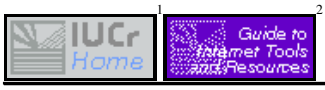

HTML and Hypertext

HTML (HyperText Markup Language) is an evolving language which is used to construct documents which can be viewed by World Wide Web browsers. It was invented by Tim Berners–Lee while at CERN, the European Laboratory for Particle Physics in Geneva. HTML has been standardized by the WWW consortium as the IETF RFC 1866, commonly referred to as HTML Version 2[3]. The W3C is the official organization maintaining, standardizing and developing HTML. W3C provides a source of all HTML–related material[4].

NCSA provides an excellent primer[5] for beginners wishing to learn HTML.

HTML was designed to indicate the logical and semantic content of a document rather than its physical appearance as print on paper or pixels on a screen. The form in which a web document appears on the user's screen is a problem that has to be resolved by the particular browser (client) software depending on the hardware available and user preferences. This particular aspect of the design specification, together with the capability of distributing text instead of images, opens up the access of HTML–based information to the physically handicapped such as the blind, where HTML–marked text can be translated into Braille or synthesized speech. With the development of technology there are also clear possibilities for machine–based language translation.

HTML is a collection of platform independent styles indicated by markup tags that define the various components of a World Wide Web document. HTML documents are plain–text (also known as ASCII) files that can be created using any text editor (*e.g.*, *Emacs* or *vi* on UNIX machines; *BBEdit* on a Macintosh; *Notepad* on a Windows machine). One can also use word–processing software by remembering to save the document as *text only with line breaks*. Some WYSIWYG editors are available and described in the WWW authoring[6] section. You may wish to try them out after having learned the basics of HTML tagging using a simple text editor. It is useful to know enough HTML to code a document before you experiment with WYSIWYG editors.

Every HTML document should contain certain standard HTML tags. The tags denote the various elements in an HTML document. HTML tags consist of a left angle bracket (<), a tag name, and a right angle bracket (>). Tags are usually paired (*e.g.*, `<H1>` and `</H1>`) to start and end the tag instruction. Some elements may include an attribute, which is additional information that is included inside the start tag. For example, you can specify the alignment of images (`top`, `middle`, or `bottom`) by including the appropriate attribute with the image source HTML code.

Each document consists of a head followed by a body. The head contains the title, and the body contains the actual text that is made up of paragraphs, lists, and other elements. Required elements are the `<HTML>`, `<HEAD>`, `<TITLE>`, and `<BODY>` tags (and their corresponding end tags `</HTML>`, `</HEAD>`, `</TITLE>`, and `</BODY>`) arranged as follows:

```
<HTML>
<HEAD>
<TITLE>An example title</TITLE>
</HEAD>

<BODY>
.
.
</BODY>
```

[1] http://www.iucr.ac.uk/iucr–top/welcome.html
[2] int.index.html
[3] http://www.sandia.gov/sci_compute/html_ref.html
[4] http://www.w3.org/pub/WWW/MarkUp/
[5] http://www.ncsa.uiuc.edu/General/Internet/WWW/HTMLPrimer.html
[6] int.authoring.html

```
</HTML>
```

Because you should include these tags in each file, you might want to create a template file with them.

The element <HTML> tells the browser that the file contains HTML–coded information. The file extension .html also indicates this an HTML document and must be used. If restricted to 8.3 filenames, use only .htm for your extension. The <TITLE> element contains the document title and identifies its content in a global context. The title is displayed somewhere in the browser window, usually at the top. The largest part of an HTML document is the <BODY>, which contains the content of the document displayed within the text area of the browser window. A number of special tags allows very wide structuration of the text: headings, paragraphs, lists, preformatted text, extended quotations, forced line breaks/postal addresses, horizontal rules. Direct use of other system resources (for example, e–mail) from the text is also possible.

Hypertext style

Writing a document in distributed hypertext is a skill that authors need to develop and Berners–Lee has given some excellent counsel on style[7] to help authors along. Too often one is confronted with texts produced as 'shovelware' in which documents prepared for distribution as printed paper are simply copied onto the W3 without further ado.

[Index[8]] – 19th July 1996 – © Yuri Grin[9] – **Not to be copied or reproduced without permission**

[7] http://www.w3.org/pub/WWW/Provider/Style/Overview.html
[8] int.index.html
[9] mailto:grin@vsibm1.mpi–stuttgart.mpg.de