"Physics of Auroral Phenomena", Proc. XXX Annual Seminar, Apatity, pp. 171 - 172, 2007

Polar Geophysical Institute

ON USAGE OF OBJECT-ORIENTED DATA STRUCTURE FOR STORAGE AND PROCESSING OF GEOPHYSICAL DATA

O.V. Martynenko[1], M.M. Gladkikh[1], I.V. Artamonov[1], D.V. Sobolev[2]

[1]*Murmansk State Technical University*
[2]*Polar Geophysical Institute*
MartynenkoOV@mstu.edu.ru

During the UAM model [Namgaladze et al., 1988, 1998] reorganisation into the Framework Upper Atmosphere Model [Martynenko et al., 2006], the object-oriented structure was developed for the processing and storage of geophysical data. The structure stores the data together with the following description: what the data are, where and how the data were obtained. It provides the opportunity to create the common unified program interface for automation of access and processing of any data irrespective of their origin, whether it is modelling or experimental. There is also a possibility to expand the description by including any additional information that characterises the stored data. The object-oriented approach that was used at structure design allows it to hold the software compatibility irrespective of the quantity of such additional characteristics: the unknown to the processing program information will be simply ignored, without any influence on the correct processing of known characteristics.

In the basis of the proposed way of storage and data processing lays the broadened concept of "data". We assume that the "data" are the numerical values of any physical characteristics of the object under study that refer to a certain place and time and are obtained by a certain method. We use the term Measurement to designate a method of obtaining the parameter value. But, it does not always correspond to the experimental measurement procedure in the common physical sense - it can also be the calculation by any model (which our system was initially created for). Correspondingly, the measured data description should include the detailed information about the physical characteristic: what units it is stored in, what place and time the stored value corresponds to, and what obtaining method was used. All this information is the mandatory data attributes in the developed structure. Besides, any information, which the system user considers to be appropriate for correct understanding of data, can be added to the structure.

The structure also considers such feature of the above-mentioned data information, that often it relates not only to one value of measured physical characteristic, but to many values at once. For example, all values of one characteristic are usually expressed in the same units and obtained by the same method. Furthermore, they are often measured in the same place and differ only by the time of measurement. On the contrary, there is often a whole set of the diverse characteristics measured simultaneously in one place, but using different methods and, of course, in different units, for example, by the experimental complex device or by the satellite equipment. For such data, the storage of repeated information with each parameter value leads to high redundancy. Therefore, the developed structure provides a method of the single record while keeping the link with all appropriate values.

At the program level the task is very close to a classical database. The differences include that very often the data are stored as extensive arrays of real numbers with relatively a small part of the text description. The common databases are intended for storage and data processing of directly opposite types: basically text and integer. Therefore the existing database software products appear ineffective for our task. It was the reason why we developed our own data storage system.

The central object of the developed structure (fig. 1) is the Model. Functionally each Model represents a method of obtaining the values of physical parameters in the spatial grid nodes, i.e. the experimental data sources fall also under this formal definition. Our usage of this term comes from original UAM model development and improvement leading to the framework system. It is possible to say that the Model is an instrument which performs the Measurement. Certainly, the Model characteristics, which can be complex enough, relate to all Measurements performed by it, and it is enough to store them all in one copy, and for measured data to store the reference to the measuring Model.

The mandatory Model properties are the set of measured parameters with all their characteristics (for example, unit measures, measure of inaccuracy, etc.) and position of nodes where Measurements were obtained. These properties are described using the formal specification rules. The Model (in our terminology) is obliged to return the value of each measured parameter in each spatial grid node after each Measurement procedure. To mitigate this requirement it is possible to set a special value of Model Attributes, which will be assigned to the node in case of data absence. But, in that case, the user must ensure that this concrete value would not appear accidentally as a result of a "correct" measurement.

The mandatory description of nodes position also comes from the origin of our system from the UAM model that uses numerical grid methods. In case of experimental data storage the grid can degenerate to singular point which, nevertheless, is characterised by a set of spatial coordinates expressed in some units, therefore the application of our structure for creation of the machine-readable data definition is quite proved. At the same time it is necessary to note, that grid dimensions not necessarily should be spatial dimensions in a strong sense, but, for example, correspond to instrument number in an array, to which measurements results will be written in the appropriate nodes.

The data itself represent multi-dimensional array of real numbers stored as a flat binary code without any explanations inside (the Data Set object on fig. 1). All internal structure of this array - its dimension, form, the description of each measurement - is stored in the description of the Model that generated this array. In the terminology defined above, the Data Set represents the result of one Measurement done by the Model. Actually it can vary from a unique value received as a one physical experiment result to the results of a whole modelling calculation - a sequence of a considerable quantity of instant distributions of the whole set of modelled physical parameters in nodes of three-dimensional or even multi-dimensional grid.

At the same time there is a possibility to supply each Data set with the special label of arbitrary structure that will distinguish it from the others. This is the alternative to creation of additional dimension in the description of the Model Spatial Grid. Such method should be applied, if such Set appears as the separate isolated object in the possible operations, and its data do not mix up with the data of other Sets and also when the amount of nodes in the given dimension is not known in advance (for example, the accumulation of Measurements during sequential starts of Model).

Usage of the described structure for data storage allows creating of the universal software for extracting and processing of geophysical data of the various content and origin. Recognition procedure automation (what kind of data this is and how they are organised) can considerably reduce the complexity of the analysis and data processing and protect from many errors.

Acknowledgments. This study was supported by the grant N 05-05-97511 of the Russian Foundation for Basic Research.

References

Martynenko O.V., M.M. Gladkikh, I.V. Artamonov. Application of the object-oriented approach in upper atmosphere modeling // Physics of Auroral Phenomena: Proceedings of 29th Annual Seminar (Apatity, 27 February –3 March 2006) – Apatity, 2006. – 314 pp. – p. 234-237.

Namgaladze, A.A., Yu.N. Korenkov, V.V. Klimenko, I.V. Karpov, F.S. Bessarab, V.A. Surotkin, T.A. Glushchenko, and N.M. Naumova (1988), Global model of the thermosphere-ionosphere-protonosphere system, Pure and Applied Geophysics, 127, 219-254.

Namgaladze, A.A., O.V. Martynenko, and A.N. Namgaladze (1998), Global model of the upper atmosphere with variable latitudinal integration step, Geomagnetism and Aeronomy International, 1, 53.

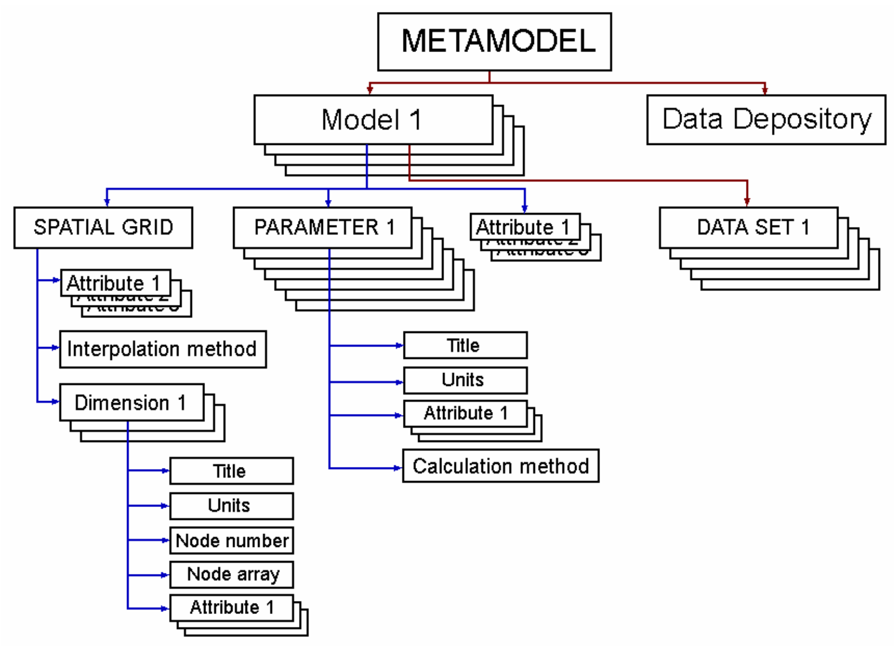

Fig. 1. Object hierarchy of the framework model.