INTRODUCTION

A model is a mathematical description of a world aspect and a data model provides means for data organization in the form of some structural principles. These structural principles are used to break all elements into smaller groups making access to and manipulation of data more efficient for end-users and applications. The concept-oriented model (COM) is a novel general-purpose approach to data model-
Concept-Oriented Query Language for Data Modeling and Analysis

(Savinov, 2009a) which is intended to solve a wide spectrum of problems by reducing them to the following three structural principles distinguishing it from other data models:

- **Duality principle** answers the question how elements exist by assuming that any element is a couple of one identity and one entity (called also reference and object, respectively)
- **Inclusion principle** answers the question where elements exist by postulating that any element is included in some domain (called also scope or context)
- **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 so that any element has a number of greater and lesser elements

Formally, the concept-oriented model is described using a formalism of nested partially ordered sets. The syntactic embodiment of this model is the concept-oriented query language (COQL). 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 are intended to generalize conventional classes and inclusion generalizes inheritance. Concepts and inclusion are used also in a novel approach to programming, called concept-oriented programming (COP) (Savinov, 2008, 2009b). Partial order relation among concepts is intended to represent data semantics and is used for complex analytical tasks and reasoning about data.

The concept-oriented model and query language are aimed at solving several general problems which are difficult to solve using traditional approaches. In particular, the following factors motivated this work:

- **Domain-specific identities.** In most existing data models elements are represented either by platform-specific references like oids or by weak identities based on entity properties like primary keys. These approaches do not provide a mechanism for defining strong domain-specific identities with arbitrary structure. Concepts solve this problem by making it possible to describe both identities and entities using only one common construct. This produces nice symmetry between two branches: identity modeling and entity modeling.
- **Hierarchical address spaces.** Elements cannot exist outside of any space, domain or context but existing data models do not support this abstraction as a core notion of the model. A typical solution consists in modeling spaces and containment like any other domain-specific relationship. The principled solution proposed in COM is that all elements are supposed to exist within a hierarchy where a parent is a space, context, scope or domain for its child elements. Thus inclusion relation between concepts turns an element into a set of its child elements. Since identities of internal elements are defined relative to the space they are in, we simultaneously get a hierarchical address space for the elements. Each element within this hierarchy is identified by a domain-specific hierarchical address like a conventional postal address.
- **Multidimensionality.** Dimension is one of the fundamental constructs which is used to represent information in various areas of human knowledge. There exist numerous approaches to multidimensional modeling which are intended for analytical processing. The problem is that there exist different models for analytical and transactional processing which 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. Data should be represented as originally existing in a multidimensional space and dimension should be used in most operations with data.