# CSN-101 (Introduction to Computer Science and Engineering)

## *Lecture 10: Computer Networking and Web Technology*

**Dr. Sudip Roy**

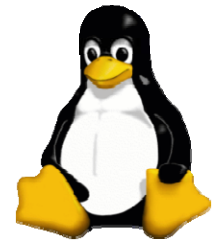*Assistant Professor*

*Department of Computer Science and Engineering*

Piazza Class Room: https://piazza.com/iitr.ac.in/fall2019/csn101

[Access Code: csn101@2019]

Moodle Submission Site: https://moodle.iitr.ac.in/course/view.php?id=45

[Enrollment Key: csn101@2019]

# Plan for Lecture Classes in CSN-101 (Autumn, 2019-2020)

| Week | Lecture 1 (Monday 4-5 PM) | Lecture 2 (Friday 5-6 PM) |
|------|---------------------------|---------------------------|
| 1 | Evolution of Computer Hardware and Moore's Law, Software and Hardware in a Computer | Computer Structure and Components, Operating Systems |
| 2 | Computer Hardware: Block Diagrams, List of Components | Computer Hardware: List of Components, Working Principles in Brief, Organization of a Computer System |
| 3 | Linux OS | Linux OS |
| 4 | Writing Pseudo-codes for Algorithms to Solve Computational Problems | Writing Pseudo-codes for Algorithms to Solve Computational Problems |
| 5 | Sorting Algorithms – Bubble sort, selection sort, and Search Algorithms | Sorting Algorithms – Bubble sort, selection sort, and Search Algorithms |
| 6 | C Programming | C Programming |
| 7 | Number Systems: Binary, Octal, Hexadecimal, Conversions among them | Number Systems: Binary, Octal, Hexadecimal, Conversions among them |
| 8 | Number Systems: Negative number representation, Fractional (Real) number representation | Boolean Logic: Boolean Logic Basics, De Morgan's Theorem, Logic Gates: AND, OR, NOT, NOR, NAND, XOR, XNOR, Truth-tables |
| 9 | Computer Networking and Web Technologies: Basic concepts of networking, bandwidth, throughput | Computer Networking and Web Technologies: Basic concepts of networking, bandwidth, throughput |
| 10 | Different layers of networking, Network components, Type of networks | Network topologies, MAC, IP Addresses, DNS, URL |
| 11 | Different fields of CSE: Computer Architecture and Chip Design | Different fields of CSE: Data Structures, Algorithms and Programming Languages |
| 12 | Different fields of CSE: Database management | Different fields of CSE: Operating systems and System softwares |
| 13 | Different fields of CSE: Computer Networking, HPCs, Web technologies | Different Applications of CSE: Image Processing, CV, ML, DL |
| 14 | Different Applications of CSE: Data mining, Computaional Geometry, Cryptography, Information Security | Different Applications of CSE: Cyber-physical systems and IoTs |

Up to MTE

# Introduction to the Internet and Web

# Internet

- It is the largest network in the world that connects hundreds of thousands of individual networks all over the world.

- The popular term for the Internet is the "information highway".

- Rather than moving through geographical space, it moves your ideas and information through cyberspace – the space of electronic movement of ideas and information.

# Internet

- No one owns it
- It has no formal management organization.
- As it was originally developed by the Department of defense, this lack of centralization made it less vulnerable to wartime or terrorist attacks.
- To access the Internet, an existing network need to pay a small registration fee and agree to certain standards based on the TCP/IP (Transmission Control Protocol/Internet Protocol) .

# The uses of the Internet

- Send e-mail messages.
- Send (upload) or receive (down load) files between computers.
- Participate in discussion groups, such as mailing lists and newsgroups.
- Surfing the web.

# What is Web?

- The **Web (World Wide Web)** consists of information organized into Web pages containing text and graphic images.

- It contains hypertext links, or highlighted keywords and images that lead to related information.

- A collection of linked Web pages that has a common theme or focus is called a **Web site**.

- The main page that all of the pages on a particular Web site are organized around and link back to is called the site's **home page**.

# How to access the Internet?

- Many schools and businesses have direct access to the Internet using special high-speed communication lines and equipment.

- Students and employees can access through the organization's local area networks (LAN) or through their own personal computers.

- Another way to access the Internet is through Internet Service Provider (ISP).

# How to access the Internet?

- To access the Internet, an existing network need to pay a small registration fee and agree to certain standards based on the TCP/IP (Transmission Control Protocol/Internet Protocol) reference model.

- Each organization pays for its own networks and its own telephone bills, but those costs usually exist independent of the internet.

- The regional Internet companies route and forward all traffic, and the cost is still only that of a local telephone call.

# Internet Service Provider (ISP)

- A  commercial organization with permanent connection to the Internet that sells temporary connections to subscribers.

- Examples:

- Prodigy, America Online, Microsoft network, AT&T Networks.

# How to access the Web?

- Once you have your Internet connection, then you need special software called a browser to access the Web.

- Web browsers are used to connect you to remote computers, open and transfer files, display text and images.

- Web browsers are specialized programs.

- Examples of Web browser: Netscape Navigator (Navigator) and Internet Explorer.

# Client/Server Structure of the Web

- Web is a collection of files that reside on computers, called **Web servers**, that are located all over the world and are connected to each other through the Internet.

- When you use your Internet connection to become part of the Web, your computer becomes a **Web client** in a worldwide client/server network.

- A **Web browser** is the software that you run on your computer to make it work as a web client.

# Hypertext Markup Language (HTML)

- The public files on the web servers are ordinary text files, much like the files used by word-processing software.

- To allow Web browser software to read them, the text must be formatted according to a generally accepted standard.

- The standard used on the web is Hypertext markup language (HTML).

# Hypertext Markup Language (HTML)

- HTML uses codes, or tags, to tell the Web browser software how to display the text contained in the document.

- For example, a Web browser reading the following line of text:

  &lt;B&gt; A Review of the Book&lt;I&gt;Wind Instruments of the 18th Century&lt;/I&gt;&lt;/B&gt;

- recognizes the &lt;B&gt; and &lt;/B&gt; tags as instructions to display the entire line of text in bold and the &lt;I&gt; and &lt;/I&gt; tags as instructions to display the text enclosed by those tags in italics.

# Addresses on the Web:IP Addressing

- Each computer on the internet does have a unique identification number, called an IP (Internet Protocol) address.

- The IP addressing system currently in use on the Internet uses a four-part number.

- Each part of the address is a number ranging from 0 to 255, and each part is separated from the previous part by period,

- For example, 106.29.242.17

# IP Addressing

- The combination of the four IP address parts provides 4.2 billion possible addresses (256 x 256 x 256 x 256).

- This number seemed adequate until 1998.

- Members of various Internet task forces are working to develop an alternate addressing system that will accommodate the projected growth.

- However, all of their working solutions require extensive hardware and software changes throughout the Internet.

# Domain Name Addressing

- Most web browsers do not use the IP address t locate Web sites and individual pages.

- They use domain name addressing.

- A **domain name** is a unique name associated with a specific IP address by a program that runs on an Internet host computer.

- This program, which coordinates the IP addresses and domain names for all computers attached to it, is called **DNS (Domain Name System ) software**.

- The host computer that runs this software is called a **domain name server.**

# Domain Name Addressing

- Domain names can include any number of parts separated by periods, however most domain names currently in use have only three or four parts.

- Domain names follow hierarchical model that you can follow from top to bottom if you read the name from the right to the left.

- For example, the domain name gsb.uchicago.edu is the computer connected to the Internet at the Graduate School of Business (gsb), which is an academic unit of the University of Chicago (uchicago), which is an educational institution (edu).

- No other computer on the Internet has the same domain name.

# Uniform Resource Locators (URLs)

- The IP address and the domain name each identify a particular computer on the Internet.

- However, they do not indicate where a Web page's HTML document resides on that computer.

- To identify a Web pages exact location, Web browsers rely on Uniform Resource Locator (URL).

- URL  is a four-part addressing scheme that tells the  Web browser:

➢ What transfer protocol to use for transporting the file

➢ The domain name of the computer on which the file resides

➢ The pathname of the folder or directory on the computer on which the file resides

➢ The name of the file

# Structure of a Uniform Resource Locators

protocol

pathname

http://www.chicagosymphony.org/civicconcerts/index.htm

Domain name

filename

http => Hypertext Transfer Protocol

# HTTP

- The transfer protocol is the set of rules that the computers use to move files from one computer to another on the Internet.

- The most common transfer protocol used on the Internet is the Hypertext Transfer Protocol (HTTP).

- Two other protocols that you can use on the Internet are the File Transfer Protocol (FTP) and the Telnet Protocol

# How to find information on the Web?

- A number of search tools have been developed and available to you on certain Web sites that provide search services to help you find information.

- Examples:

➢ Yahoo → www.yahoo.com

➢ Excite → www.excite.com

➢ Lycos → www.lycos.com

➢ AltaVista → www/alta-vista.com

➢ MSN WebSearch → www.search.msn.com

# How to find information on the Web?

- You can find information by two basic means.
- **Search by Topic** and **Search by keywords**.
- Some search services offer both methods, others only one.
- Google offers both
- ➢ Search by Topic

  You can navigate through topic lists
- ➢ Search by keywords

  You can navigate by entering a keyword or phase into a search text box.

**You must try Advanced Search options in Google**

# Web Technologies, HTML

# HTML Versions

- The history of HTML at W3C starts with HTML 3.2.

- New versions were followed a few years later by HTML 4.0, then HTML 4.01.

- HTML 4.01 is the last version of HTML, and is also the final W3C specification to define the semantics of markup.

# Structural Tags

<HTML>
   These tags enclose the entire Web page document.
</HTML>

<HEAD>
   These tags enclose the Head part of the document
</HEAD>

<TITLE>
   These tags enclose the title of the document.  This text appears in the title bar
   in the browser and on the bookmark list if someone bookmarks your web page.
</TITLE>

# Sample Structure of a Web Site

```
<HTML>
    <HEAD>
        <TITLE> John Q. Public's Web Page </TITLE>
    </HEAD>

    <BODY>
        This is John Public's Webpage!
    </BODY>
</HTML>
```

# Header Tags

Header Tags -- Used for marking sections and subsections in a document.

<H1>Header 1 -- Giant-sized and bold </H1>

<H2>Header 2 -- Large and bold </H2>

<H3>Header 3 -- Normal-sized and bold </H3>

<H4>Header 4 -- Small and bold </H4>

<H5>Header 5 -- Very Small and bold </H5>

<H6>Header 6 -- Tiny and bold </H6>

# Header Tags (cont.)

# H1 = Giant-sized and bold

## H2 = Large and bold

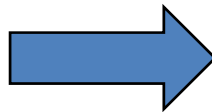### H3 = Normal-sized and bold

**H4 = Small and bold**

**H5 = Very Small and bold**

**H6 = Tiny and bold**

# Breaking Lines and Paragraphs

- <P> text </P>
  - Paragraph tag
  - Most browsers render (process) this with blank lines between each paragraph
- <BR>
  - Line break tag
  - Used when the webmaster wants a carriage return but doesn't want a blank line to follow

---

Example:
   <p>text a</p>
   <p>text b</p>
   <br>text c
   <br>text d

→

text a

text b

text c
text d

# Horizontal Rule

The <HR> tag puts a graphical line across the page.

Ex:

---

*Horizontal Rule Attributes:*

NOSHADE -- A solid line with no shading

WIDTH="xx%/xx" --      Controls the width of the line.  You may specify
                       either percentage of the width of a page or actual
                       pixel length

SIZE="xx" -- Controls the height of the line.  You need to specify the
             dimension in pixels.

ALIGN="left/center/right" -- This allows the line to be aligned to the left,
                       right, or center of the page

# Text Formatting Tags

Some basic text formatting styles:

| Tag | Result |
|---|---|
| <I> Italics </I> | *Italics* |
| <B> Bold </B> | **Bold** |
| | |
| <PRE> Preformatted Text </PRE> | `Preformatted Text` |
| <STRONG> Strong </STRONG> | **Strong** |
| <ADDRESS> Address </ADDRESS> | *Address* |
| <CITE> Citations </CITE> | *Citations* |
| <CODE> Source Code </CODE> | `Source Code` |

# Font modifications

Web creators can also change the way text looks by using the <FONT> tag

SIZE="number" - changes size of the font; 1=smallest, 7 = largest
  <FONT SIZE="7">Big</FONT> <FONT SIZE="1">Small</FONT>

Big Small

___

COLOR="color-name" - changes text color
  <FONT COLOR="red">This is red</FONT>

This is red

___

FACE="font-name" - changes font
  <FONT FACE="verdana">This is the verdana font;</FONT> <FONT FACE="chicago">this is the chicago font.</FONT>

This is the verdana font; this is chicago font.

# &lt;Font&gt; modifications (cont.)

One can combine font modifications:

&lt;FONT SIZE="7" FACE="courier" COLOR="red"&gt;Big, Courier & Red&lt;/FONT&gt;

<span style="color:red">Big, Courier & Red</span>

&lt;FONT SIZE="7"&gt;&lt;FONT FACE="courier"&gt;Big & Courier&lt;/FONT&gt; - Just Big&lt;/FONT&gt;

Big & Courier - Just Big

# Lists -- Unordered Lists

**Unordered lists:**

```
<UL>
        <LI>Item One
        <LI>Item Two
        <LI>Item Three
        <LI>Item Four
</UL>
```

➡
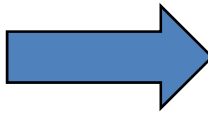
- Item One
- Item Two
- Item Three
- Item Four

---

*Unordered List Attributes:*

type="disc/circle/square"

• Disc (default)  ○ Circle  ■ Square

# Lists -- Ordered Lists

**Ordered (Numbered) Lists:**

```
<OL>
        <LI> Item One
        <LI> Item Two
        <LI> Item Three
        <LI> Item Four
</OL>
```

➡️

1. Item One

2. Item Two

3. Item Three

4. Item Four

---

*Ordered List Attributes*:

type="i/I/a/A/1"                                    (default)

| i = | i. Item One | I = | I. Item One | a = | a. Item One | A = | A. Item One | 1 = | 1.Item One |
|---|---|---|---|---|---|---|---|---|---|
| | ii. Item Two | | II. Item Two | | b. Item Two | | B. Item Two | | 2. Item Two |
| | iii. Item Three | | III. Item Three | | c. Item Three | | C. Item Three | | 3. Item Three |
| | iv. Item Four | | IV. Item Four | | d. Item Four | | D. Item Four | | 4. Item Four |

start="xx"

• This attribute lets you specify which number/letter will start the list

# Lists -- Definition Lists

## Definition Lists:

```
<DL>
    <DT>List Name One
        <DD>This is where information about List Name One would go</DD>
    </DT>
    <DT>List Name Two
        <DD>This is where information about List Name Two would go</DD>
    </DT>
</DL>
```
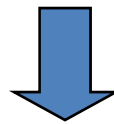
List Name One

    This is where information about List Name One would go

List Name Two

    This is where information about List Name Two would go

# Links

The anchor tag <A> is used to link one document to another or from one part of a document to another part of the same document.

Basic Links:

<A HREF="http://www.stanford.edu/">Stanford University</A>

Inter-document Links:

<A HREF="#spot">Point to 'spot' in this document</A>

Defining a point in a document:

<A NAME="spot">Spot</A>

Email links:

<A HREF="mailto:someone@somehost.com">Email someone@somehost.com</A>

# Graphics

To have a graphic appear on a webpage, web designers must to put the <IMG> tag in with the address where the graphic "lives":

<IMG SRC="http://www.someplace.com/images/fish.gif">
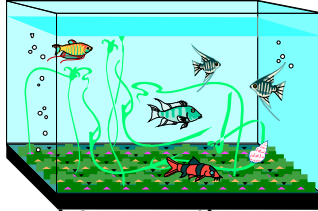


Graphics attributes:

alt="text": insert a description of the graphic for those who are using browsers that cannot process images (e.g., page readers for the blind)

width="xx/xx%": width in pixels/percentage

height="xx/xx%": height in pixels/percentage
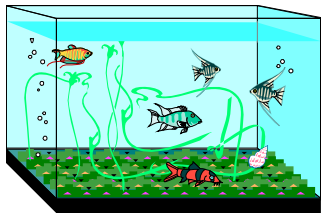
border="xx": pixel length of the border surrounding the image.

hspace="xx": places a buffer of space horizontally around the image

vspace="xx": places a buffer of space vertically around the image

align="top/middle/bottom/right/left": aligns image in relation to the text (see next 2 slides)
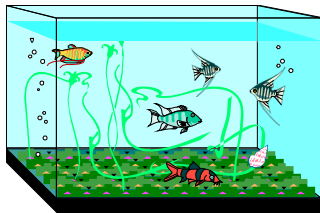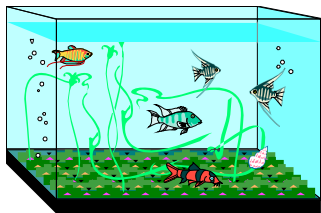
# Graphics (cont.)

\<img src="http://www.someplace.com/images/fish.gif" align="top">All about Fish

 All about Fish

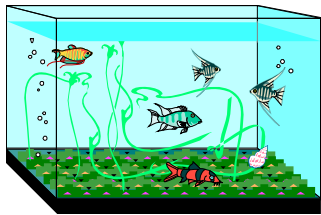\<img src="http://www.someplace.com/images/fish.gif" align="middle">All about Fish

 All about Fish

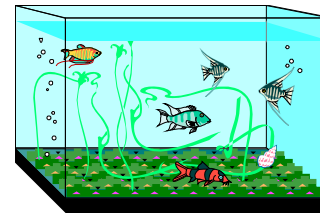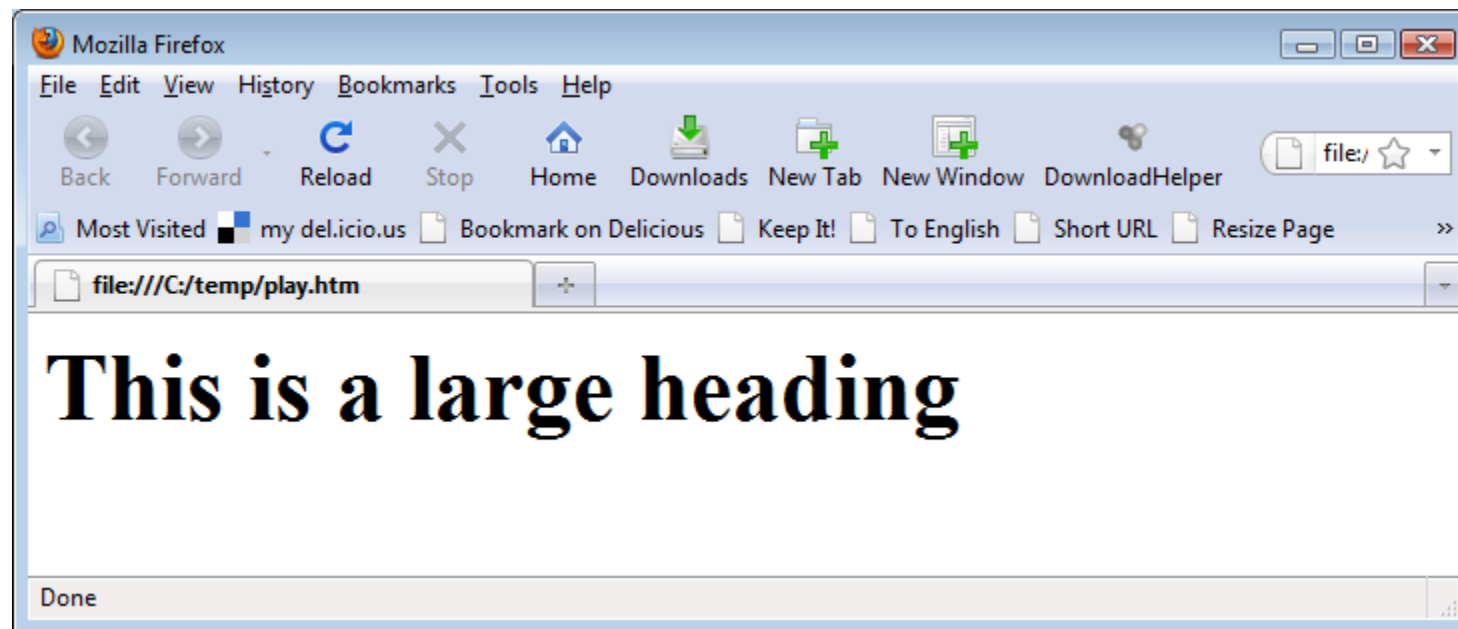\<img src="http://www.someplace.com/images/fish.gif" align="bottom">All about Fish
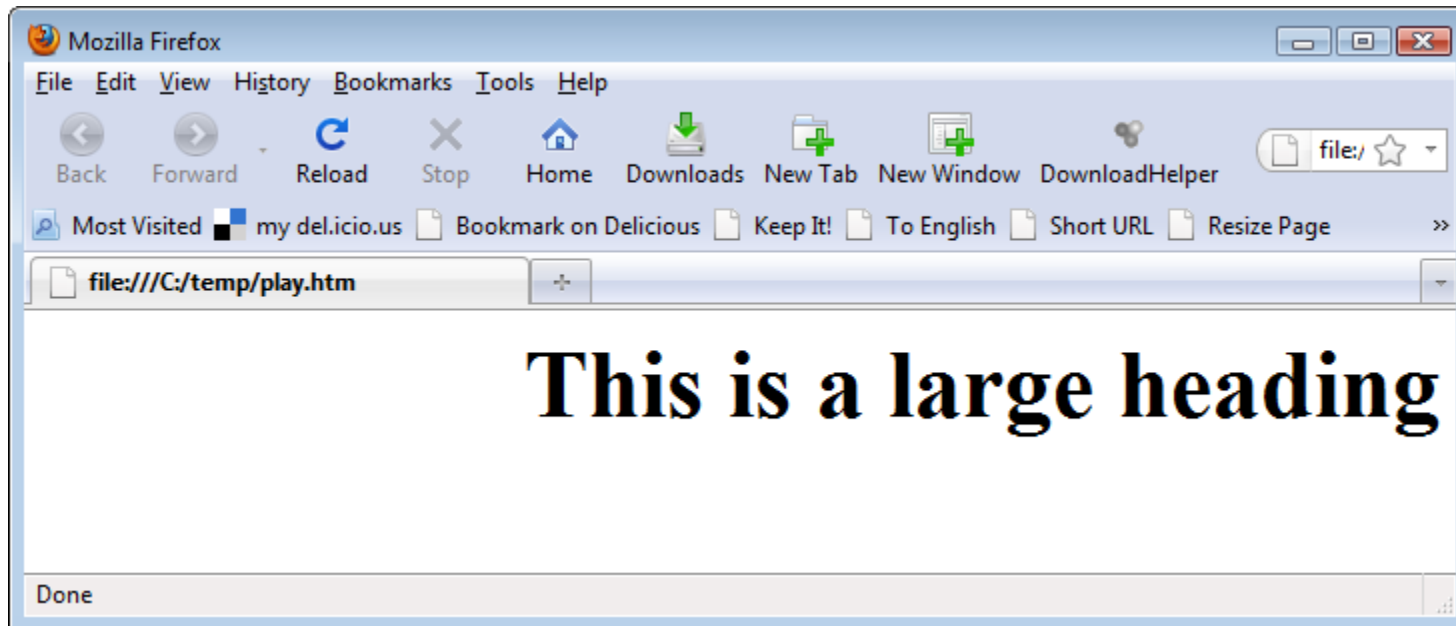
 All about Fish

# Graphics (cont.)

`<img src="http://www.someplace.com/images/fish.gif" align="left">`

`<img src="http://www.someplace.com/images/fish.gif" align="right">`

# Examples of HTML tags

| Tag | Effect |
|---|---|
| <h1> … </h1> | large heading |
| <h2> … </h2> | smaller heading |
| <p> … </p> | paragraph |
| <b> … </b> | bold text |
| <i> … </i> | italicized text |
| <u> … </u> | underlined text |

IIT ROORKEE

<h1>This is a large heading</h2>
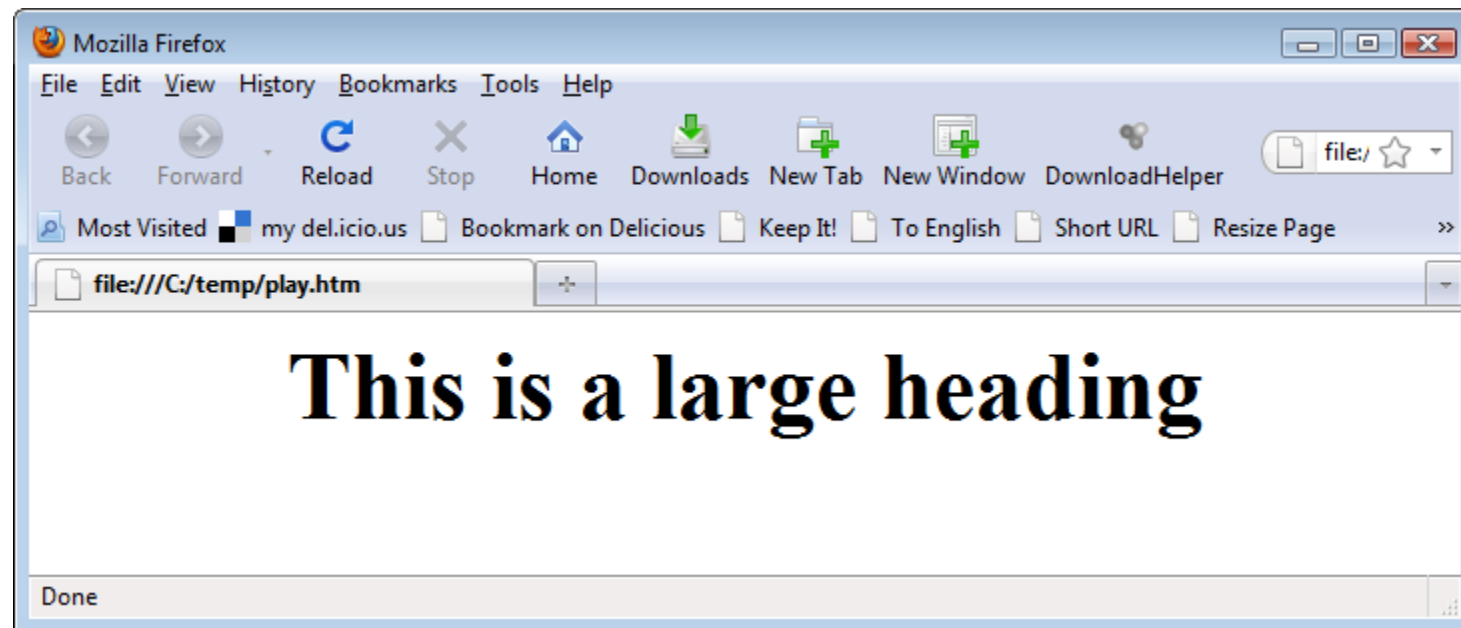
<h1 align = "right">This is a large heading</h2>
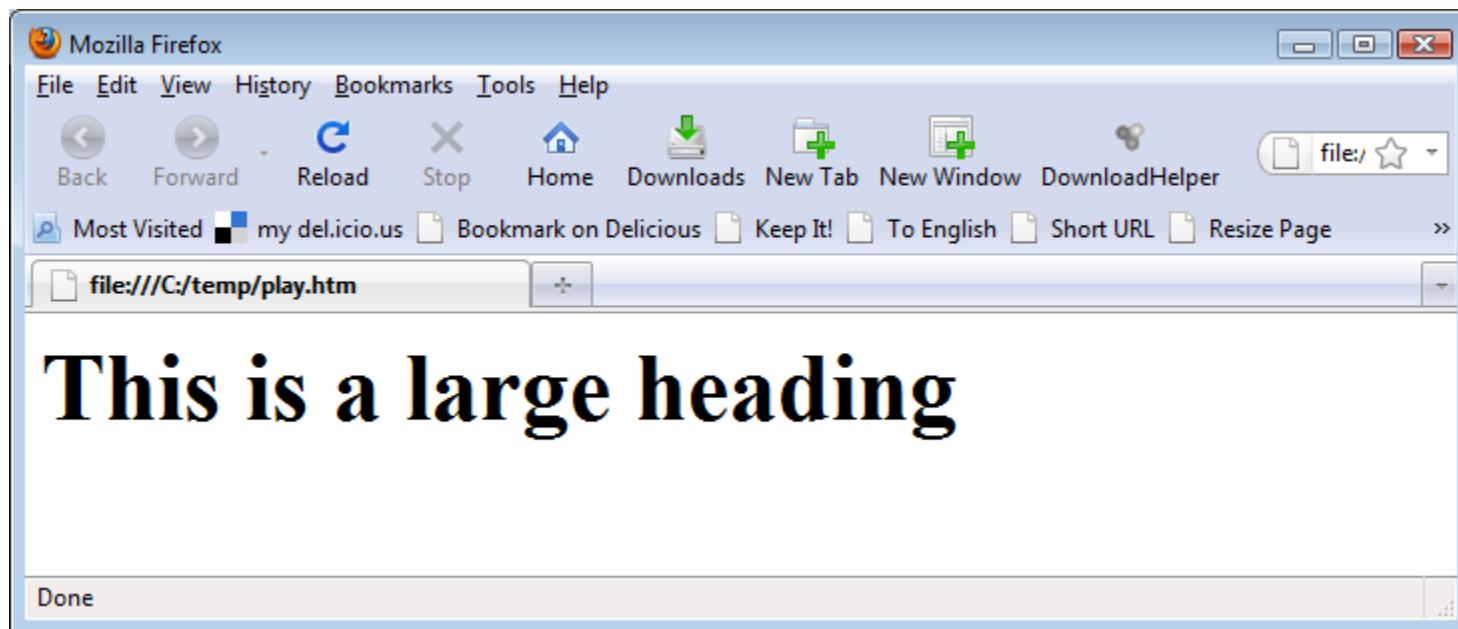
<h1 align = "center">This is a large heading</h2>

<h1 align = "left">This is a large heading</h2>

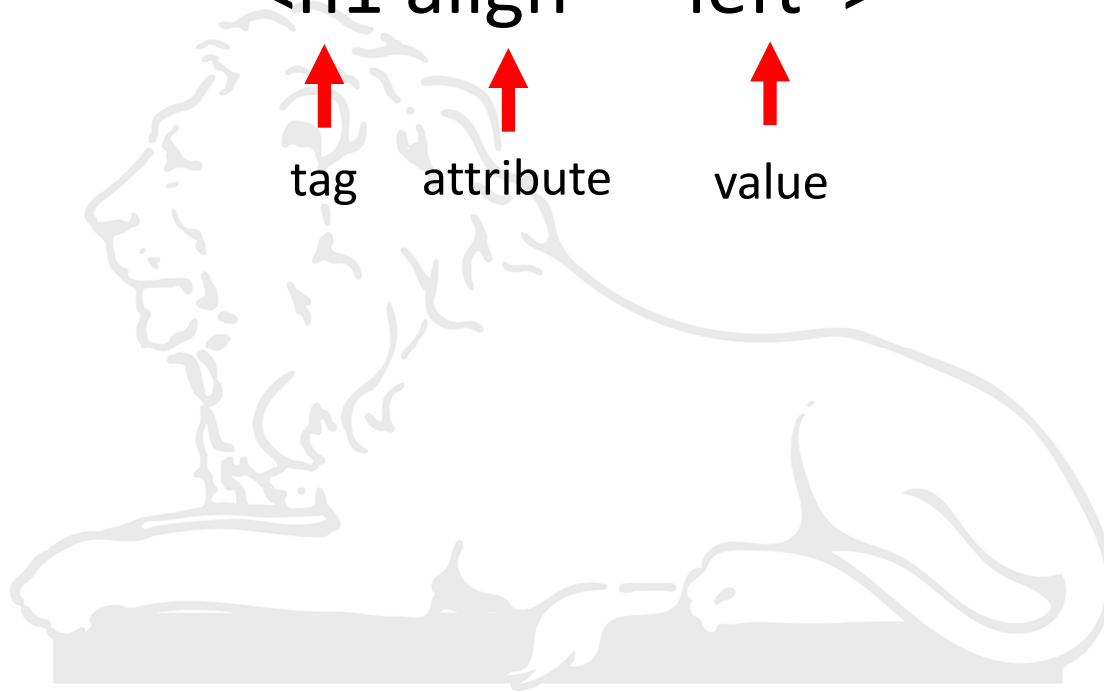attribute name = "value"
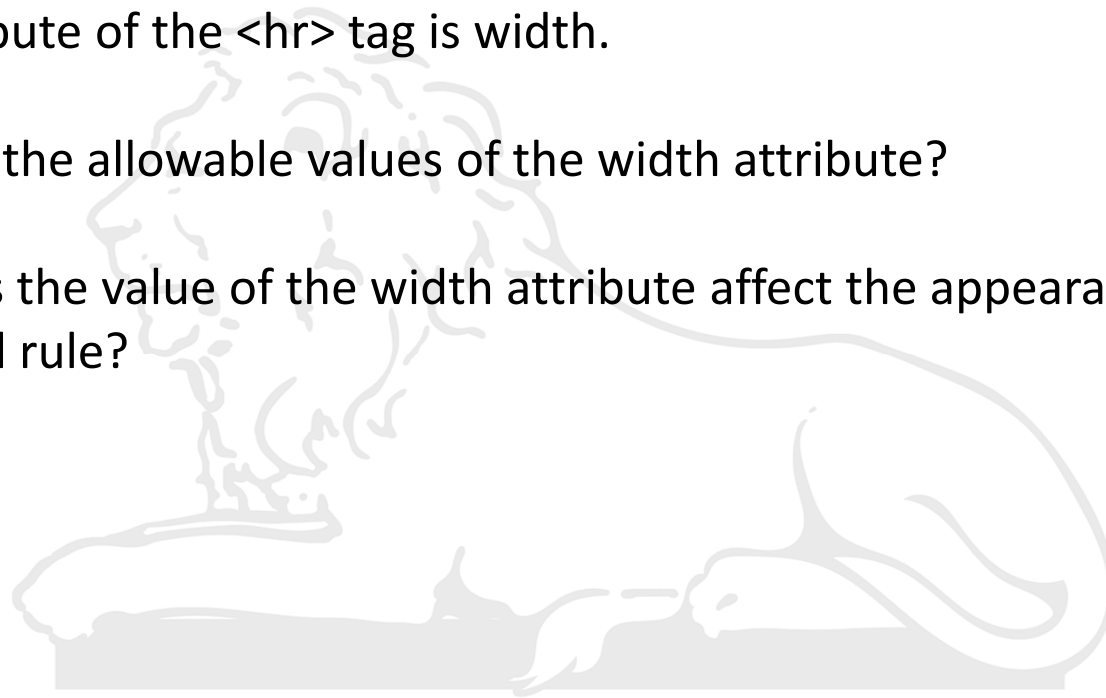
<h1 align = "left">

tag    attribute    value

# A few questions

What are the allowable values of the *align* attribute of the *h1* tag?

One attribute of the <hr> tag is width.

What are the allowable values of the width attribute?

How does the value of the width attribute affect the appearance of the horizontal rule?

# HTML Page: Example

```html
form.html
1   <!DOCTYPE html>
2   <html>
3   <body>
4      <h1 align="center">Static HTML form</h1>
5      <br><br>
6      <b>Photo: </b>
7      <br><br>
8      <img src="img1.jpg" alt="My photo" width="100" height="100" align="middle">
9      <br><br>
10
11     <b>First name: </b>
12     <input type="text" name="firstname" value="A">
13     <br><br>
14     <b>Last name:</b>
15     <input type="text" name="lastname" value="B">
16     <br><br>
17     <b>Age:</b>
18     <input type="text" name="lastname" value="20">
19
20     <br><br>
21     <b>Address:</b>
22     <input type="text" name="lastname" value="C">
23
24
25  </body>
26  </html>
27
```

# Web Programming

# Hypertext & HTML

- HyperText Markup Language (HTML) is the language for specifying the *static* content of Web pages
  - *hypertext* refers to the fact that Web pages are more than just text
    - can contain multimedia, provide links for jumping within & without
  - *markup* refers to the fact that it works by augmenting text with special symbols (tags) that identify structure and content type

HTML is an evolving standard (as new technology/tools are added)

- HTML 1 (Berners-Lee, 1989): very basic, limited integration of multimedia
  in 1993, Mosaic added many new features (e.g., integrated images)

- HTML 2.0 (IETF, 1994): tried to standardize these & other features, but late
  in 1994-96, Netscape & IE added many new, divergent features

- HTML 3.2 (W3C, 1996): attempted to unify into a single standard
  but didn't address newer technologies like Java applets & streaming video

- HTML 4.0 (W3C, 1997): current standard
  attempted to map out future directions for HTML, not just react to vendors

- XHTML 1.0 (W3C, 2000): HTML 4.01 modified to conform to XML standards

# Web development tools

- many high-level tools exist for creating Web pages

    e.g., Microsoft FrontPage, Netscape Composer, Adobe PageMill,

    Macromedia DreamWeaver, HotDog, …

    also, many applications have "save to HTML" options (e.g., Word)

    *for most users who want to develop basic, static Web pages, these are fine*

- assembly language vs. high-level language analogy
  so, why are we learning low-level HTML using a basic text editor?
    - may want low-level control
    - may care about size/readability of pages
    - may want to "steal" page components and integrate into existing pages
    - may want dynamic features such as scripts or applets

# Tags vs. elements

- HTML specifies a set of *tags* that identify structure and content type
    - tags are enclosed in `<  >`

        - `<img src="image.gif" />` specifies an image

    - most tags come in pairs, marking a beginning and ending

an HTML *element* is an object enclosed by a pair of tags
- `<title>` and `</title>` enclose the title of a page
`<title>My Home Page</title>` is a TITLE element

`<b>This text appears bold.</b>` is a BOLD element

`<p>Part of this text is <b>bold</b>.</p>`
    is a PARAGRAPH element that contains a BOLD element

*HTML document is a collection of elements (text/media with context)*

# Web rules of thumb

- HTML provides for lots of neat features,
  but just because you can add a feature doesn't mean you should!

  don't add features that distract from the content of the page

  ➢ use color & fonts sparingly and be careful how elements fit together

   e.g., no purple text on a pink background, no weird fonts

  ➢ use images only where appropriate

   e.g., bright background images can make text hard to read

   e.g., the use of clickable images instead of buttons or links can slow access

  ➢ don't rely on window or font size for layout

   e.g., font size may be adjusted by viewer, window constrained

  ➢ don't be annoying

   e.g., no pop-up windows, excessive advertising, silly music

  ➢ break large document into smaller or provide a menu (either internal or frame)

  ➢ stick to standard features and test using both IE and Netscape

  ➢ utilize style sheets to make changes easy & ensure consistency

# The *index.html* file

- The file name "index.html", or "index.htm" is reserved.

- This is the file in a directory which will be used automatically by default if a URL ends in the directory name instead of a file name.

- On most servers, the use of this file as the default helps prevent unauthorized access to the directory.

- Some servers may have a hierarchy of default file names.