Overview of lecture

- Big picture
- Text basics
- Forms
- Tables
- Miscellanea

HyperText Markup Language

- Disagreement about HTML’s role
  - Only give the content and structure of the document, leave visualization to the browser
  - Browsers vary (graphical, text based, mobile devices)
  - User preferences vary (some people like larger fonts)
  - Environment varies (screen sizes, fonts available, etc.)
  - But authors want to control what the document looks like
- Trend towards separating content from presentation
  - Cascading Style Sheets – presentation information only
  - HTML documents contain little formatting
Current state of the standards

- W3C (World Wide Web Consortium) sets standards
- Last HTML standard 4.01 (December 1999)
- XHTML 1.0 new XML-based format
  - XML (extensible markup language) – focuses on semantics and is used as general purpose format for structured data
  - A document called DTD or XML Schema defines what tags and attributes are allowed in an XML document
- For the programming assignments (and beyond) use W3C’s HTML validator service
  http://validator.w3.org/

Writing web pages

- There are many ways of presenting the same information in a web page
- Must balance conflicting considerations
  - Standards-compliance
  - Aesthetics for target platform
  - Performance
    - Large files take longer to download
    - Cost of running code on server
    - Cost of running code on client

Elements of web documents

- Text – browsers treat spaces, tabs and new lines in HTML source the same, wrap text
- Comments – between <!-- and --> not rendered
- Various tags
  - <tag attribute1="value" attr2="val2">text</tag>
  - For some tags the end tag not needed
    - In HTML they look like <tag> in XML <tag/>
  - All tags must obey proper nesting rules
    - Valid: <i><b>Bold italic text</b></i> italic text</i>
    - Invalid: <i> <b> </i> </b>
A valid web page

```html
<html>
<head>
<title>Bucky Badger's web page</title>
</head>
<body>
<h1>Welcome to Bucky's web page</h1>
<p>I am Bucky, the mascot for University of Wisconsin athletics. Please visit <a href="http://www.uwbadgers.com/football/index.html">the web page of our football team</a> and <a href="http://www.uwbadgers.com/basketball/index.html">the web page of our basketball team</a>.</p>
</body>
</html>
```

Overview of lecture

- Big picture
- Text basics
- Forms
- Tables
- Miscellanea

Tags controlling appearance

- Content-based styles
  - `<em>` emphasis
  - `<strong>` strong emphasis
  - `<code>` program code
  - `<address>` an address
  - `<cite>` citation

- Physical styles
  - `<b>` bold
  - `<i>` italic
  - `<tt>` monospaced
  - `<big>`/`<big>` bigger
  - `<small>`/`<small>` smaller
  - `<sub>`/`<sub>` subscript
  - `<sup>`/`<sup>` superscript
  - `<font>`/`<font>` don’t use it
Tags controlling text flow

- `<br>` introduces a line break (new line)
- `<p>` defines a paragraph
  - Browser leaves empty line before paragraph
  - Closing tag often omitted
- `<pre>` preformatted text
  - Line breaks and indentation not ignored
  - Browser uses monospaced font

Special characters

- Characters "<", ">", and "&" cannot appear in document text (confusing to browser)
  - `&lt;` `<` `&gt;` `>` `&amp;` `&`
- Numeric codes can be used for these and other characters (greek letters, math, etc.)
  - `&#60;` `<` `&#62;` `>` `&#38;` `&`
  - `&copy;` © `&reg;` ® `&nbsp;` non-breaking space
  - [http://www.w3.org/TR/html401/sgml/entities.html](http://www.w3.org/TR/html401/sgml/entities.html)

Other tags

- Headings `<h1>` to `<h6>`
- Lists can be nested
  - `<ul>` bulleted (unordered) lists
  - `<ol>` numbered (ordered) list
    - The "start" attribute controls the label of the first item
    - The "type" attribute (A,a,l,l,1) controls whether numbers, letters or roman numerals (upper/lower case) are used
  - `<li>` list item (closing tag often omitted)
Overview of lecture

- Big picture
- Text basics
- Forms
- Tables
- Miscellanea

About forms

- Forms are the traditional way for users to send information to a web server
  - The user fills out fields in the browser
  - The user submits the form
  - http carries the user input to the web server
  - A server side program processes the user data
  - The server sends a reply document to the client
The <form></form> tag

- Forms use the <form></form> tag
  - The “action” attribute has to URL of the server side program processing the form
  - The “method” attribute has to be get or post
- Inside this tag there are
  - Various controls collecting data from the user
    - Text fields, check boxes, file selection controls, etc.
    - Normal text

<input> controls

- The “type” attribute determines what they do
  - “text” used for text fields
  - “submit” the button submitting the form “reset” for resetting
  - “checkbox” checkboxes the users can select
  - “radio” multiple choice radio buttons
  - “hidden” submitted but not shown to user
- The “name” attribute allows the program to distinguish between controls (e.g. age or weight?)
- The “value” attribute specifies the text on the control and/or the value sent to server
Other controls

- `<textarea></textarea>` used for multi-line text
  - Must use the “name” attribute
  - Has “cols” and “rows” attributes specifying size
- The `<select></select>` tag encloses `<option>` tags and gives multiple choices to user
  - The “multiple” attribute controls whether the user can check one or more options

Overview of lecture

- Big picture
- Text basics
- Forms
- Tables
- Miscellanea

About tables

- Their original role was to display tables
- Their most prevalent use is for controlling the placement of visual elements on the page
  - Frames control placement too – don’t use them
- The table is a collection of rows
- The rows are collections of cells
- Cells on the same row/column are aligned
- Cells can contain anything (even other tables)
The **<table></table>** tag

- Defines a table
- The “border” attribute defines the width of the lines used to draw the table (in pixels)
  - Defaults to 0 which means no lines are drawn
- The “width” attribute controls table width
  - By default it is in pixels
  - It can be given as a percentage of the window
  - If not specified, the table is only as wide as needed to display cell contents

The **<tr></tr>** tag

- Defines a table row
- The “align” attribute controls horizontal alignment of text in cells – can be “left”, “right”, “center”
- The “valign” attribute controls vertical alignment of text in cells – can be “top”, “bottom”, “middle”
- The “nowrap” attribute instructs the browser not to wrap the text from within the cells

The **<td></td>** tag

- Defines a table cell
  - Has “align”, “valign” and “nowrap” attributes
  - “width” can be given as percentage of table width
  - “height” gives minimum height for cell
  - “colspan” allows a cell to span multiple columns
  - “rowspan” allows a cell to span multiple rows
Overview of lecture

- Big picture
- Text basics
- Forms
- Tables
- Miscellanea
The `<img>` tag

- Used for including images
  - Common formats: jpeg, gif, png
- Important attributes
  - `src` gives the URL of the image
  - `alt` specifies a text for non-graphical browsers
    - Graphical browsers display it if image not available
  - `width` and `height` specify image size
    - Browsers stretch or compress image to desired size
    - Default unit pixels, can use `%` for width

Other things about images

- Parts of an image can be transparent
- Images can be links to other documents
  - With image maps the target depends on the position within the image the user clicked on
- Simple animations possible with gif images
- With CSS (or the deprecated background attribute of the body tag) an image can be used as background (for the full document)

Anchors and URLs

- URLs can point to a specific fragment of a document
  - `http://server/document.html#fragmentname`
  - `<a name="fragmentname">text</a>`
- Almost all tags have the `id` attribute `<h1 id="frn">text</h1>`
- URLs can be relative
  - Makes it easy to move a group of hyperlinked documents
  - Base URL = URL of the document (`http://srv/dir/doc.html`)
  - The parts not in the relative URL taken from base URL
    - `otherdoc.html` → `http://srv/dir/otherdoc.html`