CXC-DM-005

CXC Data Model

Vol. 5

FITS Kernel Design

Jonathan McDowell
Chandra X-ray Center
October 22, 2001

Contents

1 Introduction

This document describes the CIAO DM ('Data Model') FITS kernel design. It gives the mapping between the FITS kernel and the Abstract Design.

The DM FITS kernel in its present design relies on W. Pence's CFITSIO package for access to the FITS files.

2 Datasets

A DM Dataset maps to a single FITS disk file. Access to a virtual file may be optimized by opening the FITS file with a CFITSIO filter to perform simple pre-filtering, reducing the number of rows that must be queried with the DM layer filtering.

We used to test for the first few bytes of the file being 'SIMPLE', required for a FITS file. To support CFITSIO's ability to read gzipped FITS files and networked FITS files, we instead now use the full fits_open_file routine. Unfortunately this means that anything CFITSIO thinks it can read, it will recognize as a FITS file - that isn't always what we want, because for instance we don't necessarily like its interpretation of an IRAF .imh file, and we at least want to know that's what we have. One strategy would be to try the SIMPLE test first, then try other kernels, then fall back on fits_open_file last.

3 Blocks

Each DM block corresponds to a FITS HDU. Primary and IMAGE HDUs are considered to be Image blocks. BINTABLE and TABLE HDUs are considered to be Table blocks; on writing, BINTABLE is the default but TABLE HDUs can be created if the hint TABLE=ASCII is selected. Random Groups HDUs are not supported.

Certain FITS HDUs are recognized as special. In particular, HDUs containing Good Time Intervals are considered to be part of the data subspace of the main ('most interesting') block. In an initial implementation, this rule was applied strictly and GTI HDUs were not visible as separate DM blocks. However, the demands of back compatibility and interoperability forced us to step back from this position. Now, GTI HDUs are separate DM blocks in their own right, so that -for instance - you can write header keys to them; but they are also part of the data subspace of other blocks. This means that they may be created or altered while perfoming operations on another block; caveat emptor.

At dataset open, all FITS headers are read and parsed for structural data.

3.1 Block names

The DM first looks for the keyword HDUNAME. If found, its value is the block name; otherwise it looks for EXTNAME and EXTVER and concatenates them to make the block name.

4 Table columns

In a FITS table, the TTYPE/TUNIT/TFORM keywords give the name, unit and data type of each scalar column component.

4.1 Vector columns

The DM-specific FITS keywords MTYPEn/MFORMn are used to define explicit groupings of columns and coordinate names.

In addition, the presence of a vector coordinate system forces its parent to be a vector column - see below.

4.2 Coordinates on columns

Coordinates on columns are defined using the standard WCS keywords TCTYPn, etc.

The TCTYPn keyword has a complicated syntax. If the 5th byte is the character '-', then it specifies both a name and a projection type. For instance, the value '`RA---TAN`' specifies the name RA and the projection type TAN. Otherwise, the TCTYPn just specifies the name, and the projection is assumed to be LINEAR.

Coordinate names and component names are defined using the DM-specific keyword TC-NAMn; if this is not present, the name part of TCTYPn is used.

MTYPE/MFORM pairs can be used to define vector coordinate systems. The special cases RA,DEC and xLON,xLAT are recognized as pairs even in the absence of MTYPE/MFORM, and force the implicit creation of a vector column.

4.3 FITS internal scaling

The TSCAL/TZERO scaling mechanism allows FITS files to apply a linear scaling to the internal table data; this corresponds to the idea of logical and physical coordinates. When this scaling is present, we consider the raw values to be the logical column descriptor and the scaled values to be the physical coordinate system on this descriptor.

In the current implementation, the column descriptors are always the physical column descriptor except if the scale=no directive is used; the unscaled values are not seen and the column type is forced to be floating point. In the new design, both descriptors will be implemented in a way directly analogous to the logical coordinates in images, and the #1, #2 notation will be used to access the logical (unscaled) values.

5 Images

The standard FITS BITPIX, NAXISn, etc. keywords are used to define the image data type and size.

5.1 Image coordinates

Standard WCS keywords are used. Multiple coordinate systems are supported. Physical coordinate transformations are stored using the 'P' subscript: CTYPE1P, etc. The primary world coordinate transforms are stored with no subscript: CTYPE1, etc.

The DM transforms are:

- Physical (LP) transform: logical to physical
- World (PW) transform: physical to world

However, the FITS transforms are

- Physical (LP) transform: logical to physical
- World (LW) transform: logical to world

The DM library handles conversion between the external FITS representation of the LP, LW transforms and the internal PW representation.

6 Data Subspace

The Data Subspace uses GTI extensions for filters on variables called TIME, and DM-specific keywords for other variables.

7 Header Descriptors

In general, FITS header keys map to DM header keys. However, FITS keywords which are considered 'structural' do not have corresponding DM keys. The DM also writes structural FITS keys of its own. These are described elsewhere but are summarized in this section.

Key data types are assumed to be double or long rather than float or short, unless a DFORM keyword explicitly forces the type.

7.1 Long keyword name convention

The DM supports long keyword names using a convention in which DTYPEn/DVALn pairs are written to the file:

```
DTYPE12=  'Gaussian.Coefficient'
DVAL12 =   4.32
```

defines a single DM keyword with name Gaussian.Coefficient and value 4.32; the value n=12 pairs the two keywords together but has no other significance.

7.2 Unit convention

The joint CFITSIO/DM units convention is applied, in which a FITS comment field beginning with square parentheses is recognized as a unit:

```
FLUX =  1.3E-14  / [erg/cm**2/s] Flux of source
```

7.3 Special header keys

The CHECKSUM and DATASUM keywords are automatically updated whenever a block is closed, if and only if they are already present. Therefore, you may write these keywords as standard DM string-valued keys with a blank initial value, and they will be written to the output file with the correct value at block close.