

Synthetic Environment Databases

Michael R. Welch

Orion Development Group, Inc.

1. Introduction

1.1 Purpose

A Synthetic Environment database, as used by the M&S community, is different from the conventional “software” database. This document serves to differentiate between the two similar but uniquely different concepts and to fully define what a Synthetic Environment database is in SEDRIS terms.

This document is intended to coalesce existing ideas already presented by the SEDRIS team. When completed, this document will become Chapter 3 of the Synthetic Environment Domain Description.

1.2 References

1. SEDRIS Presentation made on 2 July 1997
2. Synthetic Environment data representation and Interchange Specification (SEDRIS) Rationale Document, Version Draft 0.2, 20 June 1996
3. FAQest Transcript, STRICOM, 23-24 July 1997
4. DoD 5000.59-P, DMSO Modeling and Simulation (M&S) Master Plan, 17 October 1995
5. CCTT Synthetic Environment Database Generation Process

2. Definition of a Database

2.1 What an SE Database Is and Is Not

“Database” is an overloaded term. [3] Because of their wide usage in the computer science world, the stand-alone flat file and relational databases have lent a common definition to the term. A database is generally thought of as tables comprised of rows and columns that can be associated with each other. The contents of the resulting indexed cells are represented by numbers and alphanumeric strings. Tables can be linked by relationships between their data elements. The entire set of tables, the database, can be queried using a language-like scheme to access the data.

A synthetic environment database is not an RDBMS like Access or Oracle; a data repository with supporting database management capability. A synthetic environment database is much more! It is a complex, integrated set of objects which define and describe a natural environment. The contained data elements each describe some aspect of a geographical region

DRAFT

and the elements or events expected to occur there. Additionally, a synthetic environment database encapsulates the relationships between the data elements.

The basis for a complex synthetic environment database is a meta-data description of the multiple types of data which the database can support. Meta-data is information which describes the characteristics of data and its organization and relationships. [4] The meta-data provides a model of the data contained in a synthetic environment database rather than a physical layout of the data's format. A data model defines what the separate data objects are and how they relate to each other. A format specifies the actual bytes used for each entity and how they would be organized for storage on a disk. [3]

Databases which support visual-based simulations have been in existence for a long time. But, a synthetic environment database includes more than visual and radar (*i.e.*, non-optical sensor) information. In fact, two categories of data: warfighter systems (vehicles and weapons) and the natural environment [3] are contained in an environmental database. Since this database holds data that is related to other data, it is somewhat like a conventional database; but, the data is comprised of primitive data types used for representing environmental objects and features – not primitives like integer, float and string.

To support the diverse needs of the M&S community for a wide range of information describing the terrain surface, complex terrain features, objects with special 3-D model attributes and characteristics, and atmospheric and oceanographic features, [2] the synthetic environment database includes an integrated polymorphic representation of its various features:

- Cartographic,
- 3-D model instances,
- Terrain surfaces,
- Topologic.

Consuming applications, despite sharing a common geographical area requirement, often only need a subset of the entire data set. [2] Representational polymorphism enables the data consumer to extract the information as it relates to his needs; the database has encapsulated the data in multiple forms to serve multiple users.

2.2 Typical Data Elements

The most common element contained in current synthetic environment databases is terrain. Terrain includes a description of the surface and naturally occurring features (rivers) as well as man-made features such as permanent structures, roads, railroads and buildings. In general, the terrain does not include transient structures like vehicles. Terrain extends into the ocean to a depth of 10 meters. [3]

Terrain surface data is obtained from a source's digital elevation model. The Defense Mapping Agency (DMA) is the most used source of terrain data for DoD simulations. The DMA's digital terrain elevation data (DTED), interim terrain data (ITD), and digital feature analysis data (DFAD) are used to develop the terrain skin and terrain feature models.

DRAFT

Additionally, satellite SPOT imagery is often used to aid in correcting the alignment and positioning of terrain features.

The DTED is a digital elevation model characterized as a grid of elevation values for a specific geographic area. DTED level 1 data is based on a grid-post spacing of approximately 100-meters. DTED level 2 data has a grid post spacing of 1×1 arc-second in geographic areas between 50 degrees south latitude to 50 degrees north latitude. The grid-post spacing changes to 1×2 arc-seconds beyond the 50th parallels. The 1×2 arc-second spacing is approximately a 30-meter spacing (actually 31×38 meters) of grid posts which is considered to be sufficient fidelity for ground based simulations. [5]

The ITD was developed by DMA as an interim product for digital terrain analysis until the Tactical Terrain Data could be fielded in its final format. The ITD is the primary source data set for cultural/terrain feature models. The ITD is comprised of six thematic layers: vegetation, surface materials, slope, surface drainage, obstacles, and transportation. Each layer is contained in a segregated fileset. The filesets are divided into individual files containing data for a Global Coordinate System cell. Layers may contain point, linear, and areal features that correspond in resolution to a 1:50,000 scale map. Each feature type contains a comprehensive variety of attributes. [5]

The DFAD is another source of terrain feature and cultural data. DFAD information is used wherever the ITD information is non-existent or insufficient for certain desired cultural/terrain features. ITD is used as the primary source of cultural information except for urban cultural features. DFAD is used as the primary source for the urban areas and as supplemental, secondary source for all other features. [5]

Additional information for terrain models can be obtained from the DMA's various maps and charts of the regions of interest. [5]

- 1:50,000 Topographic Charts
- 1:100,000 Cross Country Charts
- 1:250,000 Military Engineering Geography Charts
- 1:500,000 Tactical Pilotage Charts

Beyond terrain data, a synthetic environment database contains information about 3-D models, surface textures, environmental phenomenon, object attributes and miscellaneous data.

Information for representing 3-D objects, both static and dynamic, is obtained from engineering drawings, blueprints, CAD data and photographs of the specific vehicles and personnel. Surface textures, specific images and colors are usually obtained from photos. Special effects, like detonations and weapons firing, use video clips as the source data. [5]

Environmental phenomenon, such as haze, fog, smoke and weather, are modeled by capturing their attributes of form, intervisibility, color, etc. Other object attributes are stored to indicate infrared signatures, surface material, electro-magnetic spectrum frequency, and pulse type and rate.

Finally, a miscellaneous collection of data is maintained to accommodate animations, system specific data, place names (and other labels), tables of values, etc.

3. Constraints on Making a Database

Development of a run-time synthetic environment database to support computerized simulations has a variety of constraints.

3.1 Application Requirements

Constraints levied by application requirements include the significant need for run-time access of the database. Access speeds must support human perception/reaction in real-time. This constraint may lead to subdividing the synthetic environment into a group of correlated databases.

A primary application constraint is imposed by the architecture of the trainer: stand-alone or networked. A stand-alone application controls its own environment. A networked application must provide for a variety of contingencies. The intent of a networked application is to co-join multiple training entities in the same synthetic environment. This multiplicity means that actions/positions of other entities cannot be predicted. In addition, the networked simulators may be homogeneous or heterogeneous with an inter-mixing of man-in-the-loop, SAF and CGF entities. Each of these entity variations has a potential impact on the configuration and content of the synthetic environment database.

The type of simulation (air, ground, near ground, sea, under sea) has a direct impact on the database extents, size, density and fidelity. Application specific emphasis (infantry versus aircraft) also impacts the attributes of the database. A singular simulation type (armor ground trainer) dictates a specific geographic extent, the density of detail and the fidelity of representation to achieve the required training results. These constraints would be significantly different from those imposed for a high altitude aircraft trainer. The combination of simulation types (infantry, armor, helicopter and combat support aircraft) requires a much larger database which contains environment information of differing levels of detail.

3.2 Intended Computational Platforms

Selection of computational platforms is constrained by real-world cost and budget limits. Visual database size requirements are usually determined by polygon counts – which is set by the capacity of the image generation hardware. Databases must be maintained on off-line storage devices. These devices are specified in terms of their capacity measured in M-bytes, G-bytes or T-bytes. The cost of such devices is directly proportional to the capacity; therefore, a realistic limit on overall database size is imposed by the hardware budget. The other gating size restriction is total processor memory. The more RAM which is available to the processor, the more efficiently it can perform the required operations on the synthetic environment data. But, the amount of RAM is also constrained by the hardware budget.

Overall computational capacity directly affects system performance. Trainees require interactive feedback to their input. To provide this feedback, the processor must be able to perform its computations in real-time. The amount of synthetic environment data which must be processed has a direct impact on achieving this requirement. When there is too much data to process, it can be reduced by selective thinning. Terrain surface detail is reduced through the use of “level-of-detail” polygon thinning. This thinning allows features near the eye-point to be in greater detail than features farther away. Other feature detail can be thinned in a similar manner. The end result is a visual scene with a less detailed representation of the synthetic environment, but within the capacity of the computer system to generate in real-time.

3.3 Project Budget Constraints

The budget for development of a synthetic environment database for a specific project will be affected by the availability of affordable technology and legacy or currently under-development databases.

Visions of the perfect simulation representation must always be constrained by the technology capability to generate the synthetic environment. Frequently, state-of-the-art technology may be available to create the “perfect” representation, but budget reality will always determine what can actually be accomplished.

Budgetary limits may be stretched through the use of existing synthetic environment data. Legacy databases which generally cover the project’s simulation domain may contain a good supply of initial source data which can be processed for project use. Additionally, parallel development of databases by other projects provide an opportunity to leverage a project’s time and cost budget in a positive direction.

4. Database Generation Steps

Synthetic environment database generation is a multi-step process ranging from requirements determination to compiling the run-time data sets.

4.1 Articulation and Definition of Requirements

The basic requirements are bounded by the questions: what type of training is to be performed and what geographic region is to be represented? Training type – air, ground, near ground, sea, under sea or a combination of types – determines fidelity levels while geographic region determines the extent of the environment.

Training interaction (single stand-alone entity, multiple entities of the same type or joint forces) controls what type of synthetic environment data will be required and will define possible multiple levels of detail and fidelity.

DRAFT

During the early stages of simulation specification, it is reasonable to expect a significant degree of fluidity in the definition of the synthetic environment requirements. Variety and modification should be planned for from the beginning.

4.2 Data Collection

The first step in data collection is to determine what data is available to satisfy the defined requirements. A search to identify possible raw data providers and maintainers of environmental databases with properties similar to the requirements should establish data availability. For each data source identified, the question “Is the data of the correct resolution to satisfy my requirements?” must be asked. Attempting to utilize data of improper resolution may well introduce more problems than it solves.

A variety of sources can be tapped to obtain the required environmental data. Section 2.2 describes the various types of data that might be included in a synthetic environment database:

- Digital elevation models from existing repositories such as the DMA,
- Maps and charts from government and private agencies,
- Textures and imagery from photos and video sources,
- 3-D model dimensions from engineering drawings, blueprints, CAD data and photographs
- Object attributes, observations, descriptions from textual compilations.

Additionally, existing SEDRIS transmittals may be a ready source for detail items or even significant sections of terrain.

4.3 Value Adding

Once initial data collection has been completed, the process of “adding value” relative to the planned simulation must be performed. This value adding step includes confirmation of data sources to ensure their correctness and possibly updating of the information based on the passage of time from original data recording to the time-frame of the simulation. Man-made features (and some natural features) change over time!

Where different resolution of data was used, corrections/conversions must be made to achieve a common resolution of all data that will be used together. This does not preclude the existence of detail with differing resolutions when presentation requires more or less fidelity. Anecdotal information may also be added to features contained in the database to support specific presentations (*e.g.*, labeling for 2-D maps).

4.4 Transforming and Tailoring the Data

The “close to ideal” database must now be tailored to meet the specific needs of the target simulation. Budget constraints which limit computational capacity, off-line data storage and image generation speed must be addressed by thinning details. Thinning generally reduces features and polygon counts. The level of detail is diminished while trying to achieve the maximum representation to support the simulation.

DRAFT

This tailoring of the underlying data may result in converting from grid representation to a Triangular Irregular Network (TIN) representation. It may also require a resampling of the source data to achieve a less stringent measurement interval.

4.5 “Assembling” the Database

When all the required data is collected, adjusted and tailored, it is then time to assemble the final environment. To accomplish this task, data from the various sources must be integrated into a single database. As the data is added, implied relationships must be recorded and any remaining constraints must be applied.

Models of 3-D features (buildings, bridges, etc.) must be built with the required level of detail and fidelity. To assist in movement determination, topology should be pre-computed and added to the database. This topology of face and edge association will make it easier to determine which portions of the terrain are roads, rivers, etc. As appropriate, textures for surfaces must be applied by association along with attributes that affect visibility and movement.

4.6 “Compiling” and Transmission of the Database

Finally, the database interface must be generated to support the specifically identified target platforms. Differing computer hardware will likely require that multiple instantiations of the database interface be compiled and distributed.

The database itself must be made available for interchange between the multiple users involved in the simulation. The interchange transmittal must be assembled into a predetermined format or with the use of SEDRIS it will be immediately ready for transmittal.

4.7 Database Generation and Interchange with SEDRIS

Figure 1 illustrates the use of SEDRIS transmittals to support the creation of multiple synthetic environment databases in various native formats. [1]

The left hand side of the figure shows the wide variety of source data that is encapsulated into SEDRIS transmittals by the Data Source Agencies. The source data is then extracted from the transmittals and converted for processing by Database Generation System 1 to create an integrated database in Native Format 1. Subsequently, the Format 1 database information is converted and encapsulated in another SEDRIS transmittal. This new transmittal is used to compile a CGF run-time database, a CIG run-time database, and a Sensor run-time database as well as to produce a set of maps. All of these runtime databases and maps are used by Training System 1.

Additionally, the same SEDRIS transmittal can be used to generate databases in other native formats for other training systems. Each new database generation system will convert the SEDRIS transmittal data for processing by their system. An integrated database in native format

is produced for each training system. From these integrated databases, run-time databases are compiled. The degree of correlation between the synthetic environments created in each training system should be extremely high since the same dataset was used to create each training system's run-time databases.

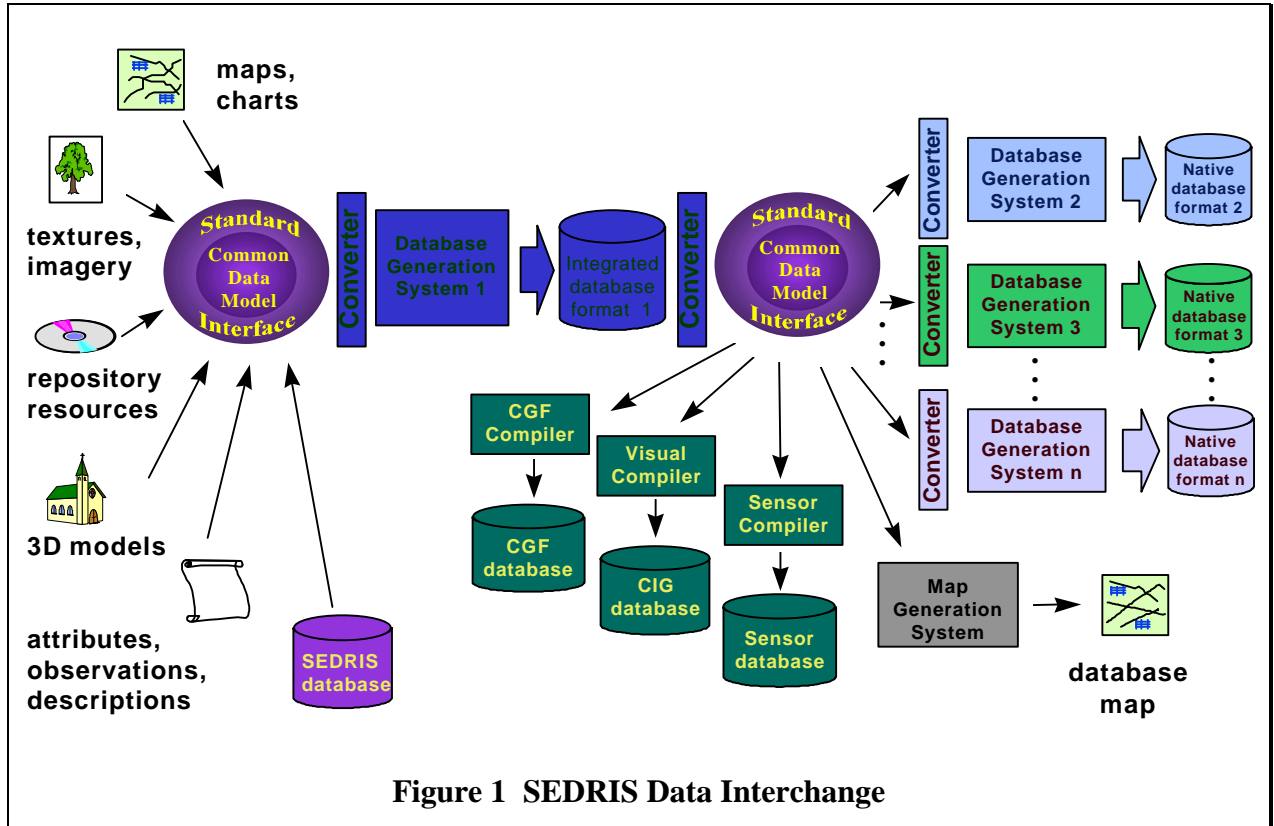


Figure 1 SEDRIS Data Interchange