

Microsoft Word to HTML Conversion Using GT_HTML

Introduction

Thanks for trying GT_HTML. GT_HTML is a set of macros for Microsoft Word for Windows 6.0 with a custom toolbar designed to let you easily create HTML documents without remembering a lot of HTML tags. Because all inserted tags have Word's hidden attribute, turning off the display of hidden text should give you a WYSIWYG display of your HTML document. Currently, only a small subset of HTML formatting tags are supported. However, the most common tags are supported, and the list is growing. Currently available features are outlined below.

We have found these macros to be useful for most simple HTML authoring, and we hope that others will too. However, we do not bond, warrant, or otherwise guarantee the suitability or functionality of this freeware product. Please direct any feedback, positive or negative to gt_html@gatech.edu.

Installation

Note: the following installation guide assumes that your executable Word for Windows file, winword.exe, is in the directory \WINWORD. You should, of course, use the directory where winword.exe is on your computer.

Copy the file GT_HTML.DOT to your \WINWORD\TEMPLATES directory. This will create a new document template called "GT_HTML". To activate the template, select Templates from the File menu. Select Attach in the dialog box, and then select GT_HTML. This template contains all the conversion macros and a redefined Toolbar. To activate the custom toolbars, select Toolbars from the View menu, and then check Toolbar 1 (GT_HTML.dot) and Toolbar 2 (GT_HTML.dot). You may need to reposition the toolbars to fit on your screen, but they should look something like this:



The template will automatically be activated when a document created with the template is opened. When working with previously created Word files that you want to convert, just open the document and follow the steps above.

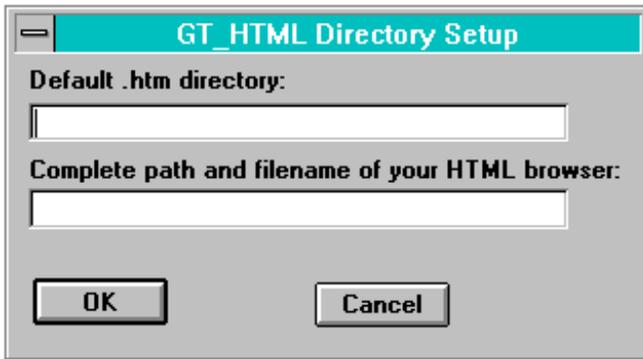
You should also copy the file SAMPLPIC.BMP to your \WINWORD directory. SAMPLPIC.BMP serves as a "place marker" for inline graphics. The file GT_HTML.CFG will automatically be placed in your \WINWORD directory when you run the SETUP macro. This file contains configuration information that several of the macros need, and should not be moved or deleted.

Getting Help for the Toolbar

Move the mouse pointer over the toolbar button of interest. A brief description of the button should appear at the bottom of the screen.

Setup

Before running any of the other macros, you should click on the **setup** button:  The following dialog box should appear:



In the first space, type in the pathname for the directory where you want to save your HTML documents. You should enter a complete, valid DOS pathname, including the drive letter. (We suggest saving both you original Word documents and your converted HTML documents in the same directory. When making changes to a document, edit the Word version, and then re-convert to HTML. This ensures that your Word “source” document and your HTML document are always current.)

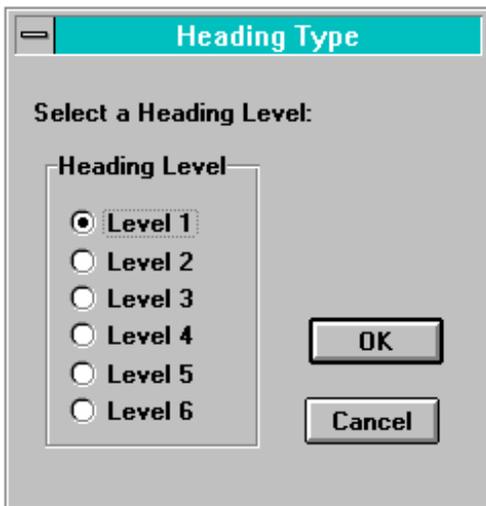
In the second space, type in the complete path and filename of your HTML browser. Currently, only MOSAIC and CELLO are supported. This will allow you to open your browser from word so that you can view your HTML documents.

Bold, Italic, and Underlined Text

Bold, italic, and underlined text should be marked using the standard Word formatting commands. They will be converted to HTML automatically.

Headings

Text to be converted to a heading should be selected with the mouse. Click on the **heading** button: . The following dialog box will appear, requesting you to select the heading level.



Make the appropriate selection, and click on "OK". The text will increase proportionately in size, and be surrounded by the appropriate HTML tags. The tags have the Hidden attribute.

Lists

Numbered and Bulleted Lists are supported. Text to be converted to a list should be selected with the mouse. Click either the **numlist** button:  or the **bullist** button:  on the Toolbar. Each paragraph in the selected text is converted to a list item. HTML list tags are inserted, and tags have the Hidden attribute.

Inserting a "Name" Tag

Name tags allow hypertext links to reference a specific point within your document. A name link should be a single word, and all name links within a document must be unique. To insert a Name tag, simply highlight the name link word and press the **name** button: . HTML name tags are inserted, and tags have the hidden attribute.

Hypertext Links

Hypertext links are created from selected text in the following way. The last continuous string of text in the selection (i.e. no spaces) is interpreted as the URL. All preceding text is interpreted as the hypertext link. For example, the word "here" in the following sentence is a hypertext link to <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/NCSAMosaicHome.html>, the NCSA Mosaic home page:

Click here for the NCSA Mosaic home page.

You would type:

Click here <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/NCSAMosaicHome.html> for the NCSA Mosaic home page.

And select with the mouse:

here <http://www.ncsa.uiuc.edu/SDG/Software/Mosaic/NCSAMosaicHome.html>

Finally, a click on the **urlink** button:  on the Toolbar will convert the selection to an HTML hypertext link. HTML tags will be inserted in red and will have the Hidden attribute.

Entering In-line Graphics

Highlight the graphic's filename and click on the **graphic** button:  on the Toolbar. The proper HTML tag will be inserted, and GT_HTML will attempt to find and insert the actual graphic you specified. If you included a pathname in your selection, GT_HTML will look there. If you included a filename only, GT_HTML will look in your default HTML directory, as defined by **setup**. If no matching graphic filename can be found, the "dummy" graphic SAMPLPIC.BMP will be inserted to approximate what the rendered document will look like. In order to insert .gif pictures into your document, you should be sure that you have installed Word's .gif format import feature. If you cannot import .gif files, see your Word documentation for information on how to install this feature.

Preformatted Text

To define a block of text as preformatted, select the block of text and click on the **pre**:  button. The proper HTML tags will be inserted, and the font will change to *courier*, to approximate the look of the rendered text.

Horizontal Rule

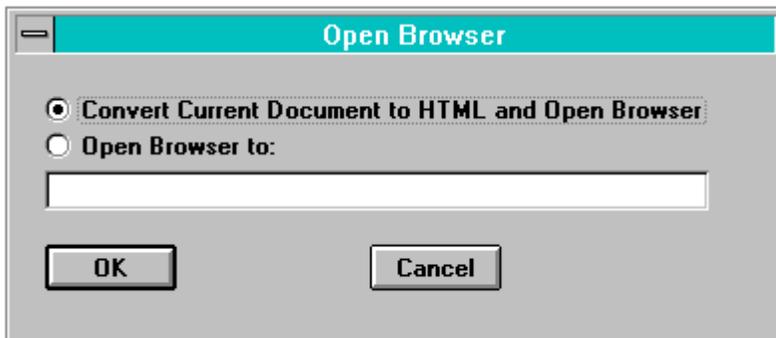
You can insert a horizontal line in your document by clicking the **hr** button: . The proper HTML tag will be inserted, and a horizontal line will appear in your document.

Opening Another Document

The **opendoc** button:  gives you an easy way to open other Word documents for editing. Highlight the name of the file, and press the opendoc button. If a complete pathname is included, Word will use that path to search for the file. If no pathname is included, Word will look first in your default HTML directory (see setup), and then in your /WINWORD directory. GT_HTML will always instruct word to search for the file with a .doc extension. This allows you to select a .htm filename and open the corresponding original Word document. For example if you select the filename *foo.htm*, Word will attempt to open the file *foo.doc* in your default HTML directory. If the file cannot be found there, Word will attempt to open the file *foo.doc* in your /WINWORD directory. If the file cannot be found, an error dialog will be displayed.

Opening a HTML Browser (Mosaic or Cello)

You can open your HTML browser from word (either Mosaic or Cello). A click on the browser button:  will cause the following dialog to be displayed:



If the top option is selected, clicking OK will convert your current document to HTML, save it in your default HTML directory, and then open your browser to that document.

You may also select the name of a url before clicking the **browser** button. This will cause the dialog to open with the second option selected, and a url in the space provided. If you have selected a complete url, the url will appear in the dialog. If you have selected only a filename, GT_HTML will assume you want to open your browser to a local file, and a url beginning with "file://" will appear in the dialog. If no pathname is provided, GT_HTML will assume the HTML file is in your default HTML directory. You should try the **browser** button a few times to see how it works.

The "Hide" Button

All of the HTML tags inserted by GT_HTML have Word's Hidden attribute. A click on the **hide** button:  will "turn off" any hidden text and paragraph and tab marks, giving you a somewhat WYSIWYG view of your document as it will be rendered. The **hide** button is a toggle, meaning that another click will turn the hidden text and marks back on.

Undo

Selecting the **undo** button:  will remove all HTML tags from the selected text.

Converting a Document to HTML

When you are ready to convert your document to HTML format, select the **htmlsave** button:  from the toolbar. **Your Word document will automatically be saved before making an HTML conversion!** A lot of file swapping goes on during the conversion process, and your changes could get lost without this automatic save. If you're not sure about current changes to your document, use Save As... to create a working copy before making the HTML conversion.

You will be prompted for a filename for the HTML document. The default is your default HTML directory as defined in **setup**, with the current filename and the extension ".htm". Simply press [Enter] to accept the default filename. Note: an intermediate form of the document is created in your HTML directory called "@HMLTEMP.DOC". Currently, this file is saved for debugging purposes. After the conversion is complete, you will be returned to the original document.

Notes on Potential Problems

"Nested" HTML tags may be a problem. Simple nested text formatting, for example text that is both bold and underlined, seems to work well. However, more complex nesting, like hypertext links in a list, have not been thoroughly tested. Nested lists, in particular, are not yet supported.