[WinWord 2.0 to HTML](#)

http://jch.vtkk.fi/default.htm

Package:

This package includes following files:

HTML.DOT	WinWord 2.0 template file including some styles for writing and some macros to handle document as HTML.
[RTFTOHTM.DLL](#)	The conversion DLL.
RTFTOHTM.LIB	The import library for the DLL.
RTFTOHTM.H	Public interface of the DLL.
HTMTOOLS.DOC	Some documentation about the converter (this file).
HTMTOOLS.HTM	Documentation converted to .HTM.
RTF2HTM.EXE	QuickWin application for converting files.
RTF2HTM.C	Source for above.
RTF2HTM.MAK	Makefile for above.
README.TXT	Quick document.

There is couple words about [installation](#) at the end of this document.
<P><HR><P>

Document template

HTML.DOT includes some HTML styles. This document template is not meant to be perfect HTML editing system, but I found it quite useful with companion of the conversion dll; RTFTOHTM.DLL. Here are some features offered by this software.
<P><HR><P>

Document HEADer

There is possible to use page header to generate <HEAD> section to html file. There is no error checking on it, so the feature should be used correctly by convention. There should be only one header in the document. You can create such by View_Header/Footer. In the Header/Footer dialog You should check 'Different First Page', and select Header. This makes 'headerf' style to the beginning of the first page.

In the header You should put only TITLE and LINK styles. You can make a paragraph to TITLE by selecting it, and clicking **T** button in the button bar or selecting a style **Title** from style list in ribbon. You can make LINKs into header by selecting text and setting its style to Link.

If you define a header into document, converter creates both HEAD and BODY sections to resulting HTML file.
<P><HR><P>

Heading 1 to Heading 6

Paragraph to be converted to a heading should be selected. It is enough if a caret is blinking within a paragraph. Then either clicking **E** button in the button bar or selecting an appropriate style from a style list.
<P><HR><P>

Emphasizing

There are some emphasis inline styles. You can use **Bold**, *Italics*, <u>Underline</u>, <u>Word</u> <u>Underline</u> by appropriate WinWord emphasis. The interpretation of **Bold** is <B>, *Italics* is <I>, <u>Underline</u> is <U>, and <u>Word</u> <u>Underline</u> is <EM>.
<P><HR><P>

Block emphasizing

There are three types of block emphasizing styles:

This is a preformatted text as HTML style <PRE>.
This style is applied by selecting this text, and clicking **P** button in the button bar.
This style can contain other HTML styles, such as anchors, emphasizing inline styles...
<P><HR><P>

> *This style is BlockQuote. This is intended to quote some larger block of text, and is emphasized in some way. This style is applied by selecting text, and clicking **B** button in the button bar.*

<P><HR><P>

This is an Address style. These styles are presented differently in different parsers. You can correct a presentation in WinWord simply by formatting style by your own taste. Paragraph written in this style is aligned right.

<P><HR><P>

Lists

There are two types of lists available. Ordered list as style Order List, and bulleted list as style Unnum List. These types You can attach to text by selecting text, and clicking a number list or bullet list button in the button bar. Each paragraph within selected text becomes a list item. The converter doesn't allow nesting these styles. If you want to nest them, you must use HTML style (below) and write a markup yourself.

This is a numbered list item one
This is item two
And item three
Style is Order List

This is a bulleted list item one
This is item two
And item three
Style is Unnum List
<P><HR><P>

HTML style

Those <P><HR><P> patterns (visible in original WinWord document) are horizontal lines, and must be a style of HTML. The style HTML allows user to embed 'plain' HTML into Word document. The converter doesn't touch to this style in any way. The style HTML allows user to enter such HTML tags or styles this converter doesn't support in other way. Here is an example.
<P><HR><P>
<FORM ACTION="http://jch.vtkk.fi/testpost" METHOD="POST">
(E-mail to:<SELECT NAME="mailinfo" SIZE=1><OPTION
SELECTED>Jorma.Hartikka@csc.fi<OPTION>jhartik@finsun.csc.fi<OPTION>Xhartik@finsun.csc.fi</SELECT>
Subject:<SELECT NAME="subjectinfo" SIZE=3><OPTION SELECTED>Testing 1...
<OPTION>Testing 2...<OPTION>Testing 3...<OPTION>Testing 4...</SELECT>)
<P><HR><P>
Name: <INPUT NAME="username" VALUE="Testusername" SIZE=30><P>
E-Mail: <INPUT NAME="usermail" VALUE="Testusermail" SIZE=40>
<PRE></PRE>
Please enter any comments on whether or not you found this server useful, and how it could be improved.<P>
<INPUT NAME="testfeedback" SIZE=60,4 VALUE="Testfeedback"><P>
Please push this <INPUT TYPE="submit" VALUE="S&ubmit"> to mail your feedback.
</FORM>
<P><HR><P>

Hypertext links

Hypertext links has specific syntax in this template, but nothing denies user to make his own... The styleurl is composed as follows: There is an url separated by space from following text. An url part of the text becomes hidden red, and button part becomes double underlined blue.

http://www.ncsa.uiuc.edu/demoweb/demo.html This is a demo page of NCSA.

By selecting whole url including a button text and clicking **L** button in the button bar makes text as following:

[This is a demo page of NCSA](http://www.ncsa.uiuc.edu/demoweb/demo.html).

You can also define an hypertext link pointing to specific part of current or exterrnal document by deffining a name in the link. Here is an example of that kind of link. Only difference to the 'ordinary' link is that a position within destination document is present.

[This is a test jump to the anchor marked 'testjump'](#).

The anchor can be made by defining a name and label. This can be made by selecting an anchor text and pressing D button in the button bar.

testjump [This is a destination for above 'testjump'](#).

Notice that although I use SmallCaps style for a name of the anchor, the case of the characters is preserved in the resulting HTML file.

`<P><HR><P>`

Inline images

Inline graphics is quite problematic in WYSIWYG point of view. I have made it in easy way (for now). It is presented simply as strike through formatting in WinWord. This allows its use as a link button in the hypertext link.

This is stand alone image: ~~JCHS.GIF~~

Here is a button image: ~~JCHS.GIF~~

Inline image can act as button. Converter cannot generate following:

<A HREF=anchor><IMG SRC=jchs.gif> The Author</A>

A button field of the hyperlink can contain either an image or plain text but not both.
`<P><HR><P>`

Undoing HTML

N button in the button bar resets style of the selected paragraph(s) back to normal. **U** button removes some character formatting properties from the selected text. It can be used to reset hyperlinks or image links back to normal.
`<P><HR><P>`

Saving a file

You can save a file as *.HTM by pressing S button in the button bar. File can be saved as *.RTF by pressing R button. Saved file gives a same name as current WinWord document and an extension .HTM or .RTF depending a save format.

Fine tuning of a resulted HTML file

The resulted HTML file should be edited with plain text editor to clean it. WinWord 2.0 RTF filter may do nasty breaks within anchors, button text, or inline, if it spans certain memory limits, so they should be corrected before publishing a file.
`<P><HR><P>`

Installation **bmk2** .

Copy HTML.DOT into WinWord directory, and RTFTOHTM.DLL somewhere in PATH (eg. WINWORD or WINDOWS).
`<P><HR><P>`

RTFTOHTM.DLL **bmk1** .

There is one exported function in the converter DLL. C-prototype is as follows:

long FAR PASCAL RTFtoHTM(LPSTR iname, int ilen, LPSTR oname, int olen);
LPSTR   ibuff;   //input buffer or filename.
int   ilen;   //length of input buffer. if zero, iname is treated as filename.
LPSTR   obuff;   //output buffer or filename.
int   olen;   //length of output buffer. if zero, oname is treated as filename.

The return value represents a length (in chars) of a resulted HTML file or return buffer.
<P><HR><P>

Basic (or Word macro) declaration is as follows:
Declare Function RTFtoHTM Lib "rtftohtm.dll" (iname$, ilen As Integer, oname$, olen As Integer) As Integer

The meaning of parameters are the same as above. If the function is to be used with Visual Basic, all the
parameters must be declared ByVal.
<P><HR><P>

Name: Jorma Hartikka Senior Programmer Analyst
Org: VTKK Government Systems ltd/Information Systems
Email: Jorma.Hartikka@csc.fi

<P><HR><P>