

## HTML

M.Skuba, *student IN-41*,  
Plokhuta T.M., *EL adviser*

Just as there is a diversity of programming languages available and suitable for conventional programming tasks, there is a diversity of languages available and suitable for Web programming.

There is no reason to believe that any one language will completely monopolize the Web programming scene, although the varying availability and suitability of the current offerings is likely to favor some over others. Java is both available and generally suitable, but not all application developers are likely to prefer it over languages more similar to what they currently use, or, in the case of non-programmers, over higher level languages and tools.

HTML (HyperText Markup Language) is the basic language understood by all WWW (World Wide Web) clients. Unmodified HTML can execute on a PC under Windows or OS/2, on a Mac, or on a Unix workstation. HTML is simple enough that nearly anyone can write an HTML document, and it seems almost everyone is doing so.

HTML was developed as part of the WWW at CERN by Tim Berners-Lee, who is now Director of the World Wide Web Consortium (W3C) at MIT's Laboratory for Computer Science. Refinement of HTML continues at W3C, with standardization via the Internet Engineering Task Force (IETF) of the Internet Society.

HTML descended from SGML (Standard Generalized Markup Language), the ISO standard language for text. SGML is in widespread use by the US Government and the publishing industry for representing documents. HTML applies SGML principles to the WWW. As such, it implements a semantic subset of SGML with similar syntax.

HTML is a markup language rather than a complete programming language. An HTML document (program) is ASCII text with embedded instructions (markups) which affect the way the text is displayed.

The basic model for HTML execution is to fetch a document by its name (e.g. URL), interpret the HTML and display the document, possibly fetching additional HTML documents in the process, and possibly leaving hot areas in the displayed document that, if selected by the user, can accept user input and/or cause additional HTML documents to be fetched by URL.

HTML applications, or what we might consider the HTML

equivalent of an application, consist of a collection of related web pages managed by a single HTTP (HTTP is the topic protocol that defines the interaction of WWW clients and servers) server. This is an oversimplification, but the model is simple, and the language is simple, and that is one of its strengths.

As HTML moves through the standardization process, and is extended by various vendors, it loses some of its simplicity, but it remains a useful language. The Web programmer generally finds HTML lacking in only two areas: its performance in certain types of applications, and the ability to program certain common tasks.

HTML is limited in its computational power. This is intentional in its design, as it prevents the execution of dangerous programs on the client machine. However, Web programmers, as they have become more sophisticated in their applications, have increasingly been hamstrung by these limits.

Tasks unable to be coded in HTML must either be executed on the server in some other language, or on the client in a program in some other language downloaded from a server. Both solutions are awkward for the programmer, often produce a sub-optimal segmentation of an application across program modules, both client and server, and reintroduce safety considerations.

Because of an HTML program's limited functionality, and the resulting shift of computational load to the server, certain types of applications perform poorly, especially in the context of clients connected to the Internet with rather low bandwidth dialup communications ( $\leq 28.8\text{Kbps}$ ). The performance problems arise from two sources:

- a) an application which is highly interactive requires frequently hitting the server across this low bandwidth line which can dramatically and, at times, unacceptably slow observed performance ;
- b) requiring all computation to be done on the server increases the load on the server thereby reducing the observed performance of its clients.