

Creating and maintaining a personal website

Christopher F Baum

Department of Economics and

Faculty Microcomputer Resource Center

Boston College

October 27, 1998



Faculty
Microcomputer
Resource
Center

Introduction

- A personal website may be used to make your teaching and/or research materials available to students, colleagues, and potential employers
- Most 'productivity' software (e.g. Microsoft Office) may be used to rough out webpages, and handle cumbersome tasks such as table creation
- A little knowledge of HTML (HyperText Markup Language), coupled with a text editor, goes a long way in improving webpages' appearance

Introduction

- In order to present complex mathematics, graphs interspersed with text, etc., Adobe Acrobat may be used to create platform-independent, downloadable 'portable documents' (PDFs) which may be viewed on screen, and printed in their original form
- Your personal website may also be used to distribute programs, data, and other research materials that you want to share with students and colleagues via FTP

October 27, 1998

Faculty
Microcomputer
Resource
Center

Outline

- How do web pages work?
- Establishing and accessing your web account
- Creating a set of draft pages
- Refining your pages
- Creating and linking PDF files
- Establishing and linking FTP-accessible files

October 27, 1998

Faculty
Microcomputer
Resource
Center

How do web pages work?

- Each page on the Web is addressed by its URL (uniform resource locator), which specifies:
- Protocol to be used for access: e.g. `http:`
- Hostname on which the object is located: e.g. `fmwww.bc.edu`, a machine connected to the Web
- Directory in which the object is located, in UNIX (forward-slash) format: e.g. `/ec-c/f98/`
- Name of the page (file): e.g. `ec761.f98.html`
- `http://fmwww.bc.edu/ec-c/f98/ec761.f98.html`

Addressing a page

- The protocol for HTML pages is `http:`, the HyperText Transport Protocol
- HTML files are plain (ASCII) text
- Other protocols are used to access other Web services: `ftp:` for downloading files, `telnet:` to connect to a remote machine interactively, `mailto:` to send a mail message to a specified user
- URLs may be case-sensitive (and on UNIX hosts generally are)

October 27, 1998

Cross-references

- Web pages refer to other pages (and other URLs in general) via links, the highlighted/underlined segments of the page that make the page hypertextual
- Links to other pages in your own directory may be relative, that is, referring to a location relative to the base page
- Links may also refer, if 'fully qualified', to URLs on other hosts, or other directories on your host

Establishing a web account

- Personal web pages in BC's domain are stored on server `www2.bc.edu`. Your homepage on that server will be referenced as `http://www2.bc.edu/~username` where that URL actually references the file `default.html` on that directory.
- When you go to the 'Student/Faculty/Staff Personal Pages' link on the Econ homepage ('Personal Web Page Server' on InfoEagle), you have the option to 'Request a WWW Account.'

Establishing a web account

- When you 'Request a WWW Account,' that user account, directory and default home page (`default.html`) will be created, and you will receive a confirmation of those specifications. The default page will merely contain 'Welcome to Jane Doe's Home Page!'
- You will be given 10 Mb space on `www2.bc.edu` for your Web files. Since most HTML files (and even many graphics) are only 5-10 Kb, this should be more than adequate for most users.

Building your home page

- For the shorthand URL (ending in ~username) to function, you must have the page `default.html` in your `www2` directory. Filenames on `www2` (which is an IBM UNIX (AIX) machine like `fmrisc`) should not contain spaces, and are case-sensitive. This implies that links to other pages on your directory must respect case: `ps1.html`, `PS1.html` and `Ps1.html` are three different files. For simplicity, use lowercase in all filenames, and use the `.html` extension throughout to signal html files.

Building your home page

- You move files between your desktop machine and `www2.bc.edu` via FTP; telnet is not enabled on this machine. On a Macintosh, use Fetch; on Windows, download the WS_FTP LE application from `www2` (not included in Eaglenet Resources).
- Alternatively, Mac users have a better option: the full BBEdit package, which includes “Open from FTP server...” and “Save to FTP server...” commands. These facilities are not included in the freeware BBEEdit Lite package.

October 27, 1998

Building your home page

- There are two options for those constructing HTML pages:
- Use a program that manages all aspects of the HTML, “hiding” the HTML details, link addresses, etc.; or
- Use a program to create the basic HTML code, but then use a text editor to modify and tweak it to meet your needs
- The second option requires more investment, but may be less frustrating in the long run.

October 27, 1998

Faculty
Microcomputer
Resource
Center

Building your home page

- Programs which support the first option (handling all the details, enabling you to remain ignorant of HTML), such as Netscape Composer, Microsoft FrontPage or Adobe's PageMill, provide an environment in which some elements of page construction will inevitably be limited by the program. It may well be easier to learn a limited amount of HTML and 'tweak' the HTML with a text editor than determine how to convince the program to do exactly what you want.

Building your home page

- The second option is supported by the 'Office' environment, in which Word (97/98) and Excel (97/98) both have 'Save as HTML' commands on their File menu. These options are especially useful in creating tabular webpages; HTML tables are simple constructions, but require lots of markup commands. A Word table or Excel sheet will automatically create the required markup.
- To tweak the page, use a text editor (not a word processor!) such as BBEdit.

Building your home page

- If you use a word processor (including Word 9x) to modify an existing webpage, you must always remember to save it as a text file (or, in Word, Save as HTML). A Web browser cannot process a word-processor file saved in its own format.
- A text editor (such as BBEdit) is much easier to use for this task, and Web browsers' ability to open a local HTML file allow you to 'preview' your pages (and test everything but the links) before uploading them to the website.

Building your home page

- As a first task, retrieve your default.html page from `www2.bc.edu`, and make some simple changes to the page with a text editor (or with Word 9x, which will import the file as HTML; remember to Save as HTML).
- Try out the procedure of web page modification by adding the line
“This page is under construction<p>”
following the “Welcome” line, and FTP the page back to `www2.bc.edu`. Check it out in Netscape!

Principles of HTML

- HyperText Markup Language (HTML) is plain ASCII text which uses “tags” to “mark up” the document—similar to *TeX/LaTeX*, or DOS Word Perfect.
- Tags are denoted by `<tag>` or `<tag>...</tag>`.
- Like TeX, the form of the source document does not matter; you may type any amount of text on each line, and the line breaks will not affect the Web browser’s processing of that HTML code. You cannot use tabs or spaces to align HTML output in any reliable fashion.

Principles of HTML

- You cannot control the user's browser preferences, so that the size of text, font, and width of screen are all up to the user's control. Generally, text which appears to be the right size on a Windows web browser will be very small when viewed in the same browser on a Macintosh (and vice versa). Some users will view the page using a larger monitor; your page should be flexibly designed so that it can be viewed comfortably with any screen width. It should avoid extreme font sizes.

Principles of HTML

- Every document should begin with `<HTML>` and end with `</HTML>` .
- Two sections should be present:
`<HEAD>...</HEAD>`
`<BODY>...</BODY>`
- The `<TITLE>...</TITLE>` in the HEAD section is used to control the text displayed on the browser's window bar; it should always be filled in with a short text phrase (for bookmark use). It is usually the only material in the HEAD section of a webpage.

Principles of HTML

- The BODY section contains the body of the document. The beginning tag may be used to set the background color for the whole page:
`<body bgcolor=00ccff>`
will create a page with a light blue background. The six hex digits represent the amount of Red, Green, and Blue, respectively, in the background color, ranging from 00 (no color) to FF (full color).
- Choices of background color may be viewed at <http://www.lynda.com/hexh.html>

Principles of HTML

- The single tag `
`, or break, is used to insert spacing in text flow; it is the equivalent of the “return” on a typewriter.
- The single tag `<p>`, or paragraph, inserts both a ‘return’ and a blank line in text. It is the equivalent of hitting ‘return’ twice on a typewriter.
- Strictly speaking, `<p>` should appear at the beginning of a paragraph, with `</p>` at the end. Very few HTML authors obey this syntax (and browsers do not enforce it).

Principles of HTML

- The single tag `<hr>`, or horizontal rule, inserts a line across the screen. It may be modified by `size=n`, where larger values of `n` imply thicker rules.
- A single-color graphic with `height=2` or `3` is probably more effective in many cases than a horizontal rule:

```
<IMG SRC="/gif-bin/990033.gif  
height="3" width="576">
```

Principles of HTML

- Within the document, headings and subheadings may be designated with `<Hn> . . </Hn>` tags, for $n=1$ (very large), 2,...6. I recommend using headings 2, 3, and 4 for most uses. Headings automatically create breaks and vertical spacing.
- Text with a particular format—e.g. listing of a computer program—may be tagged as “preformatted” with `<pre> . . . </pre>`. The contents of that block will be displayed as they appear. This should not be overused!

Principles of HTML

- Within a text block, you may specify that text is to be bold (`...`) or italic (`<i>...</i>`). Be cautious with italic text, which is often quite difficult to read in a browser. You may also specify that the text is to be larger via `<big>...</big>`. There are tags for more precise font size control, but there is no direct specification of, e.g., point size, as you have in a word processor.
- Whenever you specify a text style tag (or a heading tag), don't forget to turn it off!

Principles of HTML

- Lists of items may be created as Unordered (with bullets) or Ordered (numbered) lists.
- Unordered:

```
<UL>  
    <LI>...  
    <LI>...  
</UL>
```
- Ordered lists are generated in the same fashion, using the `...` tag. The option `START=n` may be used to control the numbering.

Refining your pages

- With the tools specified thus far, you can readily produce a decent resume/CV or course syllabus.
- Some aspects of printed pages are more challenging to create as HTML. In particular, you may want the top of your CV or syllabus to contain both flush-left and flush-right text. This must be done with HTML tables; you cannot use tabs or spaces to position material on the right margin.
- `<center>...</center>` is available to center blocks of text on the screen.

Refining your pages

- To create HTML tables, use the Save as HTML... feature of Word 9x or Excel 9x. Material within a table cell will wrap as needed, and the alignment within each column may differ, as specified in the program used to create the table. Copy and paste the table specification (`<table>...</table>`) into your HTML document.
- An alternate table creation program: BBEdit's built-in Table Builder tool, available only in the full BBEdit program (and only for the Macintosh).

Refining your pages

- Hyperlinks are inserted with the `...` syntax, where the "linkURL" specifies the Web page to be accessed if the link is hit. The text between the tags is underlined and colored to indicate that it is a hyperlink.
- The linkURL may specify a page in your own default directory, in which case it need not be 'qualified'.

Refining your pages

- If the `linkURL` refers to a file in your own directory structure, you may use 'relative addressing' such as `/courses/ec362/362ps1.html`. This implies that you have created a subdirectory in your own account named `courses` as well as a subdirectory within that directory named `ec362`. The absence of a protocol or hostname implies that the link refers to a file on the same server as your homepage.
- These subdirectories may be created with Fetch (or any FTP client).

Refining your pages

- If the `linkURL` refers to a file on another site (including other BC users' sites on `www2.bc.edu`) you must 'fully qualify' the URL with the protocol, hostname, and directory: e.g.
`http://www2.bc.edu/~baum/`
to refer to my `www2` homepage.
- A link to sites outside BC would give the protocol and hostname (e.g. `http://www.cnn.com`).
- A 'mailto' link specifies the valid email address:
`mailto:baum@bc.edu` (note absence of //)

Refining your pages

- Links may also refer to specific sections of a page or document. If the 'anchor' tag `...` is used to specify a point in the document, then a link to that point may be made within the document with `...` or from another document with `...` to link to that section of the document.

Adding graphics

- Graphics may be added to a page in either GIF or JPEG format, which may be created by many programs (including the shareware GraphicConverter for Macintosh)
- Graphics produced by digital photography or by “Draw” programs will generally be in other formats (“PICT” or “Windows Bitmap”); conversion should take the target resolution (72 dpi, 256 colors) into account

Adding graphics

- To specify that a graphic should be added on an HTML page, use tag
``
where `gph.gif` is the filename of the graphic, and "alt text" displays if the graphic is not available (or if the viewer has image loading disabled)
- Even if a single graphic is the sole item to be viewed, it is wise to create an HTML page with an image reference to present it

Adding graphics

- The placement of the graphic can be modified with the `ALIGN` option to the `IMG` tag. `ALIGN=LEFT`, `ALIGN=RIGHT`, `ALIGN=CENTER` may be used in conjunction with `<BR CLEAR=ALL>`, which forces following elements below a horizontal line (preventing graphics from colliding).
- `HEIGHT=n` and `WIDTH=n` options on the `IMG` tag, where values of `n` are in pixels, may be used to scale an image (with caution, so as not to degrade resolution).

Adding graphics

- An image can be the item enclosed in a link:
``
so that the image itself becomes the hyperlink, or “button.”
- Graphics should be used with caution; large graphics will take a long time to load over a modem connection. Scaling a graphic to a fraction of its original size will speed up download. To gauge the graphic’s size, open it with “Open Page...” in your browser, which will report its size in pixels.

Additional resources

- An excellent summary of all defined HTML tags is available at <http://werbach.com/barebones/>
- A guide to many of the resources available on the Web about website design and maintenance is available at <http://werbach.com/web/wwwhelp.html>

October 27, 1998

Faculty
Microcomputer
Resource
Center

Creating PDF files

- HTML—the current language of the Web—is not well suited to the presentation of mathematics, inline graphics, or foreign scripts beyond the simple accents of major European languages.
- Where the pagination and “look and feel” of a document must be preserved, it may be served over the Web in Adobe’s “Portable Document Format,” or PDF: e.g. IRS tax forms are downloadable as PDF files, as are many professional journals, working papers, and the like.

Creating PDF files

- PDF files may be painlessly created on either Macintosh or Windows platforms with the “full” Adobe Acrobat package. The freeware Acrobat Reader cannot create PDFs.
- The “full” package is available to academic purchasers for \$39.00 via BC Tech Products.
- “Acrobat Distiller” is the preferred mechanism for creating PDF files; the “PDF Writer” option is less reliable and creates larger, less efficient PDFs.

Creating PDF files

- Acrobat Distiller is automatically launched when “PDF” is selected in the `File->Print` dialog. The application program (Word, Excel, etc.) creates a print spool file in PostScript (`.ps`) format, then Distiller transforms that file into a PDF file. Distiller may also be used to transform a `.ps` file created by a different application (such as a “`dvi2ps`” routine in *LaTeX*, or GAUSS graphics output) into a PDF file.

Creating PDF files

- Acrobat Exchange may be used to merge PDFs (including those in portrait and landscape format). This is very useful if you are using a word processing application and a separate graphics application (Excel, DeltaGraph, etc.) and want to produce a single PDF file. Exchange can also rotate individual pages into the appropriate orientation for viewing on screen.

Serving PDF files

- Once PDF files have been created, they may be uploaded to your `www2` account, and specified in an `HREF` link as a target.
- The browser will either download the PDF file to the user's desktop, launch an "Acrobat Reader plugin" to view the PDF within the browser window, or inquire as to how the file should be handled.
- When a downloaded PDF is printed, it will have the same quality, graphics, fonts, etc. as the original document from which it was distilled.

Establishing a FTP directory

- In order to give users access to binary files (such as compiled programs, datasets in the format of a statistical package, etc.) you have two options:
- Create archive files (e.g. `.zip` archives):
reasonable if the files are not too large
- Become an “Info Provider” and use an account on `ftp.bc.edu`, BC’s anonymous FTP server, to give users access to the files. This may also be a good option if the datafiles to be served are large, given limited per-user space on `www2.bc.edu`.

October 27, 1998

Establishing a FTP directory

- When you request Info Provider status (see the Help Center for the forms), you will be given FTP access to a directory on `ftp.bc.edu`.
- Files in that directory are referenced as `ftp://ftp.bc.edu/pub/user/username/file.ext` where `username` is your University username.
- You may transfer files to that directory, create new subdirectories, etc. using FTP clients such as Fetch or WS_FTP LE.

Serving files via FTP

- Files may then be referenced on your Web pages using the `ftp:` protocol. The Web browser will use the FTP protocol to download the specified file to the user's desktop.
- Placing files on an FTP server also makes them accessible to users with either a Web browser or an FTP client: for instance, users on `bcaxp1` or `fmrisc` cannot run Netscape, but they can use the `ftp` command to download files from `ftp.bc.edu`.

Serving files via FTP

- An example of a large-scale delivery system using HTML pages linked to FTP downloads may be found at <http://ideas.uqam.ca>
- This site provides searchable access to thousands of working papers and software components via HTML-formatted bibliographic information, with FTP links to the downloadable documents or programs
- BC Econ's working papers and software archives:
<http://ideas.uqam.ca/ideas/data/bocbocoec.html>
<http://ideas.uqam.ca/ideas/data/bocbocode.html>