Hypertext Markup Language

INFO/CSE 100

Fluency in Information Technology

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

Readings and References

- Reading
  - » *Fluency with Information Technology*
    - Chapter 4, Marking up with HTML
- References
  - » World Wide Web Consortium
    - http://w3.org/
    - http://w3c.org/MarkUp/
    - http://w3c.org/MarkUp/#tutorials

History of the WWW

- http://w3.org/2004/Talks/w3c10-HowItAllStarted/?n=2
- Issues that drove development of the Web
  - » No transparent way to link documents or objects
    - References by citation or ftp address
  - » No automatic way to retrieve linked documents
    - Downloading via ftp required human intervention
  - » No standard document format
    - Each format required its own application
  - » Even with links, no way to find documents
    - except by being told where to find them, following links

---

History of the WWW

- Web beginnings
  - » 1989: Tim Berners-Lee
    - URLs, http, first browser
  - » 1993: NCSA Mosaic
    - Then Netscape, then Mozilla
  - » 1994: World Wide Web Consortium
    - http://w3.org/
    - Standards organization for Web protocols and formats
  - » 1994-5: Web crawlers and search engines
    - WebCrawler, Lycos, AltaVista, Yahoo

Web Pages in HTML

American Writers

```
<html>
    <head><title>Writer's Anecdotes</title></head>
    <body bgcolor="white" text="black"><font face="Helveti
        <img src="images/awa.png"v>
        <table>
            <tr>
                <td><a href="hdt.html">Thoreau</a></td>
                <td><a href="ed.html">Dickinson</a></td>
                <td><a href="hwl.html">Longfellow</a></td>
                <td><a href="lma.html">Alcott</a></td>
                <td><a href="sc.html">Twain</a></td>
                <td><a href="wf.html">Faulkner</a></td>
                <td><a href="rf.html">Frost</a></td>
                <td><a href="eh.html">Hemingway</a></td>
                <td><a href="js.html">Steinbeck</a></td>
            </tr>
        </table>
        <h2>Steinbeck</h2>
        <p>Steinbeck traveled to Russia several times, but
        never mastered the langauge.  Traveling with
        photographer Robert Capa in 1947 he wrote, "...I
        admit our Russian is limited, but we can say hello
        come in, you are beautiful, oh no you don't, and o
        which charms us but seems to have an application
        rarely needed, 'The thumb is second cousin to the
        left foot.' We don't use that one much."</p>
    </body>
</html>
```

Writer's Anecdotes – Microsoft Internet Explorer
File Edit View Favorites Tools Help
← Back ▾ → ▾ ⊗ ⊡ ⌂ | ⊗ Search ⓐ Favorites

American Writers

Thoreau Dickinson Longfellow Alcott Twain Fa

Steinbeck

Steinbeck traveled to Russia several times, bu
Traveling with photographer Robert Capa in 1
is limited, but we can say hello, come in, you a
which charms us but seems to have an applica
second cousin to the left foot.' We don't use tha

HTML Structure

- All HTML files use the same basic structure:
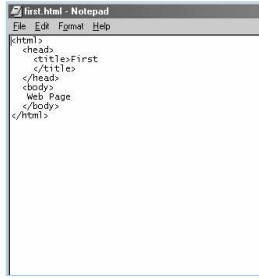
```
<html>
    <head>
        <title>Name of Page Goes Here
        </title>
        Heading content goes here
    </head>
    <body>
        Body content goes here
    </body>
</html>
```

First HTML Web Page
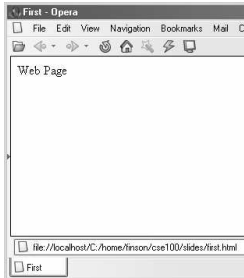
This HTML code produces this result

first.html - Notepad
File Edit Format Help
```
<html>
    <head>
        <title>First
        </title>
    </head>
    <body>
        Web Page
    </body>
</html>
```

First - Opera
File Edit View Navigation Bookmarks Mail C

Web Page

file://localhost/C:/home/finson/cse100/slides/first.html
First
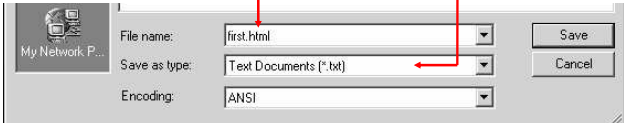
HTML Must Be Text

- Word processors (recall Chap. 2) insert formatting tags, confusing browsers
  - Create source in NotePad, etc.
  - Save in **Text** or **txt** format
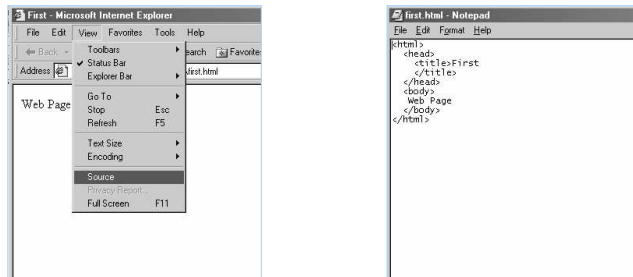 - Save with file extension **.html**

+ Avoid Confusion

My Network P...
File name: first.html   Save
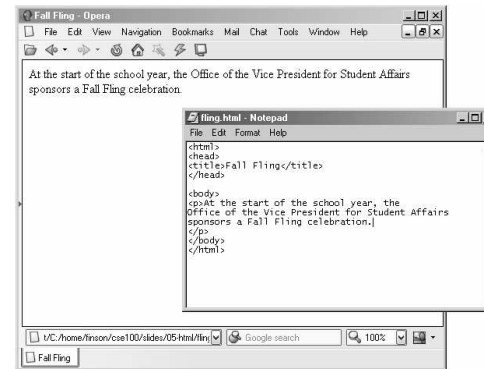Save as type: Text Documents (*.txt)   Cancel
Encoding: ANSI

The Source

The HTML code producing a page is the *source* which can always be *viewed*, either in the browser or in an editor

Add An Image

- Images are encoded many ways:
  - » GIF -- Graphics Interchange Format -- is for diagrams and simple drawings
  - » PNG -- Portable Network Graphics - GIF without patent problems
  - » JPEG -- Joint Photographic Experts Group -- is for high resolution photos, complex art
- Image tags for placing images

  `<img src="awa.png" alt="American Writers">`

  tag　attribute name　attribute value

File Locations

- The path must say how to reach the file
  - When the file is in the directory as the web page, just give the file name, `ski.jpg`
  - If the file is in a subdirectory, say how to navigate to it, `pix/ski.jpg`
  - If the file is in a superdirectory, move up using dot-dot notation, `../ski.jpg`

  The most common reason that an image is not displayed is the path is wrong … check!

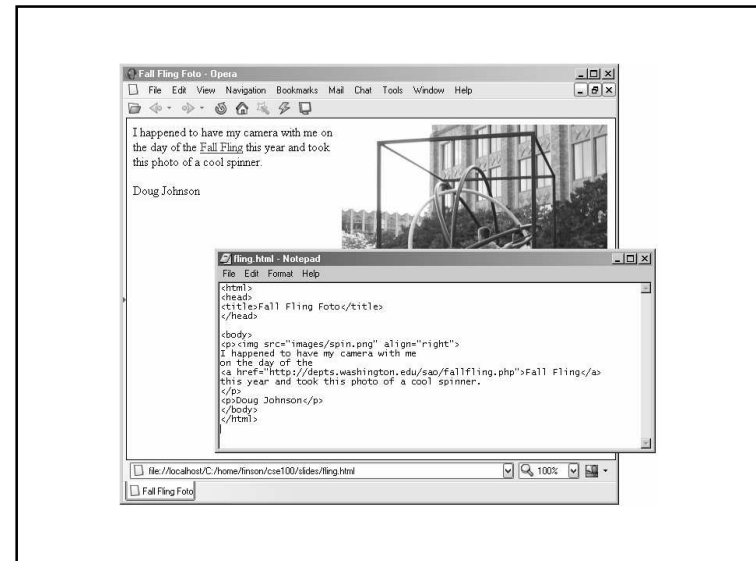## Linking to another page

- The link starts with an **a** tag…
  `<a href=" ... ">`

- The value of the **href** attribute is a URL.

  `<a href="http://www.cs.washington.edu">`

- The link encloses some text, and ends with…
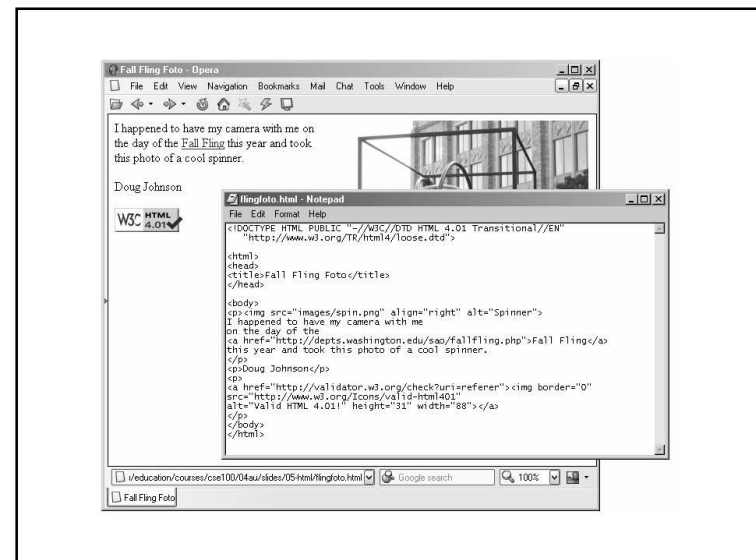
  `<a href=" ... ">CSE home page</a>`

Simple HTML

- What we've seen here is *very simple* HTML
- HTML is changing
  - » Each document should start with a "**DOCTYPE**" comment telling which version of HTML it follows
    - http://www.w3.org/QA/2002/04/valid-dtd-list.html
    - We used 4.01 Transitional
  - » There is a validator service that will check your page
    - http://validator.w3.org/

Summary

- Web pages are written in HTML
  - The files must be text
  - The file extension must be .html (or .htm)
- Tags provide formatting and other info
  - Some have attributes
  - Some need an end tag
- Use a change-and-test process
- Specify the relative path to local images
- Use a link to refer to other documents

4/6/2005   fit100-05-html © 2004-2005 University of Washington   17