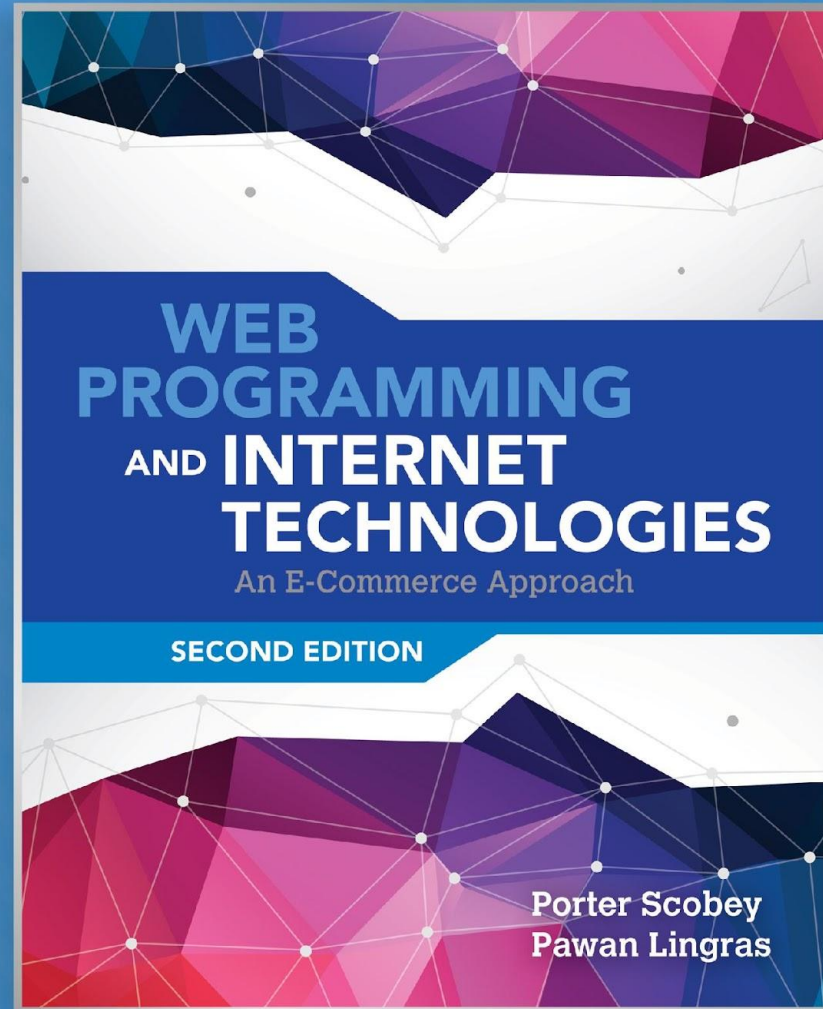


Chapter 3: HTML for Content Structure



Overview and Objectives

- Discuss briefly the history of HTML and how we got to HTML5
- Stress the importance of maintaining both a conceptual and a physical separation between the structure and presentation of a web page
- Discuss HTML tags and elements
- Discuss the basic structure of every web page
- Discuss the HTML5 DOCTYPE declaration and web page validation
- Discuss and illustrate all of the usual basic HTML markup, including headings, paragraphs, line breaks, lists, tables, images, comments, attributes and entities

What Is HTML?

- Hypertext Markup Language
- A markup language, *not* a programming language
- Uses markers called “tags” to designate the structural elements of a web page, such as:
 - headings
 - paragraphs
 - lists

Summary of HTML Versions

- HTML (1990) Tim Berners-Lee
- HTML 2 (1992)
- HTML 3.2 (1996)
- HTML 4 (1997) and HTML 4.01 (1999)
- XHTML 1.0 (2000) and XHTML 1.1 (2001)
- XHTML 2.0 (work discontinued by 2010)
- HTML5 (Fall, 2014)

HTML for Structure Only

Right up front we need to make these high-level distinctions:

- HTML's job is to describe the *structure* of a web page
- Web page *presentation* is the job of Cascading Style Sheets (CSS)
- Web page *behavior* is the job of JavaScript

A Typical HTML Tag and Element

- Here is an HTML paragraph *element*:
`<p>This is a paragraph.</p>`
- `<p>` is the *opening tag* of the element.
- `</p>` is the *closing tag* of the element.
- `This is a paragraph.` is the *content* of the element.
- Sometimes we refer to the “tag pair” `<p>...</p>`.
- Sometimes we refer to the “p tag” or simply a “p element”.

Browser Display of HTML Elements

- Every browser will have its own default way of displaying any HTML element.
- For a paragraph this might include space before and after the text, and the text font.
- Browsers also have certain “default behaviors”:
 - Reducing multiple spaces between words to a single space and removing leading and trailing spaces from element content
 - Wrapping lines as the browser window changes size

Basic Structure of Every Web Page

```
<!DOCTYPE html>
<html>
<head>
  <title>My Title</title>
</head>
<body>
  <!--An HTML comment showing where
  page content to display will go.-->
</body>
</html>
```


Some Basic Markup Tags

- `h1`, `h2`, `h3`, `h4`, `h5`, `h6` for headings
- `p` for paragraphs
- `ul` for unordered lists (• marker by default)
- `ol` for ordered lists (numbered 1, 2, 3 ... by default)
- `li` for list items in an ordered or unordered list

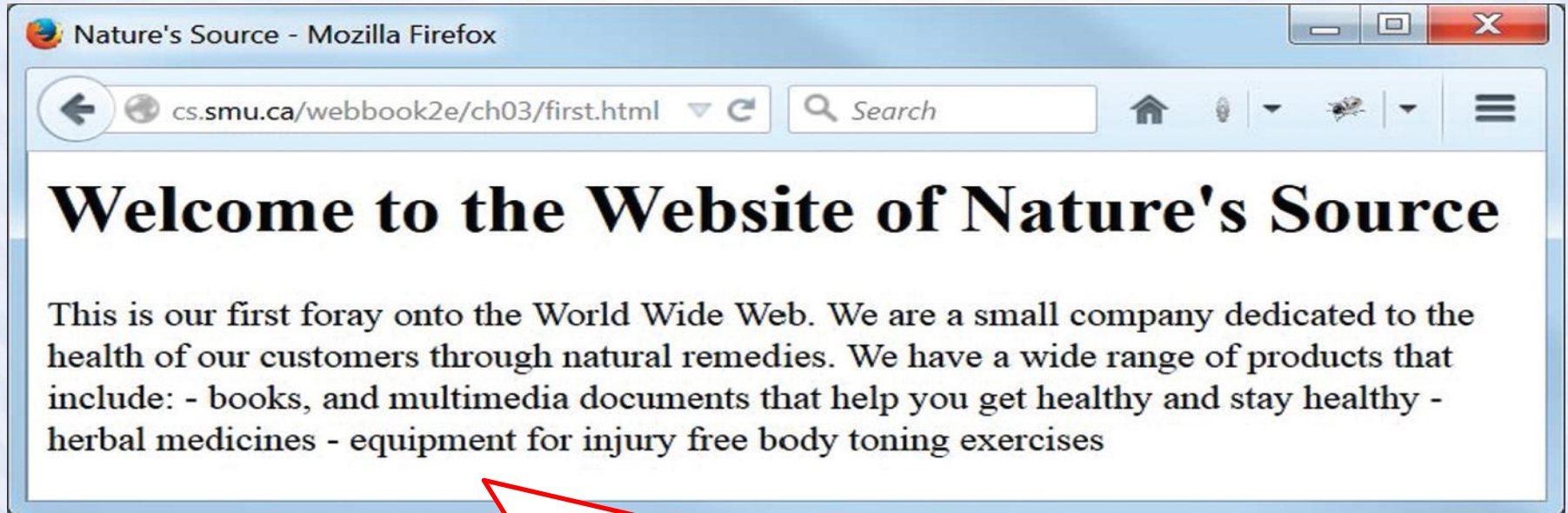
Basic HTML Markup Illustrated:

first.html

```
<html>
<head>
<title>Nature's Source</title>
</head>
<body>
<h1>
Welcome to the Website of Nature's Source
</h1>
<p>
This is our first foray onto the World Wide Web.
We are a small company dedicated to the health
of our customers through natural remedies.
We have a wide range of products that include:

- books, and multimedia documents that help you get
  healthy and stay healthy
- herbal medicines
- equipment for injury free body toning exercises
</p>
</body>
</html>
```

Basic HTML Markup Displayed: `first.html`



Note how the text from the file loses its formatting when displayed in the browser

Figure 3.2 graphics/ch03/displayFirstHtml.jpg.

Empty Elements: `
` for Line Breaks and `<hr>` for a Horizontal Dividing Line

- Not all elements have content.
- Elements with no content are called “empty elements” and have a special syntax.
- The line break element `
` is one of these.
- It causes the following text to appear on the next line but with no additional spacing.
- It has just a single tag (that is, an opening tag but no closing tag).

More HTML Markup Illustrated: second.html

```
<!-- second.html -->
<html>
  <head>
    <title>Nature's Source</title>
  </head>
  <body>
    <h1>Welcome to the Website of Nature's Source!</h1>
    <p>This is our first foray onto the World Wide Web. We are a small
    company dedicated to the health of our customers through natural
    remedies.
    <br>
    We have a wide range of products that include:</p>
    <ul>
      <li>books, and multimedia documents that help you get healthy and
      stay healthy</li>
      <li>herbal medicines</li>
      <li>equipment for injury free body toning exercises</li>
    </ul>
  </body>
</html>
```

More HTML Markup Displayed: second.html



Figure 3.4 graphics/ch03/displaySecondHtml.jpg.

What Is a “Valid” Web Page?

- A *valid* web page is one that conforms to one of several possible standards.
- The standard against which the page will be validated is determined by a DOCTYPE *declaration* in the web page.
- We use this (HTML5) DOCTYPE declaration as the first line of a file:
`<!DOCTYPE html>`

Some HTML “Boilerplate”

- The term *boilerplate* refers to markup that we can use “as is” on every page, and we need some when we attempt to validate a web page.
- The first boilerplate item is the DOCTYPE itself.
- The second is the attribute of the `html` tag:
`lang="en"`
- The third is the (empty) meta element within head element:
`<meta charset="utf-8">`

HTML5 Markup Guidelines

- Use lowercase for all tag names.
- Ensure all tag pairs are properly nested.
- Give any non-empty element both an opening tag and a closing tag.
- Use lowercase for attribute names and enclose attribute values in quotes.
- Remember: block elements may contain block or inline elements; inline elements contain only other inline elements.

HTML5 Markup Best Practice

- Make sure your page has the HTML5 DOCTYPE and all of these HTML elements in their correct order, properly nested:

```
html
  head
    meta
    title
  body
```

- Don't use any elements that are *deprecated* by the DOCTYPE against which you are validating.

A Simple Valid Web Page:

third.html

```
<!DOCTYPE html>
<!-- third.html -->
<html lang="en">
  <head>
    <meta charset="utf-8">
    <title>Nature's Source</title>
  </head>
  <body>
    <h1>Welcome to the Website of Nature's Source!</h1>
    <p>This is our first foray onto the World Wide Web. We are a small
    company dedicated to the health of our customers through natural
    remedies.
    <br>
    We have a wide range of products that include:</p>
    <ul>
      <li>books, and multimedia documents that help you get healthy and
      stay healthy</li>
      <li>herbal medicines</li>
      <li>equipment for injury free body toning exercises</li>
    </ul>
  </body>
</html>
```

How Do We Validate a Web Page?

- You can find a number of HTML validators on the web. The W3C itself provides this one:
<http://validator.w3.org/>
- You can enter the URL of the page you want validated directly into the validator and click on a button to validate it.
- Many other ways are possible: For example, the Firefox Web Developer add-on lets you validate the web page you are viewing from a dropdown menu.

Validating `third.html`

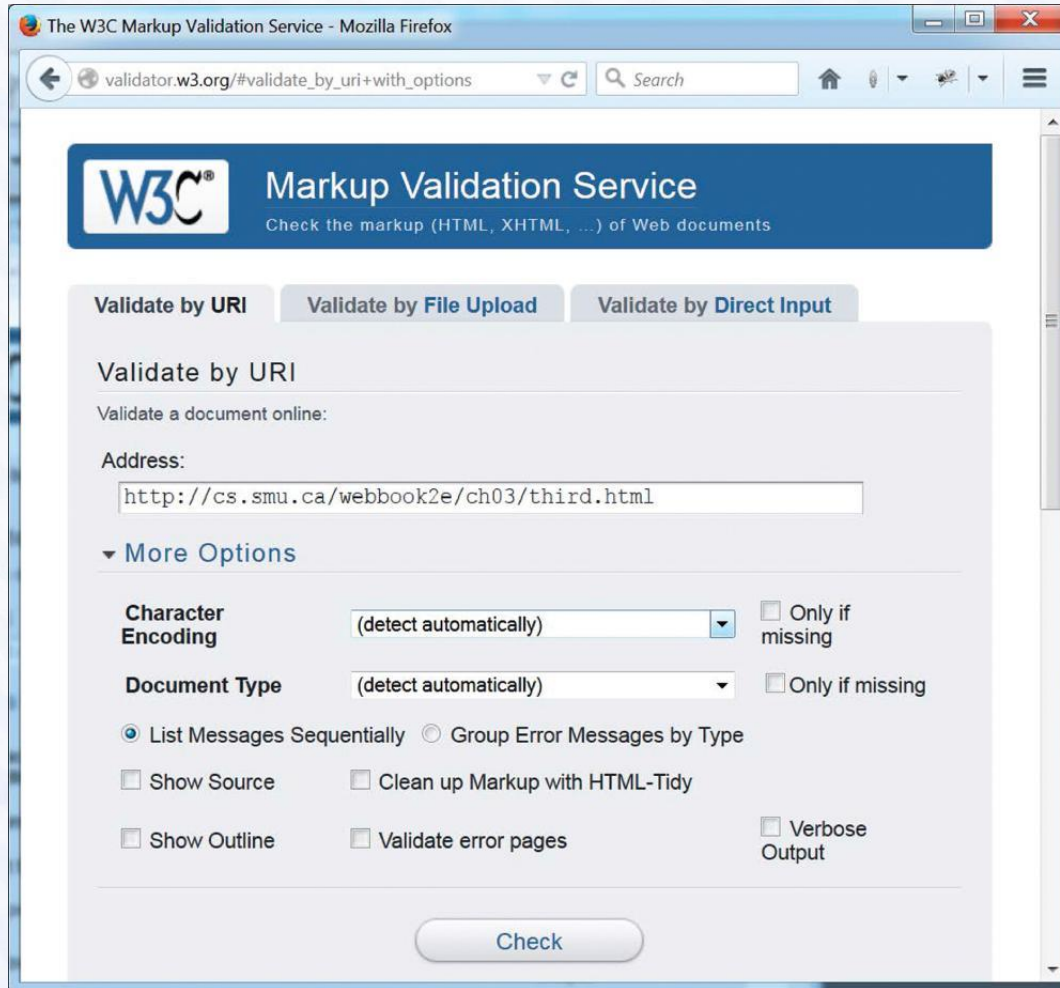


Figure 3.6 graphics/ch03/displayThirdHtmlToValidate.jpg.

Validation Report for `third.html`

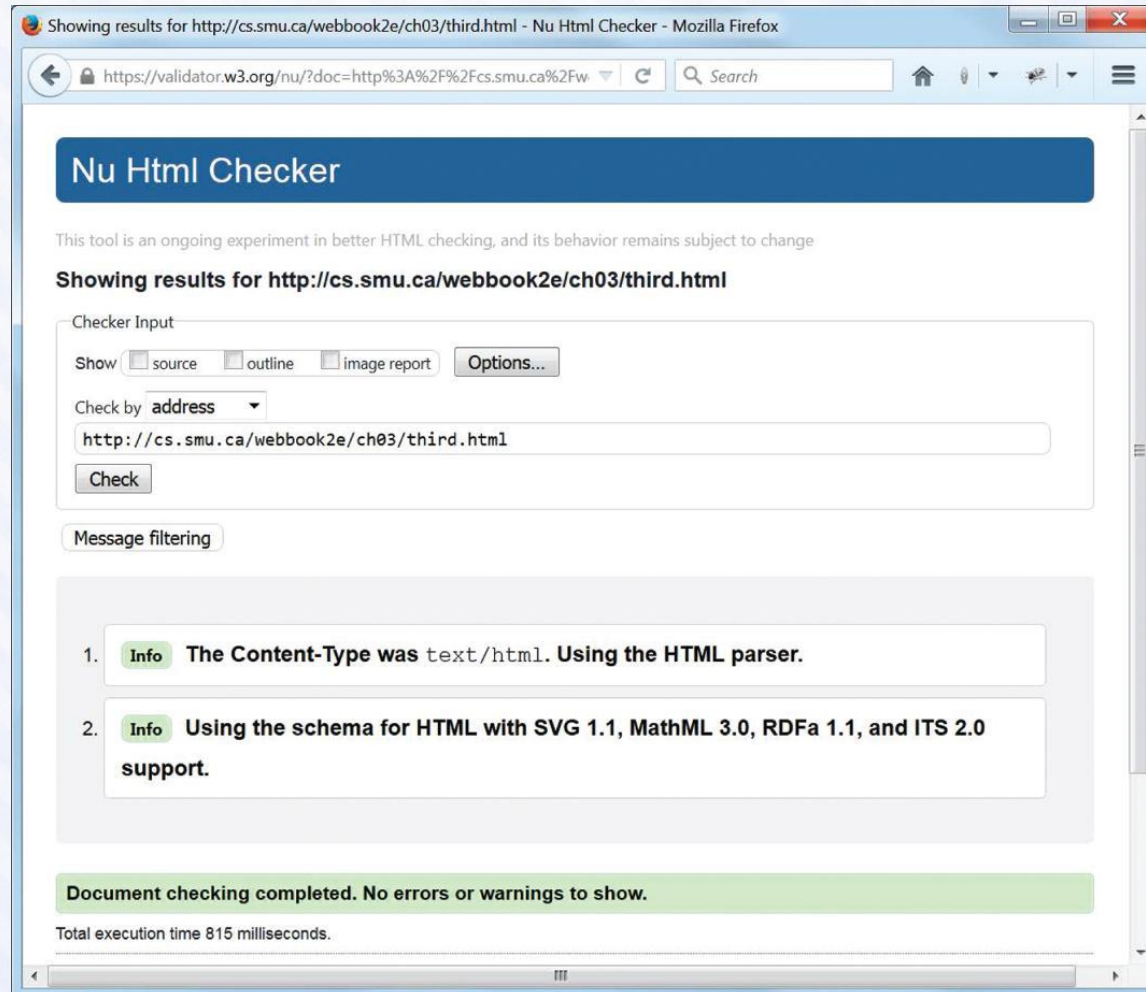


Figure 3.7 `graphics/ch03/displayThirdHtmlValidated.jpg`.

The Web Developer Add-on

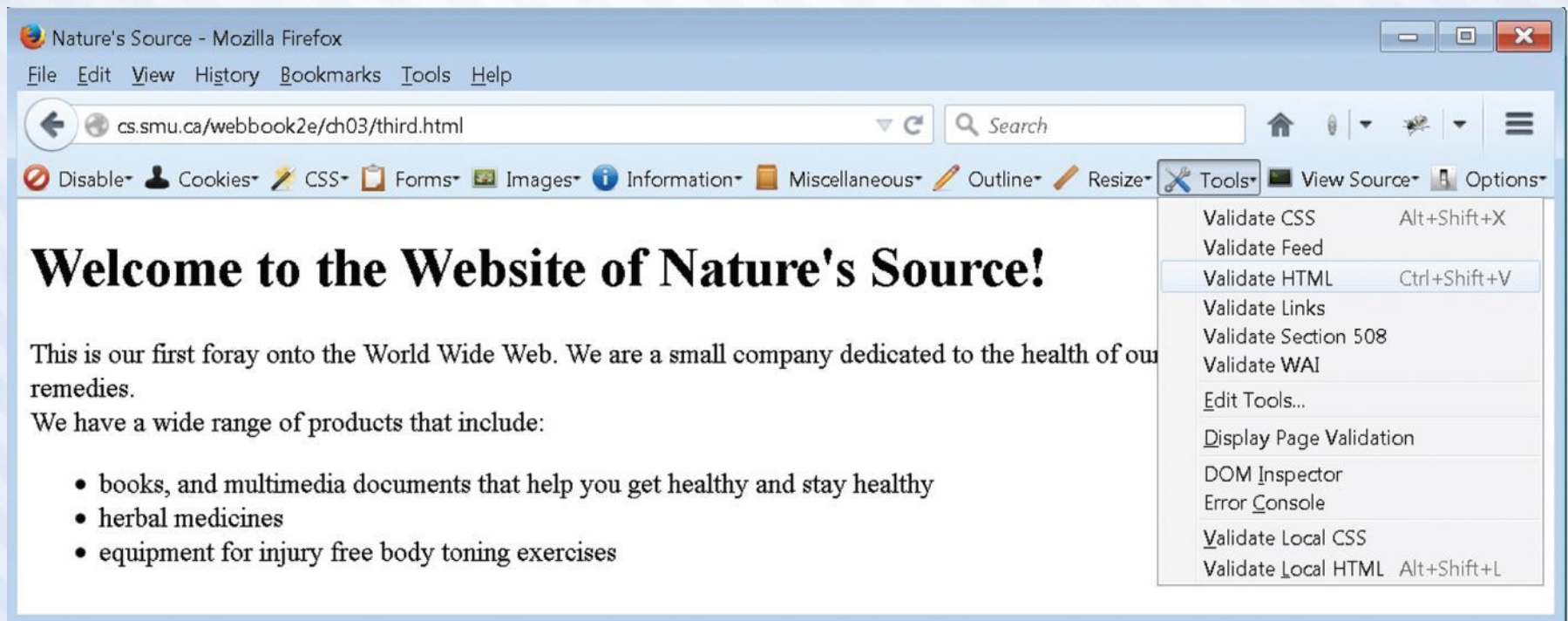


Figure 3.8 graphics/ch03/WebDeveloper.jpg.

The Empty `img` Element for Images, and Two Required Tag Attributes

- The (empty) `img` element lets us place an image on a web page.
- A typical image element:

```

```
- `src` and `alt` are called *attributes*.
- The value of `src` is the location of the image file.
- The value of `alt` is the text to display if no image.
- Attribute values are enclosed in quotes.

A Best Practice for Images

- A *pixel* is a very small area of illumination on a display screen—one of many from which an image is composed.
- The `img` tag has two optional attributes, `width` and `height`, whose values in number of pixels give the size of an image.
- A browser will scale the image to that size, but it is better to make the image the size you want in an external program and then specify the exact size in your `img` tag. Then ... no scaling required.

Block Elements vs. Inline Elements

- Some HTML elements are *block elements*, others are *inline elements*.
- Block elements (headings, paragraphs and lists, for example) flow from top to bottom of a web page, with a browser-dependent amount of vertical whitespace before and after the element.
- Inline elements (the `img` element, for example) flow from left to right, wrapping to the next line as necessary, with no horizontal spacing surrounding the element.

Tables in General

- Tables have rows and columns.
- Tables are best used to display data that lends itself to rows and columns.
- Example: A table of 28 temperatures with 4 rows (weeks) and 7 columns (week days)
- A table can be used to lay out a very simple web page, but Cascading Style Sheets are much preferred for web page layout.

HTML Basic Table Elements

A table element with two rows and two columns (two table data cells) in each row:

```
<table summary="A simple table">  
  <tr>  
    <td>top-left</td><td>top-right</td>  
  </tr>  
  <tr>  
    <td>top-left</td><td>top-right</td>  
  </tr>  
</table>
```

Additional Table Features

- Content of a `td` (table data) element is displayed left-justified by default.
- Content of a `th` (table header) element is displayed bold and centered by default.
- `td` and `th` tags can take/have `rowspan="n"` and `colspan="n"` attributes to permit one cell to span multiple rows and/or columns.
- Place a `caption` element after the opening `table` tag to give a table a caption.

Simple Table Layout in index.html from nature1 (two rows and two columns)



Figure 3.10 graphics/ch03/nature1/displayIndexHtml.jpg.

ch03/nature1/index.html

```
<!DOCTYPE html>
<!-- index.html for ch03/nature1 -->
<html lang="en">
<head>
  <meta charset="utf-8">
  <title>Nature's Source - Canada's largest specialty vitamin store</title>
</head>

<body>
  <table>
    <tr>
      <td></td>

      <td>5029 Hurontario Street Unit 2<br>
        Mississauga, ON L4Z 3X7<br>
        Tel: 905.502.6789<br>
        Fax: 905.890.8305</td>
    </tr>
    <tr>
      <td>
        <h3>Welcome to Nature's Source - Protecting your health
          naturally!</h3>

        <p>Founded in 1998, Nature's Source was created .../p>
      </td>

      <td></td>
    </tr>
  </table>
</body>
</html>
```

What Is an HTML Entity?

- Some characters have a special meaning in HTML (< starts an opening tag, for example).
- So, for example, in `small<big`, “<big” would be interpreted as the start of the opening `<big>` tag in a `big` element.
- This problem is solved by this HTML *entity*:
`<`
- A browser replaces the entity `<` with the single character `<`.

HTML Entities

- Characters that need replacing by entities are often called *metacharacters*.
- Entities can also provide special characters not appearing on keyboards.
- Some additional HTML entities:
 - `>` for `>`
 - `&` for `&`
 - `©` for `©` (the copyright symbol)

Some Other Useful HTML Character Entities

Result	Description	Entity Name	Entity Number
	non-breaking space	 	
<	less than	<	<
>	greater than	>	>
&	ampersand	&	&
"	double quotation mark	"	"
'	single quotation mark (apostrophe)	'	'
¢	cent	¢	¢
£	pound	£	£
¥	yen	¥	¥
€	euro	€	€
©	copyright	©	©
®	registered trademark	®	®

HTML Hyperlinks

- A central idea of the web is that one page can link to one or many others.
- The `a` (anchor) element provides the hyperlink (link).
- The value of the (required) `href` attribute of an `a` tag is the URL or path to the other page.
- A typical `a` element:

```
<a href="estore.html">e-store</a>
```
- By default content like `e-store` appears with blue and underlined text.
- The content of an `a` element can be an `img` element.

Simple Table Layout in index.html from nature2 (four rows and five columns) (and now including “menu item” links)

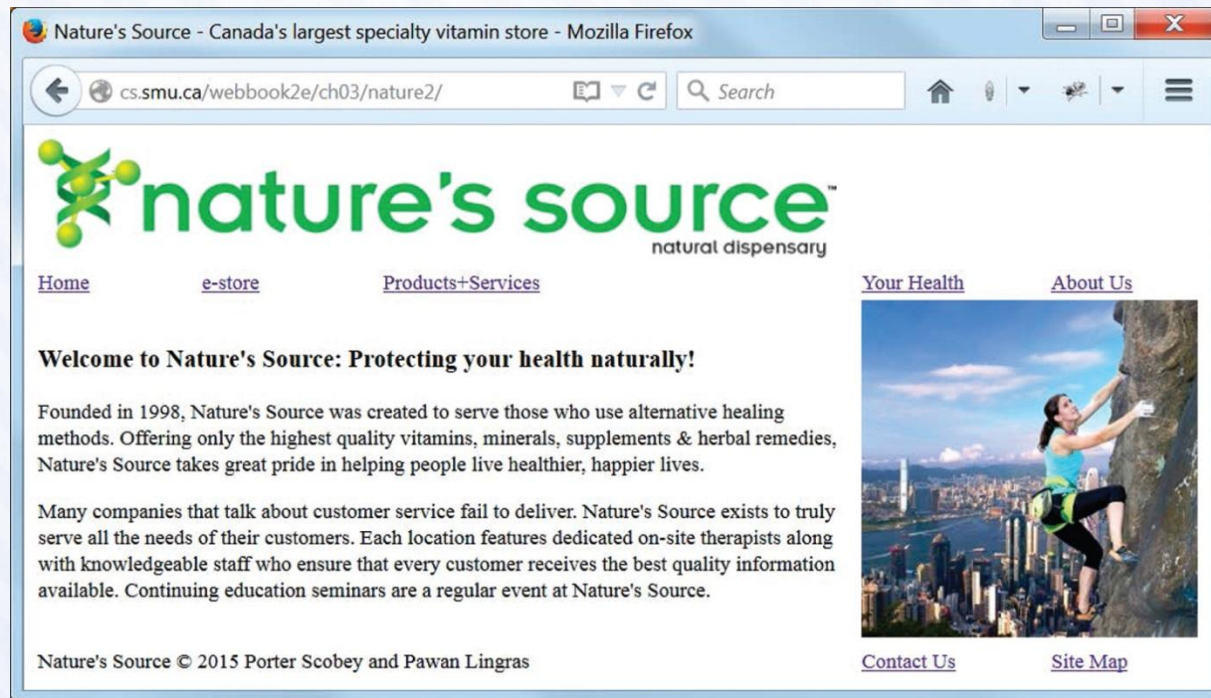


Figure 3.12 graphics/ch03/nature2/displayIndexHtml.jpg.

ch03/nature2/index.html

(1 of 2)

```
<!DOCTYPE html>
<!-- index.html for ch03/nature2-->
<html lang="en">
  <head>
    <meta charset="utf-8">
    <title>Nature's Source - Canada's largest specialty vitamin store</title>
  </head>
  <body>
    <table>
      <tr>
        <td colspan="5">
          
        </td>
      </tr>
      <tr>
        <td><a href="index.html">Home</a></td>
        <td><a href="pages/estore.html">e-store</a></td>
        <td><a href="pages/products.html">Products+Services</a></td>
        <td><a href="pages/yourhealth.html">Your Health</a></td>
        <td><a href="pages/about.html">About Us</a></td>
      </tr>
    </table>
  </body>
</html>
```

ch03/nature2/index.html

(2 of 2)

```
<tr>
  <td colspan="3">
    <h3>Welcome to Nature's Source:
    Protecting your health naturally!</h3>
    <p>Founded in 1998, Nature's Source was created ...</p>
  </td>
  <td colspan="2">
    
  </td>
</tr>
<tr>
  <td colspan="3">Nature's Source &copy; 2015 Porter Scobey and Pawan
  Lingras</td>
  <td><a href="pages/contact.html">Contact Us</a></td>
  <td><a href="pages/sitemap.html">Site Map</a></td>
</tr>
</table>
</body>
</html>
```