IDEA GROUP PUBLISHING

701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITJ2539

An Ontology of Data Modelling Languages:
A Study Using a Common-Sense Realistic Ontology

Simon K. Milton, University of Melbourne, Australia
Ed Kazmierczak, University of Melbourne, Australia

ABSTRACT

Data modelling languages are used in today's information systems engineering environments. Many have a degree of hype surrounding their quality and applicability with narrow and specific justification often given in support of one over another. We want to more deeply understand the fundamental nature of data modelling languages. We thus propose a theory, based on ontology, that should allow us to understand, compare, evaluate, and strengthen data modelling languages. In this paper we present a method (conceptual evaluation) and its extension (conceptual comparison), as part of our theory. Our methods are largely independent of a specific ontology. We introduce Chisholm's ontology and apply our methods to analyse some data modelling languages using it. We find a good degree of overlap between all of the data modelling languages analysed and the core concepts of Chisholm's ontology, and conclude that the data modelling languages investigated reflect an ontology of commonsense-realism.

Keywords: ontology; ontological analysis; data models; data modelling

INTRODUCTION

Data models have been used in information engineering environments for many decades for the precise purpose of building representations of reality. To date, there have been many different data modelling languages proposed, with the most popular being the Entity-Relationship Model (Chen, 1976) but also including the Functional Data Model (Kerschberg & Pacheco, 1976; Shipman, 1981), the Semantic Data Model (Hammer & McLeod, 1981), NIAM (Nijssen & Halpin, 1989), and Object Modelling Technique (Blaha & Premerlani, 1998). Each new modelling language has often been accompanied with claims of its superiority and at times hype when compared with the others. There has been little beyond opinion to substantiate

such claims and yet all notations purport to do similar things. We have two research questions:

Q1: How well do data models represent reality?

Q2: What are the similarities and differences between data modelling languages?

We need a theory to help us answer these questions.

The mature philosophical study of ontology has been used as a source of theory to investigate tools and techniques used in the analysis and design of information systems. A key development in the use of ontology for the study of information systems has been the work of Wand and Weber (Weber, 1997), based on Bunge's ontology (Bunge, 1977, 1979). Part of the focus of this research has been to investigate the representational power of data modelling languages (Green, 1996; Rohde, 1995; Wand, 1996; Wand & Weber, 1989, 1990, 1993; Weber, 1997). Our work is motivated by the search for semantic methods to answer the research questions above. The work of Wand and Weber, while ground breaking, is based on structural comparison of elements of grammar and concludes only presence or absence of a 'construct'. The conclusions drawn are very much based upon whether or not the data modelling language supports the ontological construct.

Our work seeks to develop semantic methods that not only detect the presence or absence of a construct, but also allow us to judge the level of agreement or disagreement between a data modelling language

The contribution of this paper is two-fold. First, we develop qualitative methods: (1) 'the method of conceptual comparison', for conceptually evaluating individual data modelling languages through ontologies and (2) 'the method for conceptual comparison', for comparing a range of data modelling languages with an ontology based on a number of individual evaluations. These methods help answer the two research questions and are detailed in the method section. Our methods are, to some extent, independent of the individual ontology chosen as the basis of comparison. As a by-product we are starting to investigate the dominant ontology within data modelling languages. Secondly, we apply the methods using Chisholm's ontology (Chisholm, 1996) to a representative range of data modelling languages.

We follow this introduction with a deeper discussion relating ontologies and data modelling languages. We then examine the realism assumed in Chisholm's ontology and relate it to that contained within Bunge's ontology (the ontology upon which BWW is based). Following this we describe Chisholm's ontology, which is the ontology used in this study, before describing the methods applied. Finally we present the results and conclude.

RELATING ONTOLOGY AND DATA MODELLING LANGUAGES

We begin by defining ontology and discussing our ontological view of data modelling languages. Our treatment of ontology stems from the philosophical tradition and a good definition of ontology is found in Honderich (1995).

Definition 1-Ontology: "Ontology, understood as a branch of metaphysics, is the science of being in general, embracing such issues as the nature of existence and the categorical structure of reality. ...Different systems of ontology propose alternative categorical schemes.