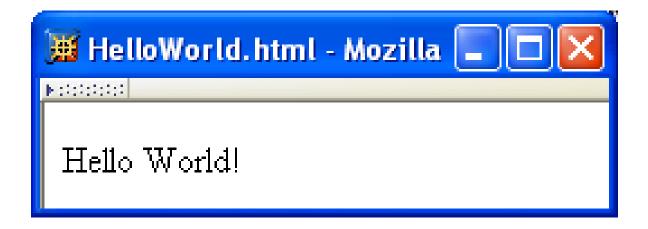
Markup Languages: XHTML 1.0

HTML "Hello World!"

Document Type Declaration

Document Instance

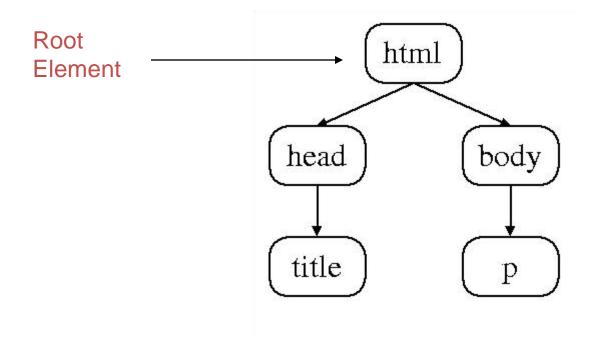
HTML "Hello World"



HTML Tags and Elements

- Any string of the form < ... > is a tag
- All tags in document instance of Hello World are either end tags (begin with </) or start tags (all others)
 - Tags are an example of markup, that is, text treated specially by the browser
 - Non-markup text is called character data and is normally displayed by the browser
- String at beginning of start/end tag is an element name
- Everything from start tag to matching end tag, including tags, is an element
 - Content of element excludes its start and end tags

HTML Element Tree



HTML Root Element

- Document type declaration specifies name of root element:
 - <!DOCTYPE html
- Root of HTML document must be html
- XHTML 1.0 (standard we will follow) requires that this element contain the xml namespace xmlns attribute specification (name/value pair)

```
<html xmlns="http://www.w3.org/1999/xhtml">
```

HTML head and body Elements

- The body element contains information displayed in the browser client area
- The head element contains information used for other purposes by the browser:
 - title (shown in title bar of browser window)
 - scripts (client-side programs)
 - style (display) information
 - etc.

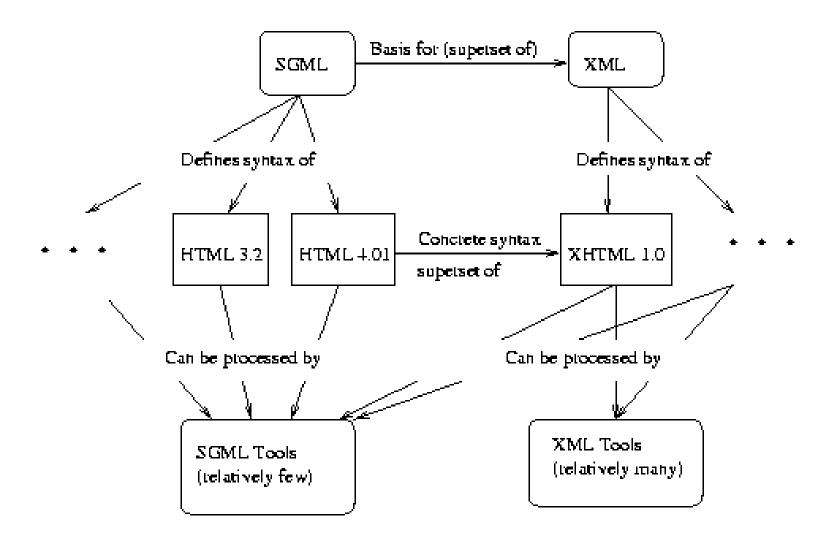
HTML History

- 1990: HTML invented by Tim Berners-Lee
- 1993: Mosaic browser adds support for images, sound, video to HTML
- 1994-~1997: "Browser wars" between Netscape and Microsoft, HTML defined operationally by browser support
- ~1997-present: Increasingly, World-Wide Web Consortium (<u>W3C</u>) recommendations define HTML

HTML Versions

- HTML 4.01 (Dec 1999) syntax defined using Standard Generalized Markup Language (SGML)
- XHTML 1.0 (Jan 2000) syntax defined using Extensible Markup Language (XML)
- Primary differences:
 - HTML allows some tag omissions (e.g., end tags)
 - XHTML element and attribute names are lower case (HTML names are case-insensitive)
 - XHTML requires that attribute values be quoted

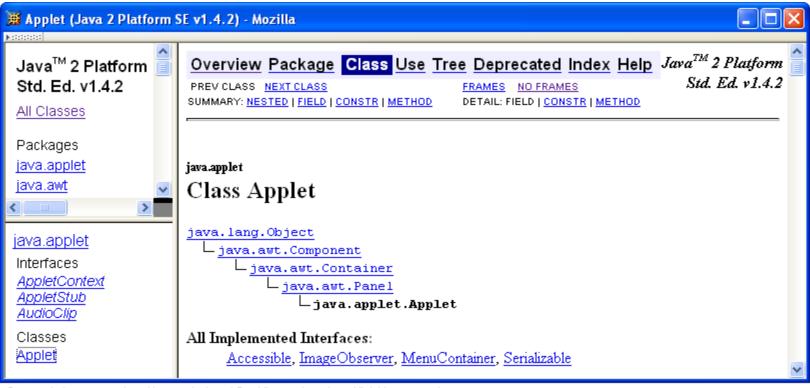
SGML and XML



HTML "Flavors"

- For HTML 4.01 and XHTML 1.0, the document type declaration can be used to select one of three "flavors":
 - Strict: W3C ideal
 - Transitional: Includes deprecated elements and attributes (W3C recommends use of *style sheets* instead)
 - Frameset: Supports frames (subwindows within the client area)

HTML Frameset



Screen shots are reproduced by permission of Sun Microsystems Inc. All rights reserved.

HTML Document Type Declarations

XHTML 1.0 Strict:
 <!DOCTYPE html
 PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">

XHTML 1.0 Frameset:
 <!DOCTYPE html
 PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
 "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">

HTML 4.01 Transitional: <!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">

XHTML White Space

- Four white space characters: carriage return, line feed, space, horizontal tab
- Normally, character data is normalized:
 - All white space is converted to space characters
 - Leading and trailing spaces are trimmed
 - Multiple consecutive space characters are replaced by a single space character

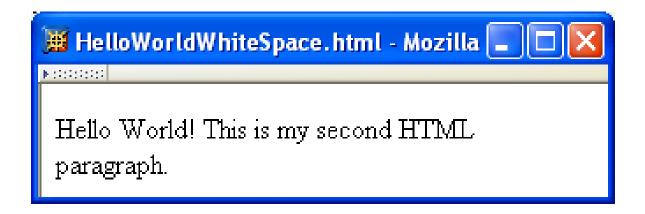
XHTML White Space

```
<body>

    Hello World!

    This is my second HTML paragraph.

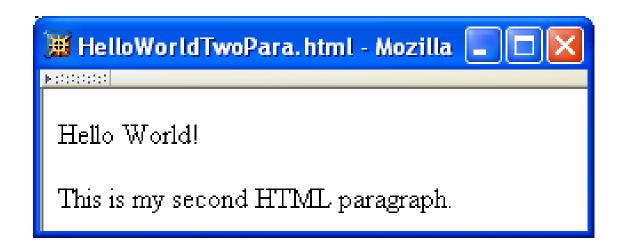
</body>
```



XHTML White Space

```
 Hello World!

 This is my second HTML paragraph.
```



```
<!DOCTYPE html
                    PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
                    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
            <html xmlns="http://www.w3.org/1999/xhtml">
              <head>
Misspelled

→ <titl>

element name
                  HelloWorldBadElt.html
                </title>
              </head>
              <body>
                >
                  Hello World!
                </body>
            </html>
```





- Browsers ignore tags with unrecognized element names, attribute specifications with unrecognized attribute names
 - Allows evolution of HTML while older browsers are still in use
- Implication: an HTML document may have errors even if it displays properly
- Should use an <u>HTML validator</u> to check syntax

Example for non-frame browsers (old)

- Since < marks the beginning of a tag, how do you include a < in an HTML document?
- Use markup known as a reference
- Two types:
 - Character reference specifies a character by its Unicode code point
 - For <, use < or < or <
 - Entity reference specifies a character by an HTML-defined name
 - For <, use <

TABLE 2.2: Example entity and character references.

Character	Entity Reference	Character Reference (decimal)
<	<	<
>	>	>
&	&	&
"	"	"
,	'	'
0	©	©
ñ	ñ	ñ
α	α	α
\forall	∀	∀

- Since < and & begin markup, within character data or attribute values these characters must always be represented by references (normally < and &)
- Good idea to represent > using reference (normally >)
 - Provides consistency with treatment of <
 - Avoids accidental use of the reserved string]]>

- Non-breaking space () produces space but counts as part of a word
 - Ex: keep together keep together ...



Non-breaking space often used to create multiple spaces (not removed by normalization)

XHTML Attribute Specifications

Example:

```
<html xmlns="http://www.w3.org/1999/xhtml" lang="en" xml:lang="en">
```

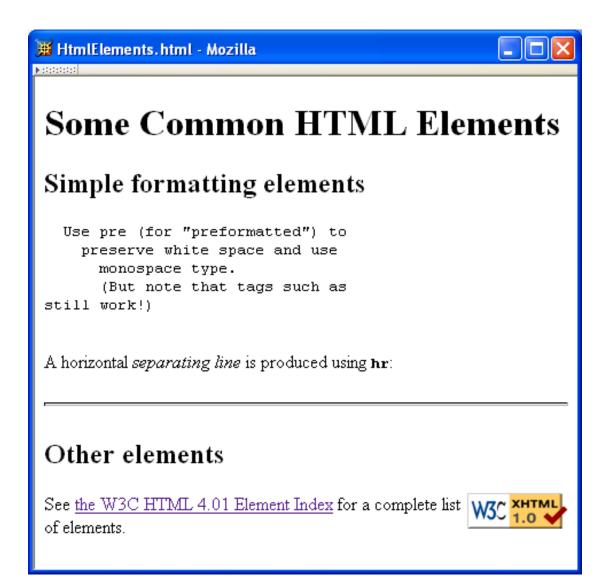
- Valid attribute names specified by HTML recommendation (or XML, as in xml:lang)
- Attribute values must be quoted (matching single or double quotes)
- Multiple attribute specifications are spaceseparated, order-independent

XHTML Attribute Values

 Can contain embedded quotes or references to quotes

```
value = "Ain't this grand!"
value = "He said, " She said", then sighed."
value = "He said, "She said", then sighed."
```

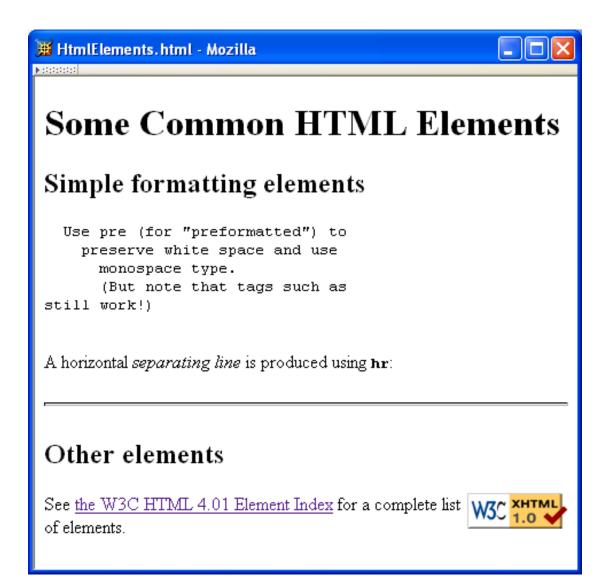
Best to normalize attribute values yourself for optimal browser compatibility



Headings are produced using h1, h2, ..., h6 elements:

```
<h1>
Some Common HTML Elements
</h1>
<h2>
```

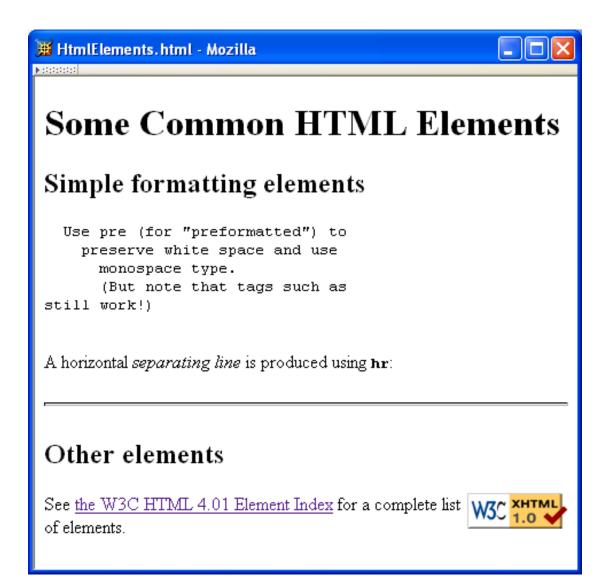
- Should use h1 f Simple formatting elements xt highest, etc. $^{</h2>}$
 - Change style (next chapter) if you don't like the "look" of a heading



 Use pre to retain format of text and display using monospace font:

- br element represents line break
- br is example of an empty element, i.e., element that is not allowed to have content
- XML allows two syntactic representations of empty elements
 - Empty tag syntax
 is recommended for browser compatibility
 - XML parsers also recognize syntax

 followed immediately by end tag), but many browsers do not understand this for empty elements

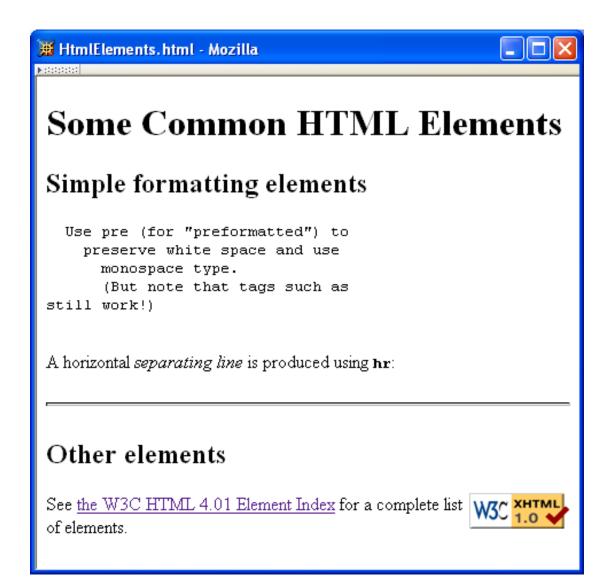


- Text can be formatted in various ways:
 - Apply style sheet technology (next chapter) to a span element (a styleless wrapper):

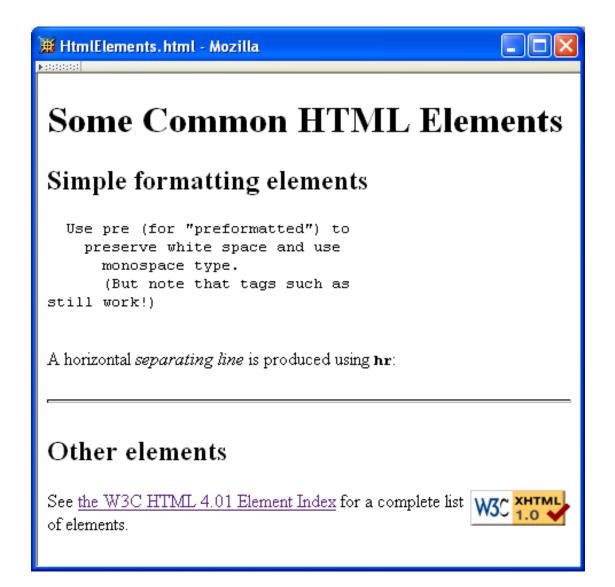
 - Use strong>hrement
 - Not recommended, but frequently used

TABLE 2.3: HTML font style elements.

Element	Font used by content
b	Bold-face
i	Italic
tt	"Teletype" (fixed-width font)
big	Increased font size
small	Decreased font size



- Horizontal rule is produced using hr
- Also an empty element
- Style can be modified using style sheet technology



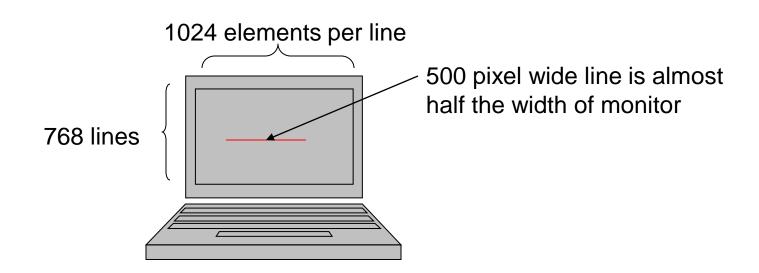
Images can be embedded using img element

```
<img
src="http://www.w3.org/Icons/valid-xhtml10"
alt="Valid XHTML 1.0!" height="31" width="88"
style="float:right" />
```

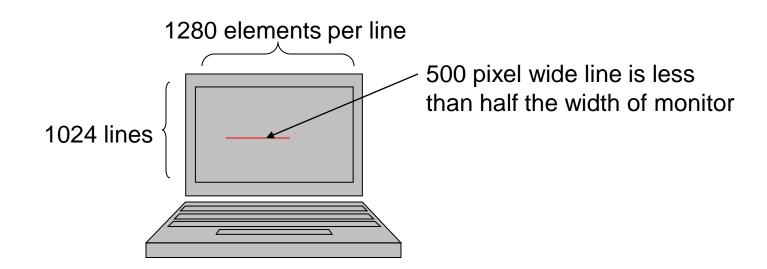
- Attributes:
 - Src: URL of image file (required). Browser generates a GET request to this URL.
 - alt: text description of image (required)
 - height / width: dimensions of area that image will occupy (recommended)

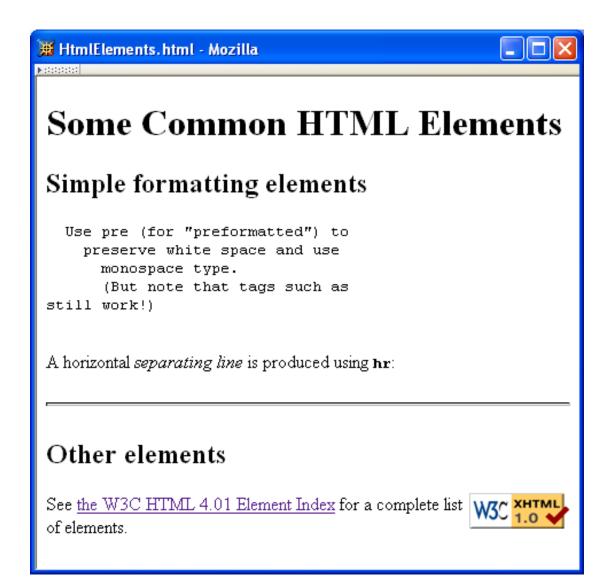
- If height and width not specified for image, then browser may need to rearrange the client area after downloading the image (poor user interface for Web page)
- If height and width specified are not the same as the original dimensions of image, browser will resize the image
- Default units for height and width are "picture elements" (pixels)
 - Can specify percentage of client area using string such as "50%"

Monitor resolution determines pixel size



Monitor resolution determines pixel size





Hyperlinks are produced by the anchor element a

```
See
<a href="http://www.w3.org/TR/html4/index/elements.html">the
W3C HTML 4.01 Element Index</a>
```

- C for a complete list of elements.
 request to URL specified in hret attribute and render response in client area
- Content of anchor element is text of hyperlink (avoid leading/trailing space in content)

Anchors can be used as source (previous example) or destination

• The f 1 to reference a destination anchor

...
near) top of client area

Comments are a special form of tag

```
<!-- Notice that img must nest within a "block" element,
such as p -->
```

Not allowed to use -- within comment

```
<!-- This is NOT
    -- a good comment.
    -->

<!-- Can't end with more than two dashes! --->
```

Nesting Elements

 If one element is nested within another element, then the content of the inner element is also content of the outer element

XHT | <tt>hr</tt> | properly |
 nested

X <tt>hr</tt>

Nesting Elements

- Most HTML elements are either block or inline
 - Block: browser automatically generates line breaks before and after the element content
 - Ex: p, div
 - Inline: element content is added to the "flow"
 - Ex: span, tt, strong, a

Nesting Elements

- Syntactic rules of thumb:
 - Children of body must be blocks
 - Blocks can contain inline elements
 - Inline elements cannot contain blocks
- Specific rules for each version of (X)HTML are defined using SGML or XML (covered later)

Consider an start tag containing attribute specification

- This is an src="valid-xhtml10.png" JRL: it is interpreted relative to the URL of the document that contains the img tag
 - If document URL is http://localhost:8080/MultiFile.html
 then relative URL above represents absolute URL http://localhost:8080/valid-xhtml10.png

TABLE 2.4: Absolute URL's corresponding to relative URL's when the base URL is http://www.example.org/a/b/c.html.

Relative URL	Absolute URL
d/e.html	http://www.example.org/a/b/d/e.html
/f.html	http://www.example.org/a/f.html
//g.html	http://www.example.org/g.html
/h/i.html	http://www.example.org/a/h/i.html
/j.html	http://www.example.org/j.html
/k/l.html	http://www.example.org/k/l.html

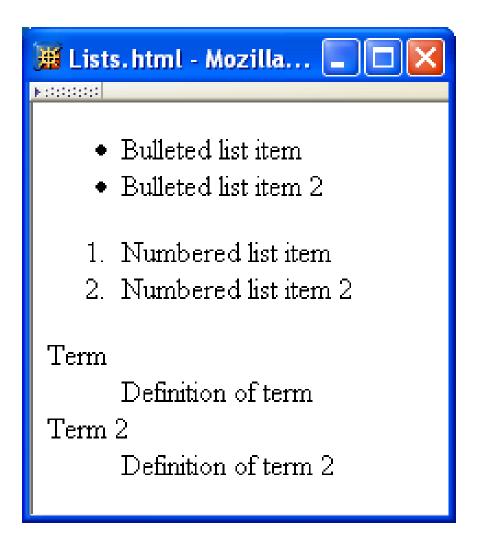
- Query and fragment portions of a relative URL are appended to the resulting absolute URL
 - Example: If document URL is http://localhost:8080/PageAnch.html and it contains the anchor element

```
then the corresponding absolute URL is

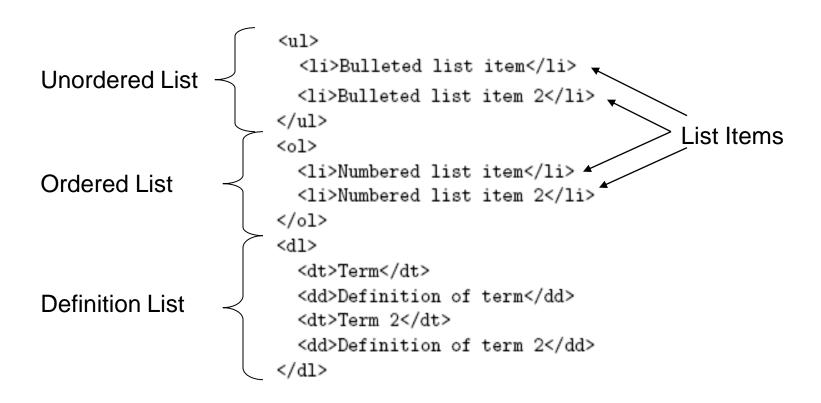
<a href="#section1">...</a>
```

- Advantages:
 - Shorter than absolute URL's
 - Primary: can change the URL of a document (e.g., move document to a different directory or rename the server host) without needing to change URL's within the document
- Should use relative URL's whenever possible

Lists

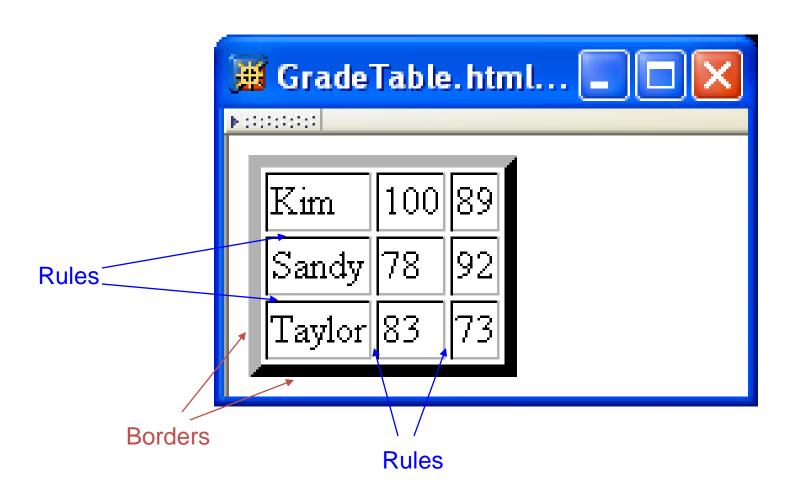


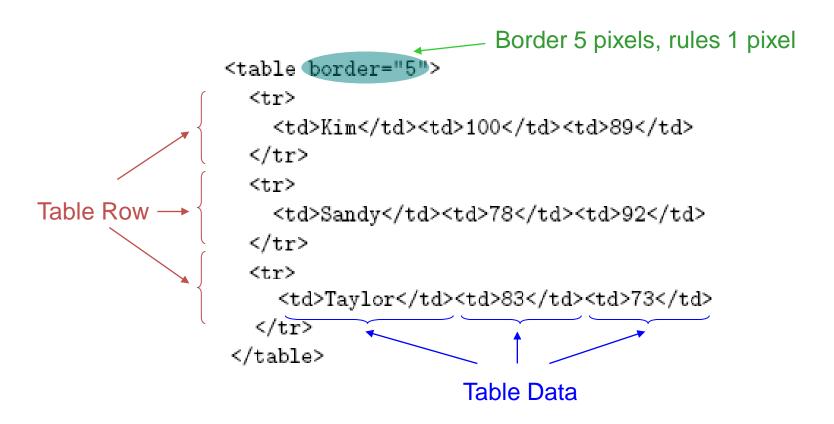
Lists

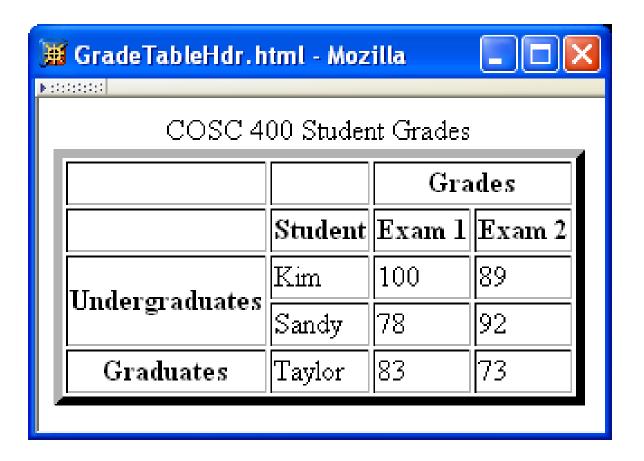


Lists

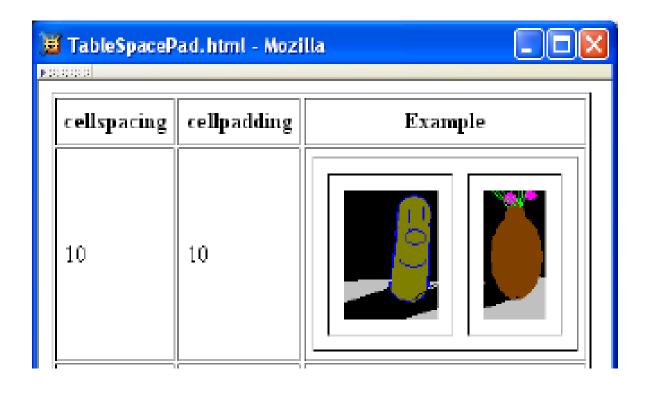




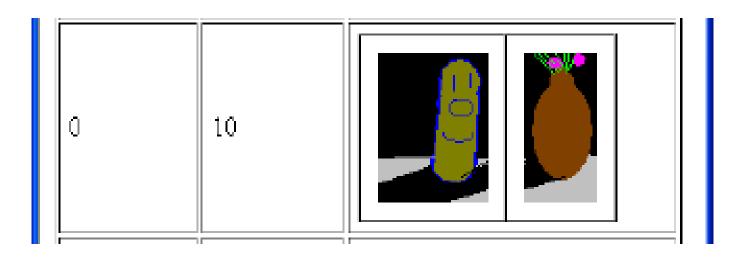




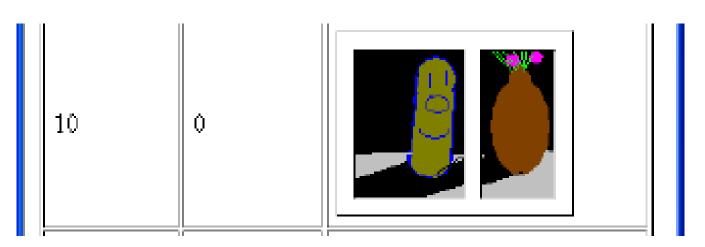
```
<caption>
 COSC 400 Student Grades
</caption>
  Grades
Table Header
 StudentExam 1Exam 2
UndergraduatesKim10089
Sandy7892
GraduatesTaylor8373
```



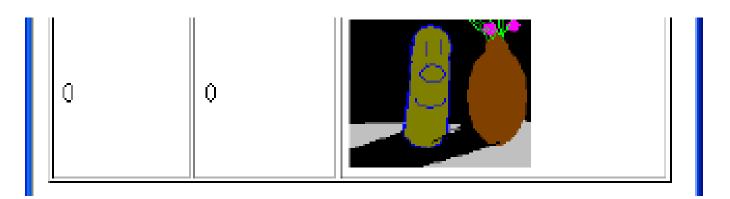
cellspacing cellpadding

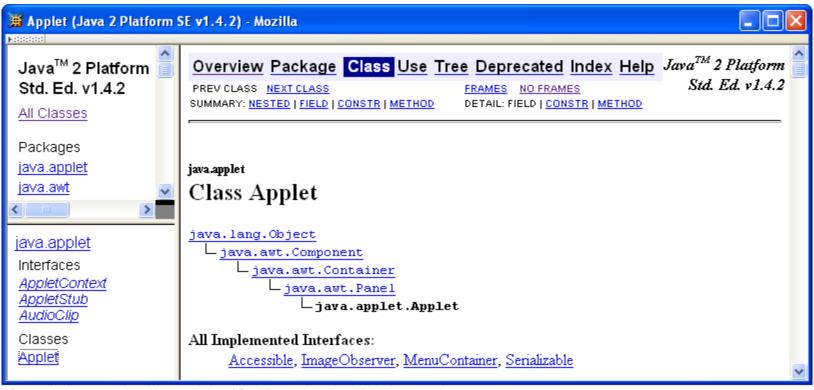


cellspacing cellpadding



cellspacing cellpadding





Screen shots are reproduced by permission of Sun Microsystems Inc. All rights reserved.

```
<!DOCTYPE html
  PUBLIC "-//W3C//DTD XHTML 1.0 Frameset//EN"
   "http://www.w3.org/TR/xhtml1/DTD/xhtml1-frameset.dtd">
<html xmlns="http://www.w3.org/1999/xhtml">
  <head>
   <title>Java 2 Platform SE v1.4.2</title>
  </head>
 <frameset rows="1*,2*">
     <frame src="overview-frame.html"</pre>
       id="upperLeftFrame" name="upperLeftFrame"></frame>
     <frame src="allclasses-frame.html"</pre>
       id="lowerLeftFrame" name="lowerLeftFrame"></frame>
    </frameset>
   <frame src="overview-summary.html"</pre>
       id="rightFrame" name="rightFrame"></frame>
  </frameset>
</html>
```

 Hyperlink in one frame can load document in another:

• \'
id/name of a frame

- User interface issues:
 - What happens when the page is printed?
 - What happens when the Back button is clicked?
 - How should assistive technology "read" the page?
 - How should the information be displayed on a small display?
- Recommendation: avoid frames except for applications aimed at "power users"

Forms

📕 LifeStory. html - Mozilla		
F (1) (1) (1) (1) (1) (1) (1) (1) (1) (1)		
Enter your name:		
Give your life's story in 100 words or less:		
Check all that apply to you: 🔲 tall 🔲 funny 🔲 smart		
Publish My Life's Story		

Forms

Each form is content of a form element

Forms

action specifies URL where form data is sent in an HTTP request

HTTP request method (lower case)

- The XHTML grammar require any child of the form element to be a block
- Many form elements are actually inline, so including a block element on top such a div or a table is a simple way to be compliant with the grammar

📕 LifeStory. html - Mozilla	
N (COCCUPATION COCCUPATION COC	
Enter your name:	
Give your life's story in 100 words or less:	
Check all that apply to you: 🔲 tall 🔲 funny 🔲 smart	
Publish My Life's Story	

Only one control inside a label element!

📕 LifeStory. html - Mozilla	
N (COCCUPATION COCCUPATION COC	
Enter your name:	
Give your life's story in 100 words or less:	
Check all that apply to you: 🔲 tall 🔲 funny 🔲 smart	
Publish My Life's Story	

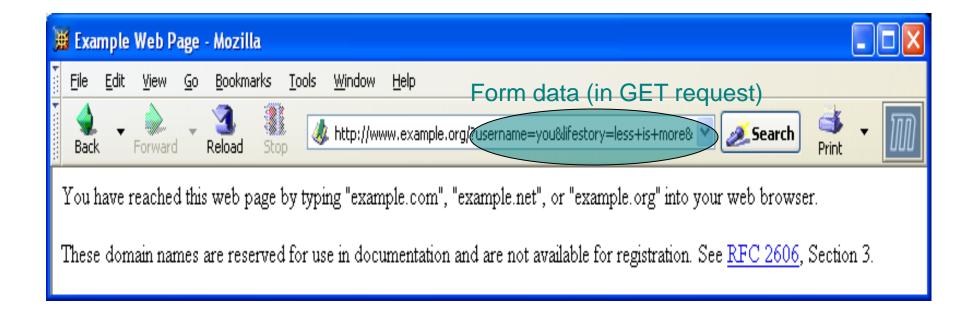
```
Check all that apply to you:
                            Checkbox control
    <label>_
    <input type="checkbox" name="boxgroup1" value="tall" />tall
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="funny" />funny
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="smart" />smart
    </label>
    <br /><br />
    <input type="submit" name="doit" value="Publish My Life's Story" />
  </div>
</form>
```

```
Check all that apply to you:
                                 Value sent in HTTP request if box is checked
    <label>
      <input type="checkbox" name="boxgroup1" value="tall" />tall
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="funny" />funny
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="smart" />smart
    </label>
    <br /><br />
    <input type="submit" name="doit" value="Publish My Life's Story" />
  </div>
</form>
```

```
Controls can share a common name
   Check all that apply to you:
    <label>
      <input type="checkbox" name="boxgroup1" value="tall" />tall
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="funny" />funny
    </label>
    <label>
      <input type="checkbox" name="boxgroup1/ value="smart" />smart
    </label>
    <br /><br />
    <input type="submit" name="doit" value="Publish My Life's Story" />
  </div>
</form>
```

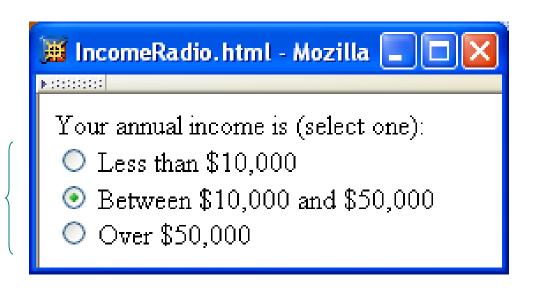
```
Check all that apply to you:
    <label>
      <input type="checkbox" name="boxgroup1" value="tall" />tall
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="funny" />funny
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="smart" />smart
    </label>
    <br /><<u>br /></u>
   <input type="submit" name="doit" value="Publish My Life's Story" />
  </div>
          Submit button: form data sent to action URL if button is clicked
</form>
```





```
Check all that apply to you:
    <label>
      <input type="checkbox" name="boxgroup1" value="tall" />tall
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="funny" />funny
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="smart" />smart
    </label>
    <br /><br />
    <input type="submit" name="doit" value="Publish My Life's Story</pre>
  </div>
                        Displayed on button and sent to server if button clicked
</form>
```

Radio buttons: at most one can be selected at a time.



```
Your annual income is (select one):<br />
<label>_____
                    Radio button control
<input type="radio" name="radgroup1" value="0-10" />
    Less than $10,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="10-50"</pre>
         checked="checked" />
   Between $10,000 and $50,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="&gt;50" />
   Over $50,000
</label>
```

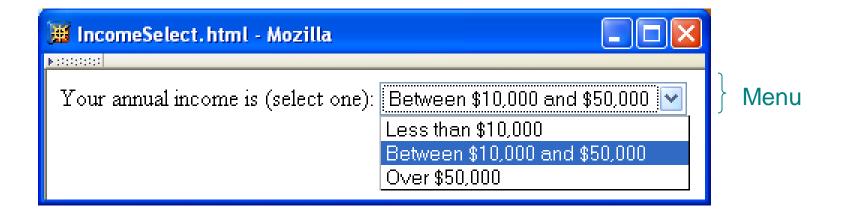
```
Your annual income is (select one):<br />
<label>
  <input type="radio" (name="radgroup1")</pre>
                                         value="0-10" />
    Less than $10,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="10-50"</pre>
         checked="checked" />
    Between $10,000 and $50,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="&gt;50" />
    Over $50,000
</label>
           All radio buttons with the same name form a button set
```

```
Your annual income is (select one):<br />
<label>
  <input type="radio" (name="radgroup1")</pre>
                                         value="0-10" />
    Less than $10,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="10-50"</pre>
         checked="checked" />
    Between $10,000 and $50,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="&gt;50" />
    Over $50,000
</label>
           Only one button of a set can be selected at a time
```

```
Your annual income is (select one):<br />
<label>
  <input type="radio" name="radgroup1" value="0-10" />
    Less than $10,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="10-50"</pre>
        checked="checked"/> This button is initially selected
    Between $10,000 and $50,000 (checked attribute also applies
                                 to check boxes)
</label><br />
<label>
  <input type="radio" name="radgroup1" value="&gt;50" />
    Over $50,000
</label>
```

```
Your annual income is (select one):<br />
<label>
  <input type="radio" name="radgroup1" value="0-10" />
    Less than $10,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="10-50"</pre>
        checked="checked">/> Boolean attribute: default false,
    Between $10,000 and $50,000 set true by specifying name as
</label><br />
                                 value
<label>
  <input type="radio" name="radgroup1" value="&gt;50" />
    Over $50,000
</label>
```

```
Your annual income is (select one):<br />
<label>
  <input type="radio" name="radgroup1" value="0-10" />
    Less than $10,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value="10-50"</pre>
         checked="checked" />
    Between $10,000 and $50,000
</label><br />
<label>
  <input type="radio" name="radgroup1" value=\( \delta gt; 50" \)</pre>
    Over $50,000
</label>
                                    Represents string: >50
```



- Other form controls:
 - Fieldset (grouping)
 - Password
 - Clickable image
 - Non-submit buttons
 - Hidden (embed data)
 - File upload
 - Hierarchical menus

