Common Sense Reasoning in Automated Database Design: An Empirical Test

VEDA C. STOREY, Georgia State University, USA
ROBERT C. GOLDSTEIN, University of British Columbia, Canada
JASON DING, Pharm2B, China

A great deal of work on automating systems design and development has been carried out, especially in the database area. Systems that semi-automate the database design process have been developed. These systems are interactive in that they may need to ask the user (usually, a database designer) for clarification. The result is that the system asks questions to the user that make the system look less intelligent than it should. This general type of problem has long been recognized with a proposed approach to overcoming it being the incorporation of common sense knowledge into a design system. The View Creation System is an expert system that plays the role of a database designer. With it, a user knowing little about database technology can express his or her database design requirements, which are represented by an entity-relationship model and then translated into a normalized relational model. The system contains a great deal of knowledge about database design, but little, if any, about the user’s application. This forces the user to specify many trivial facts that would be known by any human designer. To overcome this limitation, a Common Sense Business Reasoner is being developed that has a knowledge base containing general knowledge about the world and a reasoning tool to apply this knowledge to a database design task. An empirical study is carried out to simulate and assess the effectiveness of adding the Common Sense Business Reasoner to the View Creation System.

INTRODUCTION

Database design is the design of a database structure that will be used to store and manage data (Rob and Caronel, 2000). It consists of four phases: requirements analysis, conceptual design, logical design and physical design. Conceptual design is probably the most critical and difficult phase because the designer needs to represent the user requirements in a model, such as the entity-relationship model, that is independent of the database management system being used. It usually involves a database design expert working closely with a domain expert (end user).

Database design, in general, is a difficult problem because it involves capturing and representing users’ requirements accurately and completely in an information system, with the quality of the design dependent upon the expertise level of the designer. In an attempt to improve the quality and consistency of the designs produced, while minimizing the work on the part of the user, a number of systems have been developed for automating various aspects of the database design process (Lloyd-Williams, 1997; Lloyd-Williams and Beynon-Davis, 1992; Storey and Goldstein, 1993; Noah and Lloyd-Williams, 1998). Many of these systems assist the end user, who can be regarded as a domain expert, in designing a conceptual model.

Studies on the effectiveness of these and other support systems have been mixed. One explanation is that while these systems possess a high degree of expertise in the database design domain, they usually know very little about the application domain and the ‘real world’ (Lloyd-Williams, 1997; Storey, 1992). This causes several problems. First, communication between the user and the system is impeded because the user must spend a great deal of time entering ‘trivial’ information (e.g., SSN as a key of person). Second, systems rely completely on the user to represent their requirements in database design constructs (e.g., entity, relationship, attribute) that are unfamiliar to the user. A human expert designer, on the other hand, has some common sense knowledge about the application domain and does not need to ask the end user to articulate every detail. Furthermore, the designer can make suggestions to augment the design based upon past experience.

Although the need for the incorporation of common sense knowledge has been recognized, there do not appear to have been any formal studies undertaken to assess the feasibility or effectiveness of doing so. The objective of this research, therefore, is to:

evaluate the effectiveness of adding a common sense reasoning module to an existing database
design expert system.

The system used is a version of the View Creation System (Storey and Goldstein, 1988) that is augmented with common sense reasoning capabilities. The contributions of the research are to: empirically assess the feasibility of adding common sense knowledge to database design systems; show the extent to which the addition of common sense knowledge is feasible, and suggest areas for future research.

This paper is divided into seven sections. Section 2 reviews related research in design automation, design artifacts, and common sense reasoning. Section 3 defines common sense knowledge as it is used in this research. Sections 4 to 6 describe the experiment conducted, dealing with the hypothesis development, methodology used, and results obtained, respectively. Concluding remarks and suggestions for future research are presented in Section 7.

RELATED RESEARCH

The need for incorporating and using knowledge about the real world to make systems appear more “intelligent” has been recognized as a difficult problem for a long time (McCarthy, 1983; Hobbs and Moore, 1985). In artificial intelligence, the most ambitious project in the common sense area is CYC (Lenat, 1995), the objective of which was to build a comprehensive common sense knowledge base. Although a project such as CYC may eventually accumulate enough knowledge to provide general support to domain-specific expert systems, it does not appear to be cost effective to do so. The need for approaches to compete with CYC to hasten progress has also been noted (Minsky, 1994).

Research on software development has suggested that “stocks” of domain-specific knowledge are needed to make these systems appear more “intelligent.” This is particularly evident from research on reuse that takes advantage of stores of knowledge represented as reusable artifacts, with the most common being analysis patterns (Coad et al., 1995). Another approach is the development of ontologies. An ontology refers to a set of concepts or terms that can be used to describe some area of knowledge or build a representation of it. It is believed that ontologies could fundamentally change the way in which systems are constructed and that developers, eventually, will have libraries of ontologies from which they can choose appropriate ones (Swartout, 1999).

COMMON SENSE KNOWLEDGE AND DATABASE DESIGN SYSTEMS

Conceptual database design is the process of representing the information content of an application using constructs that are independent of the way the data is physically stored (Date, 1990). The users’ requirements are represented as the system’s basic elements; their relationships and constraints. This results in an implementation-independent conceptual data model, with the most widely used being the Entity-Relationship (E-R) model (Chen, 1993). The E-R model can then be translated into a relational data model in a relatively straight-forward manner (Teorey et al., 1986). The database design system used in this research is the View Creation System (Storey and Goldstein, 1988), although the research results are not intended to be restricted to any system.

View Creation System

The View Creation System elicits user’s requirements by asking the user to identify “things” of interest (i.e., entities) in his or her application domain and the relationships between them in the form of A verb-phrase B. The system checks for errors and inconsistencies in the user’s input. For each entity, the user is asked to specify its name, attributes, and candidate keys, and for each relationship, its cardinalsities and its attributes. The system then transforms the requirements into a set of 4NF relations. An overview of the system is shown in Figure 1.

The knowledge base of the View Creation System consists of rules and facts about database design but nothing about the application domains for which it might be used. As a result, the user routinely must enter trivial information (e.g., the key of a person entity) and answer obvious questions. Thus, its lack of common sense knowledge is a major shortcoming of the View Creation System compared to a human designer, as it is with other design systems (Storey and Goldstein, 1993).

Common Sense Knowledge

Common sense is defined as general knowledge about the way the world works; that is, everyday knowledge that everyone knows about the world. This definition is based upon that of Naïve Semantics (Dahlgren, 1995), which pro-
Related Content

**Unifying Access to Heterogeneous Document Databases through Contextual Metadata**
www.igi-global.com/chapter/unifying-access-heterogeneous-document-databases/9207?camid=4v1a

**Reconciling Information Privacy and Information Access in a Globalized Technology Society**
www.igi-global.com/chapter/reconciling-information-privacy-information-access/8007?camid=4v1a

www.igi-global.com/article/industry-practice-understanding-three-dimensions/51157?camid=4v1a

**Technology-Mediated Synchronous Virtual Education: An Empirical Study**
www.igi-global.com/article/technology-mediated-synchronous-virtual-education/178635?camid=4v1a