

# AUTOMATICALLY CONVERTING PREFORMATTED TEXT INTO REFLOWABLE TEXT FOR TV VIEWING

## BACKGROUND OF THE INVENTION

### 1. Technical Field

The present invention relates in general to displaying HTML data on data processing system display devices and in particular to displaying HTML data on a television connected to a set-top box data processing system. Still more particularly, the present invention relates to selectively converting preformatted HTML text into reflowable text for display on a television connected to a set-top box data processing system.

### 2. Description of the Related Art

Many contemporary data processing system applications employ a standard page description language, the Hypertext Markup Language (HTML), for data transfers to other data processing systems or applications. HTML data comprises standard ASCII text documents in which there are embedded "tags," text enclosed within brackets (<>). Normally, tags are utilized in pairs within the document, with a slash (/) within the second tag signifying the end of the desired formatting. For example, text contained between the tags "<B>" and "</B>" will be displayed in a boldfaced font by an HTML-compatible application.

HTML documents may be viewed as simple text documents, in which case the tags are visible, or using an HTML-compatible application, which will not display the tags. HTML-compatible applications recognize tags as structure or formatting commands, references to other elements such as graphical, motion video, or audio data, or links to other HTML documents. Thus, an HTML-compatible application will not display the tags, but will utilize the tags to format the display of other text within the HTML document.

Most of the time, any extra "white space" from spaces, tabs, or hard returns in an HTML document is stripped out by any HTML-compatible application displaying the document, such as a World Wide Web (WWW or "Web") browser. For example, multiple sequential spaces within an HTML document are collapsed and displayed as a single space by Web browsers. An exception to this elimination of extra white space results from the preformatted text tag <PRE>. Any extra white space within text surrounded by the tags <PRE> and </PRE> is retained in the final output. This allows text within an HTML document to be presented in the same formatting, with all original white space, when viewed with an HTML-compatible application as when viewed using a simple text editor.

The preformatted text tag may be utilized to indent or format lines, and thus may be employed for things such as code examples. Furthermore, text may be aligned by padding it with white spaces, such that preformatted text may be utilized to create tables in the final display. Such use of the preformatted text tag was common with HTML documents created for early versions of HTML, which did not support tables. Often, however, low-end document editors utilize the preformatted text tag for all HTML documents created with that editor. Preformatted text tags are also often used to quickly and easily convert files which were originally in some sort of text-only form to HTML documents.

For set-top box systems, preformatted text may have an undesirable effect on display of HTML documents. Television screens, the display device for set-top box systems are

typically designed, are usually narrower than the 80 character computer screens utilized with conventional data processing systems. Preformatted text within HTML documents does not "reflow," or naturally break lines of text into the maximum line length which may be displayed for a given display device. Because preformatted text is not reflowed to fit specific display devices, lines of text often will not fit a television screen. The user must therefore scroll the display to the right to view the end of a line of text, then scroll back to the left to view the beginning of the next line. This renders the user interface extremely annoying, and generally impedes a user's ability to read the HTML document.

HTML documents are one specific example of a class of information media which may generally be referred to a "compound" documents. Compound documents include text and other features, such as graphics, tables or spreadsheets. Furthermore, other markup languages are widely used, such as GML and SGML. Such other markup languages may have differently preformatted text tags or employ different approaches to precluding reflow of selected text blocks.

It would be desirable, therefore, to intelligently reflow text within compound document to fit a display screen or window while leaving special features such as graphics or tables within the document undisturbed. As a specific example, it would be desirable to improve the readability of HTML documents incorporating preformatted text tags, eliminating any unnecessary need to scroll the display of a user interface to the right and left for each line of text in the display. It would also be desirable to reflow preformatted text within an HTML document where the nature of the text renders such alteration appropriate.

## SUMMARY OF THE INVENTION

It is therefore one object of the present invention to provide a method for selectively reflowing text within a compound document to fit a display area while leaving specially formatted features of the document undisturbed.

It is another object of the present invention to provide an improved method for displaying HTML data on data processing system display devices.

It is another object of the present invention to provide an improved method for displaying HTML data on a television connected to a set-top box data processing system.

It is yet another object of the present invention to provide a method and apparatus for selectively converting preformatted HTML text into reflowable text for display on a television connected to a set-top box data processing system.

The foregoing objects are achieved as is now described. Preformatted text within HTML documents is examined to determine if reflowing the text during formatting for display within a narrow viewing area is appropriate. This determination is based on various special formatting indicia which indicate that the text is a table or other informational device in which lines of text should not be reflowed. If no special formatting indicia are detected, the preformatted text tags are removed, allowing the preformatted text to be reflowed when formatted for display. Otherwise, the preformatted text is left alone or revised according to a user preference. Preformatted text, when appropriately reflowed, is thus rendered more readable and the necessity for scrolling right and left is eliminated.

The above as well as additional objects, features, and advantages of the present invention will become apparent in the following detailed written description.

## BRIEF DESCRIPTION OF THE DRAWINGS

The novel features believed characteristic of the invention are set forth in the appended claims. The invention itself