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.

An overview of temporal database support in a number of database systems appears in http://cs.ulb.ac.be/conferences/ebiss2017/files/slides/gamper_ebiss2017.

Bibliography

[Abiteboul et al. (1995)]   S. Abiteboul, R. Hull, and V. Vianu, *Foundations of Databases*, Addison Wesley (1995).

[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.

[Armstrong (1974)]   W. W. Armstrong, "Dependency Structures of Data Base Relationships", In *Proc. of the 1974 IFIP Congress* (1974), pages 580–583.

[Atzeni and Antonellis (1993)]   P. Atzeni and V. D. Antonellis, *Relational Database Theory*, Benjamin Cummings (1993).

[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 Management of Data* (1977), pages 47–61.

[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 Using Bounded History Encoding", *ACM Transactions on Database Systems*, Volume 20, Number 2 (1995), pages 149–186.

[Codd (1970)]   E. F. Codd, "A Relational Model for Large Shared Data Banks", *Communications of the ACM*, Volume 13, Number 6 (1970), pages 377–387.

[Codd (1972)]   E. F. Codd. "Further Normalization of the Data Base Relational Model", In *[Rustin (1972)]*, pages 33–64 (1972).

[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 Domains and Keys", *ACM Transactions on Database Systems*, Volume 6, Number 3 (1981), pages

387–415.

[Gregersen and Jensen (1999)]   H. Gregersen and C. S. Jensen, "Temporal Entity-Relation-ship Models-A Survey", *IEEE Transactions on Knowledge and Data Engineering*, Volume 11, Number 3 (1999), pages 464–497.

[Jensen et al. (1994)]   C. S. Jensen et al., "A Consensus Glossary of Temporal Database Con-cepts", *ACM SIGMOD Record*, Volume 23, Number 1 (1994), pages 52–64.

[Jensen et al. (1996)]   C. S. Jensen, R. T. Snodgrass, and M. Soo, "Extending Existing De-pendency Theory to Temporal Databases", *IEEE Transactions on Knowledge and Data Engi-neering*, Volume 8, Number 4 (1996), pages 563–582.

[Maier (1983)]   D. Maier, *The Theory of Relational Databases*, Computer Science Press (1983).

[Rustin (1972)]   R. Rustin, *Data Base Systems*, Prentice Hall (1972).

[Soo (1991)]   M. Soo, "Bibliography on Temporal Databases", *ACM SIGMOD Record*, Vol-ume 20, Number 1 (1991), pages 14–23.

[Stam and Snodgrass (1988)]   R. Stam and R. Snodgrass, "A Bibliography on Temporal Databases", *IEEE Transactions on Knowledge and Data Engineering*, Volume 7, Number 4 (1988), pages 231–239.

[Tansel et al. (1993)]   A. Tansel, J. Clifford, S. Gadia, S. Jajodia, A. Segev, and R. Snodgrass, *Temporal Databases: Theory, Design and Implementation*, Benjamin Cummings (1993).