Introduction to Web Programming

Lecture 2: Basic HTML

2.1: Basic HTML

- **2.1: Basic HTML**
- 2.3: Web Standards
- 2.2: More HTML Elements

Hypertext Markup Language (HTML)

- describes the *content* and structure of information on a web page
  - not the same as the *presentation* (appearance on screen)
- surrounds text content with opening and closing **tags**
- each tag's name is called an **element**
  - syntax: ⟨*element*⟩ *content* ⟨/*element*⟩
  - example: `<p>This is a paragraph</p>`
- most whitespace is insignificant in HTML (ignored or collapsed to a single space)
- we use the latest version called HTML5

Structure of an HTML page

```
<!DOCTYPE html>
<html>
  <head>
    information about the page
  </head>

  <body>
    page contents
  </body>
</html>
```

- the **header** describes the page and the **body** contains the page's contents
- an HTML page is saved into a file ending with extension `.html`
- `DOCTYPE` tag tells browser to interpret our page's code as HTML5

Page title: ⟨title⟩

describes the title of the web page

```
<title>Chapter 2: HTML Basics</title>
```

- placed within the head of the page
- displayed in the web browser's title bar and when bookmarking the page

Paragraph: ⟨p⟩

paragraphs of text (block)

```
<p>
When you realize
you want to spend the rest of your life with somebody,
you            want the rest of your life to start
as      soon    as      possible. </p>
```

When you realize you want to spend the rest of your life with somebody, you want the rest of your life to start as soon as possible.

- placed within the body of the page
- more paragraph examples

Headings: `<h1>`, `<h2>`, ..., `<h6>`

headings to separate major areas of the page (block)

```
<h1>University of Tianjin</h1>
<h2>Department of Computer Science</h2>
<h3>Sponsored by Micro$oft</h3>
```

University of Tianjin

Department of Computer Science

Sponsored by Micro$oft

- More heading examples

Horizontal rule: `<hr>`

a horizontal line to visually separate sections of a page (block)

```
<p>First paragraph</p>
<hr />
<p>Second paragraph</p>
```

First paragraph

Second paragraph

- should be immediately closed with `/>`

More about HTML tags

- some tags can contain additional information called **attributes**
  - syntax: *<element attribute="value" attribute="value"> content </element>*
  - example: <a `href="page2.html"`>Next page
- some tags don't contain content; can be opened and closed in one tag
  - syntax: *<element attribute="value" attribute="value" /> or <element attribute="value" attribute="value">*
  - example: <hr `/>`
  - example: <img `src="bunny.jpg" alt="pic from Easter"`>

Links: <a>
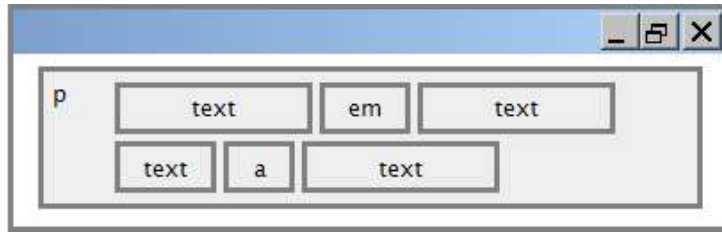
links, or "anchors", to other pages (inline)

```
<p>
  Search
  <a href="http://www.baidu.com/">Baidu</a> or our
  <a href="Lecture-2.html">This lecture</a>.
</p>
```

Search <u>Baidu</u> or our <u>This lecture</u>.

- uses the `href` attribute to specify the destination URL
  - can be **absolute** (to another web site) or **relative** (to another page on this site)
- anchors are inline elements; must be placed in a block element such as p or h1

Block and inline elements (explanation)

- **block** elements contain an entire large region of content
  - examples: paragraphs, lists, table cells
  - the browser places a margin of whitespace between block elements for separation
- **inline** elements affect a small amount of content
  - examples: bold text, code fragments, images
  - the browser allows many inline elements to appear on the same line
  - must be nested inside a block element

Images: `<img>`

inserts a graphical image into the page (inline)

```
<img src="images/gollum.jpg" alt="Gollum from LOTR" />
```


- the `src` attribute specifies the image URL
- HTML5 also requires an `alt` attribute describing the image

More about images

```
<a href="http://theonering.net/">
  <img src="images/gandalf.jpg" alt="Gandalf from LOTR"
       title="You shall not pass!" />
</a>
```


- if placed in an a anchor, the image becomes a link
- `title` attribute is an optional tooltip (on ANY element)

Line break: `<br>`

forces a line break in the middle of a block element (inline)

```
<p>Teddy said it was a hat, <br /> So I put it on.</p>
<p>Now Daddy's sayin', <br /> Where the
heck's the toilet plunger gone?</p>
```

```
Teddy said it was a hat,
So I put it on.

Now Daddy's sayin',
Where the heck's the toilet plunger gone?
```

- Warning: Don't over-use br (guideline: >= 2 in a row is bad)

- br should not be used to separate paragraphs or used multiple times in a row to create spacing

Phrase elements : ⟨em⟩, ⟨strong⟩

em: emphasized text (usually rendered in italic)
strong: strongly emphasized text (usually rendered in bold)

```
<p>
  HTML is <em>really</em>,
  <strong>REALLY</strong> fun!
</p>
```

HTML is *really*, **REALLY** fun!

- as usual, the tags must be properly nested for a valid page

Nesting tags

```
<p>
  HTML is <em>really,
  <strong>REALLY</em> lots of</strong> fun!
</p>
```

- tags must be correctly nested
  - (a closing tag must match the most recently opened tag)
- the browser may render it correctly anyway, but it is invalid HTML
  - (how would we get the above effect in a valid way?)

Comments: `<!-- ... -->`

comments to document your HTML file or "comment out" text

```
<!-- My web page, by Marc G.
     IWP  course, Winter 2018       -->
<p>IWP course is <!-- NOT --> a lot of fun!</p>
```

```
IWP course is a lot of fun!
```

- many web pages are not thoroughly commented (or at all)
- still useful at top of page and for disabling code
- comments cannot be nested and cannot contain a `--`

2.3: Web Standards

- 2.1: Basic HTML
- **2.3: Web Standards**
- 2.2: More HTML Elements

Web Standards

- It is important to write proper HTML code and follow proper syntax.
- Why use valid HTML and web standards?
    - more rigid and structured language
    - more interoperable across different web browsers
    - more likely that our pages will display correctly in the future
    - can be interchanged with other XML data: SVG (graphics), MathML, MusicML, etc.

W3C HTML Validator

```
<p>
  <a href="http://validator.w3.org/check/referer">
    <img src="http://webster.cs.washington.edu/w3c-html.png"
    alt="Validate" />
  </a>
</p>
```

- validator.w3.org
- checks your HTML code to make sure it follows the official HTML syntax
- more picky than the browser, which may render bad HTML correctly

2.2: More HTML Elements

- 2.1: Basic HTML
- 2.3: Web Standards
- **2.2: More HTML Elements**

Unordered list: `<ul>`, `<li>`

ul represents a bulleted list of items (block)
li represents a single item within the list (block)

```
<ul>
  <li>No shoes</li>
  <li>No shirt</li>
  <li>No problem!</li>
</ul>
```

- No shoes
- No shirt
- No problem!

More about unordered lists

- a list can contain other lists:

```
<ul>
  <li>Simpsons:
    <ul>
      <li>Homer</li>
      <li>Marge</li>
    </ul>
  </li>
  <li>Family Guy:
    <ul>
      <li>Peter</li>
      <li>Lois</li>
    </ul>
  </li>
</ul>
```

- Simpsons:
  - Homer
  - Marge
- Family Guy:
  - Peter
  - Lois

Ordered list: `<ol>`

ol represents a numbered list of items (block)

```
<p>RIAA business model:</p>
<ol>
  <li>Sue customers</li>
  <li>???</li>
  <li>Profit!</li>
</ol>
```

RIAA business model:

1. Sue customers
2. ???
3. Profit!

- we can make lists with letters or Roman numerals using CSS (later)

Definition list: `<dl>`, `<dt>`, `<dd>`

dl represents a list of definitions of terms (block)
dt represents each term, and dd its definition

```
<dl>
  <dt>newbie</dt> <dd>one who does not have mad skills</dd>
  <dt>own</dt> <dd>to soundly defeat
    (e.g. I owned that newbie!)</dd>
  <dt>frag</dt> <dd>a kill in a shooting game</dd>
</dl>
```

newbie
    one who does not have mad skills
own
    to soundly defeat (e.g. I owned that newbie!)
frag
    a kill in a shooting game

Quotations: `<blockquote>`

a lengthy quotation (block)

```
<p>As Lincoln said in his famous Gettysburg Address:</p>
<blockquote>
  <p>Fourscore and seven years ago, our fathers brought forth
    on this continent a new nation, conceived in liberty, and
    dedicated to the proposition that all men are created equal.</p>
</blockquote>
```

As Lincoln said in his famous Gettysburg Address:

    Fourscore and seven years ago, our fathers brought forth on this continent a new nation, conceived in liberty, and dedicated to the proposition that all men are created equal.

Inline quotations: ⟨q⟩

a short quotation (inline)

```
<p>Quoth the Raven, <q>Nevermore.</q></p>

Quoth the Raven, ″Nevermore.″
```

- Why not just write the following?
  `<p>Quoth the Raven, ″Nevermore.″</p>`

We don't use " marks for two reasons:

1. HTML shouldn't contain literal quotation mark characters; they should be written as "
2. using ⟨q⟩ allows us to apply CSS styles to quotations (seen later)

HTML Character Entities

a way of representing any Unicode character within a web page

character(s)	entity
⟨ ⟩	< >
é è ñ	é è ñ
™ ©	™ ©
π δ Δ	π δ Δ
И	И
″ &	" &

- Complete list of HTML entities
- How would you display the text & on a web page?

HTML-encoding text

```
&lt;p&gt;
  &lt;a href=&quot;http://google.com/search?q=marc&amp;ie=utf-8&quot;&gt;
    Search Google for Marc
  &lt;/a&gt;
&lt;/p&gt;
```

```
<p> <a href="http://google.com/search?q=marc&ie=utf-8"> Search Google for Marc </a> </p>
```

- To display the link text in a web page, its special characters must be encoded as shown above

Deletions and insertions: `<del>`, `<ins>`

content that should be considered deleted or added to the document (inline)

```
<p>
  <del>Midterm Exam</del> <ins>Final</ins> is on <del>Jan 20</del> <ins>Jan 27</ins>.
</p>
```

~~Midterm Exam~~ <u>Final</u> is on ~~Jan 20~~ <u>Jan 27</u>.

Abbreviations: ⟨abbr⟩

an abbreviation, acronym, or slang term (inline)

```
<p>
  Safe divers always remember to check their
  <abbr title="Self-Contained Underwater Breathing Apparatus">SCUBA</abbr> gear.
</p>
```

Safe divers always remember to check their SCUBA gear.

Computer code: ⟨code⟩

a short section of computer code (usually shown in a fixed-width font)

```
<p>
  The <code>ul</code> and <code>ol</code>
  tags make lists.
</p>
```

The ul and ol tags make lists.

Preformatted text: `<pre>`

a large section of pre-formatted text (block)

```
<pre>
    TIEI stands for
     Tianjin International Engineering Institute
     and it's located in Tianjin    (China)
</pre>
```

```
    TIEI stands for
     Tianjin International Engineering Institute
     and it's located in Tianjin    (China)
```

- displayed with exactly the whitespace / line breaks given in the text
- shown in a fixed-width font by default
- how would it look if we had instead enclosed it in `code` tags?

Web page metadata: `<meta>`

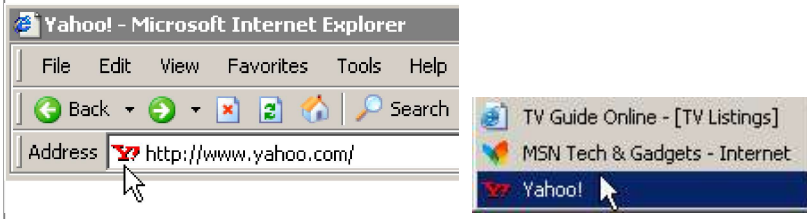*information about your page (for a browser, search engine, etc.)*

```
<meta charset="utf-8" />
<meta name="description"
  content="Authors' web site for IWP." />
<meta name="keywords" content="web, tiei" />
```

- placed **in the `head`** section of your HTML page
- `meta` tags often have both the `name` and `content` attributes
  - some `meta` tags use the `http-equiv` attribute instead of `name`
  - the `meta` tag with `charset` attribute indicates language/character encodings
- using a `meta` tag Content-Type stops validator "tentatively valid" warnings

Favorites icon ("favicon")

```
<link href="filename" type="MIME type" rel="shortcut icon" />
```

```
<link href="yahoo.gif" type="image/gif" rel="shortcut icon" />
```


- the `link` tag, placed in the `head` section, attaches another file to the page
    - in this case, an icon to be placed in the browser title bar and bookmarks