

HTML OVERVIEW

Hyper Text Markup Language is the method of marking up what is basically a simple text file with TAGS (instructions to the computer but invisible to the user, enclosed in <> brackets) to create web pages for the Internet. These pages are read by a browser program and interpreted to lay the page out on a user's screen. They can include graphics as files called by the browser when the page is loaded and then inserted on screen, and they can include links (anchors) which when clicked on will cause the browser to call other files. Although it is a good idea to have all the graphics used in a page on your own web-site it is not absolutely necessary; the links can be to any page on the Internet.

Once a page is loaded into the browser there is no further interaction with the Internet unless the user clicks on a link, or his browser has been instructed by code on the page to do something more, automatically.

Many aspects of the page can be controlled, but since they may be viewed on almost any computer, and using many different types of browser, there is no absolute control of the page layout. It can be of any width and height, and the fonts used to display it are, essentially, under the control of the user not the page designer. You may assume that Times will be the main font, Courier will be used as the monospace font and Helvetica for various heading styles, since most users will not have changed the fonts supplied with their browser to others of their own choice - however this is not guaranteed.

The window size on different browsers varies quite a lot. The window seen by Netscape Navigator and Microsoft Internet Explorer is considerably smaller than that seen by Fresco, and heaven knows what AOL does, so designs must allow for this. You can not assume, even, that the user will be using an 800 x 600 display mode, though nowadays they probably are. By putting a piece of graphics 600 pixels wide on your page, the width of the whole page will be forced to accommodate it, so if the user is set up with a smaller mode they will have to scroll sideways all the time to be able to read all the text.

The user may well be blind, so titles and alternates should be used for text to speech prog-rammes to get their teeth into. Also some people turn off the graphics display on their browsers to speed up page loading; the page should work adequately with graphics removed.

So the pages must be all things to all browsers, and watch it, the language is developing so some browsers do more than others. This document is based on the published HTML 3.2 specification of Jan 1997, but frames, which are in common use, aren't even on that spec!

This document is divided up into sections, and is subject to continuous alteration, so there is no index of page numbers.

Section 1 :	HTML DETAIL	This section: background notes on what HTML does
Section 2 :	Quick Look-up List (QLL)	All the Tags (codes) in the HTML 3.2 specification: TABLES, LISTS, FORMS, MAPS, COLOURS, EFFECTS, STYLES, SPECIAL CHARACTERS
Section 3 :	Specials	Info on how to use various provider's programmes.
Section 4 :	http://www.durus.demon.co.uk/test.html	The HTML code for a page I maintain on my site to check out code on my browser and test how other browsers handle different Tags. A useful learning document, but wotchit, it gets updated!

HTML DETAIL

The text is controlled by "tags" which are short codes enclosed between < and > brackets.

These may be single

<hr> gives a horizontal rule,

or may come in pairs

<big> gives enlarged text, </big> ends it.

When tags are indicated as to be used in pairs, it is obligatory to use the closing tag, i.e. the </big> may not be omitted.

The Quick Look-up List (QLL) gives a list of all the tags in the HTML 3.2 specification and indicates when pairs are obligatory.

A page consists of three parts: The **Identifier**, the **Header** and the **Body**.

The **Identifier** usually consists of one optional line telling the browser what generation of HTML is being used (see QLL) followed by an <HTML> tag. The whole document is then finished by </HTML>. They are blind to the user.

The **Header** is enclosed between <Head> and </Head> tags and contains the Title of the page and any computer readable <Meta ..> tags. Mostly it is blind to the user, but the <Title> is shown on some browsers, on all "history lists" and sometimes when the page is printed, and can be spoken by text to speech programmes..

The **Body** is where the page is designed in detail. It lives between <Body and </Body> tags. Read the HTML printout of "http://www.durus.demon.co.uk/test.html" at the back of this document to get the general idea. Check the printout against the actual page on the Internet.

Comments may be added anywhere within the <HTML> .. </HTML> tags (see QLL). They are inserted thus <!-- Comment -->, and are blind to the browser. It is useful to add comments to a page to break it up and make it easy to see what you did originally when you come to alter it in a year's time!

There is no limit to the length of a page, except the difficulty of handling it on the browser and the astonishing time it takes to load over the Internet.

So a page goes:	<pre> <!Doctype HTML Public "-//W3C//DTD HTML Final//EN"> <HTML> <Head> . . . All sorts of head stuff . . . </Head> <Body> . . . The main visible page stuff . . . </Body> </HTML> </pre>	<p>Identifier</p> <p>Start of HTML document</p> <p>Start of header</p> <p>End of header</p> <p>Start of body</p> <p>End of body</p> <p>End of HTML document</p>
-----------------	---	---

HTML tags are case insensitive (except in very few special cases indicated in the QLL), HTML treats multiple consecutive spaces as a single space and it ignores return characters. Use caps or lower case as you wish within tags to make them most readable on your screen when you are editing the source code. A non-breakable space which is treated as a character not as a space can be inserted by using " " (without the quotes) and
 is used to force a line break.

You can nest tags, but do NOT interlace tags :-

<i> </i>	is illegal and may cause problems,
<i> </i>	is correct and has the same effect.

Just to make it complicated, one tag pair (<p>) may have the closing tag (</p>) omitted, since it is implicit that a new paragraph must start after the end of its predecessor (this is for backwards compatibility with old versions of HTML, I believe). Browsers should not mind </p> being left out, although a few old ones get in a muddle if you do.

There is a problem with leaving out `</p>`, and that is that you find you have interlaced tags unwittingly since `<p>` is really part of a pair, so :-

```
<p> ... <big> ... <p> ... </big>
<p> ... <big> ... </p><p> ... </big>
```

is illegal since it acts as which shows the interlacing.

Different browsers will display this differently.

It could legally be written as :-

```
<p> ... <big> ... </big><p><big> ... </big>
```

or frequently

```
<p> ... <big> ... <br> ... </big>
```

might be enough for what you want.

I have learned that it is not absolutely necessary to use the `<p>` tag all the time, and it is worth avoiding it if possible on complicated pages. However it can be used to control the alignment of text when used like this :-

```
<p align=right> ... </p> Right justifies the enclosed text in this para only.
```

Note that although opening tags may take modifiers, closing tags do not, so :-

```
<body bgcolor="gray" text="black"> would be closed by </body>
```

No there are no spelling mistakes in the above line, we're speaking American here, "colour" and "grey" are wrong and will not be recognised by the browser.

In the Quick Look-up List I have indicated tags that do not take modifiers thus :-

```
<br>
<center> ... </center>
```

Note the spelling again

Tags that do take modifiers are shown thus :-

```
<hr __> HORIZONTAL RULE
```

(puts a rule line across page or table cell)

with modifiers :-

```
.. width=N° (Sets rule width in pixels)
   ="N°%" (Sets rule width in %age of page or table cell width)
.. size=N° (Sets height of rule in pixels)
.. NoShade (Makes rule solid)
```

So the definition of a horizontal rule line could be :-

```
<hr width="60%" size=8>
```

Which would produce a fairly fat raised line filling 60% of the available width.

In general, modifiers can be left out if the tag still makes sense. In the case above I did not want a solid line, so "NoShade" was omitted. Often there is a sensible default if modifiers are left out, so `<hr>` alone would produce a minimum thickness line occupying the full available width, of course!

To see more of how to build a page read Section 4 of this document, preferably with the actual page (taken off the Internet) on a computer screen in front of you to see what each part of the code does. I update <http://www.durus.demon.co.uk/test.html> every so often, so the printout may be out of date, if so, click on "View -> Source" or equivalent on your computer's menu system to get the current HTML.

While reading you will need to get used to the fact that you can no longer work in point sizes (you aren't in control of them, the browser is) and you have to start thinking in pixels if you are working in absolute sizes, or even better, if the page layout allows you to, percentages. A page which is mostly defined in percentages will tend to work on more screen resolutions, but it's tricky to get correct because images always work in pixels.