PTC® MKS Toolkit®

Text Processing

One of the basic functions of the PTC MKS Toolkit is text processing. Since most PTC MKS Toolkit utilities use text files in some way, creating and manipulating such files is a useful skill when creating almost any kind of PTC MKS Toolkit-based solution.

PTC MKS Toolkit provides several text processing utilities to handle any variety of text processing jobs, such as editors like the fully interactive vi editor, simple text formatting utilities like c, banner, or fmt, and complex formatting systems like groff and its related utilities.

The vi Screen Editor

vi is a complex and powerful text editing tool. From simple text entry to complex text processing tasks like macros, vi can handle it all. You can even call other PTC MKS Toolkit commands from within vi to perform their actions on a selected portion of text.

To help you in learning how to use vi, the ROOTDIR/sample/guide directory contains an interactive tutorial in the files browse.v, edit.v, doc.v, and program.v. To use the tutorial, simply view the files in the order listed using vi itself.

In addition to the standard vi editor, PTC MKS Toolkit also includes vi for Windows (viw), which adds Windows functionality such as mouse-controlled cut-and-paste, scrollbars, and printing to vi's standard features.

The sed Utility

When you need to perform the same editing task (such as replacing one text string with another) on multiple files, a utility like sed comes in handy.

The sed utility is a non-interactive stream editor. You do not use it in an interactive session like vi, but rather provide it with a list of editing commands either in a file or on the command line, and sed then provides those commands on a specified group of text files.

Searching for Text

Often, when dealing with a large number of text files, it's useful to be able to quickly see which files contain a specified string of text or a string of text that matches a specified pattern known as a regular expression. For such purposes, PTC MKS Toolkit contains the grep family of utilities.

The fgrep utility searches for lines in files that contain a specified string of text; the egrep utility searches for lines in files that contain a string that matches an enhanced regular expression; and the grep utility, by default, searches for lines in files that contain a string that matches a basic regular expressions. With the use of options, these utilities can simply report a list of files that contain matching strings or the number of lines in each file that contain matches. Finally, the Ultragrep graphical utility (ugrep) lets you search a selection of files that match a given pattern for lines that contain a string that matches a regular expression.

Simple Text Formatting Utilities

PTC MKS Toolkit contains several utilities designed to provide text formatting capabilities including:

- The c utility arranges text in a file into a multiple columns. The shorter the lines, the more columns c can create.
- The banner utility takes a string of text and creates large versions of the characters in the string using x's to draw the characters. This utility can be used to produce small signs or posters.
- The fmt utility is a simple text formatter that can format a text file into tidy paragraph. It has the option of creating different styles of paragraphs and, if desired, justifying lines so that they are the same length. fmt is useful for formatting short notes or e-mail messages.
- The fold utility breaks lines in a text file at a specified point or at the last blank before that point. fold can make a file that contains lines longer than the screen width more readable.
- The expand utility replaces tabs in a text file with an appropriate number of spaces to preserve the file's formatting. The unexpand utility replaces spaces with tabs in a similar manner.
- The nl utility adds line numbers to each line of a text file.
- The pr utility formats a text file with pagination and adds headers and footers.

Displaying Text Files

PTC MKS Toolkit contains several utilities for displaying text files including:

- The cat utility simply displays a text file without any breaks. cat is most useful for short files.
- The more and pg utilities display a text file one screenful at a time, pausing for input at the end of each screen. Both utilities also provide the capability to search for a text string.
- The head and tail utilities display a specified number of lines at the beginning and end of text files, respectively. This is useful when you only want to see a certain number of lines to identify the file's contents.

The groff Formatting System

The GNU groff formatting system consists of the groff utility and its related utilities. It takes a text file containing special mark-up commands and produces output in a variety formats such as PostScript® and HTML. In addition to the standard formatting commands available, groff also lets you define macro packages that let you simplify the mark-up of your text file by assigning groups of commands to a single new command. The groff formatting system includes several pre-defined macro packages for purposes such as writing man pages, papers, and memos. These groff formatting system (including groff, the related utilities, and the pre-defined provided as samples in the samples directory on the MKS Toolkit CD.

Other Text Manipulation Utilities

In addition to the utilities already described, PTC MKS Toolkit contains a variety of utilities that manipulate text files:

- The wc command displays the number of lines, words, and characters in a text file.
- The flip command converts UNIX-style text file to Window-style and vice versa.
- The csplit and split utilities break a text file into smaller text files. csplit uses specified criteria to determine where the file is split, while split simply breaks the file into smaller files of a given number of lines.
- The sort utility sorts the lines in a text file based on specified criteria.
- The spell utility spell-checks a text file using both a dictionary provided with MKS Toolkit and user dictionaries.

- The tr utility translates characters. Two strings are provided on the command line. When a character in the input file occurs in the first string; it is replaced by the corresponding character in the second string.
- The cut utility extracts specified fields or characters from each line of the input file.
- The paste utility concatenates corresponding lines from the specified input files.