

SEDRIS Data Coding Standard (SDCS)

Dr. Paul A. Birkel
The MITRE Corporation
1820 Dolley Madison Blvd.
McLean, VA 22102
(703) 883-6399
pbirkel@mitre.org

Keywords:

HLA, OMT, OMDD, FOM, SOM, SEDRIS, DDDS/DDM, JMCDM, UCDM, FACC, SNE, SE

ABSTRACT: *In order to support the unambiguous description of environmental data, the Synthetic Environment Data Representation and Interchange Specification (SEDRIS) provides both a Data Representation Model (SDRM) and a Data Coding Standard (SDCS). While the former addresses how to describe “environmental things” in terms of data modeling constructs meaningful to simulation developers (e.g., geometry, feature, image, topology, and data table), it explicitly avoids enumerating all of the “environmental things” which these data modeling constructs could be used to represent. The SEDRIS Data Coding Standard explicitly addresses the environmental “enumeration problem” in an extensible manner capable of meeting the needs of both pre-simulation and run-time data interchange using an approach compatible with major DoD initiatives in Data Dictionaries and Data Modeling (e.g. the JMCDM and UCDM). The SDCS provides a solid basis for extending the High Level Architecture (HLA) Object Model Data Dictionary (OMDD), and supporting community BOM/SOM/FOM-building activities, by providing a rigorous, extensible, internally consistent, and increasingly complete data dictionary from which to select environmental objects and attributes for use in simulation. SDCS version 2.1 includes not only support for the land, sea, air, and space domains, but also for localized fixed and mobile “things” of both military and civilian nature.*

1. Introduction

The Synthetic Environment Data Representation and Interchange Specification (SEDRIS) project is an ongoing activity sponsored by the Defense Modeling and Simulation Office (DMSO) in support of the authoritative representation of the physical environment within M&S [1]. SEDRIS is focused on solving the problems of achieving loss-less and unambiguous environmental data representation and interchange across the heterogeneous distributed simulation community.

The issue of environmental data representation and modeling is of great importance and wide interest regardless of how the simulation is constructed/operated (e.g., heterogeneous or homogeneous, distributed or not, multi-process or single). It also persists throughout the simulation life-cycle from pre-execution through initialization/execution to after-execution review and analysis.

Environmental data modeling is also an issue which transcends the bounds of M&S to include the operational military forces and their supporting information systems (i.e., C⁴ISR systems). As such it touches on DoD-wide efforts to develop and promulgate a standard data dictionary (the Defense Data Dictionary System – DDDS) and data model (the DoD Data Model – DDM).

Environmental data modeling is not just a military problem, either, but extends into the US National Information Infrastructure (NII) and transnational efforts to assess worldwide resources and environmentally damaging activities. Finally, it is a significant focus of commercial efforts to develop common environmental information processing capabilities (e.g., the Open GIS Consortium).

As a result, while the SEDRIS project is also addressing the complex problems of developing specific mechanisms for *interchanging environmental data*, the problems of *representing environmental data* are themselves quite far-ranging and potential solutions have significant ramifications outside of the DoD M&S community.

In this paper, we focus on the latter topic (environmental data representation), and in particular the problem of enumerating and specifying all of the environmental “things” which could become factors in a simulation (or other information processing application). Additional information on the SEDRIS project as a whole, and related topics not addressed herein, can be found at <http://www.sedris.org> [2]. Examples are drawn from the SEDRIS Developers Kit Version 2.0 release of January 1999.

We will address five topics:

1. How does the SEDRIS Data Representation Model (SDRM) relate to the SEDRIS Data Coding Standard (SDCS), and why are they separate?
2. What is the organizational structure of the SDCS and how do the components relate to standards external to SISO (both legacy and future)?
3. What are the extents, capabilities and contents of the SDCS?
4. How does the SDCS relate to the High Level Architecture (HLA) Object Model Data Dictionary (OMDD)?
5. What SDCS work remains to be accomplished?

We conclude with a recommendation for SISO to adopt the SEDRIS Data Coding Standard as an internal “product” for required use in BOM, SOM, and FOM development as part of the Federation Development and Execution Process (FEDEP) Model.

2. Environmental Data Representation

2.1 Data Representation Model

The SEDRIS Data Representation Model (SDRM) currently includes 332 data classes. This is clearly inadequate, taken alone, to represent the complexity of the physical environment, even in relatively abstract terms.

This apparent poverty in SEDRIS data modeling capability was intentional; it focuses attention on the repeated *types* of data elements and *inter-relationships between* different data elements. Commonalties are enhanced and differences removed, resulting in economy and clarity in data model representation.

For example, the SDRM can be conceptually reduced to five basic classes: geometry, feature, image, topology, and data table; plus the organizational concepts of *hierarchy* and *model*. There is much more to the SDRM, but given the composition of these seven concepts plus various elaborations, one can fairly quickly understand the basic SDRM.

For example, the **Feature** (abstract) class is elaborated via the **Primitive Feature** (abstract) subclass into the **Point Feature**, **Linear Feature**, and **Areal Feature** concrete classes. What is immediately apparent is that there is no place in the SDRM for buildings, roads, trees, cold fronts, warm-core eddies, or other environmental “things” with which we are familiar from everyday experience. The consequences of this observation are two-fold.

First, the concept of **Point Feature** (and the associated data model class) is defined independently of the ultimate type of feature which might be conceptually represented

as being at a spatial point location. All **Point Feature** data model objects can be composed of **Classification Data**, **Property Value**, and **Data Quality** objects (among others), regardless of what “thing” in the physical environment that object ultimately is used to represent.

In particular, the specification of a SDRM object is similar to the intent and usage of a High Level Architecture (HLA) Object Model Template (OMT) complex data type. For example, while it turns out that a **Point Feature** has no associated data fields (since as an object it merely composes/organizes other objects to form the representation of a “point feature”), a **GCC Location 3D** object contains three data fields¹, as follows:

```
float64  x;
float64  y;
float64  z;
```

The Real-time Platform Reference FOM (RPR-FOM) refers to a similar concept as a “PositionStruct” [3]. What the SDRM effectively provides is analogous to a Reference FOM for environmental data. It is a superset of a Reference FOM in that it allows for greater organizational complexity and inter-relationships than available using OMT notation. It is a subset of a Reference FOM in that it doesn’t (intentionally) “get down to the each’s”.

To bring this first observation to a close, we note that the SDRM is a Data *Representation* Model, that is, a data model about how to represent “things”, and not a data model “of the things”. Put another way, it is a *Meta-Data* Model.²

Our second observation is that a separate mechanism is required to determine what “thing” any particular **Point Feature** actually represents in a SEDRIS transmittal. Which brings us to the SDCS.

2.2 Data Coding Standard

The SEDRIS Data Coding Standard (SDCS) provides a mechanism to specify the environmental “things” which a particular data model construct is intended to represent. That is, a “tree” *could* be represented alternatively as a **Point Feature**, an **Aggregate Geometry**, a **Data Table**, a **Model**, or some combination of these. Which of these the data modeler (i.e., the data provider of the SEDRIS transmittal) chooses is orthogonal to the semantic of the “thing” that is represented – that is, *what it is the representation of*.

¹ But, as it happens, composes/organizes no other data model objects.

² Not to be confused with a metadata model!

In particular, the choice among the data models supported by the SDRM in Rumbaugh notation, a BOM/SOM/FOM in OMT notation, a Defense Information Systems Agency (DISA) conceptual data model in IDEF-1X notation, or some other data model and data modeling method is immaterial. What is important is the explicit decision to clarify the semantics surrounding the specification of the “thing”.

The provision of such a “thing” in a SEDRIS transmittal pre-simulation, or publication of such a “thing” using the RTI at run-time, must result in a shared understanding of “what the thing is and what it potentially means” to all participating applications.

3. Data Coding Standard Structure

3.1 Component Overview

What is the SDCS, then? It is a triumvirate of three inter-related sets of codes intended to provide a complete suite of capabilities to unambiguously state “what this is”, including not only its inherent/Platonic “what”, but also modifiers which may represent clarifying information or information which may change over time.

We therefore divide the SDCS into three components:

1. SEDRIS Classification/Feature Codes (SCC),
2. SEDRIS Attribute Codes (SAC), and
3. SEDRIS State Codes (SSC).

Classification/Feature codes address the “what is it” question. It is a building (“AL015”), a river/stream (“BH140”), an air warning light (“GA065”), or perhaps an ocean floor area (“OF010”).

Attribute codes (and associated values) address the “additional clarifying characteristics” question. Perhaps the building functional category is “lighthouse”, the river-stream “depth” is 1.5 meters, the air warning light color category is “red”, or the ocean floor area has a bottom material composition of “coral”.

State codes address the “how it may deviate from normality” question. The lighthouse may have sustained 30% damage, or the air warning light may be switched off (or only partially illuminated).

Working together, these three SDCS components support the unambiguous description of “what it is” for each environmental “thing” of interest. For example, the **Classification Data** component of the **Feature** object “holds” a classification/feature code and the **Property Value** com-

ponents of the **Feature** object each “holds” an attribute code (and value).³

State codes are used in conjunction with **State Related Geometry** and **State Related Feature** objects. Further discussion of the details of these SDRM objects is outside the scope of this paper.

3.2 External Standards

The SDCS was developed while looking both forward to future DISA-sponsored DoD standards for environmental data, as well as backward to legacy environmental data and data models. While DISA is currently focused on developing *conceptual*, rather than logical or physical, data models, SEDRIS is necessarily focused on the *logical* and *physical* levels. The former to capture semantics and organization; the later to support physical data interchange and the SEDRIS Transmittal Format (STF).

3.2.1 Future Standards

The most important emerging standards are those with either military, international, or commercial scope.

DISA, under the Defense Data Dictionary System (DDDS) and DoD Data Model (DDM) activities, is leading military efforts. The DDM uses only terms as specified in the DDDS; there they have a standard name, definition, and associated “DDDS Counter”. The counter is not maintained as a unique keyword; only the standard name is guaranteed to be invariant (while the DDDS continues to steadily evolve).

The DDM uses IDEF-1X notation and, although conceptual, specifies exactly each environmental “thing” included in the data model – one would, for example, find a “bridge” in the DDM, and all of the allowed relationships between bridges and other DDM data elements. This tightly constrained *conceptual* data model is considered a “work in progress”; therefore there is currently no available environmental data which is accessible according to this data model.

There are two environmental components of the DDM. The USIGS⁴ Conceptual Data Model (UCDM) addresses terrain-related (specifically mapping, charting, geodesy and imagery) data elements⁵, whereas the Joint METOC⁶

³ And the **Data Quality** component of the **Feature** object is used to hold lineage, accuracy and other feature-specific metadata.

⁴ US Imagery and Geospatial Services.

⁵ Note that NIMA has no current plans to extend their scope to include characterizations of terrain-related data

Conceptual Data Model (JMCDM) addresses the air, sea, and space data elements. These two conceptual data models overlap and are inconsistent as regards spatial referencing data elements and related items. They will eventually be harmonized.

The JMCDM is in a state of continuous evolution, without version control or a formal release process as yet. A snapshot of the JMCDM was taken 8-11-98 and used to rationalize the SDCS METOC component. Interestingly, many of the current SEDRIS METOC data requirements were not yet covered by the JMCDM despite the fact that the JMCDM is believed to be "almost complete" as regards its conceptual modeling phase. Initial discussions with the JMCDM development team led to an agreement to eventual extensions to the JMCDM to accommodate substantial portions of the SDCS METOC component.

The UCDM is also undergoing continual revision, although starting from a known base – the Digital Geographic Information Exchange Standard (DIGEST) Feature and Attribute Coding Catalog (FACC) – an international standard [4]. NIMA maintains a backward mapping from the UCDM (as it evolves) to DIGEST FACC since FACC is the current NIMA standard data-coding scheme for all of their vector-based products. Since FACC is already being addressed by SEDRIS as a current standard and NIMA has no plans to provide data according to the UCDM, the SEDRIS project has no current plans to establish specific mappings between the SDCS terrain component and the UCDM.

The primary international activity in this area is the International Organization for Standardization and the International Electrotechnical Commission (ISO/IEC) TC211 committee on Geographic Information/Geomatics (see <http://www.statkart.no/isotc211/pow.htm>). Although currently focused on defining a framework for the development of profiles, their general approach is FACC-based.

The primary commercial activity in this area is that of the OpenGIS Consortium, and their Abstract and Implementation Specifications for Simple Features (see <http://www.opengis.org/techno/specs.htm>). As the name states, both their data model and data coding scheme are relatively impoverished at this time, although the SEDRIS project has been in communication with the OGC in this area.

3.2.2 Current Standards

While there are many current standards for narrowly-defined environmental data exchange which include a data coding component, those of most importance to SEDRIS are the ones which represent significant repositories, or providers, of environmental data. Chief among these are the Master Environmental Library (see <http://mel.dmsomil>), currently using the World Meteorologic Organization (WMO) GRIB and BUFR formats (respectively: *FM 92-VIII Ext. GRIB (GRIdded Binary)*, and *Binary Universal Form for the Representation of meteorological data*) for METOC data. Unfortunately, these provide relatively restricted data coding mechanisms. Since the MEL will be supporting SEDRIS as an additional transmittal mechanism in FY99, we have ensured that the SDCS supports a superset of the GRIB/BUFR coding capability.

The other primary current environmental data coding standard is the DIGEST FACC, currently in release as Edition 2.0, and dated June 1997 (see <http://www.digest.org/overview.html>). Edition 2.0a is expected to be released February 1999, and the SDCS was baselined using a pre-release of Edition 2.0a. It is intended that the SDCS will incorporate future changes to FACC, thus ensuring upward compatibility with NIMA-provided environmental data. The DIGEST FACC committee has been briefed on the SDCS, and NIMA has agreed to submit SDCS codes for terrain-related data back to the DIGEST FACC committee for inclusion in future FACC revisions.

A related standard, IEEE 1278.1A, encompasses enumerations for military models, whose physical/structural characteristics are often included as part of pre-simulation environmental data transmittals. The Institute for Simulation and Training (IST) maintains a companion document for these enumerations [5]. These enumerations have been re-organized and included as part of the DCS. Mappings from these DCS data items to the DIS Standard are maintained for the convenience of simulation developers and bi-annual updates to the IST enumeration document will be tracked.

4. SEDRIS DCS Details

4.1 Classification/Feature Codes

The SEDRIS Classification/Feature Codes (SCC) are based on the DIGEST FACC coding scheme for features, which defines a 5 ASCII character string divided into two

elements suitable for the support of sensor models, e.g., thermal modeling and imaging.

⁶ METeorology and OCeanography.

alphabetic⁷ characters followed by three numeric characters. The first two (alphabetic) characters are interpreted as the **Category** and **Subcategory**, respectively. Subcategory divisions vary across categories. Character codes are assigned sequentially, for the most part, and have little relationship to the category/subcategory name.

The remaining three (digit) characters are interpreted as the feature index within the appropriate subcategory. By convention, features are initially assigned indices at intervals of 10 in some moderately intuitive order, and then infilled as required (consistent with the original ordering) for new features. Limited formal rules apply to the assignment of codes for new features and reverse engineering the current feature organization does not necessarily lead to complete consistency.

While the first two alphabetic characters provide a degree of hierarchical organization, within a specified subcategory (e.g., “AA” meaning “Culture – Extraction”) the organization is non-hierarchical. Features rarely have been assigned indices outside of the first 200 or so (e.g., “GB220” Air Obstruction) of the 999 possible. There are currently 57 category/subcategory combinations and 472 features, or less than ten features in each subcategory. Some linear search may be required to locate an appropriate code.

Within these constraints, two major extensions have been made in developing the SCC. First, additional categories have been added to cover other environmental domains than just terrain (more specifically, mapping & charting, the current scope of the FACC). These now include:

- A: Culture
- B: Hydrography
- C: Hypsography
- D: Physiography
- E: Vegetation
- F: Demarcation
- G: Aeronautical Information
- I: Cadastral
- **M: Mobile**
- **O: Ocean**
- S: Special Use
- **W: Weather and Space**
- Z: General

⁷ Alphabetic characters are restricted to the range from “A” to “Z”, inclusive. Although DIGEST FACC is silent on the subject of case-sensitivity, convention dictates the use of upper case only.

Additional subcategories have also been added, e.g., within the Ocean category, the following subcategories have been defined:

- **OA:** Ocean – Acoustic
- **OB:** Ocean – Biologic
- **OF:** Ocean – Floor
- **OS:** Ocean – Surface
- **OV:** Ocean – Volume
- **OZ:** Ocean – Abstract

Finally, additional codes have been assigned, as consistently with the original DIGEST rules and conventions as possible, but taking the opportunity to improve consistency where appropriate. The SCC now includes about 700 items, or a 50% expansion compared to the number of FACC feature codes.

The second major extension to FACC feature codes is that additional codes have been added which do not represent physical environment “things”, but rather groupings of environmental “things” – that is, classifications such as “terrain”, “vegetation”, “ocean surface”. This supports the **Classification Related Geometry** and **Classification Related Feature** data model objects. To emphasize this extension to the concept of specifying codes only for pure “real world” features, the SCC is intentionally named for the more general classification concept rather than the narrower feature concept.

Finally, the SSC is explicitly case insensitive (that is “s” = “S”), however the SEDRIS usage rules state that only upper case should be used for clarity.

4.2 Attribute Codes

The SEDRIS Attribute Codes (SAC) are based on the DIGEST FACC coding scheme for attributes, which defines a 3 ASCII character string⁸ divided into two alphabetic characters followed by one alphanumeric character. Occasionally, the third character may actually be numeric, but in most cases it is alphabetic. Codes are generally assigned based on mnemonic abbreviations (e.g., acronyms) of the attribute name (e.g., “BFC” for Building Function Category, or “HGT” for Height Above Surface Level). As a result they are lexicographically ordered and locating an appropriate code typically requires a linear search until the user learns the codes by rote. Limited

⁸ As was the case for FACC feature codes, alphabetic characters are restricted to the range from “A” to “Z”, inclusive. Although DIGEST FACC is silent on the subject of case-sensitivity, convention dictates the use of upper case only.

formal rules apply to the assignment of codes for new attributes.

Unlike the FACC feature codes, which simply specify a name and associated definition, FACC attribute codes also include specifications for:

- Units (e.g., feet, nautical miles, or meters-per-second),
- Format (i.e., Alphanumeric, Integer, Lexical, Real Number, or Structured Text),
- Range (e.g., 0.0 to 100.0),
- Increment (e.g., 0.1 meters), and
- Maximum number of characters (for lexical types).

As a result of this integration of attribute semantics with specific representational details, occasions are arising with increasing frequency where essentially duplicated attribute codes are being introduced in order to support greater precision in representation. For example, migration from a code requiring units in Integer format to one requiring Real Number format, such as the transition from LEN (Length/Diameter) to YLN (Length of Greater Precision). Additionally, name space collisions (and only three characters to work with) are beginning to result in non- (if not occasionally counter-) intuitive codes.

Within these constraints, three major extensions have been made in developing the SAC. First, allowing an additional character – four characters in all – has significantly increased the name space for new Attribute Codes. To support upward compatibility with current FACC attribute codes, all DIGEST FACC 2.0a attribute codes simply have an underscore (“_”) character added to their tail. New terrain-related attribute codes are fitted into a three-character sequence wherever appropriate in order to support potential migration to FACC at a future date. Where the attribute is not terrain-related, or a reasonable three-character code would not be meaningful, a “full” four-character code (i.e., one not using a terminating underscore) is selected. Therefore, whereas “THI_” might code the Thickness attribute, “TILB” would code the Temperature Inversion Layer Bottom Altitude attribute. The SAC now includes about 1400 items, or a 300% expansion compared to the number of FACC attribute codes. There are also about 18,000 enumerated values, or a 350% expansion compared to the number of FACC enumerated attribute values.

The second major extension to FACC attribute codes is that as codes are added, an effort is being made to organize their names and codes to increase lexicographic adjacency of semantically similar attributes, and thus improve ease-of-use. Thus, for example, we see the following set of new attributes:

- **WHS_**: Wave Height, Swell – Mean
- **WHSI**: Wave Height, Significant
- **WHSS**: Wave Height, Swell – Significant
- **WHV_**: Wave Height, Visual
- **WHW_**: Wave Height, Wind Induced – Mean
- **WHWA**: Wave Height, Wind Induced – Sig.
- **WHX_**: Wave Height, Maximum

While a lexicographic sort on the attribute codes will not necessarily result in immediate adjacency, a sort on the attribute names not only results in immediate adjacency, but also clarifies the semantic inter-relationships between similar terms.

The third, and final, major extension to FACC attribute codes is that SEDRIS Attribute Codes are restricted to coding for semantics, *only*. While default value-associated information like units, data type, and range are provided as suggestions to data providers and to set expectations for data consumers, the DCS provides for explicit exchange of such metadata as part of the exchange of attribute data. Thus, for example, any given transmittal may use the “BRF_” Broadcast Frequency (“Broadcast frequency of a communications device.”) attribute, but may choose to use a variety of units (e.g., Hz or MHz)⁹ and data types (e.g., int16, int32, float64) at different times and for different purposes within the same transmittal. This avoids the LEN/YLN coding problem and is consistent with the current thinking of the DIGEST FACC committee. Neither the JMCDM nor the UCDM, however, are adopting this approach.

Finally, the SAC is explicitly case insensitive (that is “s” = “S”), however the SEDRIS usage rules state that only upper case should be used for clarity.

4.3 State Codes

The SEDRIS State Codes (SSC) do not have an equivalent component in the DIGEST FACC, but rather are a specific proper subset of the SEDRIS Attribute Codes which are semantically appropriate for use in conjunction with **State Related Geometry** and **State Related Feature** SDRM objects. These objects support SDRM organizations based on State values. Example State Codes include:

⁹ Note that the DCS explicitly adopts the conventions of International System of Units, although some “anachronisms” remain for backward compatibility (see <http://physics.nist.gov/cuu/Units/index.html>).

- **DGEN:** Damage, General
- **DMAN:** Damage, Maneuver
- **DMIS:** Damage, Mission-specific
- **PALG:** Position, Aircraft Landing Gear
- **PAPR:** Position, Antenna – Primary
- **PASE:** Position, Antenna – Secondary
- **PCPH:** Position, Canopy/Primary Hatch
- **PDET:** Position, Deployable/Extension Tent
- **PDPR:** Position, Door – Primary
- **PDSE:** Position, Door – Secondary

The SSC, as a proper subset of the SAC, shares all of the aforementioned SAC capabilities. There are currently approximately 45 State Codes.

4.4 Practical Matters

The SDCS is managed as an integrated multi-table Microsoft Access® 97 database incorporating intra- and inter-table constraint checking, version control, and relational links to tables incorporating external standards. Currently, both the JMCDM (as of 8-11-98; see also <http://www.cnmoc.navy.mil/da/jmcdm.htm>) and the relevant IEEE Standard 1278.1A-associated enumerations and bit-encoded values are included [5]. The database is currently approximately 7 Megabytes in size.

Automation has been added to produce a variety of reports for inclusion in SEDRIS documentation, as well as export to Microsoft Word® and Excel® formats as required for review of specific sections of the database.

Additionally, a Visual Basic® application has been developed which directly accesses the SDCS database and automatically generates a collection of ANSI C files (both “.c” and “.h”) for use by application developers, as illustrated in Figure 1.¹⁰ This ensures both excellent on-line documentation immediately available to software developers, but also encourages the use of standard constants for all codes (and enumerations) in developed software source files that improve readability and reduce errors. It also potentially isolates developers from future changes in data codes where the standard constants remain unchanged. This allows for migration to new SDCS versions with only a straightforward recompilation using the updated ANSI C files and unchanged application source code. Similar applications could easily be developed for other target languages than ANSI C (e.g., Java or Ada95).

```

/*
 * GLOBAL VARIABLE (Label): SE_SCC_BUILDING
 *
 * Name:
 *   Building
 *
 * Description:
 *   A relatively permanent structure, roofed and usually walled and
 *   designed for some particular use. (See also AL100.)
 *
 * Rationale:
 *   NULL
 */

const SE_SCC_ID SE_SCC_BUILDING={{'A','L','0','1','5'}};

/*
 * GLOBAL VARIABLE (Label): SE_SAC_BUILDING_FUNCTION_CATEGORY
 *
 * Name:
 *   Building Function Category
 *
 * Description:
 *   Type or purpose of the building.
 *
 * Source: DIGEST Version 2.0
 *
 * Rationale:
 *   NULL
 */

const SE_SAC_ID SE_SAC_BUILDING_FUNCTION_CATEGORY={{'B','F','C','_'}};

/*
 * GLOBAL VARIABLE (Label): SE_PROP_VAL_BFC_LIGHTHOUSE
 *
 * ENUMERATOR FOR: BFC_
 *
 * Description:
 *   Light house
 *
 * Source: DIGEST Version 2.0
 *
 * Rationale:
 *   NULL
 */

const SE_UINT16 SE_PROP_VAL_BFC_LIGHTHOUSE= 82;

```

Figure 1: Example ANSI C File Extracts

5. SDCS and the OMDD

The DMSO Federation Development and Execution Process (FEDEP) Model [6] calls for the use of an Object Model Data Dictionary (OMDD) to support the specification of SOMs and FOMs. An OMDD System (OMDDS) has been prototyped [7] and is now operational (see <http://hla.dmso.mil>). How does the SDCS relate to the OMDD?

The SDCS, like the OMDD, is designed to assist M&S developers by specifying common representations of data used within a federation that may at some point be shared with other federates. Whereas the OMDD is intended to support runtime data exchange via the RTI, the SDCS specifically supports pre-runtime data exchange via SEDRIS.

¹⁰ Note that these files work correctly with standard C++ compilers/linkers on multiple platforms as well as in pure ANSI C environments.

Due to the nature of environmental data, i.e., voluminous and varied but generally slowly changing, the SDCS is currently much larger than the environmental content of the OMDD *Object Classes*, *Generic Element Lexicon*, and *Enumerated Data Types*. This is to be expected, as most types of environmental data will be unchanging during the course of a simulation execution and therefore unaddressed by the OMDD and most SOMs/FOMs.

On the other hand, the OMDD specifies *Complex Data Types* and *Generic Element Representations*, which in SEDRIS are addressed, separately, by the SDRM. In particular, the SDRM provides many well-defined complex data types either directly via the 332 defined data classes, or indirectly via the use of organized class instances in a SEDRIS transmittal. Such organizations, while currently providing significant flexibility to data providers, will need to become more formalized and constrained in the future, much as in the manner of FOMs.

Additionally, the SDCS provides a “looser coupling” between generic elements represented as attributes and their data types, than does the OMDD. Whereas the OMDD specifies only a few possible data types per generic element; the SDCS provides “recommend defaults” but defers to the actual use of the SDRM at transmittal-time to specify the exact data type (and units, etc.) used.

Finally, since the SDCS supports both *identification* and *classification*, whereas the OMDD only supports *identification*, the SDCS generally provides explicit organizational information which can only be inferred from the OMDD.

6. Next Steps

What SDCS work remains to be accomplished? In many respects, completing the SDCS enumerations (on which much work remains to be done) is but another step on the road to achieving the SEDRIS objective – loss-less and unambiguous environmental data representation and interchange. Much as the OMDD is preparatory to the work of defining a FOM, the SDCS and SDRM are preparatory to the work of specifying an analogous environmental data *product* – an exact description of a specific SEDRIS transmittal.

In order to assist in the preparation of that exact specification, additional shared information is required, e.g.:

- **Defined relationships between different portions of the SDCS:** For example, for each SCC, is there a standard set of SACs which should be populated? Or which delineates the only applicable SACs? E.g., should all “trees” have a “height” and “trunk diameter”? Is it an error if these values are not present in a

transmittal? Is it an error if a “tree” has a (e.g.) “depth”?

- **Standard encoding rules:** For example, given the choice of a narrowly-defined SCC and a broadly-defined one augmented by one or more SACs, which is to be preferred? E.g., a “lighthouse” feature *vs.* a generic “building” with a functional category attribute of “lighthouse”? Under what conditions is one or the other an error? Are some SACs preferred when multiple SACs are applicable (e.g., LEN_ and YLN_)? Similarly for SCCs?
- **Standard organizational schemes:** While conceptually there is an end to the *identification* problem, it is not clear that the *classification* problem will be so easily solved. While the SDRM provides the mechanisms to implement a wide variety of organizational schemes, the SDCS must support specific enumerations applicable to the many classification or organizational approaches in use in the environmental community.¹¹ For example, does one organize models of vehicles by national producer (US, CIS, ...), primary domain of operation (ground, sea surface, subsurface, ...), motive power (electric, internal combustion, external combustion, ...), surface structural characteristics (has moving parts, has no moveable parts, has no parts, ...)? All must be potentially supportable by the SDCS, in order to support efficient search and retrieval, yet the potential list could become quite extensive.

These are all challenges under current investigation.

7. Summary and Recommendations

The SEDRIS Data Coding Standard has the expressive power to help the SISO community focus its efforts in the area of environmental data modeling as captured in BOM, SOM, and FOM development as part of the FEDEP Model. In particular, it appears to offer the opportunity for future versions of the RPR-FOM to better codify its “environmental aspects”, as well as to support FOM (or equivalent data model) development in new simulations such as JWARS and JSIMS. The JSIMS Terrain Common Data Model is already using the SDCS, and as a result the SDCS has benefited from requirements input from the aggregated-level constructive simulation community.

Like SEDRIS as a whole, the SEDRIS Data Coding Standard is a work still in-progress; one based on focused, productive community development. It is not yet fully

¹¹ This issue is analogous to that of *interest management* at runtime.

mature however it already includes international participation, and ties to US DoD legacy and next-generation data modeling standards.

By providing a rigorous, extensible, internally consistent, and increasingly complete data dictionary from which to select environmental objects and attributes for use in simulation, the SDCS offers SISO an opportunity to “stand on the shoulders of giants” as regards environmental data dictionaries. Additionally, Version 2.1 of the SDCS includes not only support for the full range of land, sea, air, and space environmental domains, but also support for localized fixed and mobile “things” of both military and civilian nature.

We propose that the SDCS be adopted as a SISO-internal standard “product”, separate from other possible SEDRIS standardization activities, and a nomination has been prepared and submitted to the Standards Activity Committee.

8. References

- [1] Department of Defense, Under Secretary of Defense for Acquisition and Technology. "Modeling and Simulation (M&S) Master Plan", October 1995.
- [2] Horan, B. and Cox, R., “SEDRIS Past, Present, and Future – Technically Speaking”, In *Proceedings of the Spring Simulation Interoperability Workshop*, March 1999, 99S-SIW-057.
- [3] Shanks, G., “Real-time Platform Reference FOM: Version 0.6”, Simulation Interoperability Standards Organization, October 9, 1998.
- [4] Directorate of Geomatics, Department of National Defense, Canada. "The Digital Geographic Information Exchange Standard (DIGEST)", Edition 2.0, June 1997.
- [5] Brann, J., “Enumeration and Bit Encoded Values for Use with Protocols for Distributed Interactive Simulation Applications”, Institute for Training and Simulation, University of Central Florida, IST-CF-98-07, June 20, 1998.
- [6] Defense Modeling and Simulation Office, “High Level Architecture Federation Development and Execution Process (FEDEP) Model, Version 1.3”, December 9, 1998.
- [7] Hammond, J. et al., “Populating the HLA Object Model Data Dictionary – A Bottom Up Approach”, In *Proceedings of the Spring Simulation Interoperability Workshop*, March 1998, 98S-SIW-075.

Author Biography

PAUL A. BIRKEL is a senior principal scientist for The MITRE Corporation. He currently provides technical support and guidance to the DARPA Synthetic Theater of War (STOW) program in the area of environmental modeling, the DMSO Synthetic Environment Data Representation and Interchange Specification (SEDRIS) project, and the Joint Simulation System (JSIMS) SNE Working IPT. In addition he provides guidance and support as needed to other DoD M&S and C⁴ISR programs via the US Army Topographic Engineering Center (USATEC).