

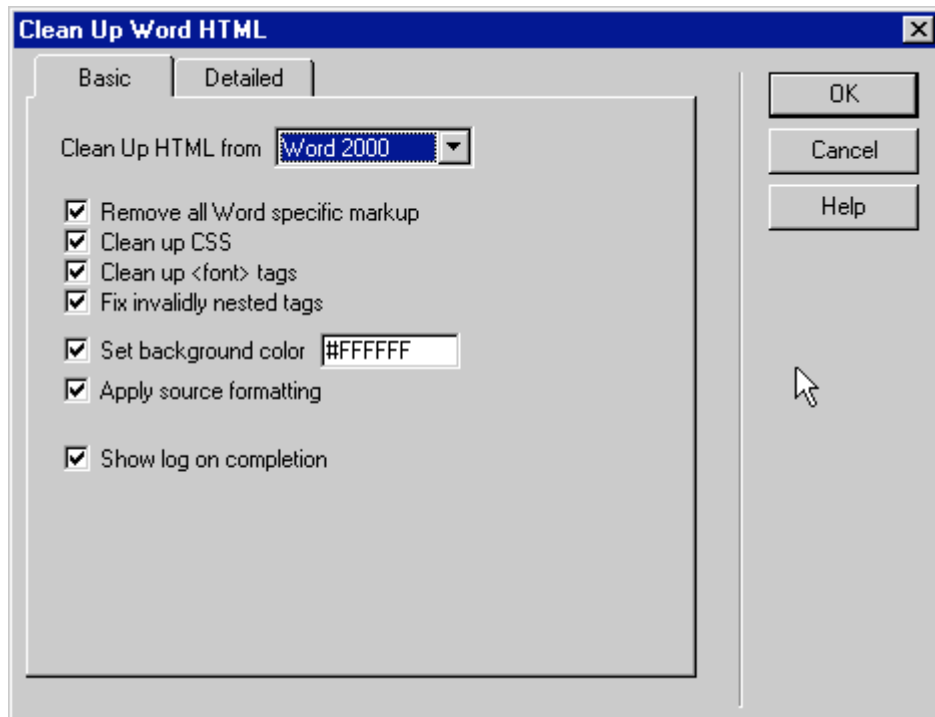
Cleaning up Word-produced HTML documents

Both WordPerfect and Word give you options to save existing word processing documents as HTML files (sometimes they simply call these web documents). Whether or not you want to pursue this depends on the documents and how they have been formatted. For example, "tables" done with tab settings and using the TAB key to move between items will have the text convert okay, but the table structure disappears. Text and information done as tables will normally convert okay. You are certainly welcome to pursue this feature, and open the resulting document in Dreamweaver and see what you think. My own bias is that normally it is less work to simply copy and paste text from the word processing document into your Dreamweaver document, particularly if you have some other features in the Dreamweaver document you want as part of the file (like navigation links or a common header at the top of the web document).

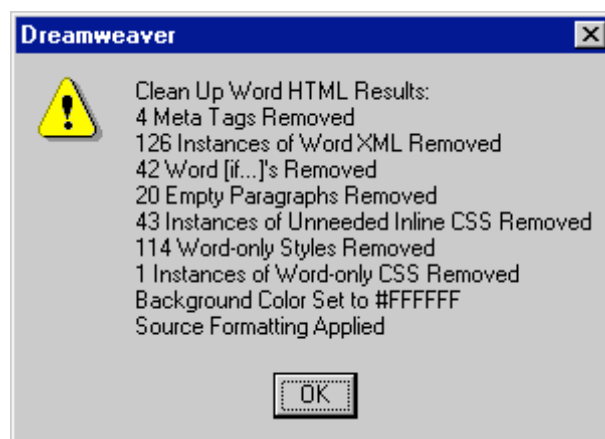
However, I did want to warn you that newer versions of Word when saving an HTML file produces horribly complex HTML code that is **larger** in kilobytes than the original Word document. Also beware that many web documents produced this way do NOT appear okay when it is posted to a server.

I experimented with a very simple Word document, of about two pages, with one line centered and in a larger font, and several paragraphs, but nothing elaborate, and saved it as an HTML document (web document). The original document was 10 kilobytes. When I made the HTML file through Word 2000, the resulting file was 20 kilobytes.

The HTML code produced by Word 2000 is so bad that Macromedia added a special option under the **Commands** menu item in the Document window - ***Clean Up Word HTML***. If this option is selected when you have open a Word-produced HTML file, this window appears:



Accepting these default settings, this is the report that Dreamweaver produced, to show things it changed.



The resulting HTML document dropped to a 13 kilobytes file. If this is the type of errors it found on a simple Word document saved as an HTML file, you can image what it might have to do to a longer document with more features.

If you take your Word documents and save them as HTML files through Word, be sure to use this "clean up" option in Dreamweaver.