CHAPTER 7

Relational Database Design

The goal of relational database design is to generate a set of relation schemas that allows us to store information without unnecessary redundancy, yet also allows us to retrieve information easily. This is accomplished by designing schemas that are in an appropriate normal form. To determine whether a relation schema is in one of the desirable normal forms, we need information about the real-world enterprise that we are modeling with the database. Some of this information exists in a well-designed E-R diagram, but additional information about the enterprise may be needed as well.

Bibliographical Notes

The first discussion of relational database design theory appeared in an early paper by [Codd (1970)]. In that paper, Codd also introduced functional dependencies and first, second, and third normal forms. BCNF was introduced in [Codd (1972)]. Armstrong’s axioms were introduced in [Armstrong (1974)].

Significant development of relational database theory occurred in the late 1970s. These results are collected in several texts on database theory including [Maier (1983)], [Atzeni and Antonellis (1993)], and [Abiteboul et al. (1995)].

[Biskup et al. (1979)] give the algorithm we used to find a lossless dependency-preserving decomposition into 3NF. Fundamental results on the lossless decomposition property appear in [Aho et al. (1979)].

[Beeri et al. (1977)] gives a set of axioms for multivalued dependencies, and proves that the authors’ axioms are sound and complete. The notions of 4NF, PJNF, and DKNF are from [Fagin (1977), Fagin (1979)], and [Fagin (1981)], respectively. See the bibliographical notes of Chapter 28 for further references to literature on normalization.

[Jensen et al. (1994)] presents a glossary of temporal-database concepts. A survey of extensions to E-R modeling to handle temporal data is presented by [Gregersen and Jensen (1999)]. [Tansel et al. (1993)] covers temporal database theory, design, and implementation. [Jensen et al. (1996)] describes extensions of dependency theory to temporal data.
[Stam and Snodgrass (1988)] and [Soo (1991)] provide surveys on temporal data
management. [Jensen et al. (1994)] presents a glossary of temporal-database concepts,
aimed at unifying the terminology. [Tansel et al. (1993)] is a collection of articles on
different aspects of temporal databases. [Chomicki (1995)] presents techniques for
managing temporal integrity constraints.


Bibliography


[Aho et al. (1979)] A. V. Aho, C. Beeri, and J. D. Ullman, “The Theory of Joins in Relational
Databases”, ACM Transactions on Database Systems, Volume 4, Number 3 (1979), pages 297–
314.


[Beeri et al. (1977)] C. Beeri, R. Fagin, and J. H. Howard, “A Complete Axiomatization for
Functional and Multivalued Dependencies”, In Proc. of the ACM SIGMOD Conf. on Manage-

[Biskup et al. (1979)] J. Biskup, U. Dayal, and P. A. Bernstein, “Synthesizing Independent
Database Schemas”, In Proc. of the ACM SIGMOD Conf. on Management of Data (1979),
pages 143–152.

[Chomicki (1995)] J. Chomicki, “Efficient Checking of Temporal Integrity Constraints Us-
ing Bounded History Encoding”, ACM Transactions on Database Systems, Volume 20, Num-


[Fagin (1977)] R. Fagin, “Multivalued Dependencies and a New Normal Form for Relational
Databases”, ACM Transactions on Database Systems, Volume 2, Number 3 (1977), pages 262–
278.

[Fagin (1979)] R. Fagin, “Normal Forms and Relational Database Operators”, In Proc. of the
ACM SIGMOD Conf. on Management of Data (1979), pages 153–160.

[Fagin (1981)] R. Fagin, “A Normal Form for Relational Databases That Is Based on Do-
mains and Keys”, ACM Transactions on Database Systems, Volume 6, Number 3 (1981), pages
Further Reading

387–415.


