



GtkSort

User Manual

Release 0.2.2

28 June 2007

Copyright © 2007 Lucas Tsatiris.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

Table of Contents

1.0.0	Introduction
1.1.0	What GtkSort Can Do
2.0.0	The GtkSort GUI And CLI
2.1.0	Working With The GtkSort GUI
2.1.1	Parameter File
2.1.2	Input File Name
2.1.3	Output File Name
2.1.4	FLR Files
2.1.5	Defining Keys For FLR Files
2.1.6	VLR Files
2.1.7	Defining Keys For VLR Files
2.1.8	Properties
2.1.9	Memory Limit
2.1.10	L2 Cache
2.1.11	DCache
2.1.12	Single Threaded GUI
2.1.13	Impose QuickSort
2.1.14	Temp. Directory
2.1.15	Sort! Button
2.1.16	Help Button
2.1.17	About Button
2.1.18	Quit Button
2.1.19	Preview Buttons
2.2.0	Working with the GtkSort CLI
3.0.0	Error Messages
4.0.0	Additional Information
	GNU Free Documentation License

1.0.0 Introduction

GtkSort is a multiprocessor external disk sorter and data manipulator for systems that support GTK+. Its source code has been ported and tested to produce valid results in Linux/x86, Linux/amd64, Linux/alpha, HP-UX/hp-pa 11.11i, Tru64 5.1B, Solaris 10/x86 and Win32 NT Class operating systems.

GtkSort processes files using multi threading in order to implement parallel algorithms. It overlaps disk I/O with sorting and reduces I/O waits. By using only sequential files it utilizes the most of the disk's I/O bandwidth. By giving exclusive read or write permissions for each processed file to only one thread, it minimizes the conflicting I/O requests that reduce the hard disk's efficiency. In order to sort records, and depending on the data type of the sort keys, GtkSort uses the standard library quick sort or its own implementation of Most Significant Digit (MSD) radix sort. GtkSort uses cache efficient algorithms keeping the cache misses at a very low rate (0.4% in version 0.2.0).

GtkSort integrates a Graphical User Interface (GUI) based on GTK+ in order to increase its friendliness against the end user. It also integrates a Command Line Interface (CLI) so it can be used in shell scripts.

GtkSort is free and open source software distributed under the terms of the GNU Public License version 2.

1.1.0 What GtkSort Can Do

- Sort large data sets much faster than the standard sort utility of the operating system.
- Sort by binary data or text data keys.
- Sort on a limited number of twelve keys.
- Support Fixed Length Record (FLR) text or binary files with fixed size keys.
- Support Variable Length Record (VLR) text files (lines of text) with delimited or fixed size keys.
- Use multiple processors and disks in parallel.
- Perform ascending or descending sort of each key.
- Preview the keys of the unsorted input and the sorted output.
- Use memory dynamically according to the limitations set by the user.
- Exploit CPU and file system specific characteristics such as the L2 cache size, the D cache size and the disk I/O block size.

The future releases will support the following:

- Multiple input files.
- More key data types.
- Unlimited number of sort keys.
- Merge multiple sorted input files into one sorted file.
- Split the sorted output in multiple files.
- Remove duplicate records from the sorted output file.
- Format the sorted output file.

- Sort Description Language (SDL).
- Distributed parallel sort among multiple computers over a network.

2.0.0 The GtkSort GUI And CLI

To invoke GtkSort, you can either use the GUI or the CLI. This chapter will describe the functionality of the GUI with references to the corresponding CLI parameters. Finally it will describe the CLI specific options.

2.1.0 Working With The GtkSort GUI

Launching GtkSort the following window will appear on the screen:

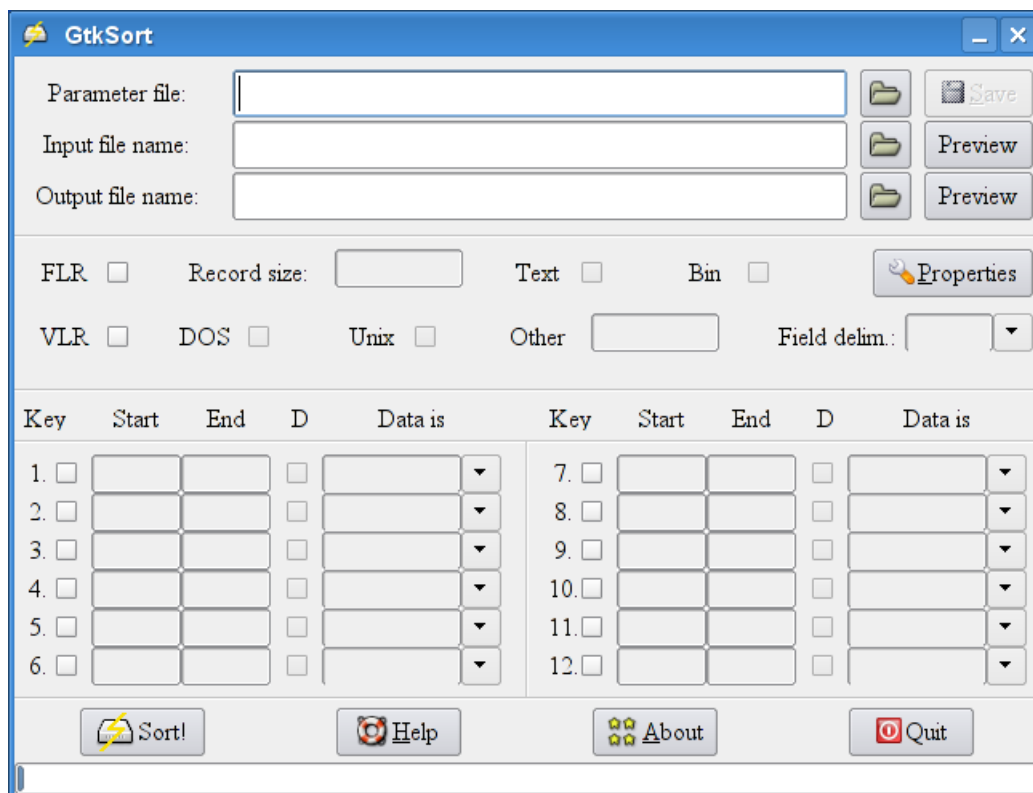



Fig. 2.1.0.1

2.1.1 Parameter File

The parameter file stores all the parameters needed by GtkSort to perform a single sort task. The parameter file can be used in later time in order to avoid the retyping and reselecting when you want to execute the same task. After typing the desired filename for the parameter file, with its full path if necessary, in the text box following the *Parameter file* label, the *Save* button will be enabled. By pressing the *Save* button all the contents of the window and the options that have been selected using the *Properties* button will be saved on the disk.

In order to load an already saved parameter file, press the file chooser button  on the right of the text box. A file chooser window will reveal (Fig. 2.1.1.1) and the desired file can be selected by double clicking on its name.

CLI parameter: -P <parameter file name>

NOTE: Parameter files cannot be created from the CLI. The -P parameter is useful only to load an already created parameter file and execute a single sort task.

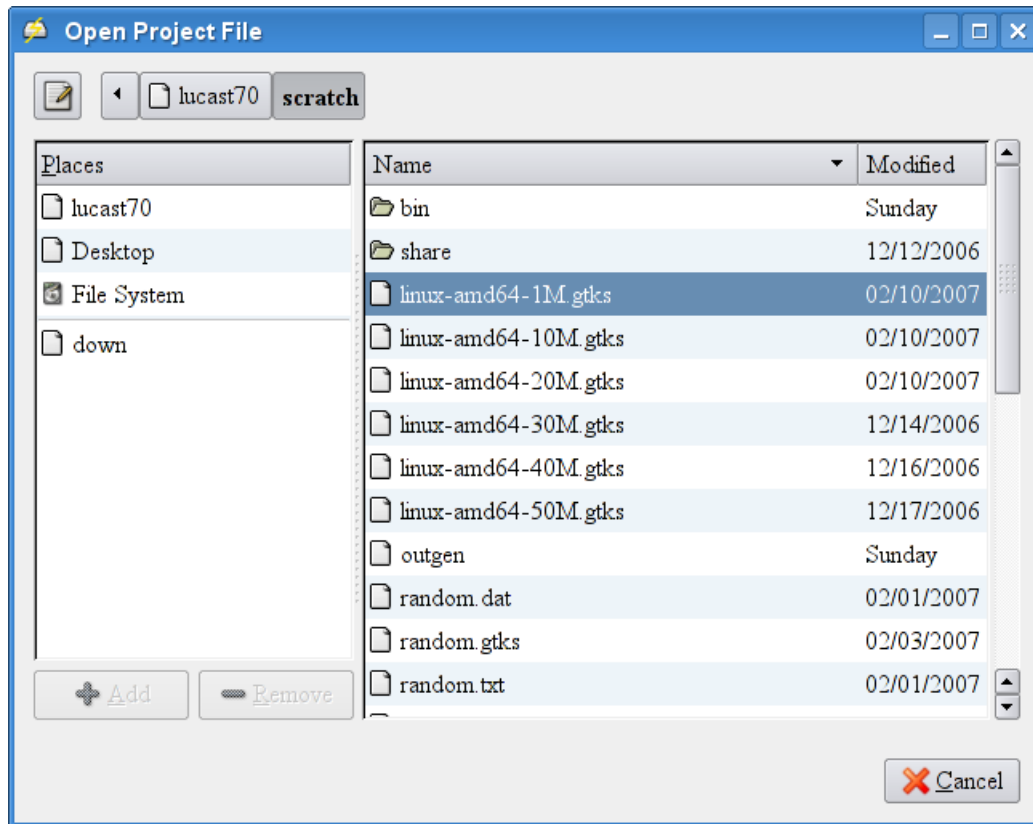


Fig. 2.1.1.1

2.1.2 Input File Name

The *Input file name* text box accepts the filename of the file to be sorted, with its full path if its necessary. Alternatively the file chooser button on the left of the text box can be used as described in **2.1.1** for the *Parameter file*.

CLI parameter: -i <input file name>

2.1.3 Output File Name

The *Output file name* text box accepts the file name that will store the sorted output of the *Input file name*. Alternatively the file chooser button on the left of the text box can be used as described in **2.1.1** for the *Parameter file*. If the file exists, it will be truncated and overwritten.

CLI parameter: -o <output file name>

2.1.4 FLR Files

The check box on the right of the *FLR* label must be checked if the input file is consisted by fixed length records. The *Record size* text box and the *Text* and *Bin* check boxes will be enabled. The *Record size* text box accepts an integer that represents the size of each record in bytes. The *Text* check box forces GtkSort to treat the input file as a text file. Text keys will be treated as NULL

terminated C-strings with maximum size the one determined by their offsets. The *Bin* check box forces GtkSort to treat the input file as a binary file. Text keys will be treated as sequences of bytes and no termination character will be considered. *Text* and *Bin* check boxes are mutual exclusive.

CLI parameters:

FLR File: -f
Record Size: -z nnnn
Binary: -b
Text: -t

2.1.5 Defining Keys For FLR Files

A key is enabled if the check box on the right of its sequence number is checked. The sequence numbers are under the *Key* label. An enabled key has the corresponding text boxes under the *Start* and *End* labels, the check box under the *D* label and the combo box under the *Data is* label also enabled. The text boxes under the *Start* and *End* labels accept integers that represent the starting and the ending offsets of the key, counting from 1. If the check box under the *D* label is checked, the key will be sorted in descending (highest-to-lowest) order. The combo box under the label *Data is* includes the supported data types a key may have. The supported data types are the following:

- *Text*: The key is plain text and will be sorted according to the 8-bit ASCII collating sequence.
- *Numtext*: The key is numeric text (numeric string) and will be sorted according to its arithmetic value.
- *Int16* and *Uint16*: The key is 16 bit signed or unsigned integer.
- *Int32* and *Uint32*: The key is 32 bit signed or unsigned integer.
- *Int64* and *Uint64*: The key is 64 bit signed or unsigned integer.
- *Float*: The key is 32 bit float.
- *Double*: The key is 64 bit float.

The keys will be processed according to their sequence number. The definitions of the disabled keys will not be saved in the *Parameter file*. Fig. 2.1.5.1 shows an example of how GtkSort sorts a text FLR file with two sort keys, one *Numtext* ascending and one *Text* descending.

CLI parameter: -k [+,-]nnnn,nnnn,[t,n,i16,u16,i32,u32,i64,u64,f32,f64]

The -k parameter can be used as many times as the number of keys. The sign in front of the start offset indicates the sort order that is + for ascending and - for descending. If no sign is given the + sign is implicit.

2.1.6 VLR Files

The check box on the right of the *VLR* label must be checked if the input file is consisted by records of variable length. The VLR check box will enable *DOS* and *Unix* check boxes, the *Other* text box and the *Field delim* combo box. GtkSort supports only text *VLR* files, that means lines of text. Binary *VLR* files are not supported yet. The *DOS* check box informs GtkSort that each line of the input file is terminated by the *CR* and *LF* characters (ASCII 13 and ASCII 10), which

is the MS-DOS style for the line termination. On the other hand, the *Unix* check box informs GtkSort that each line is terminated by the LF character only, which is the Unix style for the line termination. The *Other* text box has no functionality yet and can be ignored by the user. *DOS* and *Unix* check boxes are mutual exclusive. The *Field delim* combo box includes all the valid characters that can be used as field delimiters. Using the *Field delim* combo the user can choose one of the following:

None - *Space* - *Tab* - semi column (;) - pipe (|) - comma (,)

NOTE: GtkSort supports variable length records up to 4096 bytes.

CLI parameters:

VLR File: -v

DOS: -D

Unix: -U

Field delim.: -d [none, space, tab, semicolon, pipe, comma]

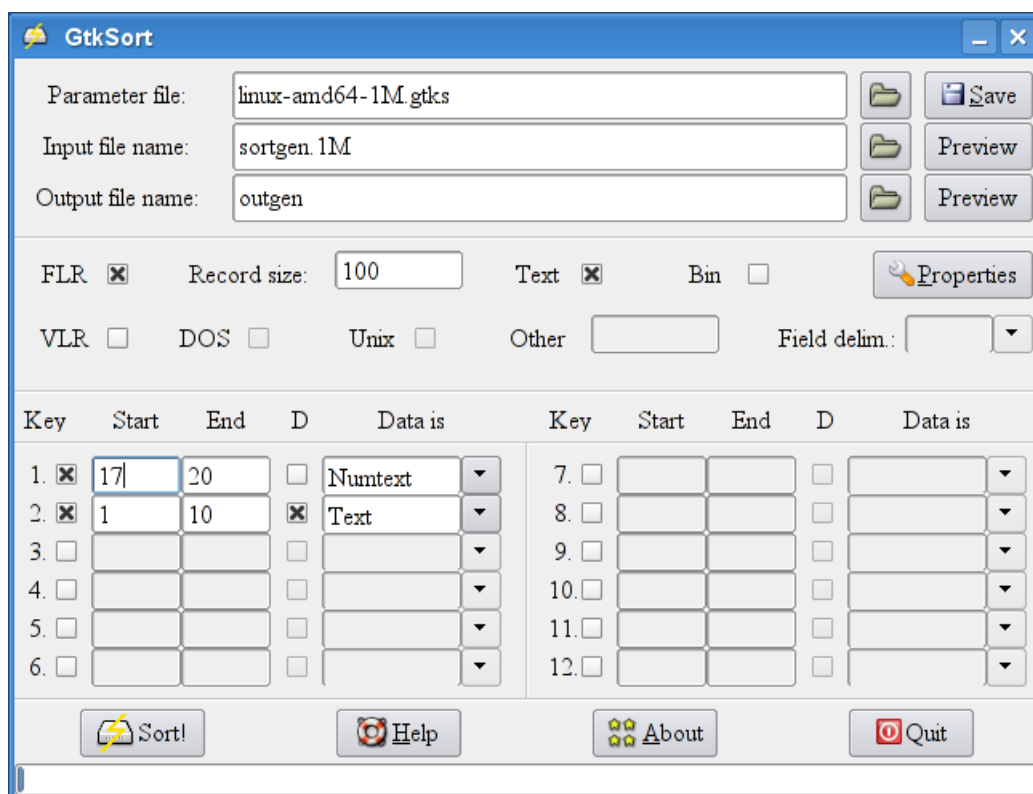


Fig. 2.1.5.1

2.1.7 Defining Keys For VLR Files

A key is enabled if the check box on the right of its sequence number is checked. Refer to **2.1.5** for more details about sort keys. If the *Field delim* combo has been set to *None*, GtkSort will expect the starting and ending offsets for each key as described in **2.1.5**. If the *Field delim* combo has been set to any other value, the *Start* label will change to *Column*, the *End* label will disappear and the text boxes below it will be disabled. In the *Column* text box, GtkSort accepts the number of the column it will use as a sort key, as it is

defined by the field delimiter character. VLR files are text only files. That means a key may have only the *Text* and *Numtext* data types. Choosing any other data type as the type of the sort key, will result an error before the execution. Fig. 2.1.7.1 shows an example of how GtkSort sorts with two keys a Unix VLR file with fields (columns) delimited by comma.

CLI parameter: -c [+,-]nnnn,[t,n]

The -c parameter can be used as many times as the number of keys. The sign in front of the column number indicates the sort order that is + for ascending and – for descending. If no sign is given the + sign is implicit.

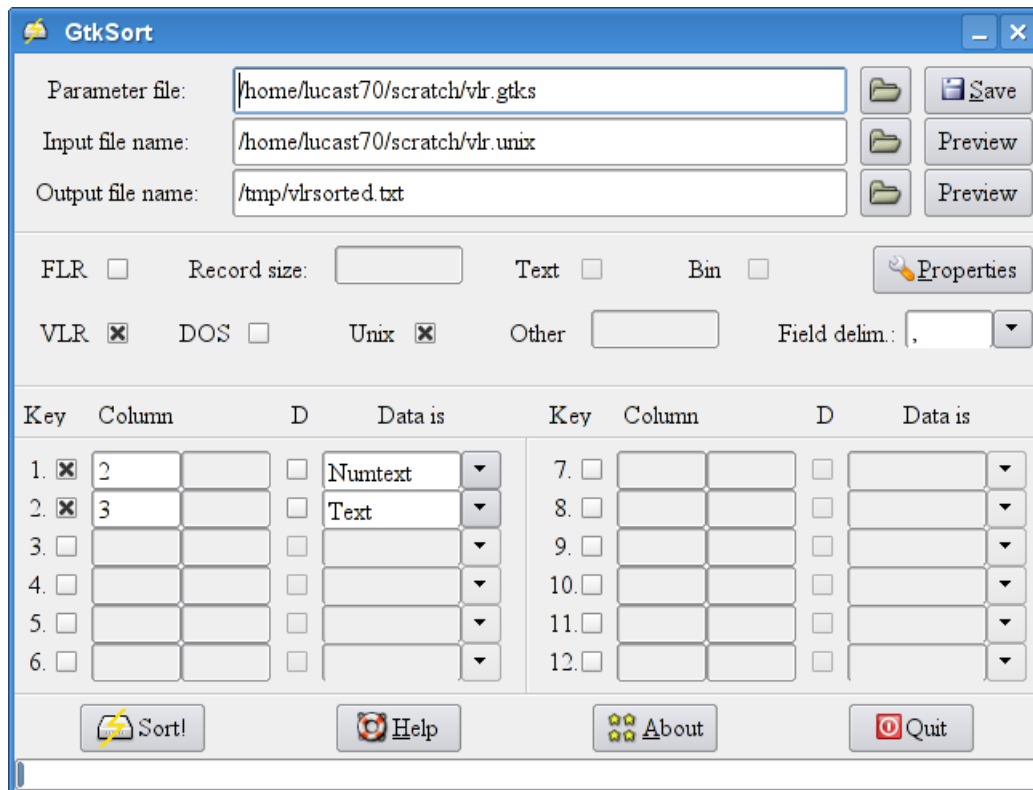


Fig. 2.1.7.1

2.1.8 Properties

Pressing the *Properties* button the properties window will appear (Fig. 2.1.8.1). The properties determine the performance and the resource consumption of GtkSort. Very conservative or very aggressive values may degrade the performance or the responsiveness or both. The default values will result an acceptable attitude in the most of the cases. Some tuning may needed in order to fit the current system's specifications or state in order to achieve the maximum performance.

2.1.9 Memory Limit

The *Memory Limit* text box is enabled when the check box on the left of the *Memory Limit* label is checked. If the check box is not checked GtkSort will use the half of the physical RAM for sorting and merging. The *Memory Limit* text box accepts the amount of memory that is available to GtkSort for sorting and merging translated to kilobytes or megabytes. The mutual exclusive *KB* and

MB check boxes define how GtkSort will treat the number in the *Memory Limit* text box. Memory limits less than 16 Megabytes will be ignored and the memory limit will be set to 16 Megabytes.

CLI parameter: -m nnnn

The number follows the -m parameter is always treated as megabytes.

2.1.10 L2 Cache

The *L2 Cache* text box is enabled when the check box on the *L2 Cache* is checked. If the check box is not checked GtkSort will use the default value of 128. The *L2 Cache* text box accepts the size of the L2 Cache in kilobytes. Refer to your CPU manual in order to find the correct value for your CPU. The command: `dmesg|grep L2`

in Linux console will return the size of L2 cache for the installed CPU.

CLI parameter: -l2 nnnn

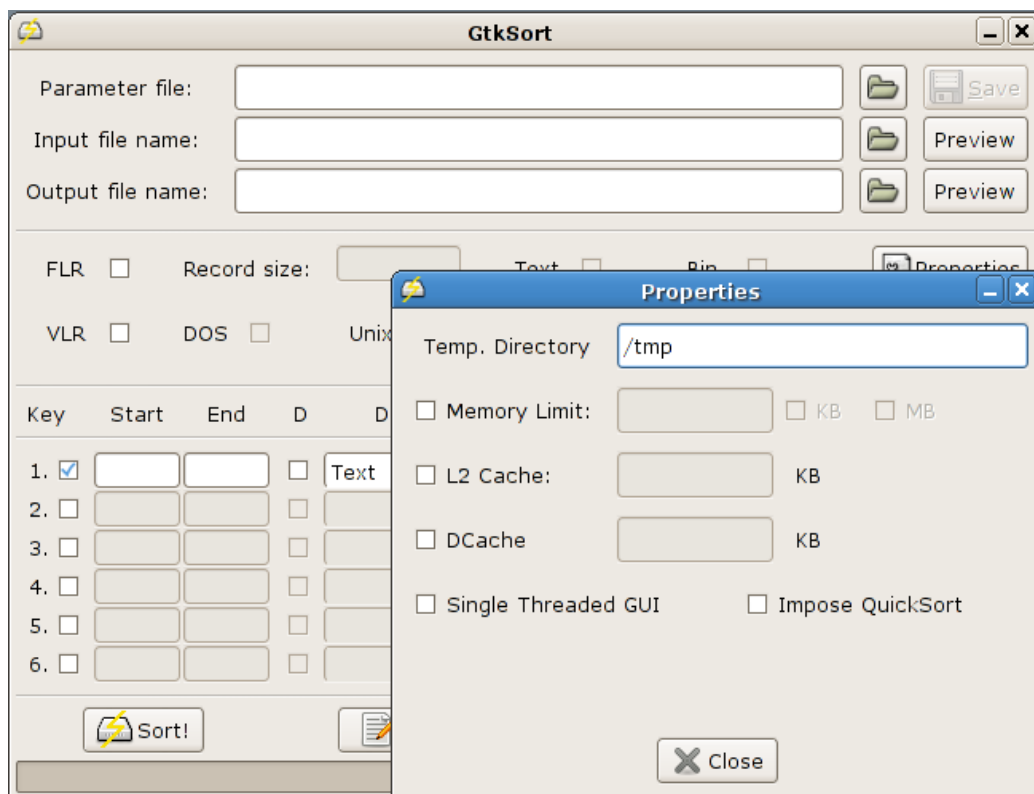


Fig. 2.1.8.1

2.1.11 DCache

The *DCache* text box is enabled when the check box on the *DCache* is checked. If the check box is not checked GtkSort will use the default value of 8. The *DCache* text box accepts the size of the D Cache in kilobytes. Refer to your CPU manual in order to find the correct value for your CPU. The command: `dmesg|grep "D Cache"`

in Linux console will return the size of D Cache for the installed CPU.

CLI parameter: -dcache nnnn

2.1.12 Single Threaded GUI

When the check box on the left of *Single Threaded GUI* label is checked, the GUI will not be responsive during the sort process. Depending on the platform, this may result better performance or not. It also solves problems caused by poor implementations of the multi-threading stack under certain operating systems.

2.1.13 Impose QuickSort


Under certain conditions, GtkSort uses its own implementation of MSD radix sort algorithm in order to improve performance. When the check box on the left of *Impose Quicksort* label is checked, GtkSort will use the quick sort implementation of the standard library of the operating system, degrading the performance. This may solve reliability problems that may appear in the future under certain architectures or operating systems.

CLI parameter: -qsort

2.1.14 Temp. Directory

The *Temp. Directory* text box accept the directory that GtkSort will use to store its temporary files during the sort/merge process. If no temporary directory is defined, GtkSort will use the system's default temporary directory.


2.1.15 Sort! Button

By pressing the *Sort!* button,  GtkSort will sort the input file and will create the output file. During the sort, the progress bar on the bottom of the main window will show the progress of the process. If the parameters are not valid, GtkSort will issue an error message and the sort process will not begin. In the case of a run-time error, GtkSort issues an error message, cancels the sort process and releases the temporary disk space it has reserved.


2.1.16 Help Button

Help button  is not functional and will be removed in the future.

2.1.17 About Button

By pressing the *About* button,  a new window will appear (Fig. 2.1.16.1) with general information about this software.

2.1.18 Quit Button

By pressing the *Quit* button,  the program exits. Quit button remains disabled during the sort process and there is no way for normal exit before its completion. By killing the program through the operating system, the temporary disk files will remain on the disk and the user will have to delete them manually.

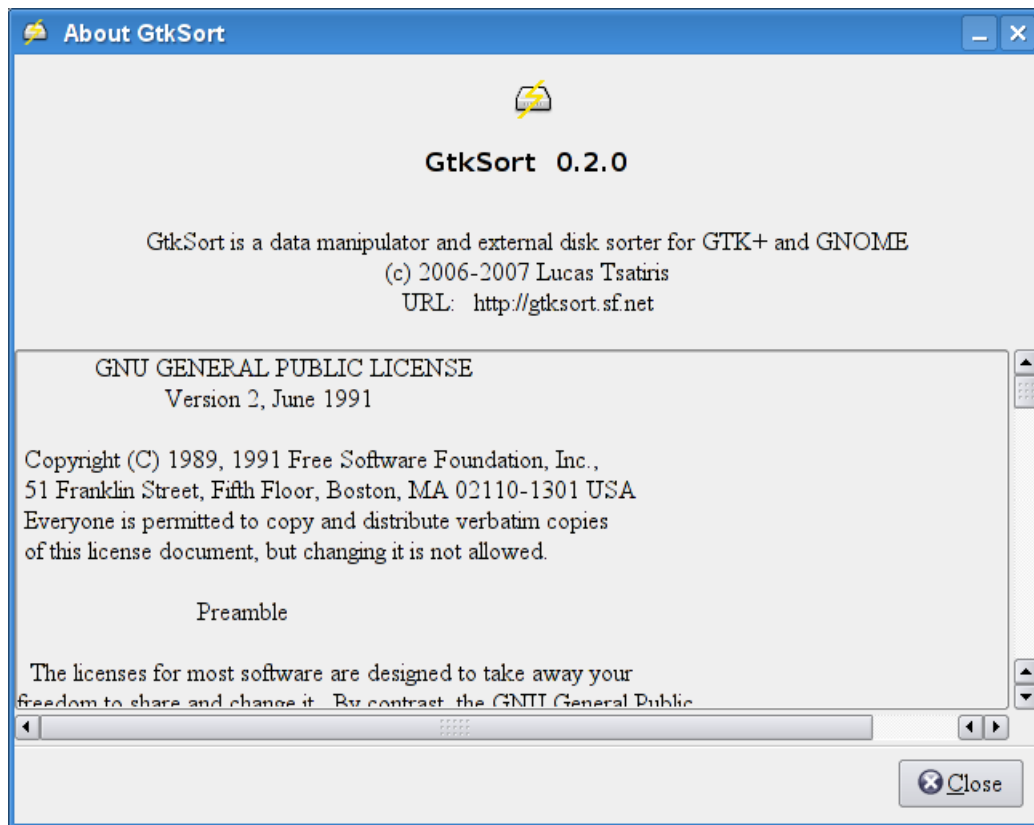



Fig. 2.1.16.1

2.1.19 Preview Buttons

The *Preview* buttons  are placed on the right of the *Input file name* and the *Output file name* text boxes. By pressing any of them, a new window will appear (Fig 2.1.18.1) displaying the contents of the sort keys of the first one thousand records, for the corresponding file. For an aligned view of the contents of the sort keys, a fixed space font can be used. By pressing the font button on the bottom of the preview window a font chooser dialog window will appear, and the desired font can be selected there.

2.2.0 Working with the GtkSort CLI

The CLI can be used from the console of the operating system. The command line argument `-h` will provide a summary of the valid command line arguments (Fig. 2.2.0.1). All the arguments are explained in the sections of the chapter 2.1.

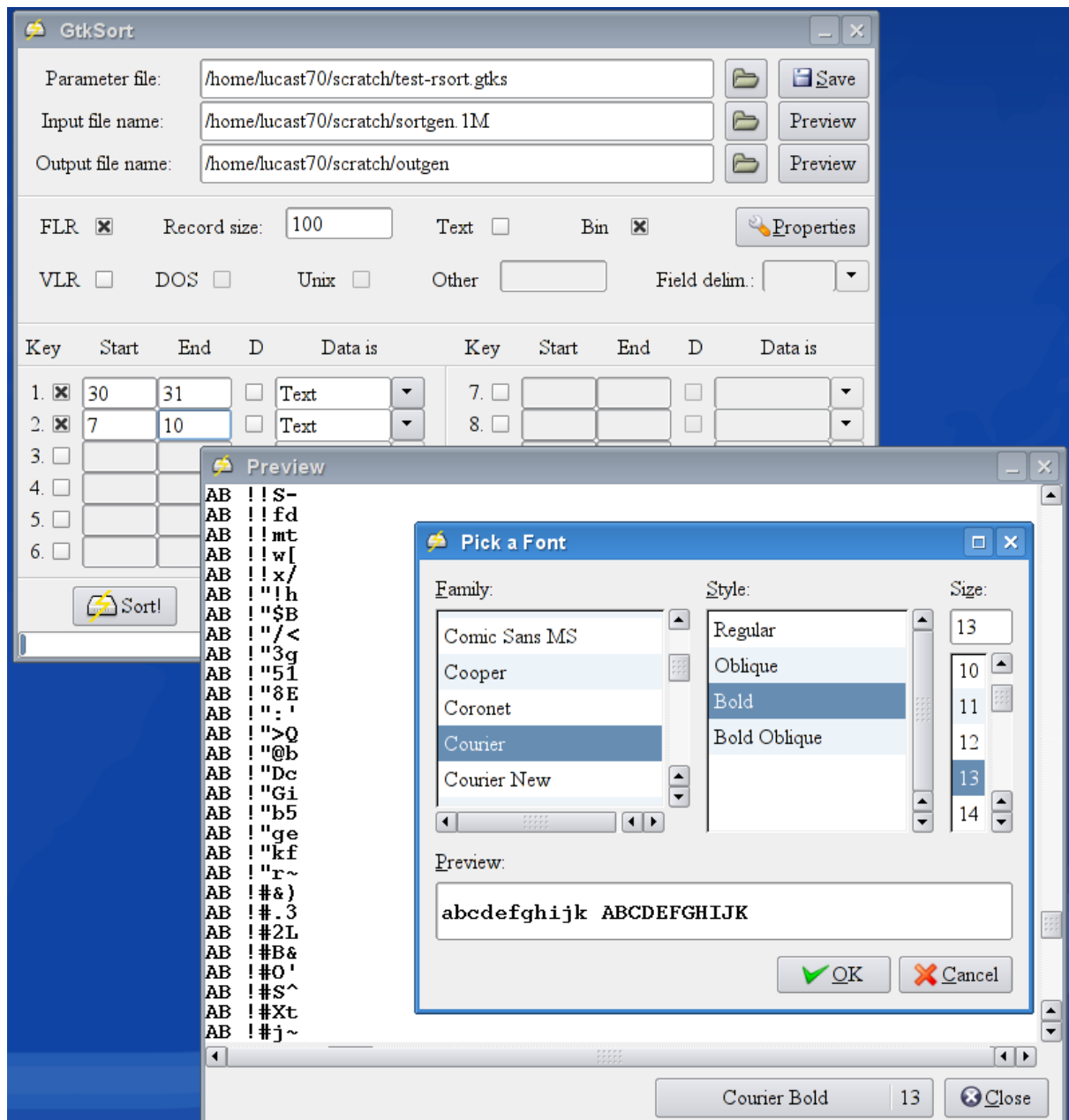


Fig 2.1.18.1

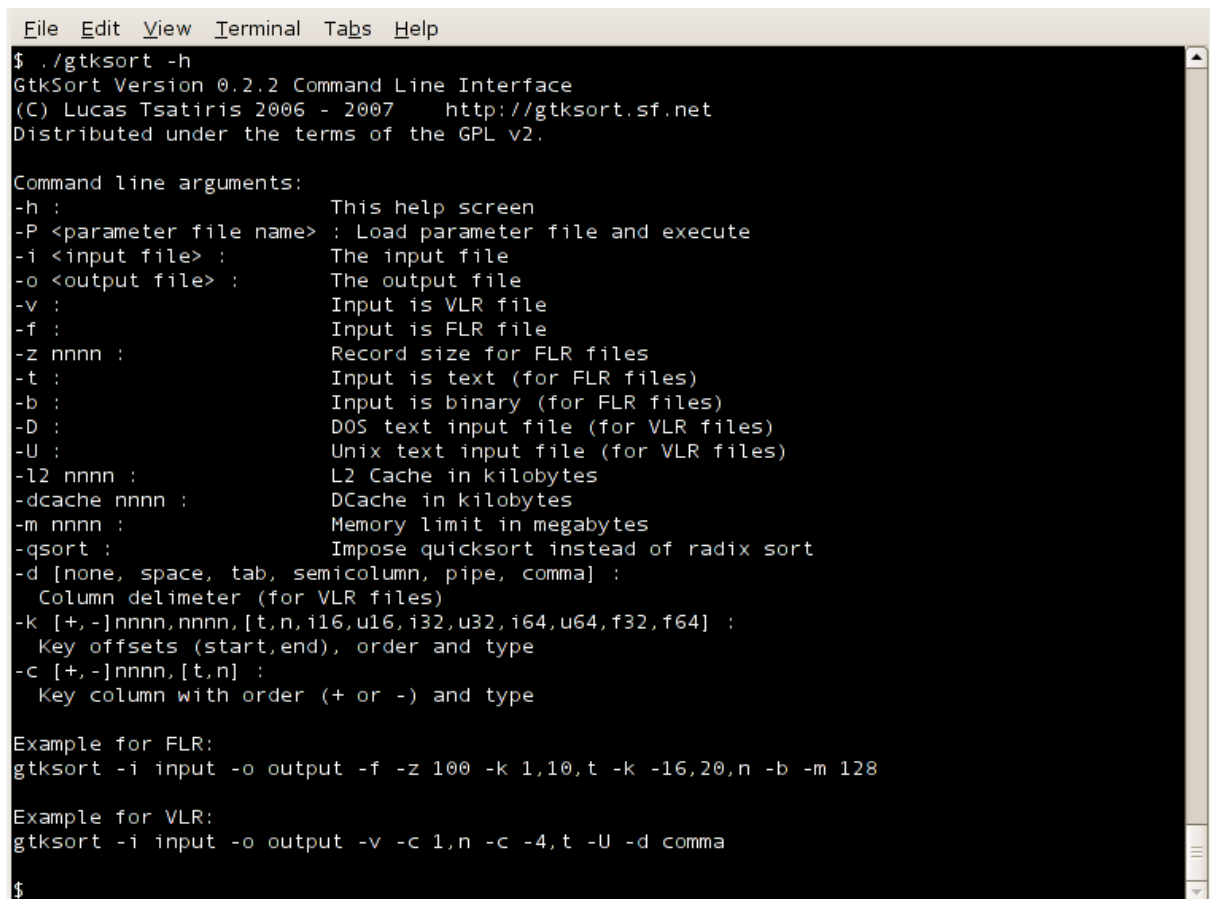
3.0.0 Error Messages

GtkSort issues error messages in various cases. The error messages come from the operating system or GtkSort itself. If they overlap, the operating system's error message has priority. The following list describes the error messages that come from GtkSort. If an error message is not described in the following list, please refer to the manual of the operating system.

Invalid parameter file

Cause: The selected file is not a GtkSort parameter file.

Action: Select a valid GtkSort parameter file.



```
File Edit View Terminal Tabs Help
$ ./gtksort -h
GtkSort Version 0.2.2 Command Line Interface
(C) Lucas Tsatiris 2006 - 2007    http://gtksort.sf.net
Distributed under the terms of the GPL v2.

Command line arguments:
-h :                This help screen
-P <parameter file name> : Load parameter file and execute
-i <input file> :    The input file
-o <output file> :    The output file
-v :                Input is VLR file
-f :                Input is FLR file
-z nnnn :           Record size for FLR files
-t :                Input is text (for FLR files)
-b :                Input is binary (for FLR files)
-D :                DOS text input file (for VLR files)
-U :                Unix text input file (for VLR files)
-l2 nnnn :          L2 Cache in kilobytes
-dcache nnnn :      DCache in kilobytes
-m nnnn :           Memory limit in megabytes
-qsort :            Impose quicksort instead of radix sort
-d [none, space, tab, semicolon, pipe, comma] :
  Column delimiter (for VLR files)
-k [+,-]nnnn,nnnn,[t,n,i16,u16,i32,u32,i64,u64,f32,f64] :
  Key offsets (start,end), order and type
-c [+,-]nnnn,[t,n] :
  Key column with order (+ or -) and type

Example for FLR:
gtksort -i input -o output -f -z 100 -k 1,10,t -k -16,20,n -b -m 128

Example for VLR:
gtksort -i input -o output -v -c 1,n -c -4,t -U -d comma

$
```

Fig. 2.2.0.1

Cannot save parameter file

Cause: Insufficient user privileges or disk space.

Action: Check the available disk space or choose a disk location where you have write permission.

Cannot load parameter file

Cause: Insufficient user privileges or file does not exist.

Action: Select an existing file name or ask your system's administrator to grant you sufficient rights on the file.

Cannot output on input file

Cause: Input and output file names are the same.

Action: Change the output file name.

File does not exist

Cause: The input file does not exist.

Action: Select an existing file.

Not enough memory

Cause: Insufficient physical memory.

Action: Lower your memory limit or add more RAM to the system.

Internal memory error 101

Cause: Corrupted heap or bad physical memory.

Action: Forward the message to the Help forum in sourceforge.net following the URL:
http://sourceforge.net/forum/?group_id=180969

No keys defined

Cause: The *Sort!* button was pressed but no key was defined at the time.

Action: Define at least one key before you press the *Sort!* button.

Invalid offsets

Cause: The *End* offset is smaller than the *Start* offset.

Action: Type an *End* offset greater or equal to the corresponding *Start* offset.

Undefined record type

Cause: Neither the *FLR* check box, nor the *VLR* check box is checked.

Action: Check the appropriate box for your data.

FLR must be text or binary

Cause: Neither the *Text* check box, nor the *Binary* check box is checked.

Action: Check the appropriate box for your data.

Cannot create thread

Cause: A thread creation has failed.

Action: Try to increase the memory limit. If the error persists forward the message to the Help forum in sourceforge.net following the URL: http://sourceforge.net/forum/?group_id=180969

VLR must be DOS or Unix

Cause: Neither the *DOS* check box, nor the *Unix* check box is checked.

Action: Check the appropriate box for your data.

Undefined record size

Cause: The *Record Size* is equal to zero.

Action: Type the record size.

Invalid data type

Cause: A key defined for a *VLR* file is not of *Text* or *Numtext* type.

Action: Choose either *Text* or *Numtext* as the data type of the sort key for the *VLR* file.

4.0.0 Additional Information

Home page: <http://gtksort.sf.net>

Other useful links:

Radix sort: http://en.wikipedia.org/wiki/Radix_sort

Sort benchmark home page:

<http://research.microsoft.com/barc/SortBenchmark>

Copyright (C) 2000,2001,2002 Free Software Foundation, Inc.
51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA
Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the

Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old

one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.