WEB TECHNOLOGIES A COMPUTER SCIENCE PERSPECTIVE

JEFFREY C. JACKSON

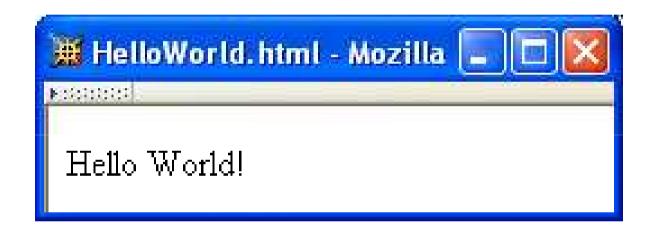
Chapter 2
Markup Languages:
XHTML 1.0

HTML "Hello World!"

Document Type Declaration

Document Instance

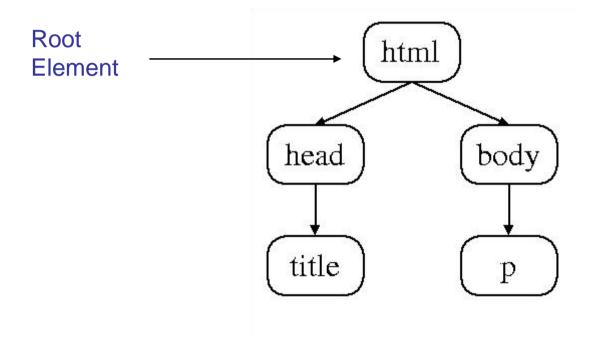
HTML "Hello World"



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</pre>
- Root of HTML document must be html
- XHTML 1.0 (standard we will follow)
 requires that this element contain 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.

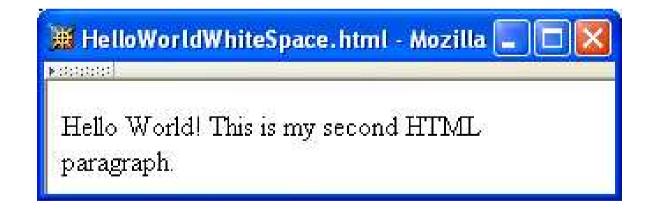
XHTML White Space

```
<body>

    Hello World!

This is my second HTML paragraph.

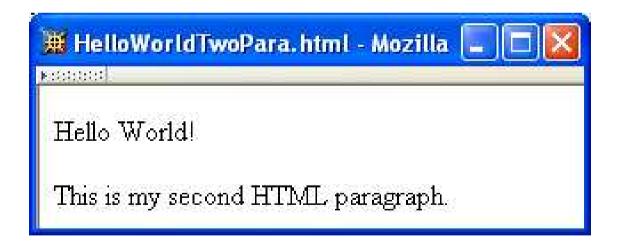
</body>
```



XHTML White Space

```
 Hello World!

 This is my second HTML paragraph.
```



Unrecognized HTML Elements

```
<! DOCTYPE html
                    PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
                    "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
            <htnl xmlns="http://www.w3.org/1999/xhtml">
              <head>
Misspelled
              → <titl>
element name
                  HelloWorldBadElt.html
                </title>
              </head>
              <body>
                >
                  Hello World!
                </body>
            </html>
```

Unrecognized HTML Elements



Unrecognized HTML Elements



- 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 HTMLdefined name
 - For <, use <

TABLE 2.2: Example entity and character references.

| Character | Entity Reference | Character Reference (decimal) |
|-----------|------------------|-------------------------------|
| < | < | < |
| > | > | > |
| & | & | & |
| 11 | " | " |
| , | ' | ' |
| 0 | © | © |
| ñ | ñ | ñ |
| α | α | α |
| A | ∀ | ∀ |

- Since < and & begin markup, within character data or attribute values these characters must *always* be represented by references (normally &1t; 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)

```
Hey, you.  Yes.  I am talking to you.

*/p>
*/p>
*/ex, you. Yes. I am talking to you.

*/p>
displays as two

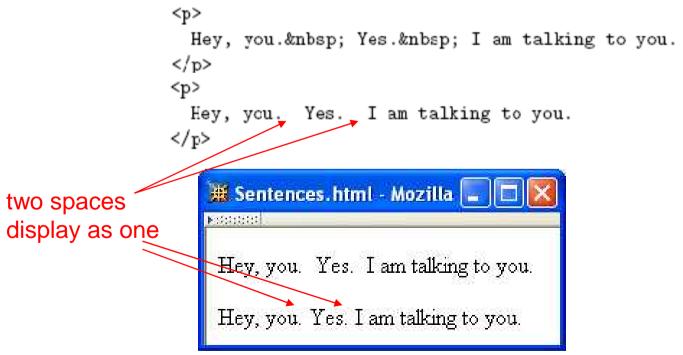
**Sentences.html - Mozilla **

**Hey, you. Yes. I am talking to you.

**Hey, you. Yes. I am talking to you.

**Hey, you. Yes. I am talking to you.
```

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



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

```
<h1>
Some Common HTML Elements
</h1>
<h2>
Simple formatting elements
</h2>
```

- Should use h1 for highest level, h2 for next highest, etc.
 - 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:

```
Use pre (for "preformatted") to
   preserve white space and use
   monospace type.
   (But note that tags such as<br />still work!)
```

Note that any embedded markup (such as

) is still treated as markup!

- 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

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

Strong element

- Strong is a logical state and bold is a physical state.
- Logical states separate presentation from the content and by doing so allow for it to be expressed in many different ways.
- Instead of rendering some text as bold you want to render it red, or a different size, or underlined, or whatever.

TABLE 2.3: HTML font style elements.

| Element | Font used by content |
|---------|-------------------------------|
| Ъ | 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)

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>
for a complete list of elements.
```

- Clicking on a hyperlink causes browser to issue GET request to URL specified in href 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

```
<a id="section1" name="section1"></a>
```

 The fragment portion of a URL is used to reference a destination anchor

```
<a href="http://www.example.org/PageWithAnchor.html#section1">...
```

 Browser scrolls so destination anchor is at (or near) top of client area

Example for Destination Anchor Element

```
<html>
<body>
<a id="a1" href="#intradocument">Click this Link</a>
Praba
Generate as much as space you want
/
<h2 id="intradocument">Intradocument Linking</h2>
The link that jumps to that location looks like this:
<a href="#p1">Home</a>
</body>
</html>
```

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

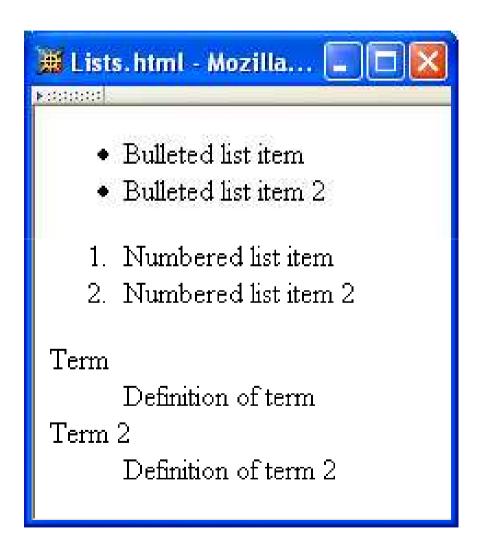
```
<!-- This is NOT
    -- a good comment.
    -->

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

Block and Inline Elements

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

Lists



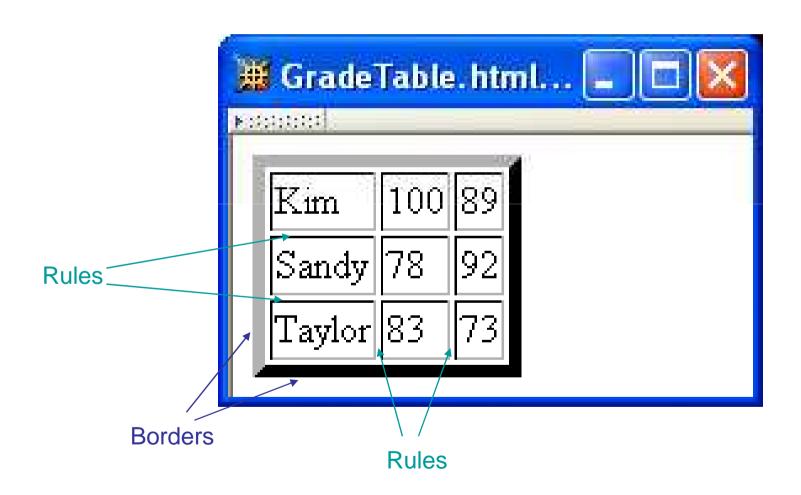
Lists

```
<u1>
                     Bulleted list item
Unordered List
                     Bulleted list item 2
                   List Items
                   <01>
                     Numbered list item
Ordered List
                     Numbered list item 2
                   <d1>
                     <dt>Term</dt>
                     <dd>Definition of term</dd>
Definition List
                     <dt>Term 2</dt>
                     <dd>Definition of term 2</dd>
                   </d1>
```

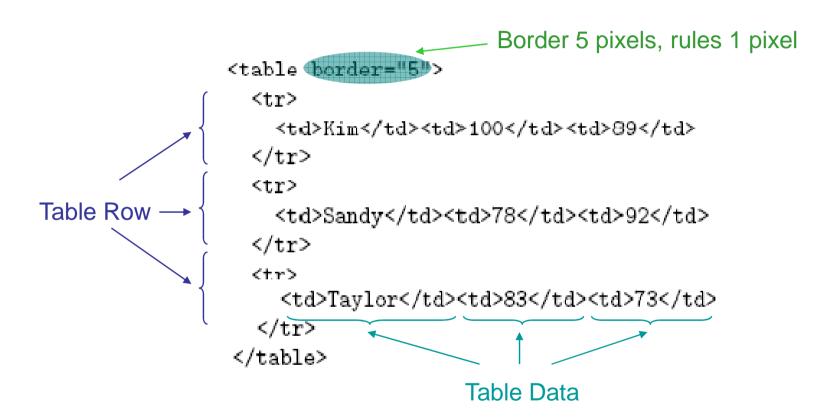
Lists



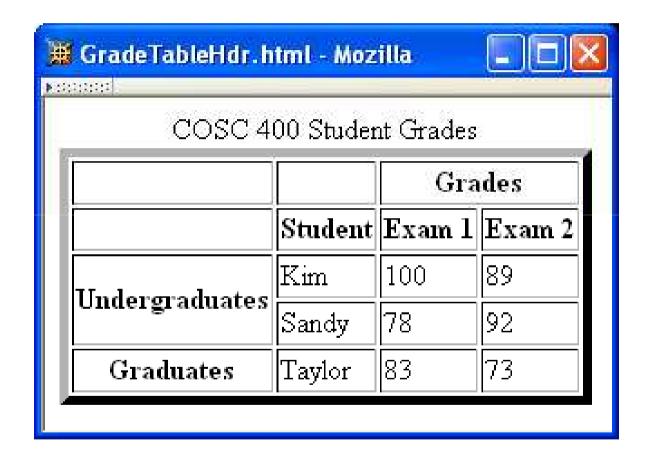
Tables



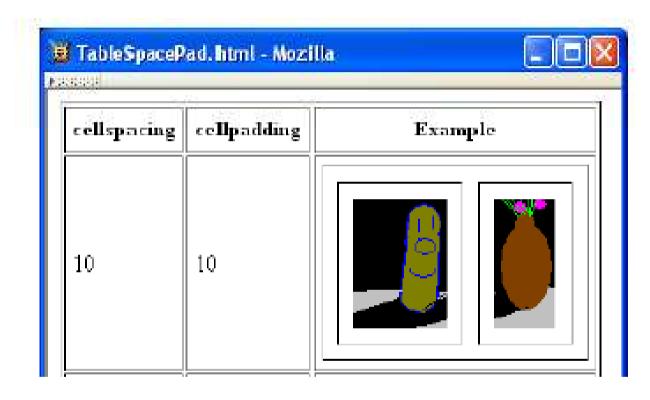
Tables



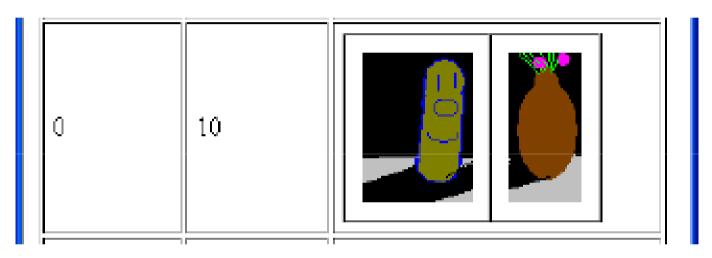
Tables



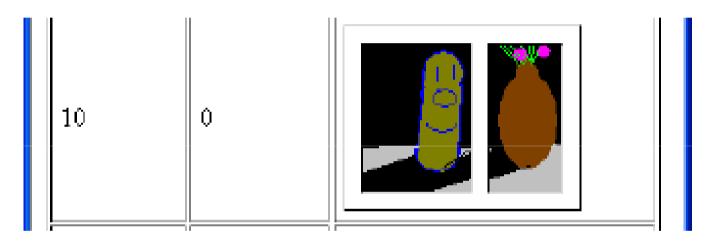
```
<caption>
 COSC 400 Student Grades
 </caption>
 \langle tr \rangle
   Grades
Table Header
 StudentExam 1Exam 2
UndergraduatesKim10089
Sandy7892
GraduatesTaylor8373
```



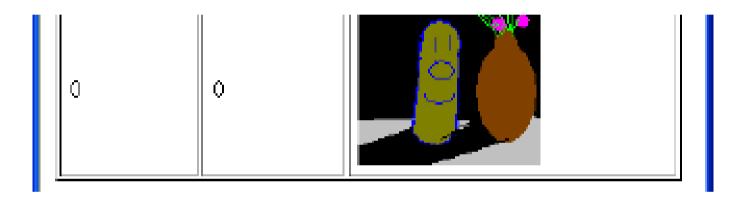
cellspacing cellpadding

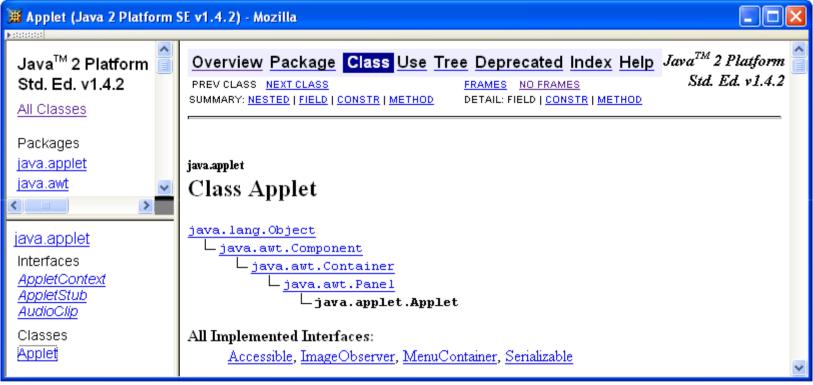


cellspacing cellpadding



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:

```
<a href="java/applet/package-frame.html" target="lowerLeftFrame">
```

 Value of target attribute specification is 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"

| Ⅲ LifeStory.html - Mozilla | |
|---|--|
| B (0.000)(0.00) | |
| 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 | |
| | |

Each form is content of a form element

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

HTTP request method (lower case)

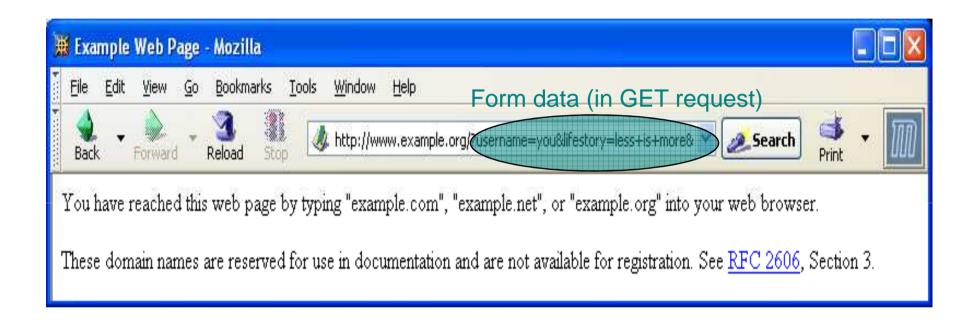
| Ⅲ LifeStory.html - Mozilla | |
|---|--|
| B (0.000)(0.00) | |
| Enter your name: | |
| Give your life's story in 100 words or less: | |
| | |
| | |
| | |
| | |
| | |
| Check all that apply to you: 🔲 tall 🔲 funny 🔲 smart | |
| Publish My Life's Story | |
| | |

| Ⅲ LifeStory.html - Mozilla | |
|---|--|
| B (0.000)(0.00) | |
| 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 | |
| | |

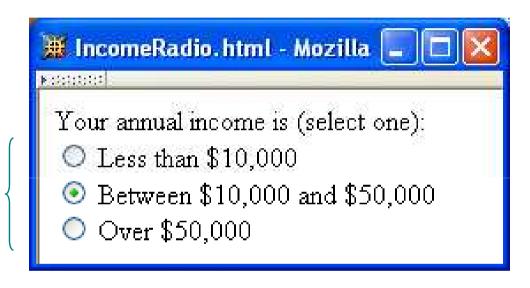
```
Check all that apply to you:
    <label>_
                            Checkbox control
    <input type="checkbox" name="boxgroup1" value="tall" />tall
    </label>
    <label>
      <input type="checkbox" name="boxgroup1" value="funny" />funny
    </label>
   <label>
      <input type="checkbox"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





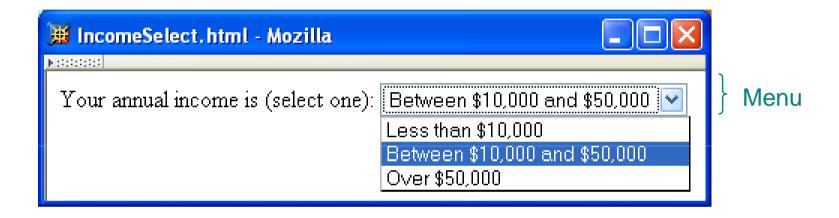
Radio buttons: at most one can be selected at a time.



```
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"</pre>
                                         value="10-50"
         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"</pre>
                                         value="10-50"
         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" />
    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

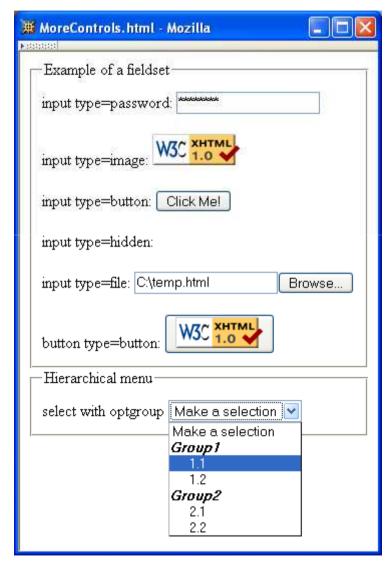


TABLE 2.5: HTML $4.01/XHTML\ 1.0$ non-deprecated form controls.

| Element | type Attribute | Control |
|----------|----------------|---|
| input | text | Text input |
| input | password | Password input |
| input | checkbox | Checkbox |
| input | radio | Radio button |
| input | submit | Submit button |
| input | image | Graphical submit button |
| input | reset | Reset button (form clear) |
| input | button | Push button (for use with scripts) |
| input | hidden | Non-displayed control (stores server- |
| | | supplied information) |
| input | file | File select |
| button | submit | Submit button with content (not an |
| | | empty element) |
| button | reset | Cancel button with content (not an empty |
| | | element) |
| button | button | Button with content but no predefined ac- |
| | | tion |
| select | N/A | Menu |
| option | N/A | Menu item |
| optgroup | N/A | Heading in a hierarchical menu |
| textarea | N/A | Multi-line text input |
| label | N/A | Associate label with control(s) |
| fieldset | N/A | Groups controls |
| legend | N/A | Add caption to a fieldset |