INTRODUCTION

Computer-based information technologies have been extensively used to help industries manage their processes and information systems become their nervous center. More specifically, databases are designed to support the data storage, processing, and retrieval activities related to data management in information systems. Database management systems provide efficient task support and tremendous gain in productivity is thereby accomplished using these technologies. Database systems are the key to implementing industrial data management. Industrial data management requires database technique support. Industrial applications, however, are typically data- and knowledge-intensive applications and have some unique characteristics (e.g., large volumes of data with complex structures) that makes their management difficult. Product data management supporting various life-cycle aspects in the manufacturing industry, for example, should not only to describe complex product structure but also manage the data of various life-cycle aspects from design, development, manufacturing, and product support. Besides, some new techniques, such as Web-based design and artificial intelligence, have been introduced into industrial applications. The unique characteristics and usage of these new technologies have created many potential requirements for industrial data management, which challenge today’s database systems and promote their evolvement.

BACKGROUND

From a database-technology standpoint, information modeling in databases can be identified at two levels: (conceptual) data modeling and database modeling, which results in conceptual (semantic) data models and logical database models. Generally, a conceptual data model is designed, and then the designed conceptual data model is transformed into a chosen logical database schema. Database systems based on logical database models are used to build information systems for data management. Much attention has been directed at conceptual data modeling of industrial information systems. Product data models, for example, can be viewed as a class of semantic data models (i.e., conceptual data models) that take into account the needs of engineering data. Recently conceptual data modeling of enterprises has received increased attention.

Generally speaking, traditional ER/EER (Entity-Relationship/Extended Entity Relationship) or UML models in the database area can be used for industrial data modeling at the conceptual level. But limited by the power of the aforementioned data models in industrial data modeling, some new conceptual data models such as IDEF1X and STEP/EXPRESS have been developed. In particular, to implement the share and exchange of industrial data, the Standard for the Exchange of Product Model Data (STEP) is being developed by the International Organization for Standardization (ISO). EXPRESS is the description method of STEP and a conceptual schema language, which can model product design, manufacturing, and production data. EXPRESS model hereby becomes a major one of conceptual data models for industrial data modeling. Much research has been reported on the database implementation of the EXPRESS model in context of STEP, and some software packages and tools are available in markets.

As to industrial data modeling in database systems, the generic logical database models, such as relational, nested relational, and object-oriented databases, have been used. However, these generic logical database models do not always satisfy the requirements of industrial data management. In nontransaction processing, such as CAD/CAM (Computer-Aided Design/Computer-Aided Manufacturing), knowledge-based system, multimedia, and Internet systems, most of the data-intensive application systems suffer from the same limitations of relational databases. Some nontraditional database models based on the aforementioned special, hybrid, or extended database models have been proposed accordingly.

MAJOR ISSUES AND SOLUTIONS

Conceptual Data Models

Much attention has been directed at conceptual data modeling of engineering information (Mannisto, Peltonen, Soininen, & Sulonen, 2001; McKay, Bloor, & de
Pennington, 1996). Product data models, for example, can be viewed as a class of semantic data models (i.e., conceptual data models) that take into account the needs of engineering data (Shaw, Bloor, & de Pennington, 1989). Recently, conceptual information modeling of enterprises such as virtual enterprises has received increasing attention (Zhang & Li, 1999). Generally speaking, traditional ER (P. P. Chen, 1976) and EER can be used for engineering information modeling at conceptual level. But, limited by their power in engineering modeling, some new conceptual data models have been developed.

IDEF1X is a method for designing relational databases with a syntax designed to support the semantic constructs necessary in developing a conceptual schema. Some researchers have focused on the IDEF1X methodology (a thorough treatment of the IDEF1X method can be found in Wizdom Systems, 1985). The use of the IDEF1X methodology to build a database for multiple applications was addressed in Kusiak, Letsche, and Zakarian (1997).

As mentioned earlier, STEP provides a means to describe a product model throughout its life cycle and to exchange data between different units. STEP consists of four major categories, namely, description methods, implementation methods, conformance testing methodology, and framework, and standardized application data models/schemata. EXPRESS (Schenck & Wilson, 1994), being the description methods of STEP and a conceptual schema language, can model product design, manufacturing, and production data, and the EXPRESS model hereby becomes a major conceptual data model for engineering information modeling.

On CAD/CAM development for product modeling, Eastman and Fereshetian (1994) conducted a review and studied five information models used in product modeling, namely, ER, NAIM, IDEF1X, EXPRESS, and EDM. Compared with IDEF1X, EXPRESS can model complex semantics in engineering applications, including engineering objects and their relationships. Based on the EXPRESS model, it is easy to implement the share and exchange of engineering information.

It should be noted that ER/EER, IDEF1X and EXPRESS could model neither knowledge nor fuzzy information. The first effort was undertaken by Zvieli and Chen (1986) to extend ER models to represent three levels of fuzziness. The first level refers to the set of semantic objects, resulting in fuzzy entity sets, fuzzy relationship sets, and fuzzy attribute sets. The second level concerns the occurrences of entities and relationships. The third level relates to the fuzziness in attribute values of entities and relationships. Consequently, ER algebra was fuzzily extended to manipulate fuzzy data. In G. Q. Chen and Kerre (1998), several major notions in the EER model were extended, including fuzzy extension to generalization/specialization, and shared subclass/category, as well as fuzzy multiple inheritance, fuzzy selective inheritance, and fuzzy inheritance for derived attributes. More recently, using fuzzy sets and possibility distribution (Zadeh, 1978), fuzzy extensions to IDEF1X (Ma, Zhang, & Ma, 2002) and EXPRESS were proposed and (Ma, Zhang, & Ma, 2001; Ma, Zhang, Ma, & Chen, 2000), respectively.

Unified modeling language (UML; Booch, Rumbaugh, & Jacobson, 1998; OMG, 2003), standardized by the Object Management Group (OMG), is a set of OO (object oriented) modeling notations. UML provides a collection of models to capture the many aspects of a software system. From an information modeling point of view, the most relevant model is the class model. The building blocks in this class model are those of classes and relationships. The class model of UML encompasses the concepts used in ER as well as other OO concepts. In addition, it also presents the advantage of being open and extensible, allowing its adaptation to the specific needs of the application, such as workflow modeling of e-commerce (Chang, Chen, Chen, & Chen, 2000) and product structure mapping (Oh, Hana, & Subh, 2001). In particular, the class model of UML is extended for the representation of class constraints and the introduction of stereotype associations (Mili et al., 2001).

With the popularity of Web-based design, manufacturing and business activities, the requirement has been put on the exchange and share of engineering information over the Web. Because of limitations of HTML in content-based information processing, the World Wide Web Consortium created eXtensible markup language (XML), a language similar in format to HTML, but more extensible. This new language lets information publishers invent their own tags for particular applications or work with other organizations to define shared sets of tags that promote interoperability and that clearly separate content and presentation. XML provides a Web-friendly and well-understood syntax for the exchange of data. Because XML affects the way data is defined and shared on the Web (Seligman & Rosenthal, 2001), XML technology has been increasingly studied, and more and more Web tools and Web servers support XML. Bourret (2001) developed product data markup language, the XML for product data exchange and integration. As to XML modeling at concept level, Conrad, Scheffner and Freytag (2000) used UML was used for designing XML DTD. Xiao, Dillon, Chang, and Feng (2001) developed an object-oriented conceptual model to design XML schema. Lee, Lee, Ling, Dobbie, and Kalinichenko (2001) used the ER model for the conceptual design of semistructured databases in. Note, however, that XML does not support imprecise and uncertain information modeling as well as it does knowledge modeling. Intro-