

RTFSyntax

by

Ross Cutler

rgc@cs.umd.edu

1-6-94

Introduction

RTFSyntax is an application for programmers that provides a service to color syntax highlight rich text. By color syntax highlighting, I mean automatically changing the color and font of the selected rich text using regular expressions previously defined in **RTFSyntax**. For example, you can use this program to make all of your Objective-C comments appear in red Helvetica 12-point italics. Because you define what you want colored using regular expression, you can completely customize how your RTF code looks.

Motivation

With NEXTSTEP 3.2, Edit and cc(1) now support RTF. You can utilize RTF in your source files by creating links from one file to another, and also by changing the fonts and colors to make the code easier to read. This program automates the latter.

Usage

To use **RTFSyntax**, you must first specify a default configuration file in the preferences. A configuration file contains a list of regular expressions and the corresponding font and color you wish to use for the matching text. A sample configuration file *Objective-C.syntax* is provided with this program. However, you can create your own easily enough by using the Configuration window. The usage is very straightforward, and the only thing that needs commenting is the regular expression syntax.

GNU's regex (version 0.12) is used to do the regular expression matching. The syntax used is extended POSIX, with the "." character matching anything but newline. For ease of entering the regular expression, the newline character is ignored; it can be entered using "\n" as you'd expect. Note that non-matching lists (e.g. [^a]) do not match newlines. Please see the GNU regex documentation (included as the file *regex.dvi*) for further information about the extended POSIX syntax.

To use **RTFSyntax**, simply highlight the RTF text you wish to color and select the "Color Syntax" service (Command-`). Note that you do not have to color the entire file at once; for speed reasons, you probably should only color the text that needs coloring. Also, in Edit, make sure you convert an ASCII file to RTF (Command-R) before you use **RTFSyntax**.

Comments

If two regular expressions find matches that intersect, the match that occurs earliest in the file is used. If they both start at the same offset in the file, then the longer one is used.

Unfortunately, IB cannot parse RTF header files. This should be fixed in the next release of NS.

RTFSyntax was implemented using the Text class to highlight text. Though this is very clean and was easy to do, it does have significant overhead. For large source files (~1000 lines), it can take about 60 seconds to highlight the entire file (using *Objective-C.syntax* on a ALR Pentium computer). I found that about 70% of the program's CPU time was being used in 4 lines of Text method calls. The other 30% was taken up by calls to regexexec().

Future Directions

The obvious enhancement to **RTFSyntax** would be increased speed. One could make it at least twice as fast by not using the Text class and modifying the RTF code directly. Also, for fixed regular expressions, compiled code generated from lex(1) could be used (ack!). I provide the source for anyone who wishes to do this.

Ideally, Edit would have instantaneous color syntax highlighting. Possibly one could subclass the Text class to provide this feature, and then graft that subclass into Edit. Please let me know if you have any ideas on how to go about doing this.

Legal Stuff

RTFSyntax is released under the GNU General Public License.