core

WEB

programming

HyperText Markup Language (HTML)

Designing Documents for the
World Wide Web

Agenda

- **Introduction to HTML**
- **Creating and publishing a Web page**
- **Validating a document**
- **Main HMTL elements**
- **Block-level HTML elements**
- **Text-level HTML element**
- **Creating hypertext links**
- **Adding images to documents**
- **Building tables**

The World Wide Web

- **Definitions**
  - The World Wide Web
    - The set of computers on the Internet that support HTTP
    - Not a separate network
  - HTTP
    - The HyperText Transfer Protocol
    - The language used by a WWW client (e.g. Netscape, Internet Explorer) to request documents from a WWW server (i.e. the program running at Web sites like amazon.com or yahoo.com)
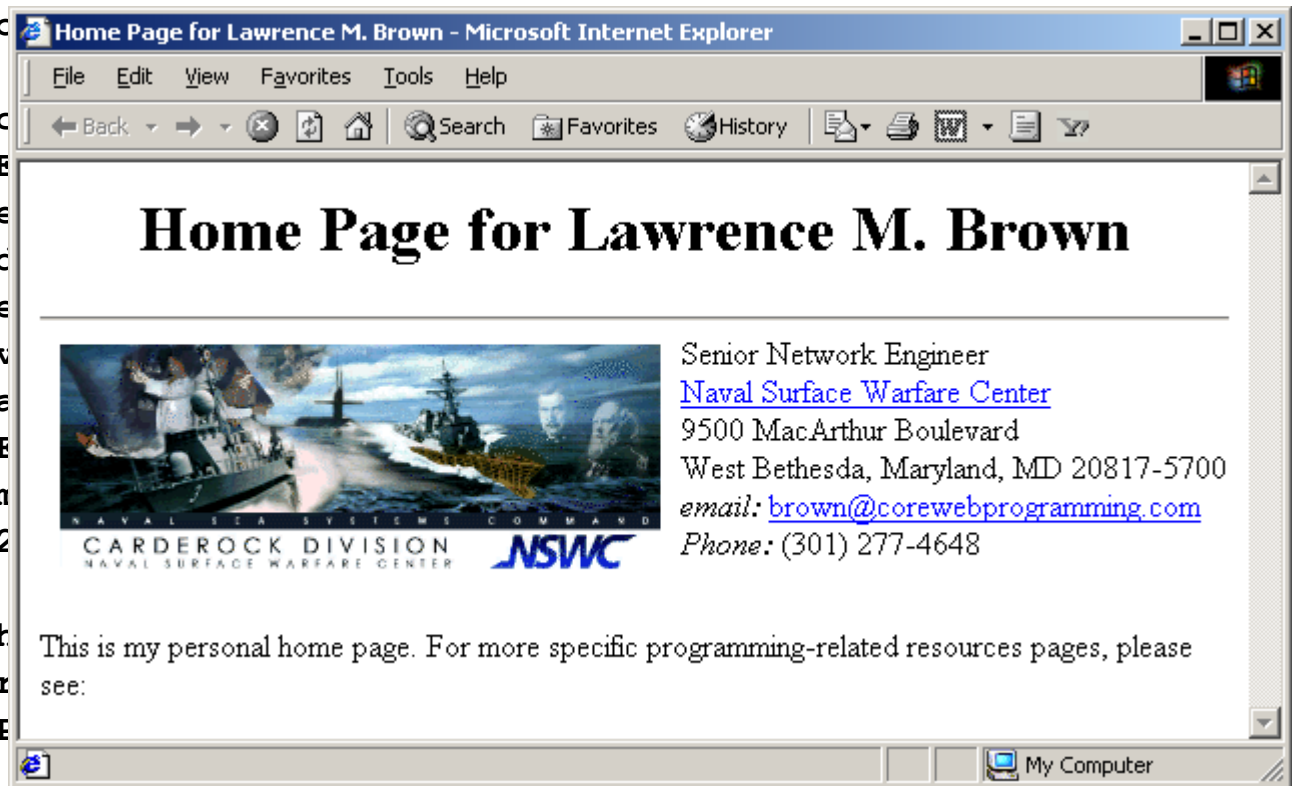 - HTML
    - The HyperText Markup Language
    - The language used to design web pages

HyperText Markup Language

- **Text Mixed with Markup Tags**
  - Tags Enclosed in Angle Brackets (<H1>Introduction</H1>)

- **What Does Markup Describe?**
  - Appearance
  - Layout
  - Content (Can't Enforce an Exact Look)

- **Changes in HTML 3.2 to HTML 4.0**
  - Standardization of frames
  - Deprecation of formatting elements (vs. style sheets)
  - Improved cell alignment and grouping in tables
  - Mouse and keyboard events for nearly all elements
  - Internationalization features

HTML Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
  <Title>Home Page for Lawrence M. Brown</Title>
</HEAD>
<BODY BGCOLOR="WHITE">
<H1 ALIGN="CENTER">Ho
<HR>
<IMG SRC="images/nswc
     HSPACE=10 VSPACE
Senior Network Engine
<A HREF="http://www.c
Naval Surface Warfare
9500 MacArthur Boulev
West Bethesda, Maryla
<I>email:</I> <A HREF
brown@corewebprogramm
<I>Phone:</I> (301) 2
<P>
This is my personal h
programming-related r
<!-- Rest of Sample P
</BODY>
</HTML>
```


Home Page for Lawrence M. Brown - Microsoft Internet Explorer

Home Page for Lawrence M. Brown

Senior Network Engineer
Naval Surface Warfare Center
9500 MacArthur Boulevard
West Bethesda, Maryland, MD 20817-5700
email: brown@corewebprogramming.com
Phone: (301) 277-4648

CARDEROCK DIVISION
NAVAL SURFACE WARFARE CENTER   NSWC

This is my personal home page. For more specific programming-related resources pages, please see:

www.corewebprogramming.com

Creating and Publishing a Web Page

1. **Create an HTML document**
2. **Place the document in a world-accessible directory (often `public_html` or `www`) on a system running an HTTP server**

   ```
   Unix> cd

   Unix> chmod a+x .    (Note the ".")

   Unix> mkdir public_html

   Unix> chmod a+x public_html
   ```

3. **Access the web page through http://hostname/~username/filename**

   - E.g. `http://www.apl.jhu.edu/~lmb/test.html`
   - If the filename is omitted, a system default filename is assumed (often `index.html`)
     - E.g. `http://www.apl.jhu.edu/~hall/` refers to the file `index.html` in hall's `public_html` directory

www.corewebprogramming.com

Creating and Publishing a Web Page, cont.

4. Validate the Document

- Check the syntax using a formal HTML validator
  - http://validator.w3.org/
  - http://www.htmlhelp.com/tools/validator/

- The version of HTML against which the document is validated is based on the DOCTYPE

  <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">

- The WWW Consortium recently added advice that Web pages include information on the character set, even though ASCII or Latin-1 is the default. The validator at http://validator.w3.org/ gives warnings if you omit this. You can ignore such warnings if you wish.

HTML Document Template

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
  <TITLE>Title</TITLE>
</HEAD>

<BODY>
<H1>Main Heading</H1>

<!-- Rest of page goes here -->

</BODY>
</HTML>
```

Goes on browser's title bar
May not appear in printouts

Main heading. Often used as title
Appears in printouts

HTML comment
Replace with body of
WWW page

Title - Netscape

File Edit View Go Communicator Help

Main Heading

Document: Done

www.corewebprogramming.com

Main HTML Elements

1. **DOCTYPE**
2. **HTML**
3. **HEAD**
   - TITLE element required
   - Optional elements:
     - BASE
     - META
     - BGSOUND
     - SCRIPT, NOSCRIPT
     - STYLE
     - LINK

Main HTML Elements (Continued)

4. BODY Element

– <BODY BGCOLOR="YELLOW">

– HTML Attributes and Attribute Values

- BACKGROUND
- BGCOLOR
- TEXT
- LINK, VLINK, ALINK
- OnLoad, OnUnload, OnFocus, OnBlur

5. Elements inside BODY element

– <BODY>
   Remaining HTML elements
   </BODY>

META Element

- **Records document information, forwards and refreshes pages**
  - NAME="author"
  - NAME="keywords"
  - NAME="description"
  - HTTP-EQUIV="refresh"

META Element, Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
  <TITLE>News Headlines</TITLE>
  <META HTTP-EQUIV="REFRESH"
      CONTENT="3600">
</HEAD>

<BODY>
<H1 ALIGN="CENTER">News Headlines</H1>

<H2>National News</H2>
Blah, blah, blah.

<H2>International News</H2>
Yadda, yadda, yadda.

</BODY>
</HTML>
```

Block-Level Elements

- **Headings**
  - H1 ... H6
  - ALIGN
- **Basic Text Sections**
  - P
  - ALIGN
  - PRE
  - WIDTH
  - ADDRESS
  - BLOCKQUOTE

www.corewebprogramming.com

Block-Level Elements, cont.

- **Lists**
  - OL
  - LI
  - UL
  - LI
  - DL
  - DT
  - DD
- **Tables and Forms (Postponed)**
- **Misc.**
  - HR
  - DIV
  - CENTER
  - MULTICOL (Netscape only)

Headings

- **Heading Types**
  - <H1 ...> ... </H1>
  - <H2 ...> ... </H2>
  - <H3 ...> ... </H3>
  - <H4 ...> ... </H4>
  - <H5 ...> ... </H5>
  - <H6 ...> ... </H6>
- **Attributes: ALIGN**
  - Values: LEFT (default), RIGHT, CENTER
- **Nesting tags**
  - Headings and other block-level elements can contain text-level elements, but *not* vice versa

Headings, Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD>
  <TITLE>Document Headings</TITLE>
</HEAD>
<BODY>
Samples of the six heading types:
<H1>Level-1 (H1)</H1>
<H2 ALIGN="CENTER">Level-2 (H2)</H2>
<H3><U>Level-3 (H3)</U></H3>
<H4 ALIGN="RIGHT">Level-4 (H4)</H4>
<H5>Level-5 (H5)</H5>
<H6>Level-6 (H6)</H6>
</BODY>
</HTML>
```

Headings, Result

Document Headings - Microsoft Internet Explorer

File   Edit   View   Favorites   Tools   Help

Samples of the six heading types:

Level-1 (H1)

Level-2 (H2)

Level-3 (H3)

Level-4 (H4)

Level-5 (H5)

Level-6 (H6)

Done   My Computer
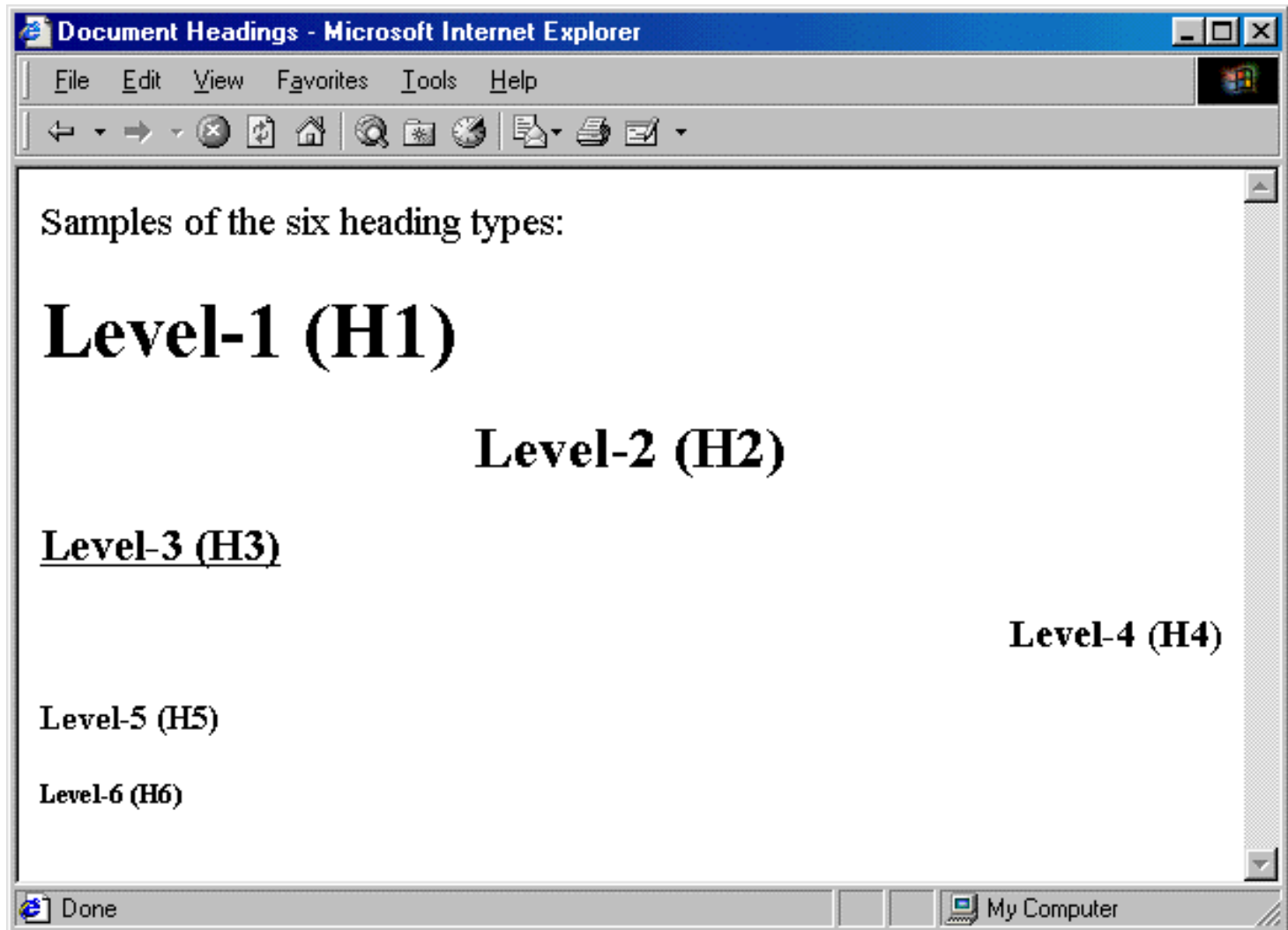
P – The Basic Paragraph

- **Attributes: ALIGN**
  - LEFT (default), RIGHT, CENTER. Same as headings.
  - Whitespace ignored (use <BR> for line break)
    - Consecutive <P>'s do not yield multiple blank lines
  - End Tag is Optional:

```
<BODY>
<P>
Paragraph 1
</P>
<P>
Paragraph 2
</P>
<P>
Paragraph 3
</P>
</BODY>
```
Fully-Specified

```
<BODY>
Paragraph 1
<P>
Paragraph 2
<P>
Paragraph 3
</BODY>
```
Equivalent with Implied Tags

www.corewebprogramming.com

Preformatted Paragraphs

- **The PRE Element**
  - \<PRE\> ... \</PRE\>

- **Attributes: WIDTH**
  - Expected width in characters. Not widely supported.

- **Problem: Special Characters**

```
<PRE>
if (a<b) {
  doThis();
} else {
  doThat();
}
</PRE>
```

Desired Character	HTML Required
<	<
>	>
&	&
"	"
Non-breaking space	

OL: Ordered (Numbered) Lists

- **OL Element**
  - <OL>
      <LI>…
      <LI>…

      ...
      </OL>

  - Attributes: TYPE, START, COMPACT

- **List entries: LI**
  - <LI ...> ... </LI> (End Tag Optional)
  - Attributes: (When inside OL) VALUE, TYPE

```
A sample list:
<OL>
   <LI>List Item One
   <LI>List Item Two
   <LI>List Item Three
</OL>
```

A sample list:

1. List Item One
2. List Item Two
3. List Item Three

Nested Ordered Lists

```
<OL TYPE="I">
<LI>Headings
  <LI>Basic Text Sections
  <LI>Lists
      <OL TYPE="A">
        <LI>Ordered
          <OL TYPE="1">
            <LI>The OL tag
              <OL TYPE="a">
                <LI>TYPE
                <LI>START
                <LI>COMPACT
              </OL>
            <LI>The LI tag
          </OL>
        <LI>Unordered
          <OL TYPE="1">
            <LI>The UL tag
            <LI>The LI tag
          </OL>
        <LI>Definition
          <OL TYPE="1">
            <LI>The DL tag
            <LI>The DT tag
            <LI>The DD tag
          </OL>
      </OL>
  <LI>Miscellaneous
</OL>
```

```
  I. Headings
 II. Basic Text Sections
III. Lists
        A. Ordered
              1. The OL tag
                    a. TYPE
                    b. START
                    c. COMPACT
              2. The LI tag
        B. Unordered
              1. The UL tag
              2. The LI tag
        C. Definition
              1. The DL tag
              2. The DT tag
              3. The DD tag
IV. Miscellaneous
```

UL: Unordered (Bulleted) Lists

- **UL Element**
  - <UL>
     <LI>…
     <LI>…
     ...
     </UL>

- **Attributes: TYPE, COMPACT**
  - TYPE is DISC, CIRCLE, or SQUARE

- **List entries: LI (TYPE)**
  - TYPE is DISC, CIRCLE, or SQUARE
    ```
    A sample list:
    <UL>
      <LI>List Item One
      <LI>List Item Two
      <LI>List Item Three
    </UL>
    ```

A sample list:

- List Item One
- List Item Two
- List Item Three

www.corewebprogramming.com

UL: Custom Bullets

```
<UL TYPE="DISC">
   <LI>The UL tag
      <UL TYPE="CIRCLE">
         <LI>TYPE
            <UL TYPE="SQUARE">
               <LI>DISC
               <LI>CIRCLE
               <LI>SQUARE
            </UL>
         <LI>COMPACT
      </UL>
   <LI>The LI tag
      <UL TYPE="CIRCLE">
         <LI>TYPE
            <UL TYPE="SQUARE">
               <LI>DISC
               <LI>CIRCLE
               <LI>SQUARE
            </UL>
         <LI>VALUE
      </UL>
</UL>
```

Unordered Lists

- The UL tag
  - TYPE
    - DISC
    - CIRCLE
    - SQUARE
  - COMPACT
- The LI tag
  - TYPE
    - DISC
    - CIRCLE
    - SQUARE
  - VALUE

Text-Level Elements

- **Physical Character Styles**
  - B, I, TT, U, SUB, SUP, SMALL, BIG, STRIKE, S, BLINK
  - FONT
    - SIZE
    - COLOR
    - FACE
  - BASEFONT
  - SIZE
- **Logical Character Styles**
  - EM, STRONG, CODE, SAMP, KBD, DFN, VAR, CITE

Text-Level Elements (Continued)

- **Hypertext Links**
  - A
    - HREF, NAME, TARGET, ...
- **Images**
  - IMG
    - SRC (required), ALT, ALIGN, WIDTH, HEIGHT, HSPACE, VSPACE, BORDER, USEMAP, ISMAP
- **Misc. Text-Level Elements**
  - BR (Explicit line break)
  - AREA (Client-side image maps)
  - APPLET (Java)
  - ...

Physical Character Styles, Example

```
...
<H1>Physical Character Styles</H1>
<B>Bold</B><BR>
<I>Italic</I><BR>
<TT>Teletype (Monospaced)</TT><BR>
<U>Underlined</U><BR>
Subscripts: f<SUB>0</SUB> + f<SUB>1</SUB><BR>
Superscripts: x<SUP>2</SUP> + y<SUP>2</SUP><BR>
<SMALL>Smaller</SMALL><BR>
<BIG>Bigger</BIG><BR>
<STRIKE>Strike Through</STRIKE><BR>
<B><I>Bold Italic</I></B><BR>
<BIG><TT>Big Monospaced</TT></BIG><BR>
<SMALL><I>Small Italic</I></SMALL><BR>
<FONT COLOR="GRAY">Gray</FONT><BR>
<DEL>Delete</DEL><BR>
<INS>Insert</INS><BR>
...
```

Physical Character Styles, Result
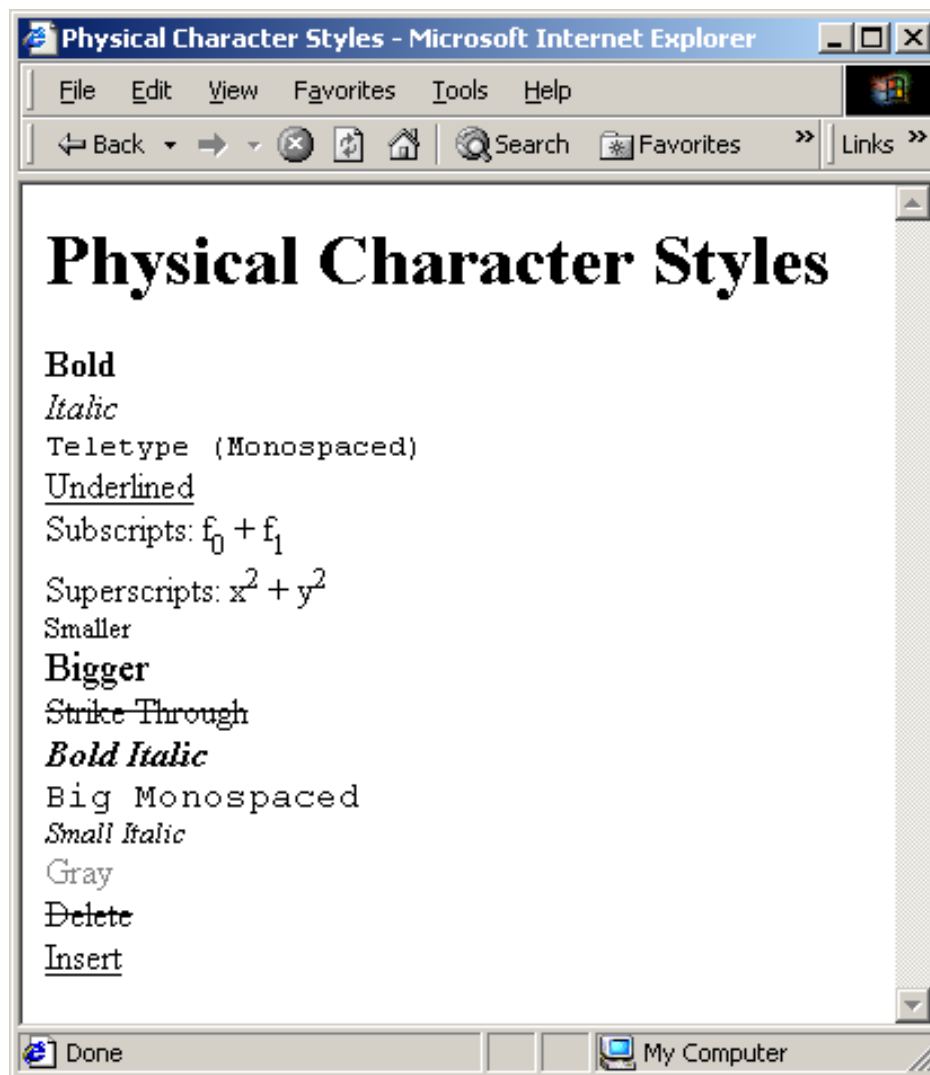
Logical Character Styles, Example

```
...
<H1>Logical Character Styles</H1>
<EM>Emphasized</EM><BR>
<STRONG>Strongly Emphasized</STRONG><BR>
<CODE>Code</CODE><BR>
<SAMP>Sample Output</SAMP><BR>
<KBD>Keyboard Text</KBD><BR>
<DFN>Definition</DFN><BR>
<VAR>Variable</VAR><BR>
<CITE>Citation</CITE><BR>
<EM><CODE>Emphasized Code</CODE></EM><BR>
<FONT COLOR="GRAY"><CITE>Gray Citation</CITE></FONT><BR>
<ACRONYM TITLE="Java Development Kit">JDK Acronym</ACRONYM>
...
```

Logical Character Styles, Result
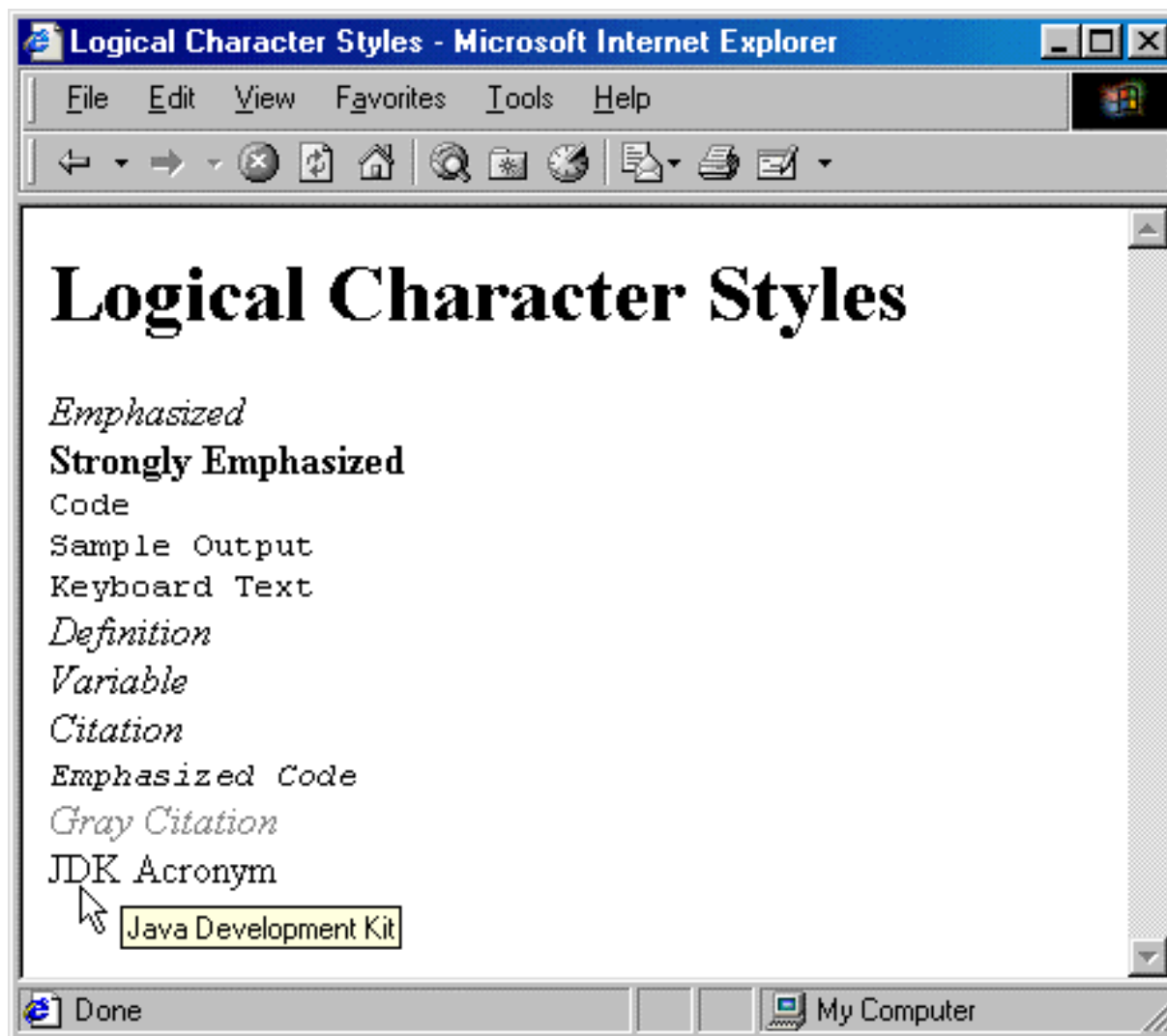
Hypertext Links

- **Links can contain images and other text-level elements (i.e., <A HREF…> ... </A>)**

- **Link to Absolute URL**

  - Use a complete URL beginning with http://

    ```
    Java is discussed in

    <A HREF="http://host/path/chapter2.html">

    Chapter 2</A>.
    ```

- **Link to Relative URL**

  - Use a filename or relative path to filename

    - Interpreted wrt location of current file

    ```
    Java is discussed in

    <A HREF="chapter2.html">Chapter 2</A>.
    ```

www.corewebprogramming.com

Hypertext Links (Continued)

- ## Link to Section
  - Use a section name (see below) preceded by #

    ```
    Images are discussed in
    <A HREF="#Section2">Section 2</A>.
    ```

- ## Link to Section in URL
  - Use absolute or relative URL, then #, then section name

    ```
    Images are discussed in
    <A HREF="chapter1.html#Section2">
    Sec. 2 of Chap. 1</A>.
    ```

- ## Naming a Section
  - Use <A NAME="..."> and do not include the pound sign

    ```
    <H2><A NAME="Section2">Images</A></H2>
    ```

www.corewebprogramming.com

IMG: Embedding Images

- ## Example
  ```
  <IMG SRC="SomeFile.gif" ALT="My Dog"
       WIDTH=400 HEIGHT=300>
  ```
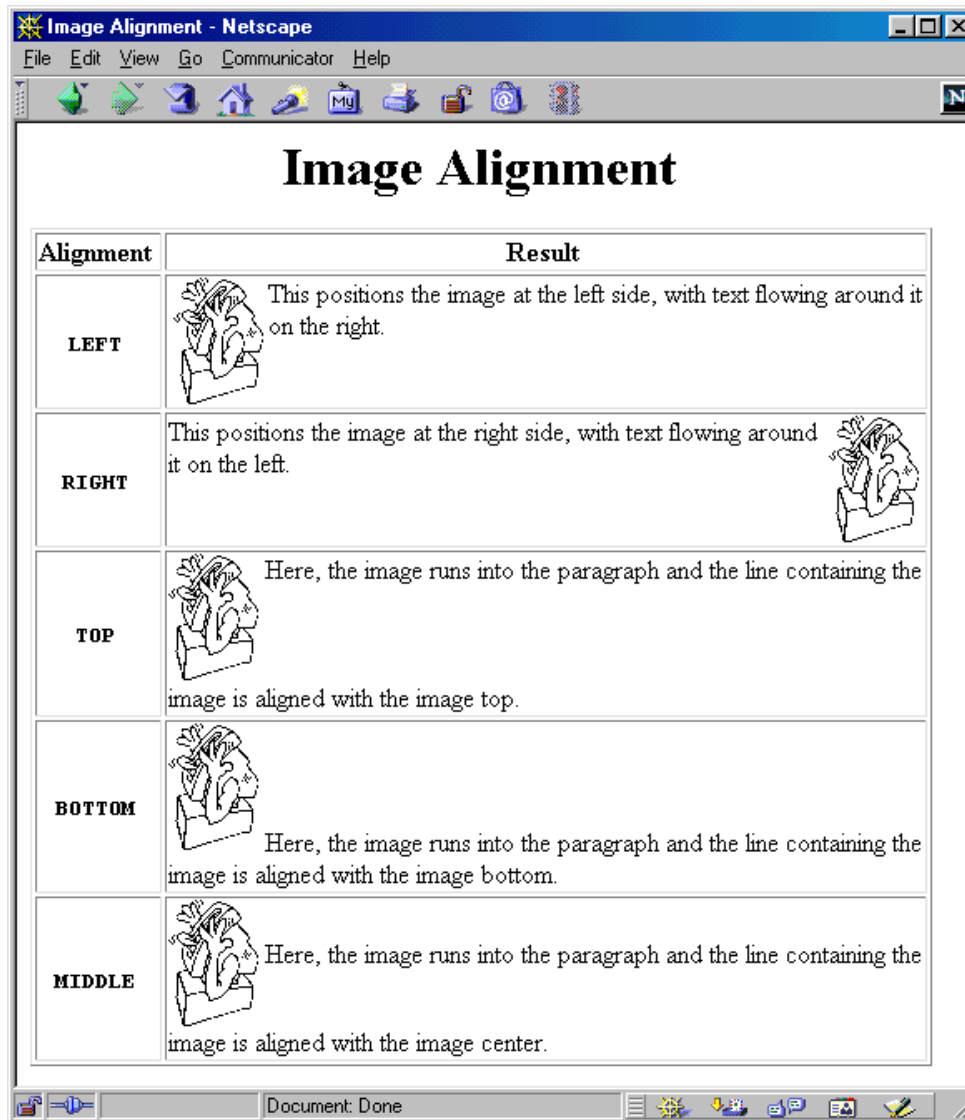
- ## Attributes:
  - SRC (required)
  - ALT (technically required)
  - ALIGN (see <BR CLEAR="ALL">)
  - WIDTH, HEIGHT
  - HSPACE, VSPACE
  - BORDER
  - USEMAP, ISMAP

Image Alignment, Example

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.0 Transitional//EN">
<HTML>
<HEAD><TITLE>Image Alignment</TITLE></HEAD>
<BODY>
<H1 ALIGN="CENTER">Image Alignment</H1>
<TABLE BORDER=1>
  <TR><TH>Alignment
      <TH>Result
  <TR><TH><CODE>LEFT</CODE>
      <TD><IMG SRC="rude-pc.gif" ALIGN="LEFT"
              ALT="Rude PC" WIDTH=54 HEIGHT=77>
          This positions the image at the left side,
          with text flowing around it on the right.
  <TR><TH><CODE>RIGHT</CODE>
      <TD><IMG SRC="rude-pc.gif" ALIGN="RIGHT"
              ALT="Rude PC" WIDTH=54 HEIGHT=77>
          This positions the image at the right side,
          with text flowing around it on the left.
    ...
</TABLE>
</BODY>
</HTML>
```
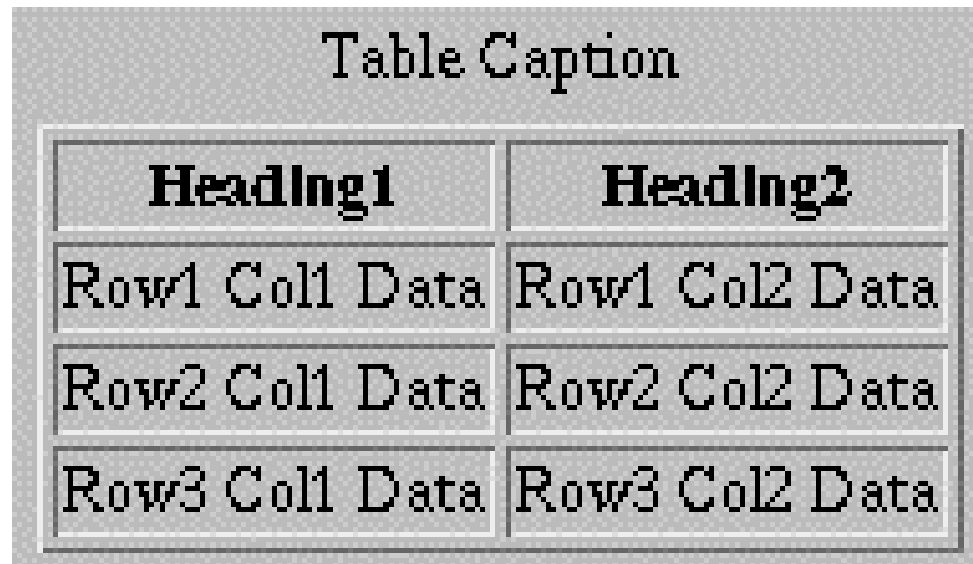
Image Alignment, Result

Tables

- ## Template

```
<TABLE BORDER=1>
   <CAPTION>Table Caption</CAPTION>
   <TR><TH>Heading1</TH>      <TH>Heading2</TH></TR>
   <TR><TD>Row1 Col1 Data</TD><TD>Row1 Col2 Data</TD></TR>
   <TR><TD>Row2 Col1 Data</TD><TD>Row2 Col2 Data</TD></TR>
   <TR><TD>Row3 Col1 Data</TD><TD>Row3 Col2 Data</TD></TR>
</TABLE>
```

Table Caption	
Heading1	**Heading2**
Row1 Col1 Data	Row1 Col2 Data
Row2 Col1 Data	Row2 Col2 Data
Row3 Col1 Data	Row3 Col2 Data

TABLE Element Attributes

- **ALIGN**
  - The ALIGN attribute gives the horizontal alignment of the table as a whole
  - Legal values are LEFT, RIGHT, and CENTER, with LEFT being the default
- **BORDER**
  - This specifies the width in pixels of the border around the table
  - This is in addition to the border around each cell (the CELLSPACING).
  - The default is zero, which also results in the visible 3D divider between cells being turned off
- **CELLSPACING**
  - This gives the space in pixels between adjacent cells. Drawn as a 3D line if BORDER is nonzero, otherwise empty space in the background color is used
  - The default is usually about 3

TABLE Element Attributes (Continued)

- **CELLPADDING**
  - CELLPADDING determines the empty space, in pixels, between the cell's border and the table element
  - The default is usually about 1
- **WIDTH**
  - This specifies the width of the table, either in pixels (<TABLE WIDTH=250>) or as a percentage of the current browser window width (<TABLE WIDTH="75%">)
- **BGCOLOR**
  - Specify the background color of the table TABLE (also legal for TR, TD, and TH)
- **BORDERCOLOR, BORDERCOLORDARK,**
- **BORDERCOLORLIGHT**
  - Non standard attributes supported by IE to specify the colors to user for the borders

TABLE Element Attributes (Continued)

- **BACKGROUND**
  - This nonstandard attribute supported by IE gives an image file that will be tiled as the background of the table
  - You might want to use style sheets instead.
- **RULES**
  - HTML 4.0 attribute that specifies which inner dividing lines are drawn
  - All are drawn if this attribute is omitted
  - Legal values are NONE, ROWS, COLS, and ALL
- **FRAME**
  - Specifies which outer borders are drawn
  - All four are drawn if this attribute is omitted
  - Legal values are BORDER or BOX (all), VOID (none), ABOVE (top), BELOW (bottom), HSIDES (top and bottom, despite the somewhat confusing name), VSIDES (left and right), LHS (left), and RHS (right)

www.corewebprogramming.com

Table **CAPTION**

- **Attribute**
  - ALIGN (Values: TOP, BOTTOM)
- **Usage**
  - An enclosing borderless table may give more flexibility than the built-in CAPTION.

TR: Table Row

- **TR is used to define each row in the table**
- **Each row will then contain TH and/or TD entries**
- **ALIGN**
  - ALIGN (legal values LEFT, RIGHT, or CENTER) is used to set the default horizontal alignment for table cells
- **VALIGN**
  - VALIGN (legal values TOP, MIDDLE, or BOTTOM) is used to set the default vertical alignment for table cells
- **BGCOLOR**
  - Sets the color for the table row, overriding any values set for the table as a whole via the BGCOLOR attribute of TABLE
- **BORDERCOLOR, BORDERCOLORDARK,**
- **BORDERCOLORLIGHT**
  - Supported only by Internet Explorer, these specify the colors to use for the row borders

Table Cells: TH and TD

- **COLSPAN**
  - COLSPAN defines a heading or cell data entry that spans multiple columns

```
<TABLE BORDER=1>
<TR><TH COLSPAN=2>Col 1&2 Heading
    <TH>Col3 Heading
  <TR><TD>Col1 Data
      <TD>Col2 Data
      <TD>Col3 Data
</TABLE>
```

Col 1&2 Heading		Col3 Heading
Col1 Data	Col2 Data	Col3 Data

Table Cells: TH and TD (Continued)

- **ROWSPAN**
  - ROWSPAN defines a heading or cell data entry that spans multiple rows; similar to COLSPAN
- **ALIGN**
  - LEFT, RIGHT, CENTER, JUSTIFY and CHAR.
  - E.g., the following aligns entries on a decimal point

    ```
    <TD ALIGN="CHAR" CHAR=".">
    ```

- **VALIGN**
  - TOP, BOTTOM, MIDDLE
- **WIDTH, HEIGHT**
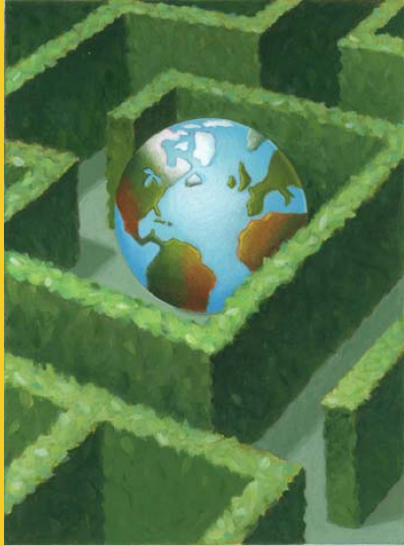 - Values in pixels only (no percentages officially allowed)
- **NOWRAP**
  - Use with caution
- **BGCOLOR, BACKGROUND**
  - Same as for TABLE and TR

www.corewebprogramming.com

Summary

- **A DOCTYPE is required to validate the document**
- **HTML document should have an enclosing HTML element, a HEAD (TITLE is required) and a BODY**
- **Documents are composed of block-level and text-level elements**
  - Text-level elements must be inside block-level elements, not vice versa
- **Hypertext links, <A HREF="…">, can be absolute or relative**
  - A link to a named section is denoted by *#section*
- **Tables are composed of main table element, <TABLE>; rows, <TR>; table headers, <TH>; and table data, <TD>**
  - Use BGCOLOR to give background colors to tables, rows, or cells
  - Use ROWSPAN or COLSPAN to join cells

www.corewebprogramming.com

Questions?