COMPSCI 111 / 111G

Mastering Cyberspace:
An introduction to practical computing

XHTML

Markup Languages

Markup

- Indicates the formatting that should be used to display the page

Markup Language

- A markup language consists of special tags which are placed in the text
- Specifies how to format the text

Hypertext Markup Language (HTML)

- Used to format web pages
- Contains hypertext information (links)
- Written in ASCII / Unicode
- Embedded format codes (tags)

Markup Languages

Heading

Paragraph

Formatting. Plain text contains all the content which will appear on the page. Using a variety of formatting styles makes the text much easier to read and can be used to emphasise points.

Emphasise

Markup Tags (HTML)

```
<H1>Formatting.</H1> Plain text contains all the
content which will appear on the page.
<P> Using a variety of formatting styles makes the text
much <EM>easier to read</EM> and can be used to
<EM>emphasise points.</EM>
```

Final Appearance

Formatting

Plain text contains all the content which will appear on the page.

Using a variety of formatting styles makes the text much *easier to read* and can be used to *emphasise points.*

Browser Wars

HTML 1.0

- Tim Berners Lee (1993)

Browsers added extra features

- Internet Explorer had unique tags
- Netscape Navigator had unique tags

Major problem

- What tags should a publisher use?
- How can this problem be resolved?

http://en.wikipedia.org/wiki/Browser_wars

Development of HTML

HTML 2.0
- Internet Engineering Task Force standard (1995)

HTML 3.2 / HTML 4.0
- W3 Consortium recommendation (1997)

HTML 4.01
- W3C recommendation (1999)

XHTML 1.0
- W3C recommendation (2000)

http://en.wikipedia.org/wiki/Html

Document Type Definition

Defines which standard is being used for the page
- We use XHTML 1.0 Strict
- Extensible Hypertext Markup Language 1.0 Strict

Should appear at the top of the file

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

http://en.wikipedia.org/wiki/XHTML

Encoding methods

Different character sets used to encode the page
- ASCII
- UTF-8
- Unicode

Need to tell the browser which encoding is used
- We will use the following at the top of the page

```
<?xml version="1.0" encoding="utf-8"?>
```

Use at the start of every file

Copy and paste the code exactly as it appears
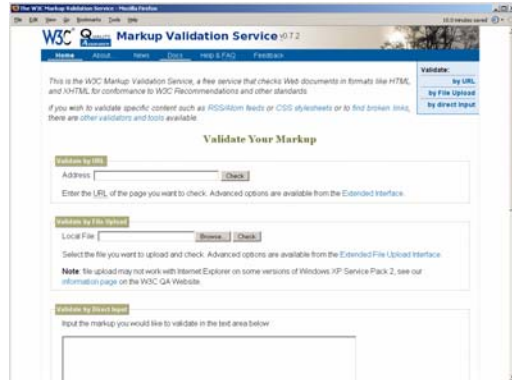- Will be provided in tests and exams
- Do not have to understand this code

```
<?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">
```

Validated Code

Online system to check correctness of code
- Provided by W3C
- http://validator.w3.org

HTML Source Code

Code used by the browser to display the page
- White space is ignored

Comments
- Ignored by the browser
- Allow you to document your code
- `<!-- Put your comment here ->`

Layout
- Use tidy layout where possible
- Make code easy to understand
- Make code easy to maintain/ modify
- Use whitespace and comments to help

Overview of tags

Markup achieved with "tags"
- Enclosed with angle brackets < … >
- Use lower case
- Come in pairs <tag> … </tag>

Tag usually applies to text between start and end tag

```
This word is in <i>italics</i>
```

Nested Tags

Tags must be correctly nested
- Cannot close an open tag until all the open tags that it affects are also closed

```
<tag>
Text only affected by "tag"

<tag2>
Text affected by both "tag" and "tag2"
</tag2>

Text only affected by "tag"
</tag>
```

Essential tags

`<html>`

- Encloses the entire document
- Specifies that the document uses html

```
<html>
.
.
.
.
.
.
.
</html>
```

Essential tags

`<head>`

- Contains information for the browser
- Does not contain any content to be displayed on the page

```
<html>
<head>
.
</head>
.
.
.
.
.
</html>
```

Essential tags

`<body>`

- Contains all the content that will appear on the page

```
<html>
<head>
.
</head>
<body>
.
.
.
</body>
</html>
```

Essential tags
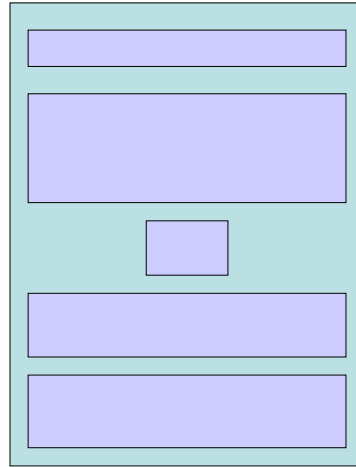
`<title>`

- Part of the head
- Specifies the title to be used by the browser
- Name of the window
- Used in navigation (bookmarks, history, etc.)

```
<html>
<head>
<title>Introduction to tags</title>
</head>
<body>
.
.
.
</body>
</html>
```

Block-level tags

Define the structure of a "block"

- Headings
- Paragraphs
- Lists
- Tables
- Preformatted text

Paragraphs

<p>
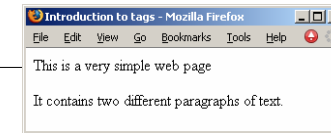
- Defines a paragraph of text

```
<html>
<head>
<title>Introduction to tags</title>
</head>
<body>
<p>This is a very simple web page</p>
<p>
It contains two different paragraphs
of text.
</p>
</body>
</html>
```
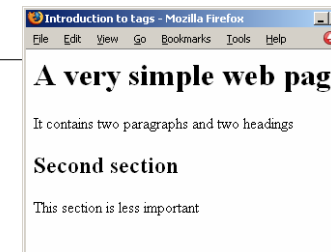
Introduction to tags - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

This is a very simple web page

It contains two different paragraphs of text.

Headings

Six levels of headings

- **<h1>**   First level heading

- **<h2>**   Second level heading

- **<h3>**   Third level heading

- **<h4>**   Fourth level heading

- **<h5>**   Fifth level heading

- **<h6>**   Sixth level heading

Headings

```
<html>
<head>
<title>Introduction to tags</title>
</head>
<body>
<h1>A very simple web page</h1>
<p>It contains two paragraphs and two
headings</p>
<h2>Second section</h2>
<p>This section is less important</p>
</body>
</html>
```

Introduction to tags - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

A very simple web page

It contains two paragraphs and two headings

Second section

This section is less important

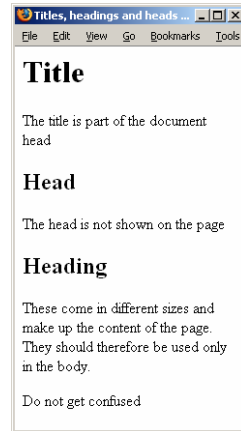# Example of Head, Heading and Title

```
<html>
<head><title> Titles, headings and heads</title></head>
<body>
   <h1>Title</h1>
   The title is part of the document head

   <h2>Head</h2>
   The head is not shown on the page

   <h2>Heading</h2>
   <p>These come in different sizes
   and make up the content of the page.
   They should therefore be used only
   in the body.</p>
   <p>Do not get confused</p>
</body>
</html>
```


Title

The title is part of the document head

Head

The head is not shown on the page

Heading

These come in different sizes and make up the content of the page. They should therefore be used only in the body.
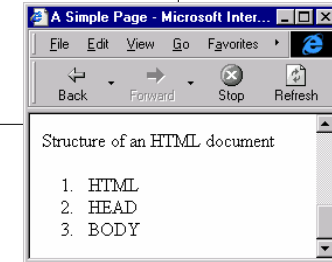
Do not get confused

Ordered Lists

Ordered Lists
- Automatically numbered
- ** … **   Contains the entire list
- ** … **   Used for each list item

```
<p>Structure of an HTML document</p>
<ol>
<li>HTML</li>
<li>HEAD</li>
<li>BODY</li>
</ol>
```
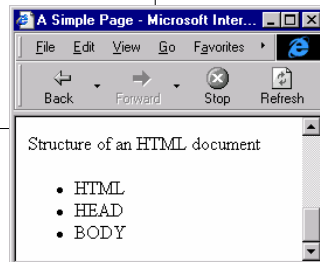

Structure of an HTML document

1. HTML
2. HEAD
3. BODY

Unordered Lists

Unordered Lists
- Bullet Points
- ** … **   Contains the entire list
- ** … **   Used for each list item

```
<p>Structure of an HTML document</p>
<ul>
<li>HTML</li>
<li>HEAD</li>
<li>BODY</li>
</ul>
```
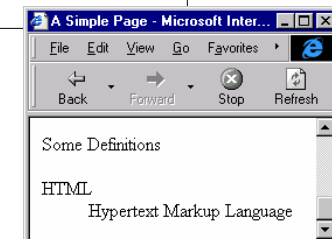

Structure of an HTML document

- HTML
- HEAD
- BODY

Definition Lists

Terms and explanations
- **<dl> … </dl>**   Contains the entire list
- **<dt> … </dt>**   A definition term
- **<dd> … </dd>**   A definition description

```
<p>Some definitions</p>
<dl>
<dt>HTML</dt>
<dd>Hypertext Markup Language</dd>
</dl>
```


Some Definitions

HTML
        Hypertext Markup Language

Inline tags

Appear within the blocks
- Apply to words within paragraphs etc.

Common inline tags
- Line Breaks
- Images
- Hypertext References

Line break

`<br>`
- Forces a line break
- Same as hitting the Enter key
- Use `<br></br>`

Backwards compatibility
- Use "hack" that older browsers will accept
- `<br />`

Attributes

Some tags require additional information
- Properties or Attributes of the tag
- `<tag property="value"> … </tag>`

Images

Pages may contain images
- But images are not plain text
- Can't be inserted directly into HTML page

Solution
- Store the image on the internet (or locally on disk)
- Tag contains the address of the image
- Web browser loads image when required
- Only use images the browser understands
- GIF, JPG, PNG

Image tag

`<img>`
- Insert an image at this location

`src`
- The source file of the image
- Attribute that specifies the file name

`alt`
- Attribute to specify alternate text
- Displayed if the image can't load
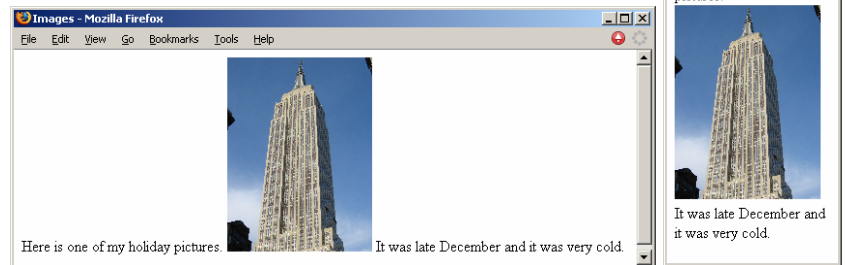- Important for people with visual impairment

`<img src="filename" alt="description"></img>`

`<img>` example

An image is inserted inline, so it is used inside a block-level element (a paragraph in this example).

```
<p>
Here is one of my holiday pictures.
<img src="Empire.jpg" alt="The Empire State Building">
</img>
It was late December and it was very cold.
</p>
```

Hypertext reference

A link to another resource on the WWW
- References to other documents
- Pages, images, files, sections

`<a>`
- **Anchor tag**

`href`
- Attribute used to specify the destination of the link
- URL

`<a href="… url …">text appearing</a>`

URLs

Fully specified
- Protocol
- Host name
- Path
- File

Relative
- Omit the first parts
- Path and file
- File

```
http://www.cs.auckland.ac.nz/courses/compsci111/index.html
/couses/compsci111/index.html
lectures/index.html
index.html
```

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>
<head>
<title>A sample page</title>
</head>
<body>
<h1>Example</h1>
<p>This is a complete xhtml web page.  You can verify that all the code is
correct using the <a href="http://validator.w3.org">W3C Validator</a>.</p>
<h2>Images</h2>
<p>If your code is correct, then you can include an image to show that the
page is validated.</p>

<p>
Author: Andrew Luxton-Reilly<br></br>
Date: 01/04/06<br></br>
</p>
<p><img src="valid-xhtml10.png" alt="Validated xhtml logo"></img></p>
</body>
</html>
```

Example page