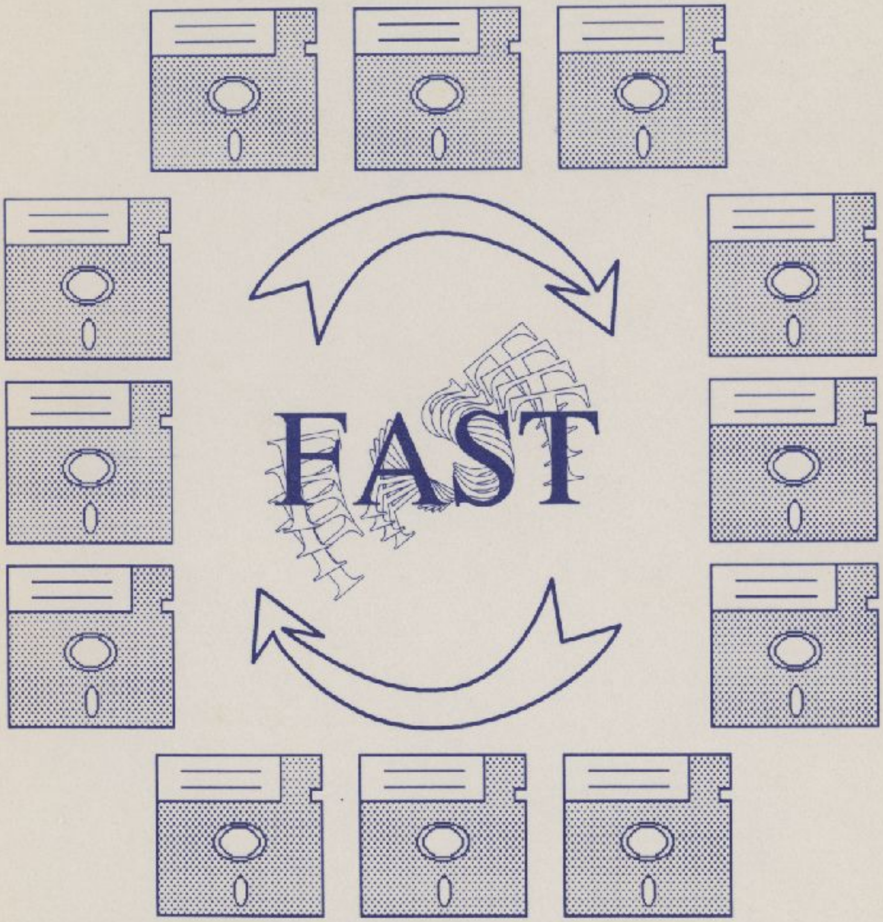
TI-SORT

The incredible SORTware!



This document describes the capabilities of
TI-SORT version 1.0

Liabilities

INSCEBOT will not be responsible for any losses resulting from the product described in this document. Actual liabilities will be limited to the original purchase price.

Licensing

This product may not be duplicated for resale or distribution without the express permission of INSCEBOT Inc.

Warranty

INSCEBOT warranties this product for a period of 90 days from date of purchase. Replacement will be accomplished by return mail with the only cost being postage. After 90 days there will be a $5.00 replacement fee.

Upgrades to the current version of TI-SORT will be provided at a nominal 40% of original cost with the return of the original diskette.

Contents

Section 1
Introduction

Section 2
Getting Started

Section 3
Operating Instructions

1.4 GLOSSARY

NESTED SORT — The function of sorting on secondary fields when there are duplicate values found in the primary sort field. TI-SORT will perform a nested sort to 8 levels if desired.

Function Keys — Function keys are those that function when Function and a key are depressed simultaneously. Normally expressed as (Fn) where n is a number between 0-9 or the quit key.

Section 2

GETTING STARTED

2.1 BEFORE YOU BEGIN

Please make a backup copy of your TI-SORT diskette and keep it in a safe place.

2.2 LOADING PROCEDURE

There are three loaders provided on your TI-SORT diskette as follows:

EXTENDED BASIC The "LOAD" file will automatically load TI-SORT from Drive #1 when the system is powered up with the EXTENDED BASIC cartridge in the cartridge slot.

LOAD and RUN File "TISORT" may be specified with any loader that supports the LOAD and RUN options. Note that this functions only from drive #1.

PROGRAM RUN File "TISORT0" may be specified as the program image to load. This function will allow loading from any drive.

2.3 ON-LINE HELP

Once TI-SORT is loaded and you have progressed past the entry display, HELP functions are available by depressing the "Function" key and the "7" key simultaneously (F7). The help displays are sensitive to the display and location on the display where the cursor is located. The intent is to provide help on the particular aspect of TI-SORT that you are currently involved in. Once the sorting process has started, help displays are no longer available.

As an added assistance, you may always return to the display following the title screen by depressing Function ESCAPE (F9). Once the sorting process is started, this is no longer possible.

If the "K" key is held down while the SORTING selection is hi-lited, the first 40 characters of the selected sort fields will be scrolled on the bottom of the display. This allows a visual verification of the desired sort key.

Section 3

OPERATING INSTRUCTIONS

3.1 MENU SELECTIONS

Once you are past the title display, several menu-oriented displays will guide you through the selection process of TI-SORT. The first display allows the selection of the type of files you wish to sort, or the ancillary functions of sort order selection and disk cataloging.

The selections are made by depressing the SPACE bar until the desired selection is hi-lited and then depressing the ENTER key. Alternately the first character of the selection may be depressed to hi-lite the field.

Please refer to section 4 for detailed descriptions of the file types which TI-SORT will support.

3.1.1 TI-BASE DATA-BASES

This selection means that there is an existing TI-BASE data-base which you wish to sort.

3.1.2 FIXED RECORDS

This selection requires that the data you wish to sort on is located in the same physical place in each record. Note that this does not require a fixed record type -- only that the data is consistent.

3.1.3 DELIMITED RECORDS

This selection indicates that each field is separated by a delimiter; i.e., possibly each field is separated by a comma to delineate the fields.

Note that the specific delimiter may be entered on the next display.

3.1.4 BASIC FILES

These are files which have been created by either BASIC or EXTENDED BASIC. See section 5 for limitations on DISPLAY type BASIC files.

3.1.5 ORDER

The sort may be specified as either in ascending or descending order. The current selection is displayed in parentheses.

The order is changed by depressing the ENTER key while this selection is hi-lited.

3.1.6 CATALOG

A directory of any file device may be obtained.

3.2 FILE SELECTIONS

Once you have selected the type of file that you are going to sort, TI-SORT will prompt you for the names of the input/output files and the device to be used for scratch files. These should be entered as "device.filename" or in the case of the scratch file as "device"; i.e., DSK1.TEST. Once the first entry is made, the following entries assume default values which may be changed if desired.

If errors are made in entering data, there is a redo option following the last selection.

3.2.1 INPUT FILES

You must enter the name of the data-base or file which you wish to have sorted. Note that for TI-BASE data-bases you need only enter the data-base name with no extensions, while for other files you must specify the complete file name.

3.2.2 DELIMITERS

If you selected a delimited type file, the delimiter may be entered here. Any single character delimiter may be used. The delimiter must be unique in the record; i.e., if a comma (,) is selected as a delimiter, the record must not contain any other commas.

3.2.3 SCRATCH FILES

The system will automatically create a file named "SCRATCH" on the device you specify, to use during the sorting process. Normally a space equal to the file being sorted is sufficient. Slightly more space may be required if the input file is a variable type, as the scratch file needs fixed records equal to the maximum record size to function. The scratch device need not be on the same device as the input file. The scratch file will be removed upon sort completion.

Note that a scratch file will be created, but not used, if the file being sorted is small enough to fit in memory (approximately 16k or less).

3.2.4 OUTPUT FILES

The resultant sorted file may be rewritten over the original file or directed to another file or device. In the case of TI-BASE data-bases, a new data-base will be created if the name and/or device is different than the original. This allows you to generate multiple data-bases that are sorted on different fields.

3.3 BUILDING STRUCTURES

If the file type that you selected was not a TI-BASE data-base, then a structure file describing the record content must exist or be created. For those of you who are familiar with TI-BASE, this will not need any explanation. For those of you who have not been exposed, an explanation is in order.

For each field in the record you must supply a name, data type and width. Additionally, if the data type is numerical, you must supply the significant number of decimals to accommodate in the sorting process.

The field types which are recognized are as follows:

(C)haracterstrings of any characters

(N)umerical + -.0-9 Numerical data. *Note that the width must include the sign and decimal point.*

(D)ate in the form mm/dd/yy

ENTER will advance fields in the description process. Directional arrows allow positioning anywhere on the display. The process is terminated by depressing Function EXE (F8) as defined on the TI-BASE overlay.

Note that for non TI-BASE files there are options for using an existing structure file, or for saving a newly created or modified structure for future use.

3.4 SELECTING SORT ITEMS

Once you have completed a functional structure, the selection display allows you to specify the fields which you want to sort on. The SPACE bar allows positioning to the desired field, and an ENTER will select this field for sorting. A number indicating the sort field order will appear to the left of the fields as they are selected.

Nested sorting is performed on the fields in the order in which they are selected. When all fields have been selected, an additional ENTER following the last field selected will initiate the sorting process.

Note that if an error is made in field selection, depressing the "R" key will restore the display to its initial condition and the selection process may be repeated.

3.5 SORTING

The sort process will read the input records into memory until memory is full. Then that block of records will be sorted and written to the intermediate scratch file. This process will be repeated until the input file has been exhausted. The blocks on the scratch file will then be merged to the output file in a sorted order.

Section 4

FILE DESCRIPTIONS

4.1 TI-BASE

TI-BASE data-bases consist of 2 files for each data-base. The files are distinguished by a terminator of /S for the structure file and a /D for the data file. The structure file is described in detail in the TI-BASE documentation. The data consists of fixed length records, where each data field appears in the same place in each record. The record length is equal to the sum of the field widths.

i.e.

```
AAAAABBBBBBCCCCDD
AAAAABBBBBBCCCCDD
    .
    .
AAAAABBBBBBCCCCDD
```

4.2 FIXED RECORDS

Fixed records expect the data to appear in the same place in each record. The file type may be either fixed or variable as long as the data to be sorted on is consistent and appears in the same location relative to the beginning of the record.

i.e.

```
AAAABBBBBCCCCCCDD
AAAABBBBBCCCCCCCCCCCDDD
     .
     .
AAAABBBBBCCDDD
```

In the above example, only fields A and B may be used to sort on in this format. Fields C and D are not consistent in their location in each field.

4.3 DELIMITED RECORDS

Delimited records have fields which are variable in length and terminated with a specific character as a delimiter. The delimiting character may not appear anywhere else in the data; i.e., using a comma as a delimiter.

AAA,BBBB,CC,DDDD
AA,BBB,CCCCC,DD
.
.
AAAAAA,BBBBBBBB,CCCCCCCCCC,DDDDDDD

It does not matter if the file type is variable or fixed as long as the data follows the required format.

4.4 BASIC FILES

The data format which TI-SORT expects is dependent on how it was specified within BASIC. The files will be either INTERNAL or DISPLAY.

4.4.1 INTERNAL RECORDS

Numeric data is written 9 characters per field with the first character being a character count of 8 followed by an 8 character radix 100 floating point value. Note that when creating the structure for this type of data that the width should be specified as 8.

String data is written as a character count with a following character string. When creating the structure always specify the maximum string length expected for the width.

i.e. numeric, string, numeric

NrrrrrrrNstring1Nrrrrrrr
NrrrrrrrNstringnumber2Nrrrrrrr
.
.
NrrrrrrrNstring-3Nrrrrrrr

4.4.2 DISPLAY RECORDS

Both numeric and string data are in ASCII form with the spacing and separators dependent on the BASIC format used.

Numeric data always has at least one preceding and following blank.

BASIC Printlist Separators

- , places data in the next print field. (note that each print field equals 14 spaces)
- ; places data adjacent to prior data.
- : places data on next line.

WARNING

TI-SORT will not correctly sort display data which have strings with imbedded blanks.

The reason is that TI-SORT cannot distinguish imbedded blanks from blanks between fields.

Section 5

SORTING PHILOSOPHY

5.1 SORT METHODOLOGY

The methodology utilized by TI-SORT is to read as many records into memory as possible, sort this block of records, and then write the block out to a scratch file. This cycle continues until the input file has been completely read. The blocks of sorted data on the scratch file are then merged into the output file.

The sorting method used is called quick-sort. It was developed by C.A.R. Hoare and is described on page 347 of Fundamentals of Data Structures by Horowitz. The methodology employed by this sort is to successively partition the original file into subfiles and sort each subfile separately.

5.2 LIMITATIONS

TI-SORT will sort up to 32767 records.

TI-BASE records may be any size that TI-BASE supports.

Non TI-BASE files are limited to a record size of 255 characters.

Up to 17 fields may be defined in a record.

Nested sorts up to 8 levels may be performed.

Sufficient scratch space must be present. Required space may be calculated by multiplying the maximum record size by the number of records.

During the sort process, the sort fields are assembled into a "key". The length of this key is limited to 255 characters.

5.3 BASELINE SORTING SPEEDS

TI-SORT will operate in one of three modes dependent on the data being sorted.

Mode 1 If the input file will totally fit in the memory available (approximately 16k), then a scratch file will be created but not used. The data will be read in, sorted in place, and rewritten to the output file. The approximate number of records may be computed by dividing the record size into 16k; i.e., using 80 character records yields approximately 200 records.

Mode 2 If the input data will not totally fit in memory, then subfiles of sorted data will be written to the scratch file. When the input file is completely read in, then the subfiles from the scratch file will be merged into the output file. The number of subfiles may be computed by dividing the total number of records by the number of mode 1 records; i.e., 500 80 character records will require 3 subfiles.

Mode 3 In order to merge subfiles in mode 2, there must be room for one record from each subfile in available memory. In the examples we have used, if more than approximately 200 subfiles are created, then some additional disk work is involved in the merge process. The approximate mode 3 threshold may be determined by (number of mode 1 records) squared. If we use 255 characters per record then the number of mode 1 records is approximately 64. Squaring this, we arrive at 4096 as the threshold for entering mode 3.

Note that the transition between modes is transparent to the user and the only outward indication will be a slower merging speed.

Some baseline speeds for comparison purposes are as follows:

500 80 char records on a Horizon RAM Disc. = 90 seconds

500 80 char records on SS/SD floppy drive = 8 min. 11 sec.

The sorting speed may be enhanced by having the scratch file on a different device from the output file. This minimizes mechanical head movement.

INSCEBOT

P.O. Box 291610
Pt. Orange, FL 32129