CGS 3175: Internet Applications
Fall 2007

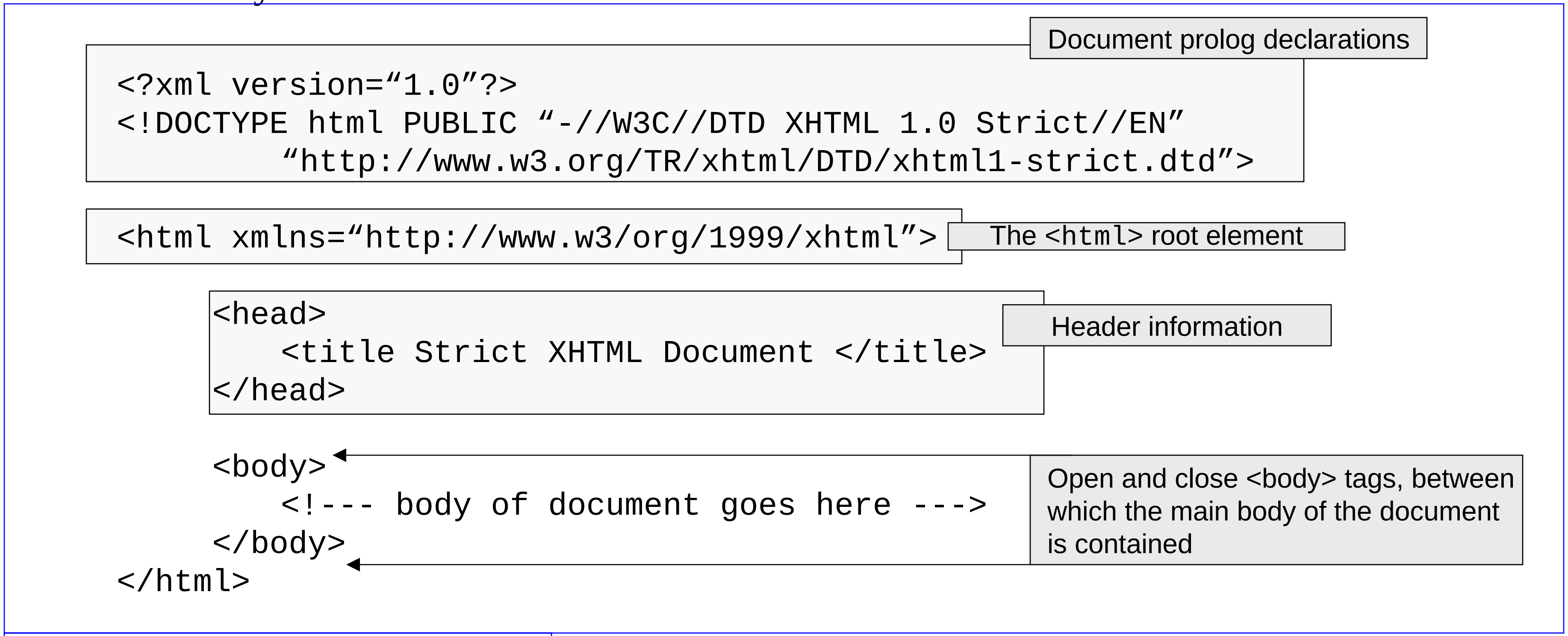
XHTML – Part 2 – Basic Formatting

Instructor : Dr. Mark Llewellyn
markl@cs.ucf.edu
HEC 236, 407-823-2790
http://www.cs.ucf.edu/courses/cgs3175/fall2007

School of Electrical Engineering and Computer Science
University of Central Florida

How XHTML Documents Are Structured

- XHTML documents are comprised of a simple three-part framework:
  1. Document prolog
  2. Header section
  3. Body of document.



```
<?xml version=“1.0”?>
<!DOCTYPE html PUBLIC “-//W3C//DTD XHTML 1.0 Strict//EN”
        “http://www.w3.org/TR/xhtml/DTD/xhtml1-strict.dtd”>

<html xmlns=“http://www.w3/org/1999/xhtml”>

      <head>
         <title Strict XHTML Document </title>
      </head>

      <body>
         <!--- body of document goes here --->
      </body>
</html>
```

Document prolog declarations

The <html> root element

Header information

Open and close <body> tags, between which the main body of the document is contained

XHTML Strict Document Framework

Document Framework Elements

- The elements that make up the framework of XHTML documents do not produce any output in a browser window. Instead, they provide information to the program about the document.

- The elements that makeup the framework are: `<html>`, `<head>`, `<title>`, and `<body>`. We'll look at each of these more closely.

The <html> Element

- The <html> element is the root element of an XHTML document, within which every other element in the document is contained (recall our discussion on nesting of elements).
- The document begins with the <html> start tag and ends with the </html> end tag. The header and body information of the document are contained in the root element.
- As you may have noticed in the example on page 37 in the previous section of notes, the <html> start tag defines an attribute called xmlns with a value of http://www.w3.org/1999/xhtml. This attribute declares the XHTML namespace and is requires in all XHTML documents. If you omit it, it will be inserted automatically by the parser, however, it is recommended that you insert it into the document.

An Aside on Namespaces

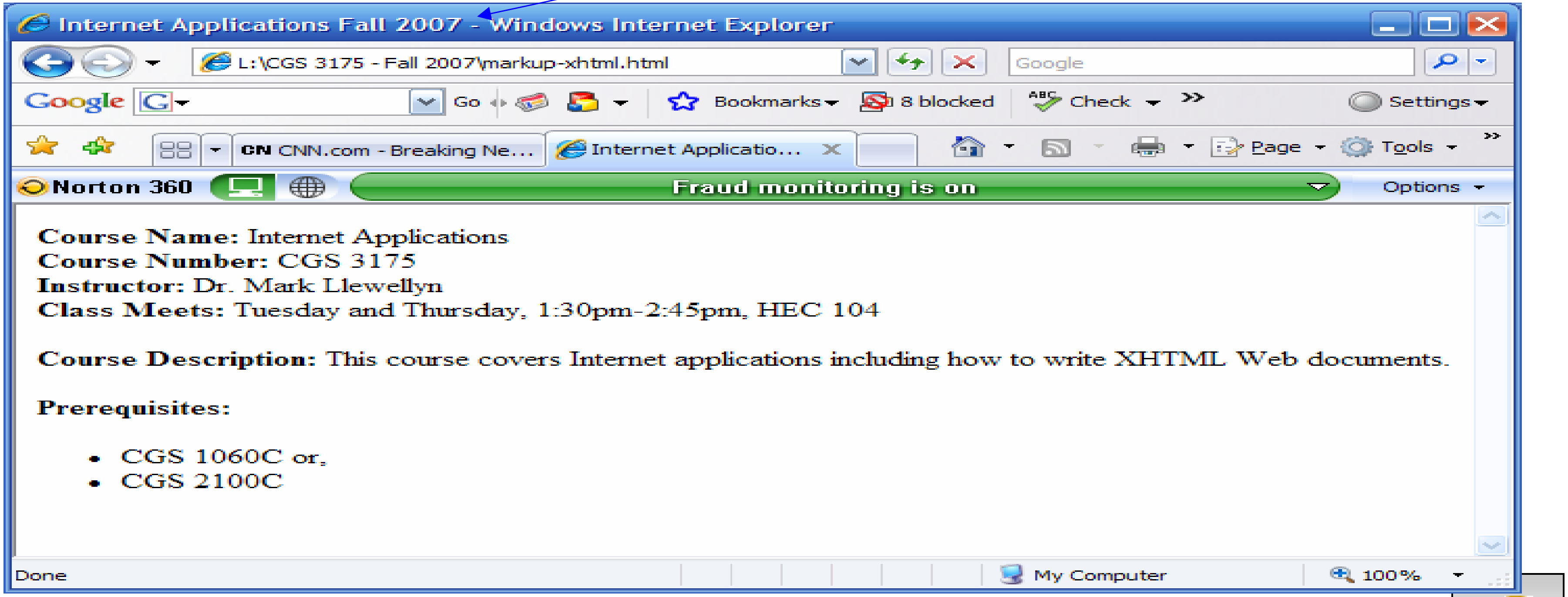
- The `<html>` start tag defines an attribute called `xmlns` with a value of `http://www.w3.org/1999/xhtml`.
- The `xmlns` attributes stands for XML Namespace, and the value of this attribute defines the namespace that the element names for XHTML belong to.
- Namespaces in XML are collections of element and attribute names, XML allows you to create your own language and your own element and attribute names. This can lead to confusion if your language elements and/or attribute names are the same as someone else's names but have different meanings. Namespaces in XML allows you to specify which namespace the element and attribute names belong to.
- The `xmlns` attribute tells the program parsing the document that the elements and attribute names contained within the `<html>` element belong to the XHTML namespace.
- The value of the `xmlns` attribute, `http://www.w3.org/1999/xhtml`, resembles a Web address, but is simply used to uniquely define the namespace. If you go to this page on the Internet, you will find an informational page provided by the W3C. This is not a requirement of namespace values, it is simply a courtesy provided by the W3C.
- The xmlns attribute for the `<html>` element is required for XHTML documents and is specified as a fixed value in each of the three XHTML DTDs. While you can safely omit the attribute from your `<html>` start tag, since the parser will automatically add it, it is recommended to include it.

The <head> Element

- The start <head> tag comes directly after the <html> start tag in an XHTML document. This element must be placed inside the <html> element.
- It contains information about the document that is mainly used by programs, such as keywords for search engines and link information that defines the relationship this document has to other documents.
- The <head> tag also contains the required <title> element. The following is a list of elements that can be contained in the <head> element – all but the <title> element are optional.
  - <title> - defines the title of the document.
  - <base> - defines the document base URL, which is used for relative links in the document.
  - <link> - defines the relationship of this document to other documents.
  - <meta> - defines additional information about the document, including the document's content type and special instructions for browsers and search engines.
  - <script> - defines links to scripts used within the document, such as Javascript.
  - <style> - defines links to style sheets to be used within the document, such as CSS.

The <title> Element

- The <title> element is the only required element within the <head> element. It must be contained within the open and close tags of the <head> element.
- There can only be one <title> element per document. It defines the title of the document that is displayed in the title bar of the browser window as well as the name of bookmarks to that page.
- Most search engines use the content of the <title> element as the text to display on their results pages as well.



Course Name: Internet Applications
Course Number: CGS 3175
Instructor: Dr. Mark Llewellyn
Class Meets: Tuesday and Thursday, 1:30pm-2:45pm, HEC 104

Course Description: This course covers Internet applications including how to write XHTML Web documents.

Prerequisites:

- CGS 1060C or,
- CGS 2100C

The <body> Element

- The <body> element contains the content and all of the markup elements of the document.
- The body of the document is contained between the open <body> tag and the ending </body> tag.
- All of the other elements we will cover are contained within the <body> element.

A Warning About Formatting Elements In XHTML

- Previous versions of HTML included various formatting attributes that would allow you to set the background color for the document and define text attributes in the body element, such as bgcolor.
- XHTML Strict uses style sheets to define these formatting attributes, (recall, that this is designed to separate the content from the presentation), so they are not included in the specification.
- If you are using XHTML Transitional or Frameset, you can include additional formatting attributes within the `<body>` element. However, keep in mind the warnings about future compatibility. Style attributes using only in XHTML Transitional and Frameset are:
  - bgcolor – sets background color of the document.
  - text – sets color of the text in the document.
  - link – sets the color of hyperlinks, default color is blue.
  - vlink – used to set the color of hyperlinks that have been viewed by the user.
  - alink – used to set the color of hyperlinks that are currently active.

Basic Formatting Elements In XHTML

- Now that you understand how XHTML documents are structured, let’s start building some Web pages. We’ll start with basic formatting elements, show some examples of how to use each of the elements, and create a few documents to illustrate how they look in a Web browser.

- We’ll start by looking at block-level formatting elements.

Block-level Formatting Elements – Summary Chart

Element Name	Formatting Style
`<p>...</p>`	Paragraph element
`<br />`	Line break (empty element)
`<h1>...</h1>` to `<h6>...</h6>`	Heading elements (1 is largest, 6 is smallest)
`<hr />`	Horizontal rule (empty element)
`<div>...</div>`	Section divider

Block-level Formatting Elements

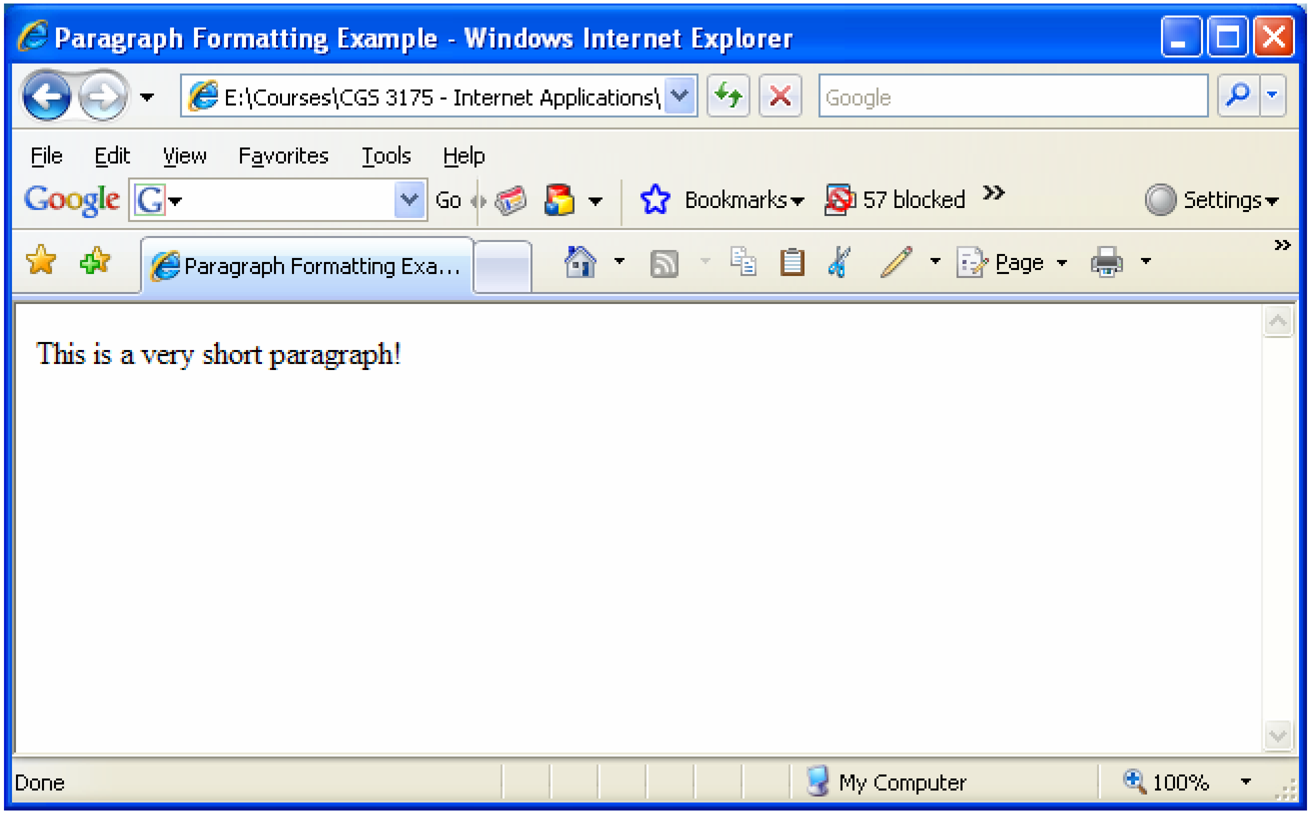
- Documents are broken into logical sections based on the document content to make it easer for users to read.
- The elements shown in the table on the previous page are used to break documents into logical chunks and to label the main content headings.
- These elements are referred to as block-level elements because they describe blocks of content.
- We’ll examine each of these elements separately.

The <p> Element

- The <p> element divides content into paragraphs.
- The <p> tag designates the beginning of a paragraph, and the </p> tag ends the paragraph.
- Most browsers will automatically insert a double line return (carriage return) around the paragraph element.
- Example:

```
<p>This is a very short paragraph.</p>
```

The <p> Element



Paragraph Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File Edit View Favorites Tools Help
Google
Go
Bookmarks
57 blocked
Settings
Paragraph Formatting Exa...
Page
Done
My Computer
100%

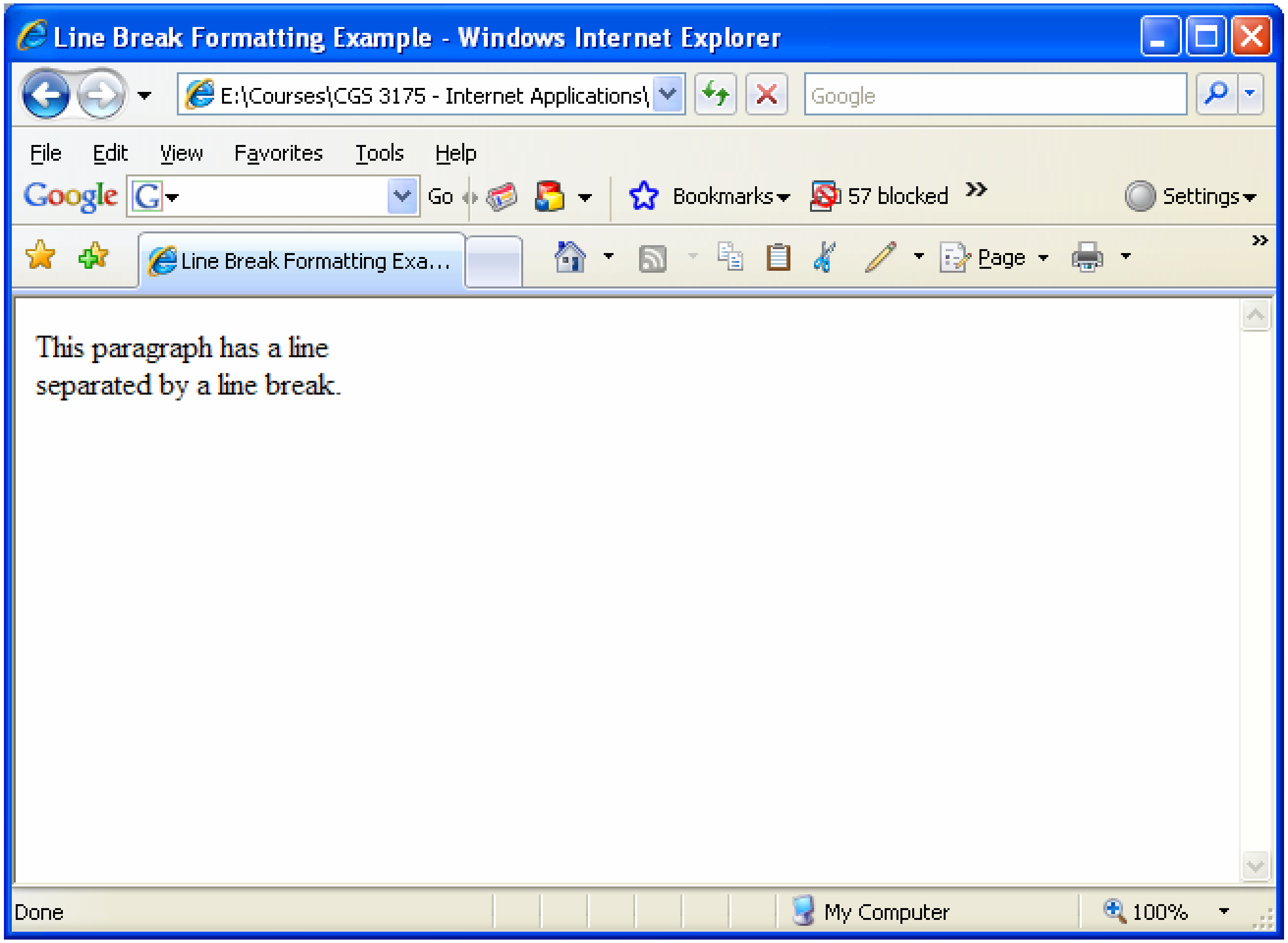
This is a very short paragraph!

The `<br />` Element

- The `<br />` element is the line break element. Similar to the `<p>` element, it is used to break up sections of text.
- The `<br />` element causes the browser to create a single line return (single carriage return).
- The `<br />` element is an empty element and must end with `/>` in order to conform to the rules of a well-formed document.
- Example:

```
<p> This paragraph has a line <br />
    separated by a line break.</p>
```

The <br /> Element


Line Break Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File Edit View Favorites Tools Help
Go
Bookmarks
57 blocked
Settings
Line Break Formatting Exa...
Page
Done
My Computer
100%

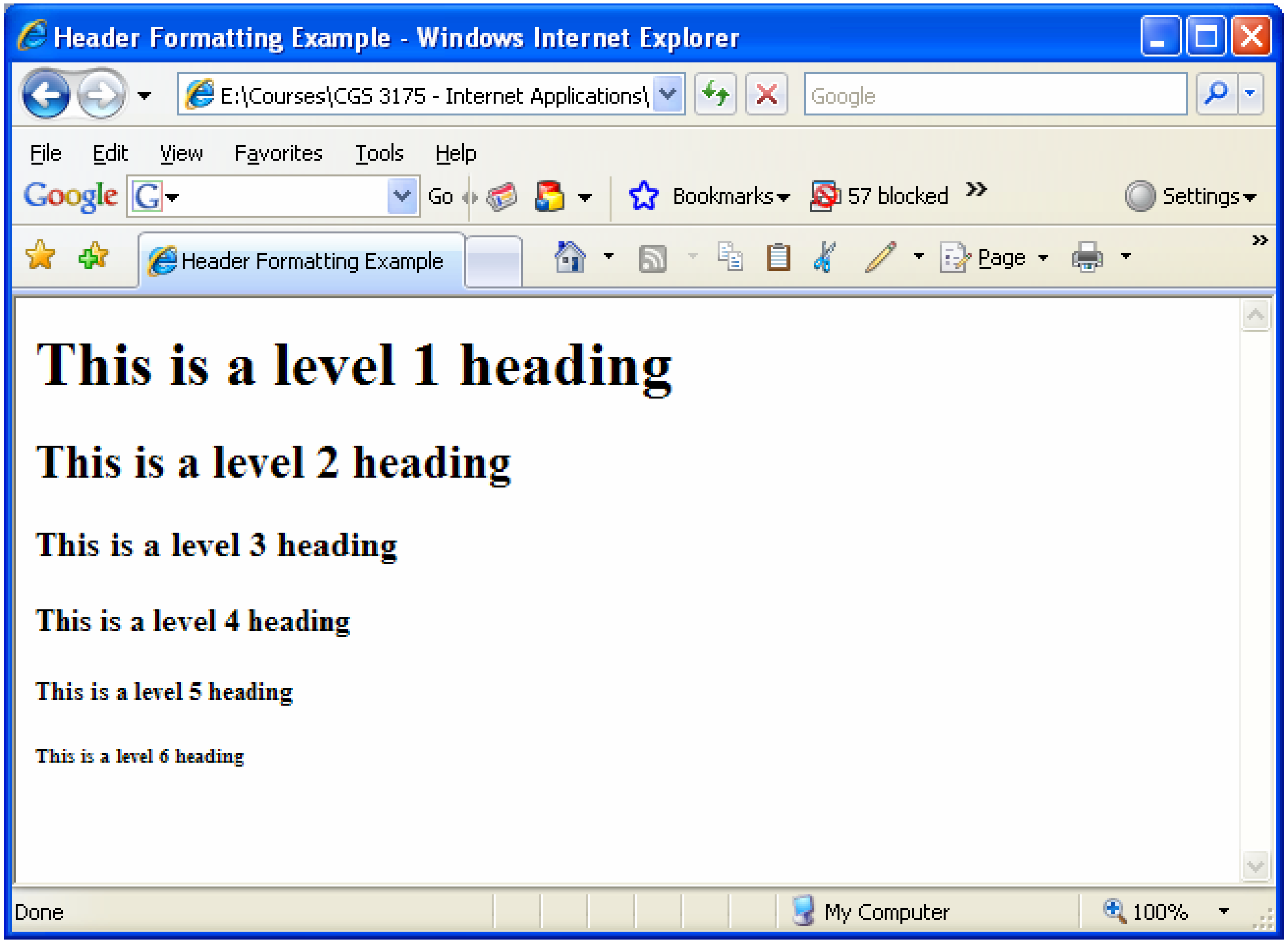
This paragraph has a line
separated by a line break.

The <h1> ... <h6> Elements

- These elements are the heading elements. They are used to label section headings of a document.
- There are six heading levels: <h1>, <h2>, <h3>, <h4>, <h5>, and <h6>.
- The <h1> head element should be used to label the top-most heading, and the rest of the elements should be used for subheads, much like a table of contents hierarchy.
- The browser will display the font for each of these levels differently, starting with a larger font for <h1> and progressively getting smaller as the heading number increases.
- Example:

```
<h1> This is a level 1 heading</h1>
<h2> This is a level 2 heading</h2>
<h3> This is a level 3 heading</h3>
<h4> This is a level 4 heading</h4>
<h5> This is a level 5 heading</h5>
<h6> This is a level 6 heading</h6>
```

The <h1> . . . <h6> Elements



Header Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File Edit View Favorites Tools Help
Go
Bookmarks
57 blocked
Settings
Header Formatting Example
Page
Done
My Computer
100%

This is a level 1 heading

This is a level 2 heading

This is a level 3 heading

This is a level 4 heading

This is a level 5 heading

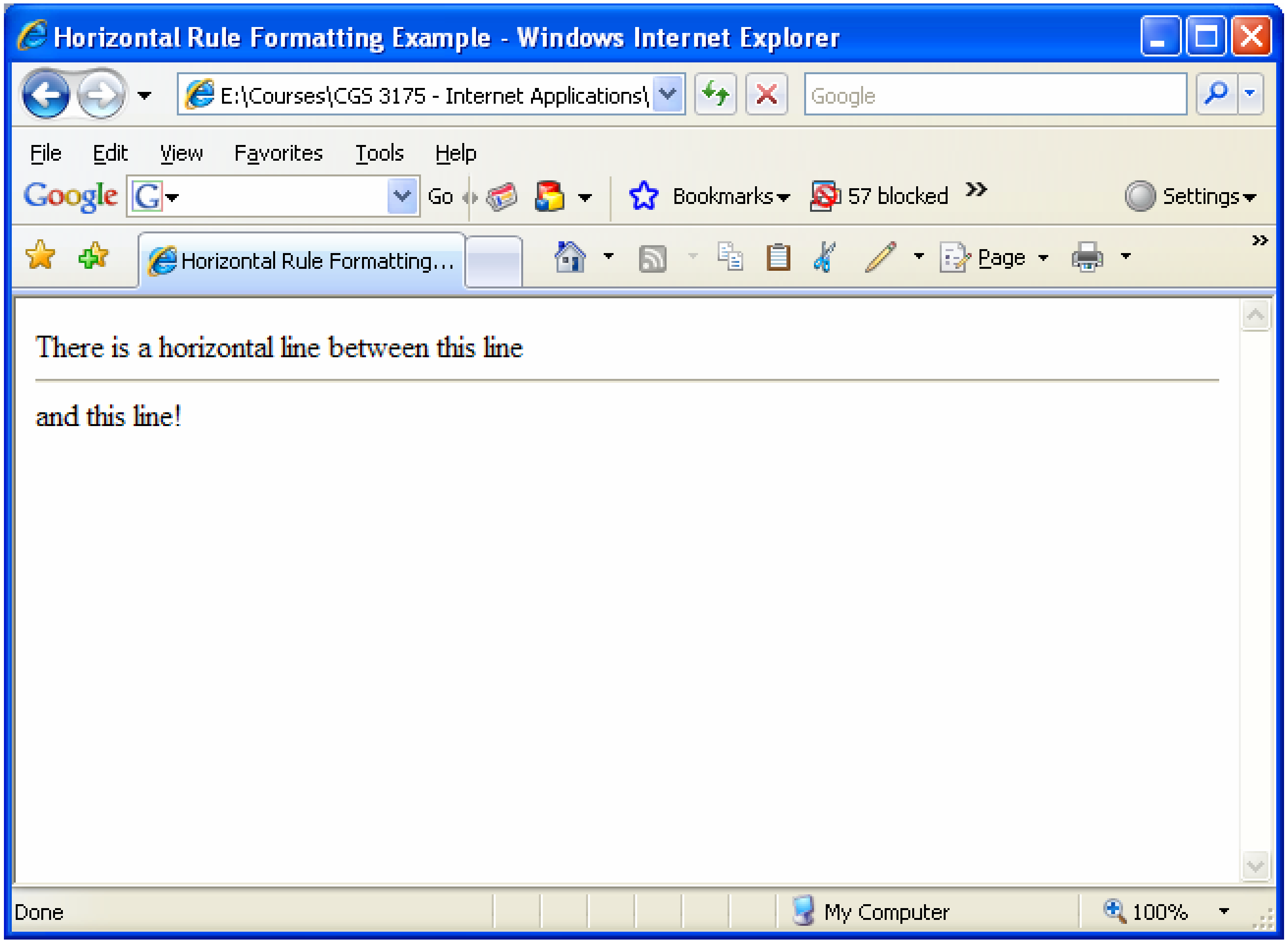
This is a level 6 heading

The <hr /> Element

- The `<hr />` element is the horizontal rule element, used to create a visible horizontal line on the Web page to indicate a section break or change in content.
- XHTML Transitional and Frameset provide a set of attributes that can be used with this element to customize the rule.
- Example:

```
There is a horizontal line between this line <hr /> and this line.
```

The <hr /> Element



Horizontal Rule Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File Edit View Favorites Tools Help
Go
Bookmarks
57 blocked
Settings
Horizontal Rule Formatting...
Page
Done
My Computer
100%

There is a horizontal line between this line

---

and this line!

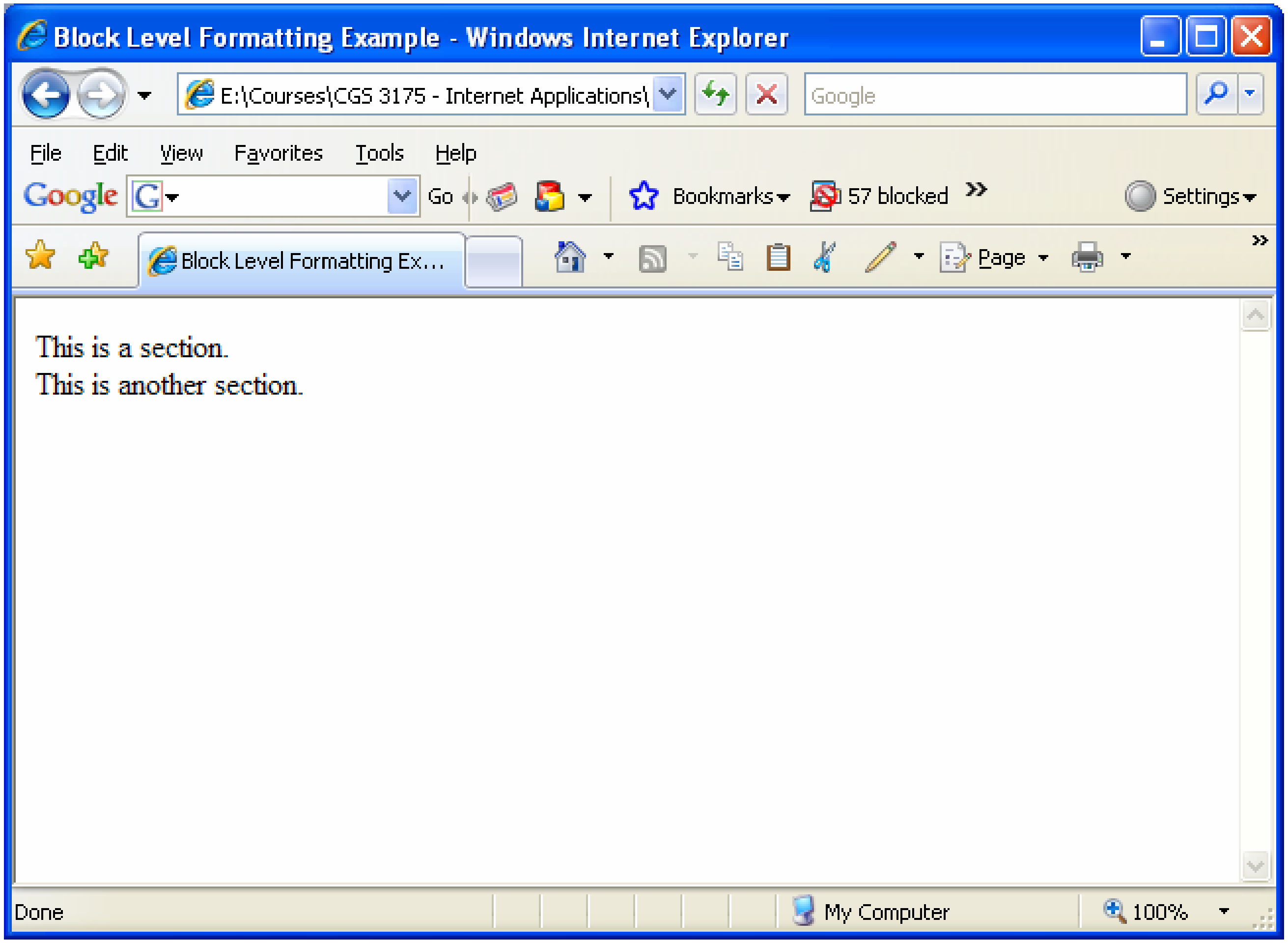
The <div> Element

- The `<div>` element is used to divide sections of content. This element is used to label sections of the document, and can contain any number of other elements.
- This element can use the id and class core XHTML attributes to identify the various sections of the document to be used with parser programs.
- Example:

```
<div> This is a section.</div>

 <div> This is another section.</div>
```

The <div> Element



Block Level Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File Edit View Favorites Tools Help
Go
Bookmarks
57 blocked
Settings
Block Level Formatting Ex...
Page
Done
My Computer
100%

This is a section.
This is another section.

Text Formatting Elements

- Text formatting elements are referred to as character-level elements because, unlike the block-level elements, which describe blocks of content, these elements describe the text itself.

- Character-level elements describe the formatting of words or phrases as opposed to sections or paragraphs.

- There are two basic groups of text formatting elements: presentation styles, and logical styles.
  - Presentational styles describe how the text should be displayed, in bold type or italics, for example.
  - Logical styles describe the meaning of the style more than the actual format.

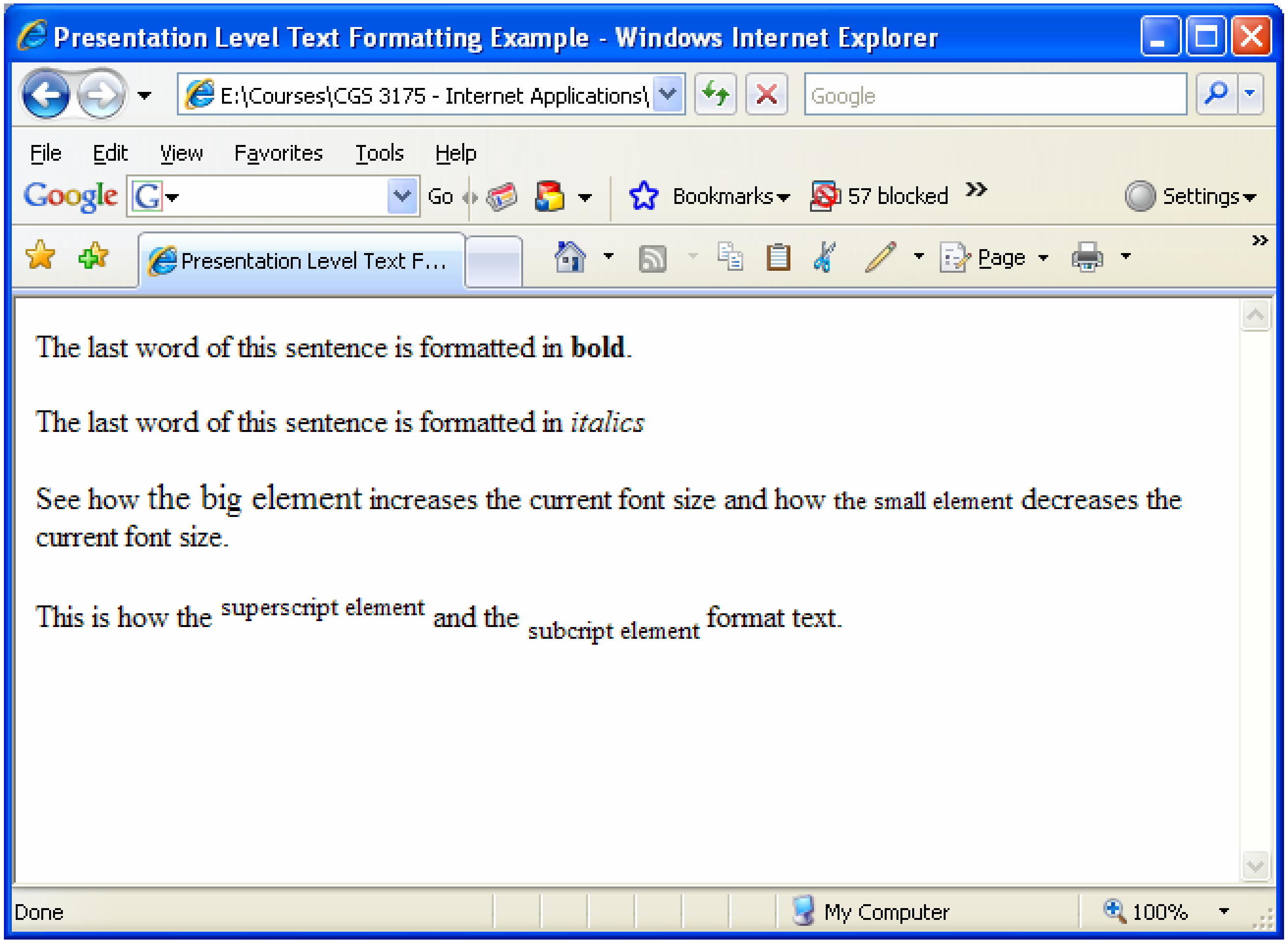
Presentational Text Formatting – Summary Chart

Element Name	Formatting Style
`<b>...</b>`	Bold font style
`<big>...</big>`	Increases current font size
`<i>...</i>`	Italic font style
`<small>...</small>`	Decrease current font size
`<sub>...</sub>`	Subscripted text
`<sup>...</sup>`	Superscripted text

Presentational Text Formatting – Example

```
<?xml version="1.0" encoding="UTF-8"?>
<!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>
     <title>Presentation Level Text Formatting Example</title>
   </head>
   <body>
      <p>The last word of this sentence is formatted in <b>bold</b>.</p>
      <p>The last word of this sentence is formatted in <i>italics</i></p>
      <p>See how <big> the big element</big> increases the current font
         size and how <small> the small element</small> decreases the
         current font size.
      </p>
      <p>This is how the <sup>superscript element</sup> and the
         <sub>subscript element</sub> format text.
      </p>
   </body>
</html>
```

Presentational Text Formatting – Example


Presentation Level Text Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File
Edit
View
Favorites
Tools
Help
Go
Bookmarks
57 blocked
Settings
Presentation Level Text F...
Page
Done
My Computer
100%

The last word of this sentence is formatted in **bold**.

The last word of this sentence is formatted in *italics*

See how the big element increases the current font size and how the small element decreases the current font size.

This is how the $^{\text{superscript element}}$ and the $_{\text{subcript element}}$ format text.

Logical Text Formatting Elements

- Logical text formatting describe the meaning of the style more than the actual format.

- Initially, browsers were left to determine the presentation of these tags as they saw fit, but over time certain standards were developed, and these are unlikely to change in the foreseeable future.

- For example, if you want a certain type, like bold, you should use the `<b>` elements, but the `<strong>` element will give you the same effect, because most browsers will interpret `<strong>` as `<b>.`

- The table on the next page lists the most commonly used logical elements.

Logical Text Formatting – Summary Chart

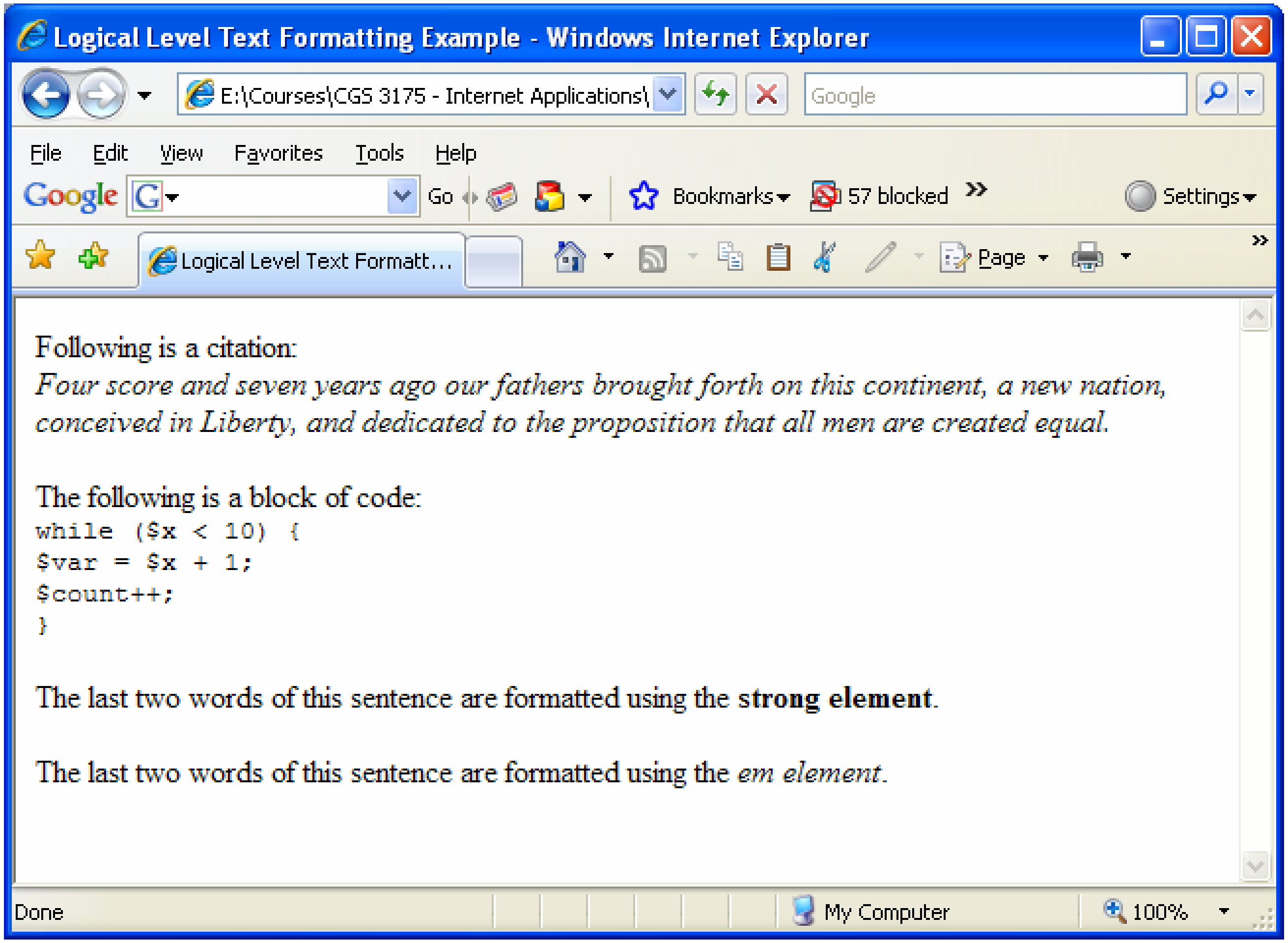
Element Name	Formatting Style
`<cite>...</cite>`	Defines a citation
`<code>...</code>`	Presents compute code examples
`<em>...</em>`	Emphasis. In most browsers, this is italics
`<strong>...</strong>`	Emphasis. In most browsers, this is bold

Logical Text Formatting – Example

```
<?xml version="1.0" encoding="UTF-8"?>
<!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>
     <title>Logical Level Text Formatting Example</title>
   </head>
   <body>
            <p>Following is a citation: <br />
               <cite>
              Four score and seven years ago our fathers brought
              forth on this continent, a new nation, conceived in Liberty,
              and dedicated to the proposition that all men are created equal.
               </cite>
        </p>
            <p> The following is a block of code: <br />
               <code>
              while ($x &lt; 10) { <br />
                 $var = $x + 1; <br />
                 $count++; <br />
              } <br />
           </code>
            </p>
            <p>
               The last two words of this sentence are formatted using the <strong> strong
element</strong>.
            </p>
            <p>
               The last two words of this sentence are formatted using the <em> em element</em>.
            </p>
   </body>
</html>
```

See page 31 for more details

Presentational Text Formatting – Example


Logical Level Text Formatting Example - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\
Google
File Edit View Favorites Tools Help
Google Go Bookmarks 57 blocked Settings
Logical Level Text Formatt...
Page
Done
My Computer
100%

Following is a citation:
Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal.

The following is a block of code:

```
while ($x < 10) {
$var = $x + 1;
$count++;
}
```

The last two words of this sentence are formatted using the **strong element**.

The last two words of this sentence are formatted using the *em element*.

Character Entities In XHTML

- The character of the less than symbol shown in the example on page 29 is written as `&lt;`.
- Certain characters in XHTML have special meaning to parser, like less than < and greater than >. These characters identify the beginning and ending of a tag, so if you want to add these characters as literal values, you must use the character entity code for them.
- A character entity is written in the following syntax: `&code;`. It begins with an ampersand (&) character, then the code for the entity, then a semicolon (;).
- Hundreds of symbols can be referenced and included on Web pages using entities. Some of the more popular symbols can be referenced by their abbreviations, like less than (`lt`) and greater than (`gt`), but they can also be referenced using their decimal value in the ASCII Table.
- Some of the most common character entity codes are shown in the table on the next page.

Some Common Character Entity Codes

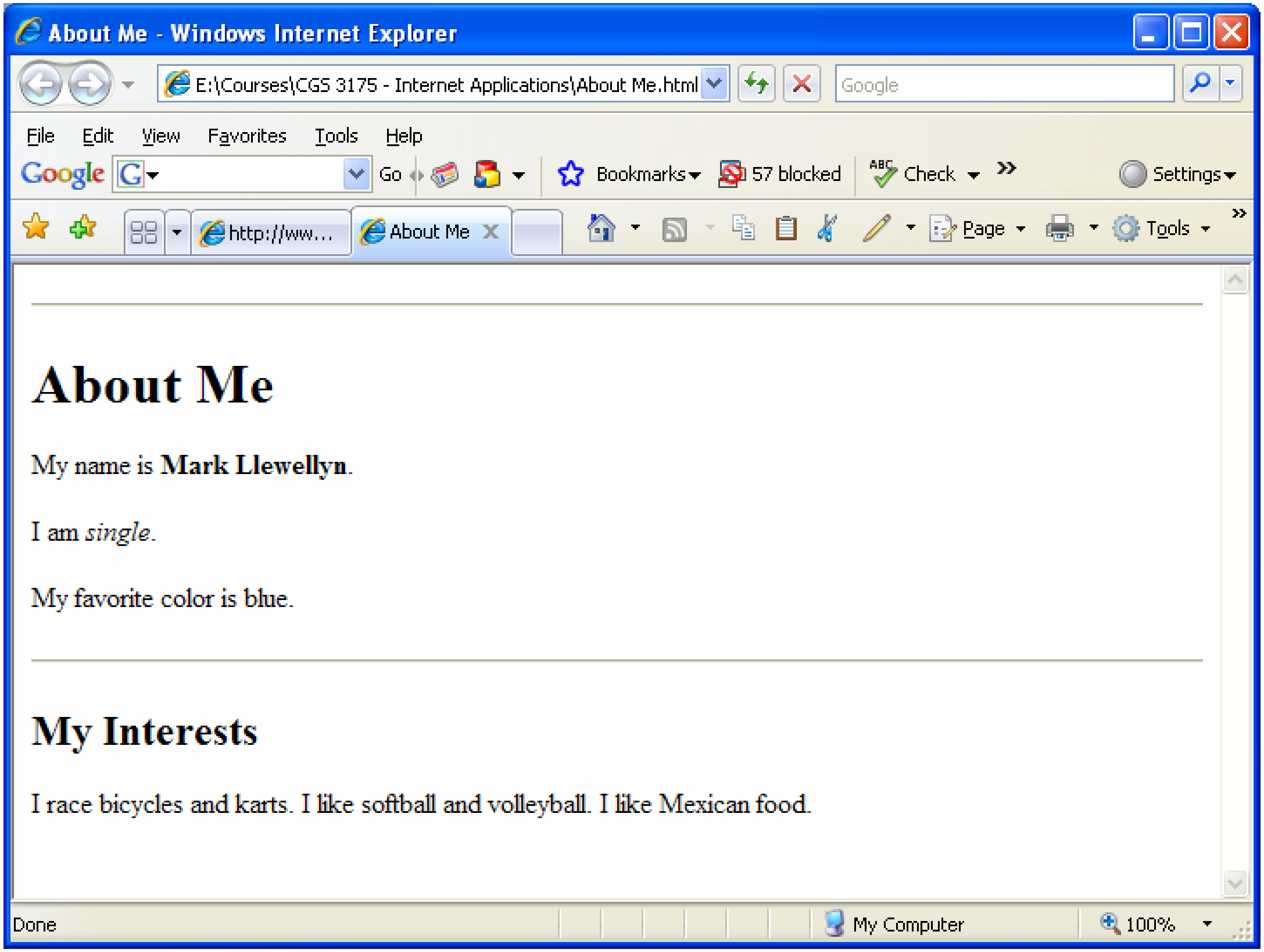
Symbol	Description	XHTML Code
>	Greater than	`&gt;   or   &62;`
<	Less than	`&lt;   or   &60;`
®	Trademark	`&trade;   or   &174;`
©	Copyright	`&copy;   or   &169;`
¢	Cent sign	`&cent;   or   &162;`

Some Practice

1. Create an XHTML document when viewed with a browser will produce a page very similar to the one shown on page 34. (Change the information so it is about you and not me!)

2. Create an XHTML document when viewed with a browser will produce a page exactly like the one shown on page 35.

Practice Problem #1 – Output



About Me - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\About Me.html
Google
File Edit View Favorites Tools Help
Google G Go Bookmarks 57 blocked Check Settings
http://ww... About Me
Page Tools
Done
My Computer
100%

About Me

My name is **Mark Llewellyn**.

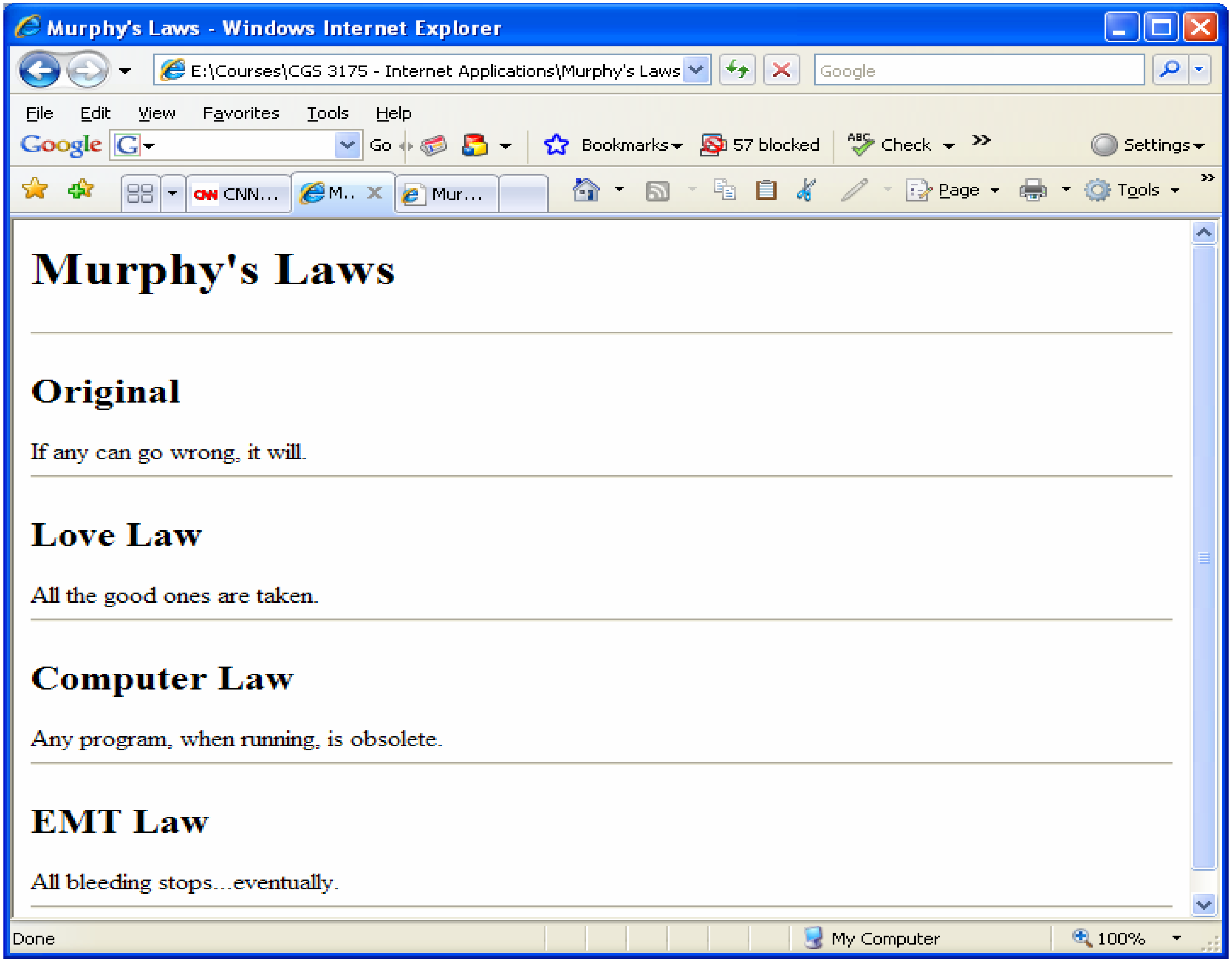
I am *single*.

My favorite color is blue.

My Interests

I race bicycles and karts. I like softball and volleyball. I like Mexican food.

Practice Problem #2 – Output



Murphy's Laws - Windows Internet Explorer
E:\Courses\CGS 3175 - Internet Applications\Murphy's Laws
Google
File Edit View Favorites Tools Help
Go Bookmarks 57 blocked Check Settings
CNN... M.. Mur...
Page Tools
Done
My Computer
100%

Murphy's Laws

Original

If any can go wrong, it will.

Love Law

All the good ones are taken.

Computer Law

Any program, when running, is obsolete.

EMT Law

All bleeding stops...eventually.