

Data Model Tools

Using ASCII Files in CIAO 4.0

CIAO users are familiar with the flexible filtering and binning capability that the Data Model tools provide with FITS files. In CIAO 4.0, the same tools also work on ASCII (text) files containing tables, using a new part of the software library called the "ASCII kernel." The ASCII kernel allows easy text file manipulation by the tools `dmlist`, `dmcop`, `dmstat`, `dmtcalc`, and `prism`. The majority of the other DM-specific tools also support ASCII input; refer to the Limitations section for exceptions.

The ASCII Kernel is being released as a beta in CIAO 4.0.

Filtering Data

Given a basic text file, `dmlist` may be used to select a range of data from two of the columns:

```
21.0  41.3  21.8
22.0  41.1  20.2
23.0  43.8  17.3
24.0  12.3  11.1
```

```
unix% dmlist "sample.dat[col3=11:20] [cols col2,col3]" data,clean
```

```
#  col2                col3
           43.80           17.30
           12.30           11.10
```

By default, unnamed columns are referred to as "col1", "col2", etc. If column names are provided, they may be used in the filter:

```
unix% cat input.txt
# ROW      time                ccd_id energy                pi
  1  87.5272969157      0      14707.57031250      1008
  2  87.5272969157      0      13968.8378906250      957
  3  87.5272969157      0      15152.52343750      1024
  4  87.5683369190      7       268.5079650879       19
  5  87.5683369190      7      1101.3159179688       76
  6  87.5683369190      7      2045.5782470703      141
```

...

```
unix% dmlist "input.txt[time=10:100] [cols energy]" data
```

```
-----
Data for Table Block input.txt
-----
```

```
ROW      energy
  1      14707.57031250
  2      13968.8378906250
  3      15152.52343750
  4       268.5079650879
  5      1101.3159179688
  6      2045.5782470703
  7      13929.0205078125
  8       3547.0227050781
  9      1672.4017333984
```

...

ASCII to FITS; FITS to ASCII

The DM creates FITS format output by default; the kernel option must be specified every time to make the output be a text file.

To copy a FITS file to a simple text output format, e.g. to be used by another piece of code:

```
unix% dmcopu input.fits "output.txt[opt kernel=text/simple] "
```

To create a basic FITS file from ASCII input:

```
unix% dmcopu input.txt output.fits
```

DM filter syntax may be included when creating a new file. Here dmcopu is used to create an output file of filtered data in FITS or text format:

```
unix% dmcopu "sample.dat[col3=11:20] [cols col2,col3]" filtered.fits
```

```
unix% dmcopu "sample.dat[col3=11:20] [cols col2,col3]" "filtered.txt[opt kernel=text/sin
```

About the Kernel

Output Text Formats: kernel option

There are three output formats allowed with the ASCII kernel: text/raw, text/simple, and text/df. The output is specified by including the "kernel" option in the output file name.

- [opt kernel=text/raw]

Simple text table format consists of free-format columns with no header keywords. The format understands only two datatypes: numbers (treated as double precision) and text strings. All columns are scalar and are given the default names "col1", "col2", etc.

- [opt kernel=text/simple]

This format is similar to text/raw, but has an optional header defining the column names. In its simplest form, the header consists of a single line of whitespace-separated column names preceded by the comment character; see the "comment" and "colnames" options below. The text/simple option is compatible with the SM plotting program.

- [opt kernel=text/df]

Data Text Format (DTF) is a pseudo-FITS format with support for headers and data subspaces. Free format tables are the default, but fixed-format fields are also supported.



Additional Options

There are several other options that may also be used to qualify a text file. Multiple options are specified as a comma-separated list. You can use these options to allow CIAO to read tables in text files with slightly different formats from the default, for example by skipping header lines or changing the field separator.

[opt sep=:], [opt sep=:,white], [opt sep=":;"]

Define the given character (e.g. ":", used here, or "/") , to be the separator for data fields. The "sep" option defines each instance of the character as a new field. This example represents four fields, with the second one being empty:

```
14 . 1 : : 23 . 2 : 15 . 1
```

If the "white" qualifier is included, the separator is treated as whitespace. This means that if you have multiple separator characters next to each other, they only count as one separator. The same example - "14.1::23.2:15.1" - then represents only three fields.

More than one character may be defined as the separator. For instance, [opt sep=":;"] defines both ":" and ";" as separators.

[opt skip=3]

Skip the given number of lines (e.g. 3) at the beginning of the file. This helps handle some formats with fixed headers.

[opt comment=#]

Lines that begin with the given character (e.g. "#", the default) prior to the first data line will be treated as comments. There is one special comment line which the "colnames" option controls, as described in the next item.

[opt colnames=first]

The first comment-character line is treated as a space-separated list of column names. The value "first" is the default; other possible values are "last" (the last comment-character line prior to the first data line) and "none" (none of the lines are treated as a colnames definition).

The defaults for each output format are:

- text/raw - comment='#',sep="\t\r",white,colnames=none,skip=0
- text/simple - comment='#',sep="\t\r",white,colnames=first,skip=0
- text/df - sep="\t\r",white,colnames=none,skip=0

Limitations

The ASCII kernel was developed to allow CIAO users to use the familiar DM syntax in manipulating and filtering text files; it is not intended as a replacement for the FITS kernel in pipelines.

- The kernel does not always work well with other tools, e.g. `dmextract`, `acis_process_events`, etc.
- Support for images is incomplete. Although some image operations will work, there is no official support for ASCII images in this release.
- To create a dataset with more than one block or write header keywords, you have to use the DTF flavor. Therefore, the simple and raw formats cannot be used with any CIAO tools which require multiple blocks and header keywords.
- Files larger than 2 GByte are not supported in the ASCII kernel.
- Header lines longer than 1024 bytes are not supported. Data lines may be arbitrarily long.



March 2008