

Concept–Oriented Model

Alexandr Savinov
SAP AG, Germany

INTRODUCTION

In recent years, we observe a significant expansion of data analytics driven by such factors as very low cost of data acquisition, storage and processing. To make business decisions, users have to carry out complex analysis which is fed by data. Data is fuel and prerequisite for any kind of analysis starting from simple charts and ending with complex data mining processes. However, data analysis is getting more and more difficult with the explosion of data volume and the variety of data sources which differ in their underlying data models, representation formats and conventions.

In order to represent and analyze data from different data sources a unified model is needed. One such model, called the concept-oriented model (COM), is described in this article. COM (2009a) is a general purpose unified data model aimed at describing many existing views and patterns of thoughts currently used in data modeling. As a unified model, its main goal is to significantly decrease differences and incongruities between various approaches to data modeling such as transactional, analytical (Savinov, 2011b), multidimensional (Savinov, 2005a), object-oriented (Savinov, 2011a), conceptual and semantic (Savinov, 2012c). The motivation behind COM and its practical benefits are similar to those for the Business Intelligence Semantic Model (BISM) (Russo, Ferrari, & Webb, 2012) introduced in Microsoft SQL Server 2012. Like BISM, COM also aims at creating a unified model for all user experiences by significantly simplifying typical analysis tasks and facilitating self-service analytics.

The creation of this model was motivated by the following problems which are difficult to solve within traditional approaches:

- **Domain-Specific Identities:** Identification in most existing models is based on two approaches: primitive references like oids or surrogates, and identifying properties like primary keys. These approaches do not provide a mechanism for defining strong *domain-specific* identities with arbitrary structure, and the focus in existing models is made on describing entities. The goal of COM is to make identities and entities equal elements of the model both having some structure.
- **Hierarchical Address Spaces:** Most models focus on describing properties of elements without providing containment relation as a basic construct. A typical solution consists in modeling spaces and containment like any other domain-specific relationship. The goal of COM is to model containment as a core notion of the model by assuming that elements cannot exist outside of any space, domain or context.
- **Multidimensionality:** There exist numerous approaches to multidimensional modeling which are intended for analytical processing. The problem is that analytical and transactional models rely on different assumptions and techniques. The goal of COM in this context is to rethink dimensions as a first-class construct of the data model which plays a primary role for describing both transactional and analytical aspects.
- **Semantics:** Existing models represent data semantics separately from data itself. Normally semantics is supported at the level of conceptual model or using some semantic model. Logical data models have

DOI: 10.4018/978-1-4666-5202-6.ch045

rather limited possibilities for representing semantic relationships and it is a strong limiting factor for the effective use of data. The goal of COM here is to make semantics integral part of the logical data model so that the database can be directly used for reasoning about data.

To solve these problems, COM introduces three main structural principles which distinguish it from other data models and naturally account for various typical data modeling issues:

- **Duality Principle:** Answers the question *how* elements exist by assuming that any data element is composed of two parts: one identity and one entity (also called reference and object, respectively). Thus an element is defined as an identity-entity couple.
- **Inclusion Principle:** Answers the question *where* elements exist by postulating that any element is *included* in some domain (also called scope or context). Thus all elements exist in a hierarchy.
- **Order Principle:** Answers the question what an element *is*, that is, how it is defined and what is its meaning by assuming that all elements are partially ordered. Thus any element has a number of greater and lesser elements which define its semantics.

Syntactically, COM is described by the concept-oriented query language (COQL) (Savinov, 2011b). This language reflects the principles of COM by introducing a novel data modeling construct, called *concept* (hence the name of the approach), and two relations among concepts, *inclusion* and *partial order*. Concepts generalize conventional classes and are used for describing domain-specific identities. Inclusion relation generalizes inheritance and is used for describing hierarchical address spaces. Partial order relation among concepts is intended to represent data semantics and is used for complex analytical tasks and reasoning about data.

Formally, COM relies on the notion of nested partially ordered sets (nested posets). Nested poset is a novel formal construct that can be viewed either as a nested set with partial order relation established on its elements or as a conventional poset where elements can themselves be posets. An element of a nested poset is defined as a couple consisting of one identity tuple and one entity tuple.

C

BACKGROUND

As a unified model, COM is able to describe many wide-spread mechanisms and data modeling patterns existing in other model.

The support for hierarchies in COM makes it very similar to the classical hierarchical data model (HDM) (Tsichritzis & Lochovsky, 1976) and therefore COM can be viewed as a reincarnation of the hierarchical model using new principles. In both models data exist within a hierarchy where any element has a unique position. The main difference of COM is that it proposes to use domain-specific identities as hierarchical addresses. Another difference is that inclusion relation is simultaneously used to model inheritance as it is done in object data models.

Hierarchies in COM make it similar to object data models (Dittrich, 1986; Bancelhon, 1996). The difference is that inclusion relation in COM generalizes inheritance and is used to model both containment and inheritance (reuse). The assumption that inheritance is a particular case of containment is one of the major contributions of the concept-oriented approach. From this point of view, containment in the hierarchical model and inheritance in object models are particular cases of inclusion relation in COM. The treatment of inclusion in COM is very similar to how inheritance is implemented in prototype-based programming (Lieberman, 1986; Stein, 1987; Chambers, Ungar, Chang, & Hölzle, 1991) because in both approaches parent elements are shared parts of children.

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/concept-oriented-model/107253

Related Content

Business Intelligence and Organizational Decisions

Thomas H. Davenport (2010). *International Journal of Business Intelligence Research* (pp. 1-12).

www.irma-international.org/article/business-intelligence-organizational-decisions/38936

Analyzing Economic Indicators of Disaster Resilience Following Hurricane Katrina

Mark L. Dottore and Christopher W. Zobel (2014). *International Journal of Business Analytics* (pp. 67-83).

www.irma-international.org/article/analyzing-economic-indicators-of-disaster-resilience-following-hurricane-katrina/107070

A Systematic Literature Review on Hospitality Analytics

João Paulo Rodrigues, Maria José Sousa and Ana Brochado (2020). *International Journal of Business Intelligence Research* (pp. 47-55).

www.irma-international.org/article/a-systematic-literature-review-on-hospitality-analytics/256929

A Study on a Combined Model in Business Intelligence for Improving Electronic Insurance

Mahdi Bazargani and Elnaz Namazi (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 994-998).

www.irma-international.org/chapter/a-study-on-a-combined-model-in-business-intelligence-for-improving-electronic-insurance/142662

Measuring Agreement Among Ranks: Sustainability Application

Kathleen Campbell Garwood and Alicia Graziosi Strandberg (2016). *International Journal of Business Intelligence Research* (pp. 45-62).

www.irma-international.org/article/measuring-agreement-among-ranks/161673