Data Modeling for Non-Standard Data

Elizabeth Rose
Carroll College

This paper discusses the role of the semantic data models in addressing the concerns of non-standard data such as text, music, design information and pictures. It begins with a review of three prominent models and then discusses the general features used to compare semantic data models. Next, the evolution of database systems is considered. Descriptions of modern systems that are not transaction-oriented are provided. Finally, semantic data model features are linked to specific database environments. General research directions for the 1990’s are discussed.

Introduction

Data models are standards or patterns for representing the structure of and relationships inherent in data. Semantic data models, a subject of many publications in the ‘80’s, have been conceived largely to address a variety of problems or limitations perceived with the traditional record-based models which serve as the foundation for database management systems used widely today. This paper addresses specifically how these semantic models facilitate the modeling of certain types of non-standard data.

The name “semantic data models” captures their intent which is to provide more meaning, that is, to represent reality more fully. Researchers do not focus on the concerns of efficient storage and processing but concentrate instead on what the data means to the people that will use it. Some models are presented merely as design tools for simplifying schema design and database usage. Other researchers propose semantic data models as next generation models that will someday serve as the foundation for commercial database systems.

What follows is a brief summary of three semantic data models. SDM, proposed by Hammer and McLeod in 1981, provides flexibility and multiple ways of representing data and many later models borrow from its rich variety of constructs. FDM, developed by Kerschberg and Pacheco in 1976 has a more limited set of primitives aiming more at simplicity and ease of use. Both models are highly prominent in the literature and appear to be an important foundation for study.

The third model SAM*, studied by Su beginning in 1983, is selected from 15-20 other models that have gained stature. SAM* is one of the semantic data models developed to support full-fledged computer-aided design (CAD) database applications. The extent to which one should tailor model features to specific database environments, i.e. the tradeoff between flexibility and ease of use, is the subject of current debate.

After a review of three particular models, the compilation of a generalized list of features of semantic data models is discussed. Such an enumeration of model characteristics is used as a basis for comparison of semantic models by many notable researchers. Semantic data models are also briefly compared to object-oriented data models, another category of alternate database models that is widely studied today.
Hammer and McLeod (1981) describe their semantic data model (SDM) as “a database description and structuring formalism that is intended to allow a database schema to capture more of the meaning of a database than is possible with contemporary database models.”

Objects in an SDM database are organized into classes. Classes can be related to other classes by 1) “subclass connections” (generalization) or by 2) “group connections” (association or membership). Subclasses are established by defining subsets of attribute values, by defining set operations on existing subclasses or by dynamic user definition. Subclasses are always eventually decomposable into specific object instances. Groups, on the other hand, are specified by description or enumeration of member classes. The base units of a group are higher-order entity types. Attributes of a group are attributes of classes of objects and not attributes of individual elements. The constituents of a grouping class may themselves be viewed as classes. Attributes assigned to classes or members can be explicit or they can be derived from other information in the database. SDM offers a rich collection of derivation primitives including statistical, Boolean, and ordering computations.

An example class description given by Hammer and McLeod (1981) is shown in Figure 1. OIL_TANKERS has a subclass connection and inherits attributes from its superclass, SHIPS. SDM includes many rules to govern conflicts...