Lecture 5
HTML

Boriana Koleva
Room: C54
Email: bnk@cs.nott.ac.uk

Overview

- Origins and evolution of HTML and XHTML
- Basic Syntax
- Standard document structure
- Basic text markup
- Images
- Hypertext links
- Lists

Origins and Evolution of HTML

- Hypertext Markup Language
- Developed for the delivery of hypertext on the WWW
- Built using SGML
- ASCII "Markup Language"

Recent Versions

- HTML 4.0 – 1997
  - Introduced many new features and deprecated many older features
- HTML 4.01 - 1999 - A cleanup of 4.0
- XHTML 1.0 - 2000
  - Just 4.01 defined using XML, instead of SGML
- XHTML 1.1 – 2001
  - Modularized 1.0, and drops frames
- XHTML 2.0 – development abandoned
- HTML 5.0
  - Working Draft (not a W3C Recommendation yet)
  - HTML and XHTML syntax

HTML versus XHTML

- HTML has lax syntax rules, leading to sloppy and sometime ambiguous documents
- XHTML syntax is much more strict, leading to clean and clear documents in a standard form
- Even minor syntax errors will prevent a document labelled as XML from being rendered fully, whereas they would be ignored in the HTML syntax
- HTML compatible with most legacy Web browsers

Editing (X)HTML

- Creating HTML documents
  - Text editors (e.g. Notepad, Emacs, Crimson Editor)
  - Source code editors (e.g. Notepad++, WebTide)
  - Authoring tools (e.g. Microsoft Expression Web, Adobe Dreamwaver, KompoZer)
- HTML files have .html extension
- The filename of your homepage should be index.html
  - If a browser does not request a specific file in a directory, normal Web server response is to return index.html
  - http://www.crg.cs.nott.ac.uk/~bnk/Teaching/WPS/

Basic Syntax

- Elements are defined by tags (markers)
- Tag format:
  - Opening tag: <name>
  - Closing tag: </name>
- The opening tag and its closing tag together specify a container for the *content* they enclose
  - E.g. <p> Hello </p>
- Not all tags have content
  - E.g. <hr>

Basic Syntax 2

- The container and its content together are called an *element*
- If a tag has attributes, they appear between its name and the right bracket of the opening tag
  - E.g.
- Comment form: <!-- … -->
- Browsers ignore comments, unrecognizable tags, line breaks, multiple spaces, and tabs
- Tags are just suggestions to the browser, even if they are recognized by the browser

HTML Document Structure

- An HTML document is composed of 3 parts:
1. a line containing HTML version information, e.g.:
```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
```
   Note: `<!DOCTYPE html>` for HTML5
2. a declarative header section
   - Delimited with the `<head>` tag
   - The `<title>` tag is used to give the document a title
     - normally displayed in the browser's window title bar
3. a body containing the document's actual content
   - Delimited with the `<body>` tag
- After document type declaration, the remainder of an HTML document is contained by the `html` element

Basic HTML Document

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">

<html>
  <head>
      <title> … </title>
      <meta http-equiv="Content-Type"
            content="text/html;charset=utf-8" >
  </head>
  <body>
      …
  </body>
</html>
```

Basic Text Markup

- Paragraph Elements: `<p>`
  - Text is normally placed in paragraph elements
  - The browser puts as many words of the paragraph's content as will fit in each line
  ```
  <p>
    Greetings!
  </p>
  ```
  - Simple HTML example
    http://www.cs.nott.ac.uk/~bnk/WPS/simple.html

Basic Text Markup 2

- Line breaks: `<br>`
- Horizontal rules: `<hr>`
- Headings
  - Six sizes, 1 - 6, specified with `<h1>` to `<h6>`
    - 1, 2, and 3 use font sizes that are larger than the default font size
    - 4 uses the default size
    - 5 and 6 use smaller font sizes
  - Headings example
    http://www.cs.nott.ac.uk/~bnk/WPS/headings.html

Basic Text Markup 3

- Blockquotes: `<blockquote>`
  - To set a block of text off from the normal flow and appearance of text
  - Browsers often indent, and sometimes italicize
  - http://www.cs.nott.ac.uk/~bnk/WPS/blockquote.html
- Font Styles and Sizes (can be nested)
  - Boldface: `<b>`
  - Italics: `<i>`
  - Smaller: `<small>`
  - Larger: `<big>` (not supported in HTML5)
  - Monospace: `<tt>` (not supported in HTML5)

Basic Text Markup 4

Example: `The <big> sleet <big> in <big> <i> Crete </i><br /> lies </big> completely </big> in </big> the street`

Display: The sleet in *Crete* lies completely in the street

- Subscripts: `<sub>` Superscripts: `<sup>`

Example: `x<sub>2</sub><sup>3</sup>`

Display: x_2^3

Basic Text Markup 5

- Character Entities

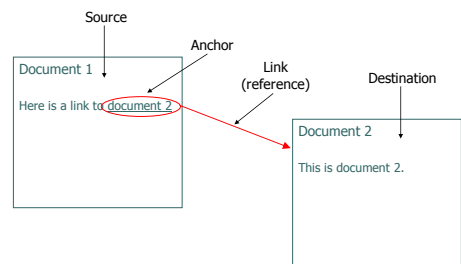| Character	Entity	Meaning	
&	`&`	Ampersand	
<	`<`	Less than	the only named character entity references in HTML5
>	`>`	Greater than	
"	`"`	Double quote	
'	`'`	Single quote (apostrophe)	
$\frac{1}{4}$	`¼`	One quarter	
$\frac{1}{2}$	`½`	One half	
$\frac{3}{4}$	`¾`	Three quarters	
°	`°`	Degree	
(space)	` `	Nonbreaking space	

Images

- All modern web browsers can display images inline (i.e. embedded in the text)
- GIF (Graphic Interchange Format)
  - 8-bit color (256 different colors)
- JPEG (Joint Photographic Experts Group)
  - 24-bit colour (16 million different colours)
- Portable Network Graphics (PNG)
  - Relatively new
  - Designed for transferring images on the Internet

Images 2

- Images are inserted into a document with the `<img>` tag with the `src` attribute
- The `alt` attribute is required by HTML
  - (in HTML5 can be omitted when textual desc. not available)
  - Non-graphical browsers
  - Browsers with images turned off
  - `<img src = "logo.jpg"`
  - `alt = "University of Nottingham Logo" />`
- The `<img>` tag has other optional attributes, including `width` and `height` (in pixels)

http://www.cs.nott.ac.uk/~bnk/WPS/image.html

Linking on the Web

Source

Anchor

Link (reference)

Destination

Document 1

Here is a link to document 2

Document 2

This is document 2.

Hypertext Links

- Hypertext is the essence of the Web!
- A link is specified with the `href` (hypertext reference) attribute of `<a>` (the anchor tag)
- The content of `<a>` is the visual link in the document
  `<a href="target.html">This is a link</a>`
- Relative addressing of targets is easier to maintain and more portable than absolute addressing

http://www.cs.nott.ac.uk/~bnk/WPS/link.html

Targets within Documents

- If the target is not at the beginning of the document, the target spot must be marked
- Target labels can be defined in many different tags with the id attribute
  `<h1 id = "baskets"> Baskets </h1>`
- The link to an `id` must be preceded by a pound sign (#)
  - target is in the same document,
  `<a href = "#baskets"> Baskets </a>`
  - target is in a different document
  `<a href = "myAd.html#baskets"> Baskets </a>`

Image Hyperlinks

- Links can include images in their content

```
<a href = "c210data.html">
   <img src = "smallplane.jpg"
    alt = "Small picture of an airplane " >
      Info on C210 </a>
```
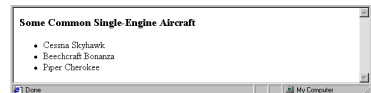

Unordered Lists

- The list is the content of the `<ul>` tag
- List elements are the content of the `<li>` tag
  ```
  <h3> Some Common Single-Engine Aircraft </h3>
  <ul>
      <li> Cessna Skyhawk </li>
      <li> Beechcraft Bonanza </li>
      <li> Piper Cherokee </li>
  </ul>
  ```

  Some Common Single-Engine Aircraft
  - Cessna Skyhawk
  - Beechcraft Bonanza
  - Piper Cherokee

  Done   My Computer
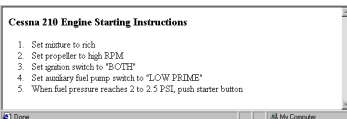

Ordered Lists

- The list is the content of the `<ol>` tag
- Each item in the display is preceded by a sequence value

```
<h3> Cessna 210 Engine Starting Instructions </h3>
<ol>
 <li> Set mixture to rich </li>
 <li> Set propeller to high RPM </li>
 <li> Set ignition switch to "BOTH" </li>
 <li> Set auxiliary fuel pump switch to
      "LOW PRIME" </li>
 <li> When fuel pressure reaches 2 to 2.5
      PSI, push starter button </li>
</ol>
```

Cessna 210 Engine Starting Instructions
1. Set mixture to rich
2. Set propeller to high RPM
3. Set ignition switch to "BOTH"
4. Set auxiliary fuel pump switch to "LOW PRIME"
5. When fuel pressure reaches 2 to 2.5 PSI, push starter button

Done   My Computer

Nested Lists

- Any type list can be nested inside any type list
- The nested list must be in a list item

```
<ol>
  <li> Single-Engine Aircraft
    <ol>
      <li> Tail wheel </li>
      <li> Tricycle </li>
    </ol> <br>
  </li>
  <li> Dual-Engine Aircraft
    <ol>
      <li> Wing-mounted engines </li>
      <li> Push-pull fuselage-mounted engines </li>
    </ol>
  </li>
</ol>
```

1. Single-Engine Aircraft
   1. Tail wheel
   2. Tricycle
2. Dual-Engine Aircraft
   1. Wing-mounted engines
   2. Push-pull fuselage-mounted engines

http://www.cs.nott.ac.uk/~bnk/WPS/nested_lists.html

Definition Lists (for glossaries)

- List is the content of the `<dl>` tag
- Terms being defined are the content of the `<dt>` tag
- The definitions themselves are the content of the `<dd>` tag

```
<dl>
  <dt> 152 </dt>
  <dd> Two-place trainer </dd>
  <dt> 172 </dt>
  <dd> Smaller four-place airplane </dd
</dl>
```

152
Two-place trainer
172
Smaller four-place airplane

Validation

- W3C HTML Validation Service
  - http://validator.w3.org/

Syntactic Differences between HTML & XHTML

- Case sensitivity
- Closing tags
- Quoted attribute values
- Explicit attribute values
- id and name attributes
- Element nesting

Summary

- Origins and evolution of HTML and XHTML
- Basic syntax and standard document structure
- Basic text markup
- Images
- Hypertext links
- Lists (unordered, ordered, definition)
- Validation
- HTML vs. XHTML syntax