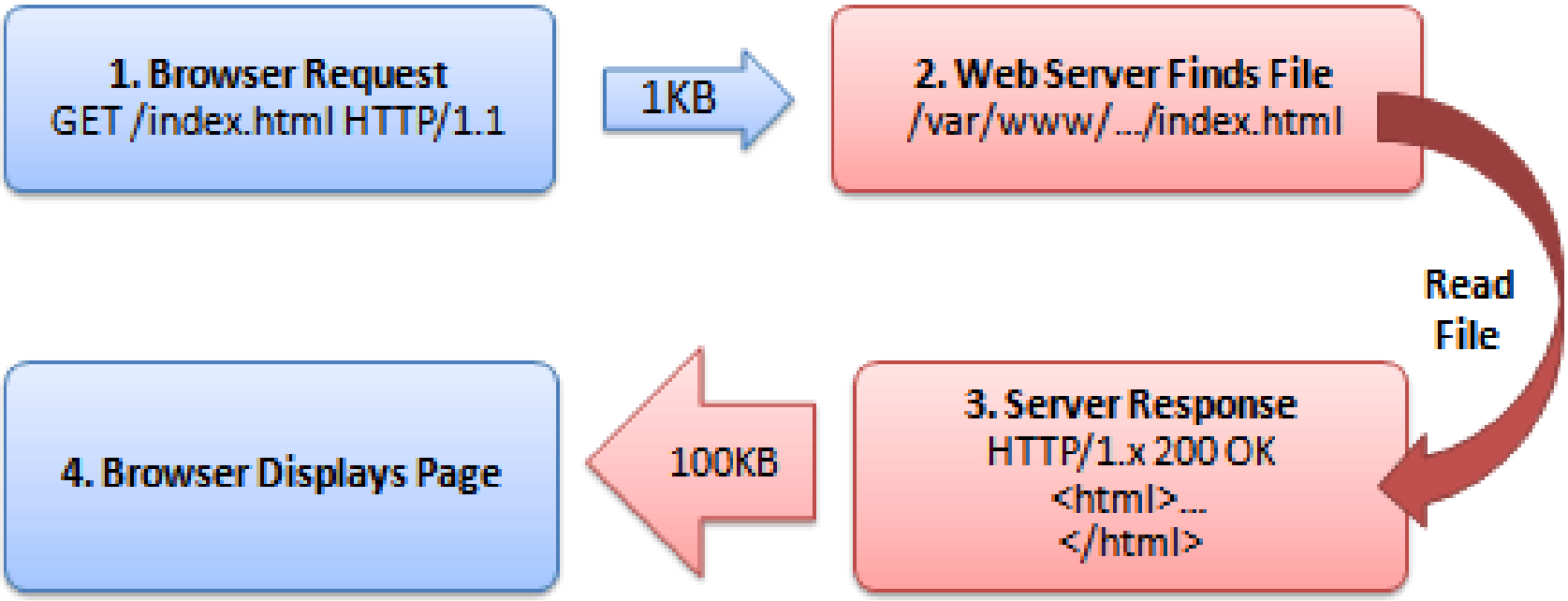
Hypertext Documents

- The key to hypertext is the use of hyperlinks (or links) which are the elements in a hypertext document that allow you to jump from one topic to another
- A link may point to another section of the same document, or to another document entirely
- A link can open a document on your computer, or through the Internet, a document on a computer anywhere in the world
- As well as links within text, images and other objects can be linked to other objects, pages or text within pages

Web Servers and Web Browsers

- A Web page is stored on a Web server, which in turn makes it available to a local or wide area network
- A web browser requests a web page via the Uniform Resource Locator (URL) in the address bar
- The response from the server is returned as mark-up text in the form of HTML
- The Web browser interprets these HTML instructions and renders the web page to screen accordingly

HTTP Request and Response



http://stackoverflow.com/questions/4109689/how-does-a-client-browser-generate-a-request-to-be-sent-to-a-server

HTML: The Language of the Web

- A Web page is a text file written in a language called Hypertext Markup Language
- A markup language is a language that describes a document's structure and content
- Markup has been around in varying forms for decades
- Early word processors used document markup to define the look ad style of a document
- HTML is not a programming language or a formatting language
- Styles are format descriptions written in a separate language from HTML that tell browsers how to render each element. Styles are used to format your document

HTML History

- A group of Web developers, programmers, and authors called the World Wide Web Consortium, or the WC3, created a set of standards or specifications that all browser manufacturers were to follow
- The WC3 has no enforcement power
- The recommendations of the WC3 are usually followed since a uniform approach to Web page creation is beneficial to everyone
- However, in the mid 1990's Microsoft and Netscape slugged it out in the 'browser wars', with HTML compatibility and features as the weapons
- Browser compatibility across devices (PC's, Phones, Tablets) still an ongoing issue

HTML Versions

HTML Version	Year Of Release
HTML	1991
HTML+	1993
HTML 2.0	1995
HTML 3.2	1997
HTML 4.01	1999
XHTML 1.0	2000
HTML 5	2012
XHTML5	2013

http://allwebprograming.blogspot.com.au/2013/12/what-is-html-hyper-text-markup-language.html

Evolution of Markup

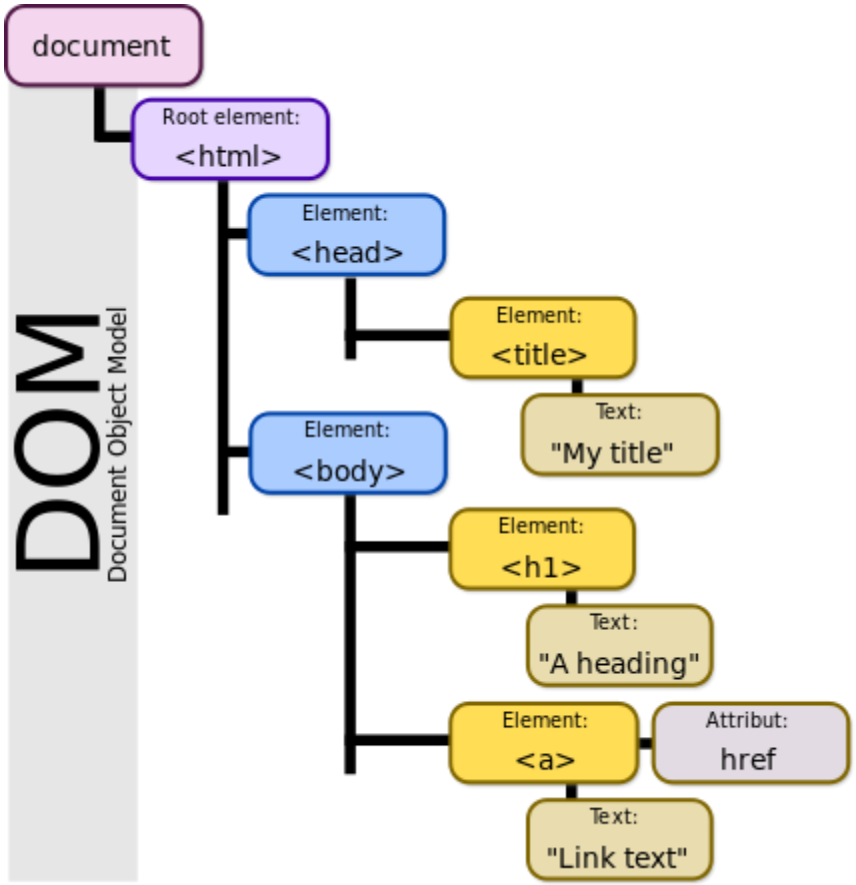
- Older features of HTML are often deprecated, or phased out, by the W3C. That does not mean you cannot continue to use them—you may need to use them if you are supporting older browsers
- Modern Web development is focusing increasingly on two main languages: XML and XHTML
- XML (Extensible Markup Language) is a metalanguage that allows developers to create their own descriptive markup language
- XML on its own is actually extremely easy
- Making XML useful, by transforming logical structure into dynamically formatted document, that is the hard bit

Evolution of Markup

- XHTML (Extensible Hypertext Markup Language) is a stricter version of HTML and is designed to confront some of the problems associated with the different and competing versions of older HTML
- XHTML is also designed to better integrate HTML with XML
- HTML will not become obsolete anytime soon (some browsers still support back to HTML version 3.2)
- There was a 12 year gap between XHTML 1.0 and HTML5

HTML Document Object Model (DOM)

- A HTML document is broken down into different pieces'
- Each piece is an 'object'
- An Element is a paired set of tags, such as
  ** **this is a paragraph** **
- An attribute is attached to an Element to extend functionality, such as
- <p **align**="center">this is a paragraph



http://www.makeuseof.com/tag/javascript-web-development-using-document-object-model/

Marking Elements with Tags

- A two-sided tag's opening tag ( ) and closing tag ( ) should completely enclose its content
- HTML allows you to enter element names in both uppercase or lowercase letters (lower case is considered standards compliant)
- A one-sided tag contains no content; general syntax for a one-sided tag:
- <element /> - such as <br /> or <hr />
- Also known as an empty element

Element Attributes

- As mentioned, Elements can contain attributes that control the behaviour, and in some cases the appearance, of elements within the page
- Attributes are inserted within the tag brackets
- <element **attribute1="value1" attribute2="value2" …**/> for one-side tags
- <element **attribute1="value1" attribute2="value2" …**>content</element> for two-sided tags

Whitespace and HTML

- HTML file documents are composed of text characters and white space
- White space is the blank space, tabs, and line breaks within the file
- HTML treats each occurrence of white space as a single blank space
- In HTML you can force multiple whitespaces with
- Best to use CSS to control formatting as forced whitespaces make for unpredictable output

HTML Basic Structure - <head>

- An HTML document is divided into two parts: the head and the body
- The <head> element contains information about the document, for example the document title or the keywords
- The content of the head element is not displayed within the Web page
- The title element contains the page's title; a document's title is usually displayed in the title bar

HTML Basic Structure - <body>

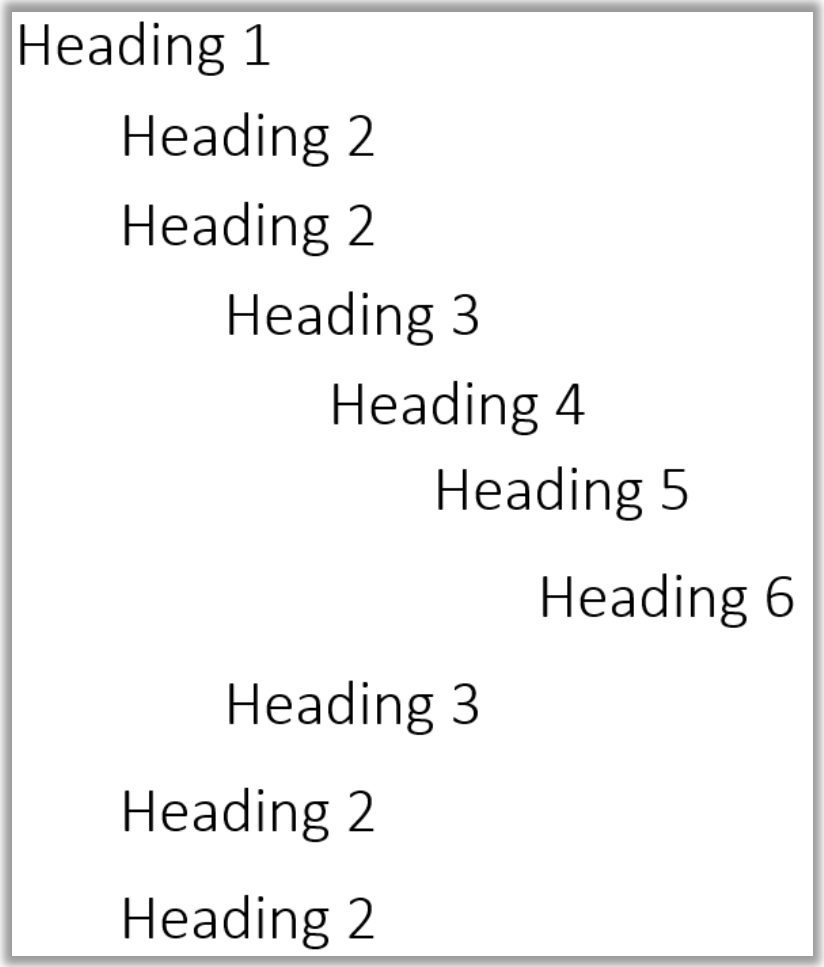
- The body element contains all of the content to be displayed in the Web page
- The body element contains all the other elements and attributes that tells the browser how to render the content to screen
- Sections of content can be sub-divided into logical sections inside the body of the HTML document using the … element

Displaying the HTML File

- As you work on a Web page, you should regularly view it with your Web browser to verify that the file contains no syntax errors or other problems
- You should alwats view the results using different browsers to check for compatibility
- If cross browser compatibility is important, you may need to check browsers on different operating system platforms, such as Windows, Apple, Linux, iOS and Android

Understanding Structure

- HTML has 6 Heading levels (1-6)
- Each logically sits within the one above it
- Lazy developers often use the different heading numbers to increase/decrease font size
- In fact, the headings describe the logical structure of the document in exactly the same way as a word processor does with heading styles



Styling with CSS

- Use Cascading Style Sheets to control the appearance of an element, such as text alignment
- Styles specified as attributes in a tag are also referred to as inline styles
- The text-align style tells the browser how to horizontally align the contents of an element
- Presentational attributes specify exactly how the browser should render an element
- Styles are extremely important for creating consistency, so that headings, text and other repeating formatting looks the same throughout a series of web pages
- CSS styling should be default rather than trying to use HTML's limited display formatting options

Understanding 'Logical' and 'Physical' Elements

- A logical element describes the nature of the enclosed content, but not necessarily how that content should appear
- A physical element describes how content should appear, but doesn't indicate the content's nature
- You should use a logical element that accurately describes the enclosed content whenever possible, and use physical elements only for general content
- When you start integrating XML and XHTML via XSLT, the true value of the above statement will become much clearer

Special Characters

- Occasionally you will want to include special characters in your Web page that do not appear on your keyboard

£ ®

- HTML supports the use of character symbols that are identified by a code number or name

&code

Some Common Examples

Symbol	Code	Name	Description
©	&#169;	&copy;	Copyright symbol
®	&#174;	&reg;	Registered trademark
•	&#183;	&middot;	Middle dot (bullet)
°	&#176;	&deg;	Degree symbol
			Nonbreaking space, used to insert consecutive blank spaces
<	&#60;	&lt;	Less than symbol
>	&#62;	&gt;	Greater than symbol
&	&#38;	&amp;	Ampersand

- Always check which markup environment you are using HTML, XML, XSLT) and its support for special characters

Tools of the Trade

- What You See Is What You Get (WYSIWYG) development tools give you a visual representation of the code you are developing (such as Microsoft's Visual Studio and Adobe's Dreamweaver)
- HTML or Text Editors give you a code view of the work
- Most editors can render the HTML code inside the editor for quick previews / testing
- In this unit, we will be using code only HTML editors to develop our code
- There are 100's of editors available, many for free

Conclusion

- This been a very quick intro into the basics of the World Wide Web and HTML
- Over the coming weeks we will dive into HTML, CSS, XML and then XSLT
- We will then bring all of these technologies and more together into integrated web pages and websites across assignment one and two
- This unit has heaps of code examples with more being added all the time
- Play with them, dissect them and find out what makes them tick
- Example based learning with be critical to your future employability in the ICT industry

ANY QUESTIONS?

- If so, please ask now
- Or email me your question
- Or post your question to the discussion board