Long Documents

While lengthy documents are somewhat discouraged on the Web, there are several methods of handling them that maximize the value of "hypertext".

HTML.edit Limitations (and How to Get Around Them)

HTML.edit documents are composed of three fields: header, body and footer. Due to a limitation in the development environment, there is a 30,000 character limit to each field. Conceivably, one could create a 90,000 character document in HTML.edit by dividing the text between the three fields (imported documents automatically overflow from body to footer). This would be much larger than the majority of web documents.

Dividing Large Documents

HTML.edit supports automatically dividing documents on import using a customizable separator character. This feature is by default set to a Form Feed character (ASCII 12), but turned off.

To turn on the feature, go to the Preferences card and select Import/Export Preferences. Make sure the "New Document on…" checkbox is checked, and select either "Form Feed" or "Other" (which uses a character in the small box as separator). I usually use insert "§" symbols (option-6) in the source text before importing the document into HTML.edit.

On import, HTML.edit will create a new document each time it encounters the separator character. This could be a chapter or section marker, for example.

Actual Practice

To test this out, I used a text copy of Melville's Typee (a tropical island adventure), which has many small chapters. In BBEdit, I opened the text copy and case-sensitive replaced all instances of uppercase "CHAPTER" with "§CHAPTER". I then turned on the New Document On... feature, then imported the document.

After the entire book had imported, I then went to the HTML.edit Index page, created a new label called "Typee" and set all the imported documents to that label. I then created a table of contents document, and used the Table of Contents command (under the Anchor menu) to build a table of contents for the entire book. This entire process took less than ten minutes.

Back to Document Structure, on to Header, or return to Design Issues.