MPS115/116 Presentation Lecture 7: An introduction to HTML

Dr Simon Willerton

1 What is HTML?

HTML

HTML stands for HyperText Markup Language. Rememeber that LaTeX is referred to as a markup language. LaTeX commands annotate text to make it display in specified ways. HTML is another markup language.

- It is used mainly for creating webpages.
- It allows easy formatting of paragraphs and headings and also commands specifically designed for webpages (e.g. hyperlinks).
- As with LaTeX, HTML files are really just annotated text files.

A minimal HTML file I

```
<!DOCTYPE html>
<html>
    <body>
        <h1>My first webpage</h1>
        <p>Hello! I am a webpage.</p>
    </body>
</html>
```

By copying and pasting this into Notepad (or any other text editor) and saving it as `index.html` you will create a basic webpage, which will open in a web-browser.

Notice the similarities to LaTeX environments: each *tag* comes in an opening and closing pair, e.g., `<p>` and `</p>`.

A minimal HTML file II

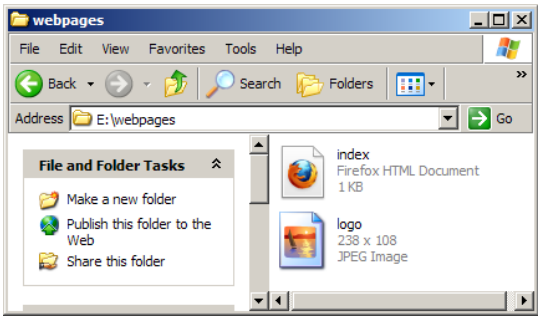
- 'The material to be displayed is between the `<body>` and `</body>` tags. (This part is called the *body* of the document.)

- Headings come in six different weights, `<h1>` to `<h6>`; the top level is `<h1>`.
- Paragraphs are built up using the `<p>` and `</p>` tags.

2 Folders and filepaths

Images

Suppose you have webpage `index.html` and in a folder called `webpages`. You may want to include an image on your page, called `logo.jpg`, say. Start by saving this in the webpages folder.



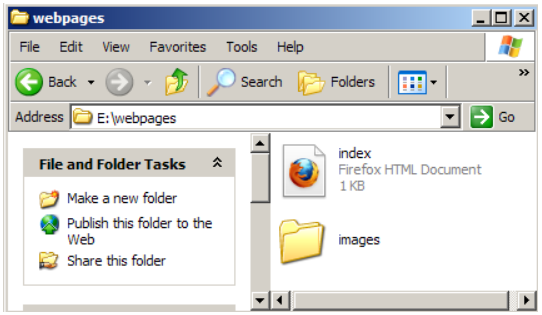
You can then include the image on your page with the command

```
<img src="logo.jpg">
```

(Note that the this tag is unusual in that it doesn't come in a pair.)

Subfolders

Suppose you decide you'd rather keep the images in a subfolder of `webpages` called `images`, to reduce clutter.



All you need to do is create the folder, move the `logo.jpg` file, then update your HTML to read

```
<img src="images/logo.jpg">
```

Here, `images/logo.jpg` is called the *relative URL* of the image.

URL

Uniform Resource Locator

This tells the computer where to find a file (such as `logo.jpg`).

- A *relative* URL starts from the folder in which the page itself sits. In the above case this is `images/logo.jpg`.
- An *absolute* URL starts from anywhere on the internet. In the above case this could be `http://maths-skills.group.shef.ac.uk/images/logo.jpg`.

3 Hyperlinks

The anchor tag

One big reason that the world-wide web has been so successful is the provision of *hyperlinks* in webpages. These allow users to click on words or images to get to new pages. To create a hyperlink we use the *anchor* tag.

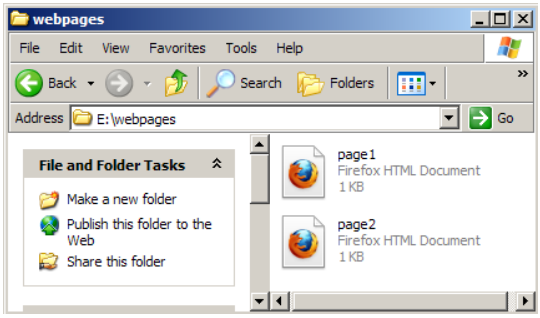
```
<a href="url">link text</a>
```

Here

- `href` stands for *hypertext reference*,
- `url` is the URL of the file to link to,
- `link text` is the text that will be clicked on.

Relative URL hyperlinks I

Suppose we have two webpages, `page1.html` and `page2.html`, both sitting in the same folder.

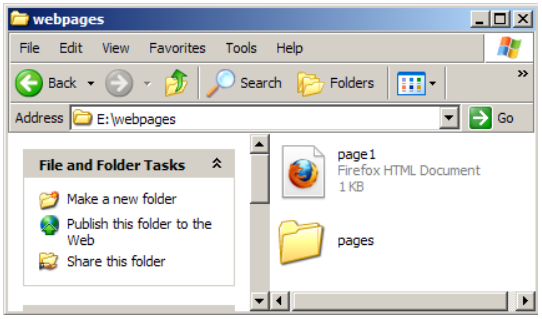


To create a hyperlink in page1 which points to page2, we use

```
<a href="page2.html">Click for page 2</a>
```

This creates the text Click for page 2 as a link: we've used a *relative* URL.

Relative URL hyperlinks II



If `page2.html` was inside a subfolder called `pages`, we would instead do

```
<a href="pages/page2.html">Click for page 2</a>
```

Absolute URL hyperlinks

Linking to external webpages, such as the BBC website, is also easy. Here, we use an *absolute URL*:

```
<a href="http://www.bbc.co.uk">BBC webpage</a>
```

It is best to use relative URLs within your own site, only using absolute URLs when you need to link to another website. Note the necessary "`http://`" part!

4 Filenames

Important!

When choosing names for HTML files or pictures you should

- only use lowercase letters (and/or numbers)
- and never use spaces.

Filenames must be identical in any links in the HTML code. Doing otherwise can lead to broken links or pictures not displaying. Even if you get away with it on your computer, it can break when your website goes live.

- Instead of using a space, use an underscore, '`_`', or a hyphen '`-`'.
- In this module, to be consistent, we use underscores.
- Rather than '`My page 1.html`' you should use '`my_page_1.html`'.

File extensions

The type of a file is usually indicated by its *file extension*. This is the part of filename after the dot, e.g. `.jpeg` or `.png`.

By default, Windows no longer shows these file extensions. You can change the default setting in Windows. (Find out how by looking on the web.)

Make sure you consistently use lower case for file extensions. We have seen problems in project with filenames like `file1.JPG` rather than `file1.jpg`, which stops a picture from appearing.

5 Activity time

Activity (Spot the difference, with a difference)**.** The circulated HTML code almost corresponds to the code for the webpage available under Week 8 at the course website. Try to spot as many places as possible that the code is different from what's on screen.

6 Further things

Viewing the page source

It is possible to look at the source code for any HTML webpage you visit. Some pages have complicated HTML, often generated by software. On others you will find very plain HTML.

To view the source, usually you can right-click on the page and select 'view source' (or similar).

Structuring web pages

Some web pages are easier to navigate than others. A menu or navigation bar can help. Another option is a site-map, which lists the entire content of the site.

Try to structure websites logically, and avoid links to pages that don't appear on a menu.

When marking student websites, sometimes it's hard to be sure every page has been read.

7 Semester 1 mini-project reminder

Submission deadline

The first mini-project must be submitted by 2.00pm tomorrow (Wednesday).

Please make sure you submit, as it counts towards your module grade.

If you are stuck, remember to make use of the discussion board.

If you haven't yet started, there's still time (just!).

About the Week 8 Computer Lab

In this week's computer lab we'll start looking at HTML and create some basic webpages.