Chapter 2.13
Architecture for Symbolic Object Warehouse

Sandra Elizabeth González Císaro
Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina

Héctor Oscar Nigro
Universidad Nacional del Centro de la Provincia de Buenos Aires, Argentina

INTRODUCTION

Much information stored in current databases is not always present at necessary different levels of detail or granularity for Decision-Making Processes (DMP). Some organizations have implemented the use of central database - Data Warehouse (DW) - where information performs analysis tasks. This fact depends on the Information Systems (IS) maturity, the type of informational requirements or necessities the organizational structure and business own characteristic.

A further important point is the *intrinsic structure of complex data*; nowadays it is very common to work with complex data, due to syntactic or semantic aspects and the processing type (Darmont et al., 2006). Therefore, we must design systems, which can to maintain data complexity to improve the DMP.

OLAP systems solve the problem of present different aggregation levels and visualization for multidimensional data through cube's paradigm. The classical data analysis techniques (factorial analysis, regression, dispersion, etc.) are applied to individuals (tuples or individuals in transactional databases). The classic analysis objects are not expressive enough to represent tuples, which contain distributions, logic rules, multivaluate attributes, and intervals. Also, they must be able to respect their internal variation and taxonomy maintaining the dualism between individual and class.

Consequently, we need a new data type holding these characteristics. This is just the mathematical concept model introduced by Diday called Symbolic Object (SO). SO allows modeling physic entities or real world concepts. The former are the tuples stored in transactional databases and the latter are high entities obtained from expert's analysis, automatic classification or some particular aggregation taken from analysis units (Bock & Diday, 2000).

The SO concept helps construct the DW and it is an important development for Data Mining (DM): for the manipulation and analysis of aggregated information (Nigro & González Císaro, 2005). According to Calvanese, data integration is a central problem in the design of DWs and Decision Support Systems (Calvanese, 2003; Cali, et al., 2003); we make the architecture for Symbolic Object Warehouse construction with integrative goal. Also, it combines with Data Analysis tasks or DM.

This paper is presented as follows: First, Background: DW concepts are introduced. Second, Main Focus divided into: SOs Basic Concepts, Construing SOs and Architecture. Third, Future Trends, Conclusions, References and Key Terms.

Background

The classical definition given by the theme's pioneer is "a Data Warehouse is a subject-oriented, integrated, time-variant, and non-volatile collection of data in support of management's Decision-Making Process" (Inmon, 1996). The fundamental purpose of a DW is to empower the business staff with information that allows making decisions based on consolidated information. In essence, a DW is in a continuous process of transformation as regards information and business rules; both of them must be considered at design time to assure increase robustness and flexibility of the system.

Extraction, Transformation and Load (ETL) constitute the fundamental process in the DW. It is liable for the extraction of data from several sources, their cleansing, customization and insertion into a DW (Simitsis, et al., 2005). When complex data is involved, this process becomes difficult, because of the integration of different semantics (especially with text data, sound, images, etc) or complex structures. So, it is necessary to include integration functions able to join and to merge them.

Metadata management, in DW construction, helps the user understand the stored contents. Information about the meaning of data elements and the availability of reports are indispensable to successfully use the DW.

The generation and management of metadata serve two purposes (Staudt et al., 1999):

1. To minimize the efforts for development and administration of a DW
2. To improve the extraction from it.

Web Warehouse (WW) is a major topic widely researched and developed (Han & Kamber, 2001), as a result of the increasing and intensive use in e-commerce and e-business applications. WW tools and applications are morphing into enterprise portals and analytical applications are being extended to transactional systems. With the same direction, the audiences for WW have expanded as analytical applications have rapidly moved (indirectly) into the transactional world ERP, SCM and CRM (King, 2000).

Spatial data warehousing (SDW) responds to the need of providing users with a set of operations for easily exploring large amounts of spatial data, as well as for aggregating spatial data into synthetic information most suitable for decision-making (Damiani & Spaccapietra, 2006). Gorawski & Malczok (2004) present a distributed SDW system designed for storing and analyzing a wide range of spatial data. The SDW works with the new data model called cascaded star model that allows efficient storing and analyzes of huge amounts of spatial data.

MAIN FOCUS

SOs Basic Concepts

Formally, a SO is a triple s = (a, R, d) where R is a relation between descriptions, d is a description and "a" is a mapping defined from Ω(discourse

Related Content

The Soprano Extensible Object Storage System
Jung-Ho Ahnand Hyoung-Joo Kim (2002). *Journal of Database Management (pp. 15-24).*
www.irma-international.org/article/soprano-extensible-object-storage-system/3273

Data Model of FRDB with Different Data Types and PFSQL
Aleksandar Takaciand Srdan Škrbic (2008). *Handbook of Research on Fuzzy Information Processing in Databases (pp. 407-434).*
www.irma-international.org/chapter/data-model-frdb-different-data/20362

Meronymic Relationships
Veda C. Storey (1991). *Journal of Database Administration (pp. 22-36).*
www.irma-international.org/article/meronymic-relationships/51092

Modeling and Optimization of Multi-Model Waste Vehicle Routing Problem Based on the Time Window
Hongjie Wan, Junchen Ma, Qiumei Yu, Guozi Sun, Hansen Heand Huakang Li (2023). *Journal of Database Management (pp. 1-16).*
www.irma-international.org/article/modeling-and-optimization-of-multi-model-waste-vehicle-routing-problem-based-on-the-time-window/321543

Web Content Management and Dynamic Web Pages-A Tutorial
Esther Gelleand Viktor Schepik (2003). *Web-Powered Databases (pp. 55-87).*
www.irma-international.org/chapter/web-content-management-dynamic-web/31424