

Diss. ETH Nr. 7325

Datenbankentwurf im Dialog – Integrierte Beschreibung von Strukturen, Transaktionen und Konsistenz

ABHANDLUNG

zur Erlangung des Titels eines
Doktors der Technischen Wissenschaften
der
**EIDGENÖSSISCHEN TECHNISCHEN HOCHSCHULE
ZÜRICH**

vorgelegt von
JÜRIG REBSAMEN
dipl. Math. ETH
geboren am 27. November 1953
von Zürich

Angenommen auf Antrag von
Prof. Dr. C. A. Zehnder, Referent
Prof. Dr. N. Wirth, Korreferent

ADAG Administration & Druck AG

Zürich 1983

Abstract

A comprehensive and successful database definition is achieved by combining various description methods and design techniques. A high-level data model should assist the designer in developing and describing data structures with corresponding consistency-preserving transactions. Furthermore, the conceptual schema should include all of the semantic integrity constraints perceived in the real world. This thesis presents some contributions to the mentioned concepts and integrates them by means of an interactive design process.

Data description is based here on a relational model extended to cover some important and frequently used integrity constraints. The specification of identification keys and relationships between entity sets is accomplished by means of structures inherent in the data model. The formulation of further constraints within the database design process bases on the database programming language Modula/R, an extension of Modula-2 including a data type "relation" and the first order predicate calculus for relational expressions.

Conceptual data structures are designed interactively by stepwise refinement. Each definition phase offers its own, often two-dimensional form of presentation. A continuous iteration capability throughout the entire process allows the designer to develop and evaluate variants of data structures. The design system automatically transforms the schema into the appropriate representation and detects incomplete data structure definitions and inconsistencies within the definition.

To arrive at an overall conceptual description of the data base, behavioral aspects of the data structure, transactions have to be considered on the logical level as well. Interactive definition and graphic presentation support this transaction modelling. Particularly the propagation of operations from an entity set to related ones can be visualized and controlled already at definition time.

Data modules are introduced as an additional software component which separates the user from the database management system. In order to perform the integrity checking, the data modules are tailored to the data structures and transactions of each individual schema; the user looks upon his database as a set of abstract data types. It is shown that the Modula/R source text of the data modules can be generated automatically for the most part.

Finally the data definition system "Gambit" is presented as an implementation of the above concepts within the database system "LIDAS". Modern dialog techniques and a meta database containing the information on the user database have facilitated the implementation on a powerful personal computer, the system Lillith.

Keywords: semantic integrity constraint, conceptual design, stepwise refinement, interactive data definition, meta database, database system architecture, transaction, behavior modelling, data module, program generation.

Zusammenfassung

Die vorliegende Dissertation behandelt Konzepte für den logischen Datenbankentwurf und befasst sich mit deren Integration in einem rechnergestützten interaktiven Prozess. Für einen modernen, umfassenden Entwurf wird postuliert, dass das konzeptionelle Datenmodell sowohl Datenstrukturen als auch Transaktionen beschreiben kann, weiter die Formulierung sämtlicher für die Anwendung erforderlicher Konsistenzbedingungen im Rahmen der Datenbankbeschreibung ermöglicht und schliesslich eine bequeme Entwurfstechnik bietet.

Die Datenbeschreibung stützt sich auf ein erweitertes Relationenmodell, mit dem die häufigsten Konsistenzbedingungen, nämlich Identifikationsschlüssel und Beziehungsregeln zwischen Relationen, durch modellinhärente Strukturtypen implizit formuliert werden können. Für die Definition von selteneren Bedingungen erweist sich der Einsatz von programmiersprachlichen Konzepten im Rahmen der Datenbankbeschreibung als handliche Lösung. Verwendet werden dazu Ausdrücke und Anweisungen der Datenbankprogrammiersprache Modula/R, einer Erweiterung von Modula-2 um den Datentyp Relation und einen Prädikatenkalkül erster Ordnung.

Die Definition von Datenstrukturen erfolgt interaktiv und nach dem Grundsatz der schrittweisen Verfeinerung. Auf jeder Abstraktionsebene stehen dem Entwerfenden spezielle, teilweise zweidimensionale Darstellungstechniken zur Verfügung. Das Entwurfssystem erleichtert den Übergang zwischen den Abstraktionsebenen durch automatische Transformation der Darstellung, es übernimmt die Überwachung von Definitionslücken, und es macht den Entwerfenden laufend auf Inkonsistenzen aufmerksam. Die systematische Einbettung von Iteration in den Dialog ermöglicht die Entwicklung und den Vergleich von Varianten.

Eine vollständige Datenbeschreibung auf konzeptioneller Ebene umfasst auch die prozeduralen Aspekte, nämlich Transaktionen. Die Modellierung erfolgt auch hier mit Vorteil im Dialog und wird durch eine graphische Darstellung verdeutlicht. Damit kann insbesondere die Fortpflanzung gewisser Operationen von einer Entitätsmenge zur anderen sichtbar gemacht und bereits zur Entwurfszeit geplant werden.

Die Überprüfung der zur Definitionszeit festgelegten Konsistenzbedingungen erfolgt in anwendungsspezifischen Datenmodulen, welche auf dem Standard-Datenbankverwaltungssystem aufbauen und dem Benutzer ein Sortiment abstrakter Datentypen anbieten. Der Modula/R-Programmtext der Datenmodule kann aufgrund der Definitionen in weiten Teilen automatisch erzeugt werden.

Der letzte Teil der Arbeit beschreibt die Realisierung der besprochenen Konzepte im Datendefinitionsteil "Gambit" des Datenbanksystems "LIDAS". Der Einsatz moderner Dialogtechniken sowie einer Metadatenbank zur Verwaltung der Definitionen vereinfacht die Implementierung auf einem leistungsfähigen Arbeitsplatzrechner, dem System Lillith.