The Synthesis and Diagnosis Rules for Conceptual Data Warehouse Design Based on First-Order Logic

Opim Salim Sitompul
Department of Computer Science
University of Sumatera Utara
Kampus USU Medan-20155, Indonesia
opim@usu.ac.id

Shahrul Azman Mohd Noah
Department of Information Science
National University Malaysia
43600 UKM, Bangi, Malaysia
sann@fsm.ukm.my

Abstract – Designing and implementing a data warehouse is a highly complex engineering task that asks for methodological support. However, devising a design methodology is almost useless unless it is supported with a CASE tool that could assist the designer in specifying and implementing the data warehouse design requirements. Nowadays, many of the developed CASE tools for data warehouse design exhibit little intelligent capabilities either during the process of design synthesis or design diagnosis. This research proposed a set of rules for conceptual data warehouse design within the context of the aforementioned design approach. The rules are based on the first-order logic and are applied to an existing ER model in order to build the data warehouse conceptual model. The implementation of the proposed rules could guide users during a design session and assist them in diagnosing inconsistencies in the model.

Keywords: Data warehouse design, automated tool, first-order logic.

1 Introduction

The use of automated tools in producing designs of information systems have become more popular and demanding because of the increasing need for better information systems. One such popular and increasing in-demand applications is the data warehouse system. As a result, effort in building automated data warehouse tools have received great attentions from researchers as can be found in [1-10]. However, almost all of these tools lack elements of intelligence, especially in term of their capability to interactively guide users during a design process. Among the aforementioned research works, only one of them claimed that the tool developed has an intelligent capability whereby the tool is equipped with a description-logic reasoning server as a background inference engine [7].

In order to contribute to the development of CASE tool for conceptual data warehouse design, we proposed a set of synthesis and diagnosis rules, which is based on the first-order logic to be used to develop the multidimensional model for data warehouse. The rules are applied to an existing ER model, which has been translated into a specification language model and used as the initial problem domain model. The synthesis and diagnosis rules work together to build and to preserve the integrity of the evolving data warehouse model during a design session.

2 The Transformation-Oriented Approach

The transformation-oriented approach that is used in this research progressively transforms an ER model into multidimensional model in five stages. For each stage there will be various transformation rules applied on the input in order to obtain the desired output. Some of the transformation rules would be in the form of syntax rules, whereas others may take the form of synthesis and diagnosis rules.

Figure 1 shows the five stages of the transformation process, consisting of: translation of ER model into specification language model, transformation of specification language model into problem domain model, expansion of the problem domain model, transformation of the problem domain model into multidimensional model, and refinement of the multidimensional model.

![Figure 1. The transformation-oriented approach](image-url)