PDS4 NucSpec Local Data Dictionary Users' Guide

Small Bodies Node

September 24, 2020

1 Introduction

The Nuclear Spectroscopy (NucSpec) discipline dictionary was originally created by the PDS Small Bodies Node and the Dawn Gamma Ray and Neutron Spectrometer (GRaND) team, in support of the PDS4 migration of the Dawn GRaND archive. It was designed to be easily expandable to support other nuclear spectroscopy archives in the PDS. This version of the Users' Guide pertains to the pre-release version of the dictionary.

1.1 Purpose of this Users' Guide

The users' guide is intended to provide help in using the NucSpec discipline dictionary to include metadata specific to nuclear spectroscopy in PDS4 Product_Observational labels.

1.2 Audience

This user's guide is for any data provider who submits nuclear spectroscopy observations to the PDS. Familiarity with basic PDS4 concepts is helpful.

1.3 Applicable Documents

PDS4 Standards Reference: https://pds.jpl.nasa.gov/datastandards/documents/sr/

PDS4 Information Model: https://pds.nasa.gov/datastandards/documents/im/

2 Overview of the NucSpec Discipline Data Dictionary

NucSpec is intended to provide classes and attributes for metadata specific to nuclear spectroscopy data. The Small Bodies Node holds the stewardship of NucSpec, with Jesse Stone (jstone@psi.edu) as the current point of contact.

Section 3 of this Guide covers how to include the classes and attributes of the dictionary in a PDS label file. Section 4 goes into more detail on the structure of the classes and attributes and how they are intended to be used. See section 5 for detailed specifications of each class and attribute.

3 How to Include the NucSpec Dictionary in a PDS4 Label

3.1 Data Dictionary Files

PDS4 dictionaries appear in several forms, typically having the same filename with different extensions. It can either be an ingest file, or a schema file coupled with a schematron file. The ingest file, with xml extension, is used for authoring the dictionary and often for ingesting it into tools, while the schema (.xsd) and schematron (.sch) files, which are compiled from the ingest file, are used to actually validate a product label. For released dictionaries, all these can be obtained from the PDS4 released schema page at https://pds.jpl.nasa.gov/datastandards/schema/released/.

Note: In its pre-release form, NucSpec is available on GitHub at https://github.com/sbn-psi/pds4-dictionary-grns.

3.2 Including the schema file in a label

In order to use the schema file, the Product_Observational element of your product label will need to have references to the dictionary added to it, as follows:

```
<Product_Observational
   xmlns="http://pds.nasa.gov/pds4/pds/v1"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xmlns:nucspec="http://pds.nasa.gov/pds4/nucspec/v1"
   xsi:schemaLocation="http://pds.nasa.gov/pds4/pds/v1
      https://pds.nasa.gov/pds4/pds/v1/PDS4_PDS_1900.xsd
      http://pds.nasa.gov/pds4/nucspec/v1
      https://pds.nasa.gov/pds4/nucspec/v1/SBN_NUCSPEC_1A00_0000.xsd">
```

This example assumes that the nucspec is the only dictionary in your label. If you have multiple dictionaries, you will need to make other modifications.

3.3 Including the schematron in a label

In order to use the schematron file, the xml prolog of your product label will need to have references to the dictionary added to it, as follows:

```
<?xml-model
href="https://pds.nasa.gov/pds4/nucspec/v1/SBN_NUCSPEC_1A00_0000.sch"
schematypens="http://purl.oclc.org/dsdl/schematron"?>
```

3.4 Including the data dictionary elements

The data dictionary defines XML elements that can be used in a `Discipline_Area`. A minimal example of the discipline area follows. A more detailed example for each class is given in the later sections of this Guide.

```
<Discipline_Area>
  <nucspec:NucSpec_Observation_Properties>
    <nucspec:Energy_Calibration>
      <nucspec:Polynomial>
        <nucspec:Polynomial_Term>
          <nucspec:order>value</nucspec:order>
          <nucspec:coefficient>value</nucspec:coefficient>
        </nucspec:Polynomial_Term>
      </nucspec:Polynomial>
    </nucspec:Energy_Calibration>
  </nucspec:NucSpec_Observation_Properties>
</Discipline_Area>
```

4 Organization of Classes and Attributes

The overall organization of classes and attributes for all of NucSpec is shown in Figure 1. The classes and subclasses appear in the label in the order they appear in the diagram reading from left to right. The dictionary contains three major classes under Nucspec_Observation_Properties, which specifies properties pertaining to a nuclear spectroscopy observation. These three major classes are outlined in the subsections below. Additional major classes may be added to the dictionary in the future to support new nuclear spectroscopy archiving or migration projects.

4.1 The Energy Calibration Class

Specifies methods and data used to determine the pulse height in energy units (e.g. MeV) for each channel of a pulse height spectrum.

This class currently includes support for polynomial conversion. Additional subclasses may be added to the dictionary in the future to describe other types of conversion.

Here is an example of how this class may be included in a label. This example has a polynomial of 7x^2 + 8x + 9, and references the GRaND BGO processing document:

```
<nucspec:Energy_Calibration>
 <nucspec:Polynomial>
   <nucspec:Polynomial_Term>
     <nucspec:order>0</nucspec:order>
     <nucspec:coefficient>9</nucspec:coefficient>
   </nucspec:Polynomial_Term>
   <nucspec:Polynomial_Term>
     <nucspec:order>1</nucspec:order>
     <nucspec:coefficient>8</nucspec:coefficient>
   </nucspec:Polynomial_Term>
   <nucspec:Polynomial_Term>
     <nucspec:order>2</nucspec:order>
     <nucspec:coefficient>7</nucspec:coefficient>
   </nucspec:Polynomial_Term>
 </nucspec:Polynomial>
 <nucspec:Calibration_Reference>
 <Internal_Reference>
   <lid_reference>
     urn:nasa:pds:dawn-grand:grand-bgo-processing
   </lid_reference>
   <reference_type>data_to_document</reference_type>
 </Internal_Reference>
 </nucspec:Calibration_Reference>
</nucspec:Energy_Calibration>
```

4.1.1 Polynomial

The pulse height in energy units (e.g. MeV) is determined for a given channel number using the polynomial: H(x) = a0 + a1 x + a2 x^2 + …, where H is pulse height with energy units and x is channel number.

Each <Polynomial_Term> subclass defines one term in the polynomial using the attributes <order> and <coefficient>.

For example, a polynomial such as 9 + 8x + 7x^2 would be described with the following:

```
<nucspec:Polynomial>
  <nucspec:Polynomial_Term>
    <nucspec:order>0</nucspec:order>
    <nucspec:coefficient>9</nucspec:coefficient>
  </nucspec:Polynomial_Term>
  <nucspec:Polynomial_Term>
    <nucspec:order>1</nucspec:order>
    <nucspec:coefficient>8</nucspec:coefficient>
  </nucspec:Polynomial_Term>
  <nucspec:Polynomial_Term>
    <nucspec:order>2</nucspec:order>
    <nucspec:coefficient>7</nucspec:coefficient>
  </nucspec:Polynomial_Term>
</nucspec:Polynomial>
```

4.1.2 Calibration Reference

This class provides a reference to a document that describes the energy calibration.

The calibration reference consists of a single <Internal_Reference> class, which will contain the LID of the document that describes the energy calibration, along with the reference type of data_to_document, which indicates that the product referenced by the LID is a document.

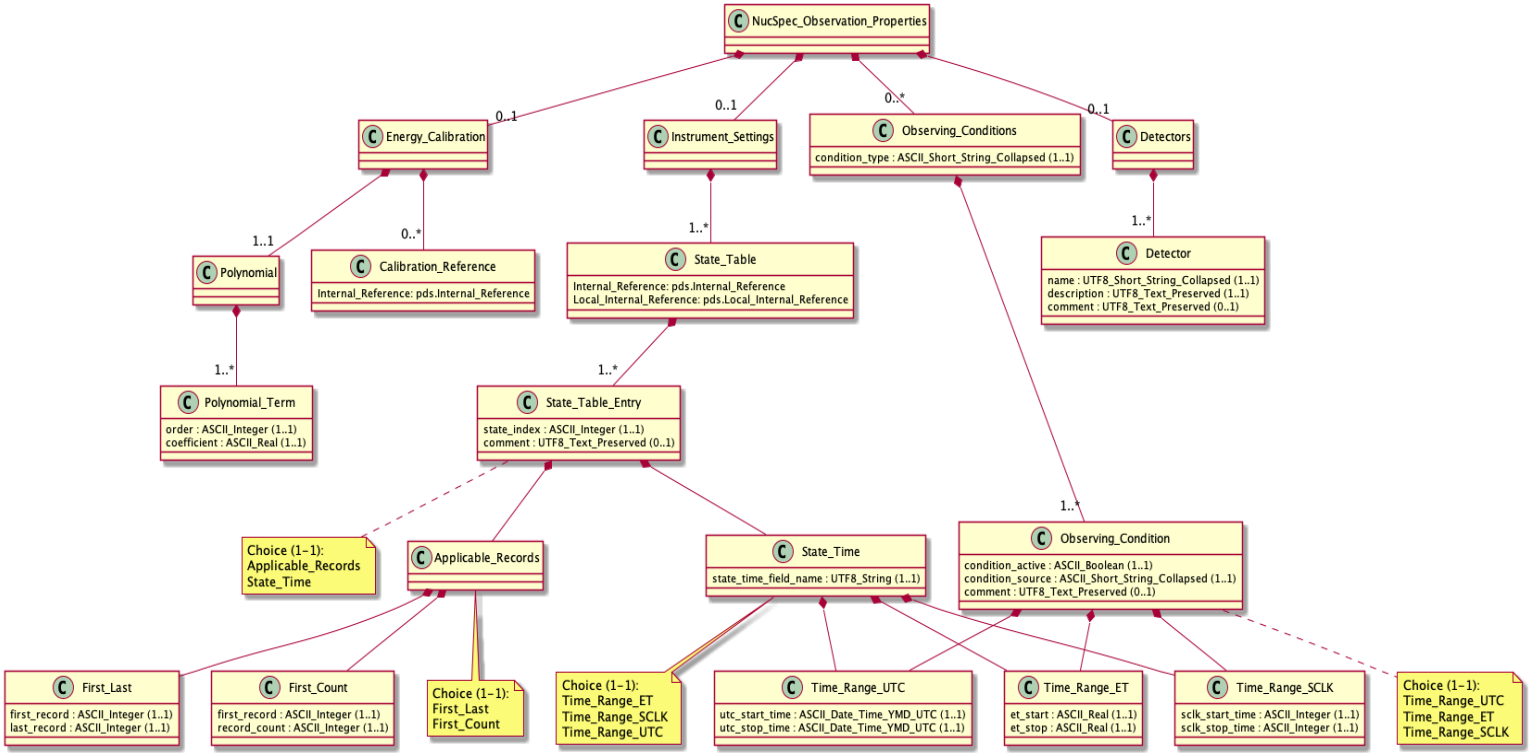

Figure 1. Organization of classes and attributes in the NucSpec dictionary

4.2 The Instrument Settings Class

The instrument settings class allows the documentation of the settings for the instrument during a specific observation. This may include calibration data such as the temperature of an instrument, or a flag that indicates the operating mode of the instrument.

The settings for an instrument are stored in a separate state table, which must also be a PDS4 product, and may be associated with individual observation either by directly indicating which records in the data product are affected by each record in the state table, or by indicating the timespan at which each state in the state table is applicable. The affected records will then have a specified time which falls within the timespan for the state table record.

4.2.1 The State Table Class

The State Table class indicates the location of the table of instrument states. This table must be a PDS4 product and is referenced with an Internal Reference. The Internal Reference will contain the LID or LIDVID of the state table, and a Reference Type that indicates that the state table is a date product.

The State table class will also contain a list of State Table Entries, one for each record in the state table, which will describe how records in the state table are associated with records in the data product.

An example that uses Applicable Records to associate two rows of the state table with two different ranges of rows in the data product.

```
<nucspec:Instrument_Settings>
 <nucspec:State_Table>
   <Internal_Reference>
     <lidvid_reference>
       urn:nasa:pds:example_bundle.data.state_table::1.0
     </lidvid_reference>
     <reference_type>nucspec_product_to_state_table</reference_type>
   </Internal_Reference>
   <nucspec:State_Table_Entry>
     <nucspec:state_index>1</nucspec:state_index>
     <Internal_Reference>
       <Local_Internal_Reference>
         <local_identifier_reference>table</local_identifier_reference>
         <local_reference_type>
           state_table_to_data_table
         </local_reference_type>
       </Local_Internal_Reference>
     </Internal_Reference>
     <nucspec:Applicable_Records>
       <nucspec:First_Last>
         <nucspec:first_record>1</nucspec:first_record>
         <nucspec:last_record>2</nucspec:last_record>
       </nucspec:First_Last>
     </nucspec:Applicable_Records>
   </nucspec:State_Table_Entry>
   <nucspec:State_Table_Entry>
     <nucspec:state_index>2</nucspec:state_index>
     <Internal_Reference>
       <Local_Internal_Reference>
         <local_identifier_reference>table</local_identifier_reference>
         <local_reference_type>
           state_table_to_data_table
         </local_reference_type>
       </Local_Internal_Reference>
     </Internal_Reference>
     <nucspec:Applicable_Records>
       <nucspec:First_Last>
         <nucspec:first_record>3</nucspec:first_record>
         <nucspec:last_record>4</nucspec:last_record>
       </nucspec:First_Last>
     </nucspec:Applicable_Records>
   </nucspec:State_Table_Entry>
 </nucspec:State_Table>
</nucspec:Instrument_Settings>
```

An example that uses State Time to associate two rows in the state table with two ET ranges in the data table, matching on the ET_MID field.

```
<nucspec:Instrument_Settings>
 <nucspec:State_Table>
   <Internal_Reference>
     <lidvid_reference>
       urn:nasa:pds:example_bundle.data.state_table::1.0
     </lidvid_reference>
     <reference_type>nucspec_product_to_state_table</reference_type>
   </Internal_Reference>
   <nucspec:State_Table_Entry>
     <nucspec:state_index>1</nucspec:state_index>
     <Internal_Reference>
       <Local_Internal_Reference>
         <local_identifier_reference>table</local_identifier_reference>
         <local_reference_type>
           state_table_to_data_table
         </local_reference_type>
       </Local_Internal_Reference>
     </Internal_Reference>
     <nucspec:State_Time>
       <nucspec:state_time_field_name>ET_MID</nucspec:state_time_field_name>
       <nucspec:Time_Range_ET>
         <nucspec:et_start>479865543</nucspec:et_start>
         <nucspec:et_stop>479866000</nucspec:et_stop>
       </nucspec:Time_Range_ET>
     </nucspec:State_Time>
   </nucspec:State_Table_Entry>
   <nucspec:State_Table_Entry>
     <nucspec:state_index>2</nucspec:state_index>
     <Internal_Reference>
       <Local_Internal_Reference>
         <local_identifier_reference>table</local_identifier_reference>
         <local_reference_type>
           state_table_to_data_table
          </local_reference_type>
       </Local_Internal_Reference>
     </Internal_Reference>
     <nucspec:State_Time>
       <nucspec:state_time_field_name>ET_MID</nucspec:state_time_field_name>
       <nucspec:Time_Range_ET>
         <nucspec:et_start>479866000</nucspec:et_start>
         <nucspec:et_stop>479867111</nucspec:et_stop>
       </nucspec:Time_Range_ET>
     </nucspec:State_Time>
   </nucspec:State_Table_Entry>
 </nucspec:State_Table>
</nucspec:Instrument_Settings>
```

4.2.1.1 The State Table Entry Class

The state table entry class describes how an individual record in the state table is associated with records in the data product. A state may apply either to a range of records in the data product, or it may apply to all of the data records that fall within a certain timespan.

If the entry applies to a range of records, then those records are described using the Applicable Records class.

If the entry applies to all records within a certain time, then the timespan is described using the State Time class.

4.2.1.1.1 The Applicable Records Class

The applicable records class describes a contiguous range of records to which a state in the state table applies. This may be described with either a starting record and a number of records in the range, or with a start and end record.

4.2.1.1.2 The State Time Class

The State Time class describes a timespan during which a given state in the state table applies. It also indicates which field in the data product will be examined to determine if the state falls within that timespan. This field must be of the same type as the timespan described.

There are currently three ways to describe a timespan: by UTC (Time_Range_UTC), by spacecraft clock (Time_Range_SCLK), or by event time (Space_Time_ET). This will determine which subclass must be used to describe the timespan. Each subclass will provide a start time and an end time.

4.3 The Detectors Class

The Detectors class describes which specific detectors were active when producing the data product. It will contain one Detector entry for each detector that was active during the observation.

For example, in order to indicate that the secondary detector was active during the observation, one might use the following snippet:

```
<nucspec:Detectors>
   <nucspec:Detector>
       <nucspec:name>Secondary Detector</nucspec:name>
       <nucspec:description>
         The backup detector for the instrument
       </nucspec:description>
       <nucspec:comment>
         Active because there was a problem with the primary detector.
     </nucspec:comment>
   </nucspec:Detector>
</nucspec:Detectors>
```

4.3.1 The Detector Class

The Detector class describes a single detector that was active at the time the data product was generated. It consists of a mandatory name and description, along with an optional comment.

4.4 The Observing_Conditions Class

The Observing_Conditions class indicates any conditions, whether or internal, that may affect the observations in some way.

The Observing_Conditions class will also contain a list of Observing_Condition elements, which indicate time ranges at which this particular Observing_Condition was active or inactive, along with sources for the contidition, if known.

4.4.1 The Observing_Condition Class

This class will contain either an activity indicator (condition_active) or a name for the source for the condition (condition_source), along with a time range for which this condition was in effect. It may be expressed as UTC, ET, or SCLK.

5 Definitions

Applicable_Records – class

The records to which this state applies.

Member of: State_Table_Entry

Cardinality: Single, Choice (choose between Applicable_Records and State_Time)

Members:

- First_Last
- First_Count

Calibration_Reference – class

Contains a reference to a document that describes the energy calibration for the product.

Member of: Energy_Calibration

Cardinality: Single, Optional

Members:

- Internal_Reference

coefficient – attribute

The coefficient of a single term of the polynomial relating channel number to pulse height.

Member of: Polynomial_Term

Cardinality: Single, Required

Data Type: Real

comment – attribute

Any notes about the state of the detector at the time of observation.

Member of: Detector

Cardinality: Single, Optional

Data Type: String

condition_active - attribute

Indicates if the condition is in effect during this time period.

Member of: Observing_Condition

Cardinality: Single, Optional

Data Type: ASCII_Boolean

Valid Values: true, false

condition_source - attribute

Describes where the condition is coning from. This may be an object that is emitting particles that interfere with the observation, or the reason for a data glitch.

Member of: Observing_Condition

Cardinality: Single, Optional

Data Type: String

condition_type - attribute

Describes the type of condition that occurred during the observation.

Member of: Observing_Conditions

Cardinality: Single, Required

Data Type: String

Valid Values: SEP, Solar Flare, Gamma-ray Burst, Electron Burst, Data Glitch

description – attribute

A description for the detector.

Member of: Detector

Cardinality: Single, Required

Data Type: String

Detector – class

Specifies a single detector which was active at the time of observation.

Member of: Detectors

Cardinality: Multiple, Required

Members:

- Detectors

Detectors – class

Specifies which detectors were active on an instrument at the time of observation.

Member of: NucSpec_Observation_Properties

Cardinality: Single, Optional

Members:

- name
- description
- comment

Energy_Calibration – class

Specifies methods and data used to determine the pulse height in energy units (e.g. MeV) for each channel of a pulse height spectrum.

Member of: NucSpec_Observation_Properties

Cardinality: Single, Required

Members:

- Polynomial
- Calibration_Reference

et_start – attribute

The earliest ephemeris time, as defined by the referenced field, for which this state is applicable.

Member of: Time_Range_ET

Cardinality: Single, Required

Data Type: Real

et_stop – attribute

The latest ephemeris time, as defined by the referenced field, for which this state is applicable.

Member of: Time_Range_ET

Cardinality: Single, Required

Data Type: Real

First_Count – class

Defines the records to which this state applies by identifying the first record (inclusive), and the number of contiguous records, including the first.

Member of: Applicable_Records

Cardinality: Single, Optional

Members:

- first_record
- record_count

First_Last – class

Defines the first and last records for which this state applies.

The record numbers are one-based and inclusive.

Member of: Applicable_Records

Cardinality: Single, Optional

Members:

- first_record
- last_record

first_record – attribute

The first record for which the state applies, inclusive.

Member of: First_Count, First_Last

Cardinality: Single, Required

Data Type: Integer

Minimum Value: 1

Rules: When used in First_Last, first_record must be less than last_record.

Instrument_Settings – class

This class specifies instrument settings. Instrument settings may be driven via a state table, which may be associated with specific observations within the data product, either by

directly indicating those records, or indicating the observation times at which certain states apply.

Member of: NucSpec_Observation_Properties

Cardinality: Single, Optional

Members:

- State_Table

last_record – attribute

The last record for which this state applies, inclusive.

Cardinality: Single, Required

Data Type: Integer

Minimum Value: 1

Rules:

- last_record must be greater than first_record
- last_record must be less than or equal to the number of rows in the data table referenced in State_Table_Entry

name – attribute

The name of the detector.

Member of: Detectors

Cardinality: Single, Required

Data Type: String

NucSpec_Observation_Properties – class

This class specifies properties pertaining to a nuclear spectroscopy observation.

Member of: Discipline_Area

Cardinality: Single, Required

Members:

- Energy_Calibration
- Instrument_Settings

- Detectors

Observing_Condition - class

Describes the time range when an observing condtion is described, along with a source and/or an activity indicator.

Member of: Observing Conditions

Cardinality: Multiple, Required

Members:

- condition_active
- condition_source
- Time_Range_SCLK
- Time_Range_UTC
- Time_Range_ET

Observing_Conditions - class

Describes a conditon that affects the obsevation in some way, whether internal or external. Includes a list of time ranges, along with activity indicators.

Member of: Nucspec_Observation_Properties

Cardinality: Multiple, Optional

order – attribute

The order of a single term of the polynomial relating channel number to pulse height.

Member of: Polynomial_Term

Cardinality: Single, Required

Data Type: Integer

Polynomial – class

The pulse height in energy units (e.g. MeV) is determined for a given channel number using the polynomial: H(x) = a0 + a1 x + a2 x^2 + …, where H is pulse height with energy units and x is channel number.

Member of: Energy_Calibration

Cardinality: Single, Required

Members:

- Polynomial_Term

Polynomial_Term – class

Describes a single term in the polynomial function.

Member of: Polynomial

Cardinality: Multiple, Required

Members:

- order
- coefficient

record_count – attribute

The number of consecutive records, including the first, for which this state applies.

Member of: First_Count

Cardinality: Single, Required

Data Type: Integer

Minimum Value: 1

Rules: The last_record, determined by first_record + record_count - 1, must be less than or equal to the number or records in the table reference by State_Table_Entry.

sclk_start_time – attribute

The earliest spacecraft clock time, in ticks since spacecraft clock start, for which this state is applicable.

Member of: Time_Range_SCLK

Cardinality: Single, Required

Data Type: Integer

sclk_stop_time – attribute

The latest spacecraft clock time, in ticks since spacecraft clock start, for which this state is applicable.

Member of: Time_Range_SCLK

Cardinality: Single, Required

Data Type: Integer

state_index – attribute

The state_index points to a row of a look up table containing the instrument settings.

Member of: State_Table_Entry

Cardinality: Single, Required

Data Type: Integer

State_Table – class

Specifies an additional look-up table that provides additional information about specific observations within the product.

Member of: Instrument_Settings

Cardinality: Single, Required

Members:

- Internal_Reference
- State_Table_Entry

State_Table_Entry – class

Associates individual observations in the data product with a record in the state table, providing additional information on those observations.

Member of: State_Table

Cardinality: Multiple, Required

Members:

- state_index
- Local_Internal_Reference
- Applicable_Records or State_Time

Rules:

- Local_Internal_Reference must reference a data table that exists in the product.
- When Applicable_Records is used, Local_Internal_Reference must refer to a table with a record count.

State_Time – class

The time during which this state was applicable. The state table entry is applicable to records in the referenced data table if the time of the data record, as given by state_time_field_name, is between the start time and stop time given in State_Time.

Member of: State_Table_Entry

Cardinality: Single, Choice (choose between Applicable_Records and State_Time)

Members:

- state_time_field_name
- Time_Range_ET or Time_Range_SCLK or Time_Range_UTC

Rules: state_time_field_name must reference a field that exists in the data table referenced in State_Table_Entry

state_time_field_name – attribute

The name of the field in the referenced data table that contains the time.

Member of: State_Time

Cardinality: Single, Required

Data Type: String

Rules: state_time_field_name must refer to a field that exists in the data table referenced in State_Table_Entry.

Time_Range_ET – class

Ephemeris time, as defined by the referenced field, for the first and last science data records for a temporally contiguous block of records with the same state. These records need not be contiguous within the file, but there must be no records with a different state between the start and stop time.

Member of: State_Time, Observing_Condition

Cardinality: Single, Choice (choose between Time_Range_ET, Time_Range_SCLK, and Time_Range_UTC)

Members:

- et_start

- et_stop

Time_Range_SCLK – class

Spacecraft clock time in ticks since spacecraft clock start for the first and last science data records for a temporally contiguous block of records with the same state. These records need not be contiguous within the file, but there must be no records with a different state between the start and stop time.

Member of: State_Time, Observing_Condition

Cardinality: Single, Choice (choose between Time_Range_ET, Time_Range_SCLK, and Time_Range_UTC)

Members:

- sclk_start_time
- sclk_stop_time

Time_Range_UTC – class

Universal Coordinated Time for the first and last science data records for a temporally contiguous block of records with the same state. These records need not be contiguous within the file, but there must be no records with a different state between the start and stop time.

Member of: State_Time, Observing_Condition

Cardinality: Single, Choice (choose between Time_Range_ET, Time_Range_SCLK, and Time_Range_UTC)

Members:

- utc_start_time
- utc_stop_time

utc_start_time – attribute

The earliest Universal Coordinated Time for which this state is applicable.

Member of: Time_Range_UTC

Cardinality: Single, Required

Data Type: Datetime

utc_stop_time – attribute

The latest Universal Coordinated Time for which this state is applicable.

Member of: Time_Range_UTC

Cardinality: Single, Required

Data Type: Datetime