



# INTERNATIONAL JOURNAL OF PURE AND APPLIED RESEARCH IN ENGINEERING AND TECHNOLOGY

A PATH FOR HORIZING YOUR INNOVATIVE WORK

## MULTIVALUED DEPENDENCIES AND FIFTH NORMAL FORM OF RELATIONAL DATABASES

PRASHANT SINALKAR<sup>1</sup>, DR. HEMANT MAHALLE<sup>2</sup>, DR. AVINASH MANWAR<sup>3</sup>

1. Government Vidarbha Institute of Science and Humanities, Amravati, India.
2. Phulsing Naik College, Pusad, India.
3. Sant Gadge Baba Amravati University, Amravati, India.

Accepted Date: 27/02/2014 ; Published Date: 01/05/2014

**Abstract:** A relational database is collection of tables of data items which may have additionally defined relationships among them. The objective of database normalization is to isolate data so that it can be propagated through the database using defined relationships. Multivalued dependency includes the well-known functional dependencies as a special case, is defined for relational database. By using this concept, a Forth Normal Form (4NF) for relational schemata is defined. We try to represent, how the concept of multivalued dependency is used in Fifth Normal Form(5NF).

**Keywords:** Relational Database, Database Normalization, Multivalued Dependency.



PAPER-QR CODE

Corresponding Author: MR. PRASHANT SINALKAR

Access Online On:

[www.ijpret.com](http://www.ijpret.com)

How to Cite This Article:

Prashant Sinalkar, IJPRET, 2014; Volume 2 (9): 662-665

## INTRODUCTION

A relational database is a database that has a collection of tables of data items, of all which is formally described and designed according to the relational model.<sup>[1]</sup> A relational database was first defined in 1970 by Edger Codd and now it is one of the prominent choice in storing data.<sup>[8]</sup>

Database normalization is the process of organizing the fields and tables of a relational database to minimize redundancy and dependency. Normalization usually involves dividing large tables into smaller tables and defining relationships between them.<sup>[9]</sup> Edgar Codd define First Normal Form (1NF) in 1970, Second Normal Form (2NF) and Third Normal Form (3NF) in 1972.<sup>[2]</sup> Ronald Fagin introduced Fourth Normal Form (4NF) in 1977<sup>[3]</sup> and Fifth Normal Form in 1979.<sup>[4]</sup>

In database theory, multivalued dependency is a full constraint between two set of the attributes in a relation.<sup>[3]</sup> A multivalued dependency is a special case of tuple-generating dependency as well as join dependency.

A table is in Fourth Normal Form (4NF) if and only if, for every one of its non-trivial dependencies  $X \twoheadrightarrow Y$ ,  $X$  is superkey – that is  $X$  is either a candidate key or a superset of thereof.<sup>[3]</sup> Second Normal Form (2NF) and Third Normal Form (3NF) are concerned with the functional dependencies whereas Fourth Normal Form (4NF) is concerned with multivalued dependency.

Fifth Normal Form (5NF) is a level of database normalization design to reduce redundancy in relational databases recording multivalued facts by isolating semantically related multiple relationships. A table is said to be in the Fifth Normal Form (5NF) if and only if every non-trivial join dependency in it is implied by the candidate keys.<sup>[4]</sup>

This paper represents dealing of Fifth Normal Form (5NF) with multivalued dependency. An example illustrated how decomposition of relational database proceed from Fourth Normal Form (4NF) to Fifth Normal Form (5NF), and try to study violation of multivalued dependency when projection and joins are allowed.

## 2. MULTIVALUED DEPENDENCIES AND FOURTH NORMAL FORM (4NF)

In Fourth Normal Form (4NF), a record type should contain two or more independent multivalued facts about an entity. Also, the record must satisfy Third Normal Form (3NF). Consider,  $C(\text{COMPANY}, \text{WORK}, \text{FUND})$ , where COMPANY may have several WORK from several FUND. COMPANY and WORK have many-to-many relationship, and COMPANY and FUND have many-to-many relationship. Using Fourth Normal Form (4NF), these two relationships should be represented as  $C1(\text{COMPANY}, \text{WORK})$  and  $C2(\text{COMPANY}, \text{FUND})$

Now, suppose we extend above example to include PROJECT as follows :

- i. COMPANY have WORK for certain PROJECT.
- ii. COMPANY gets FUND for certain PROJECT.

If there is no functional dependency between the WORK and FUND than an COMPANY have against PROJECT, then we could treat this as two independent many-to-many relationships of the form (COMPANY, PROJECT)  $\rightarrow\rightarrow$  WORK and (COMPANY, PROJECT)  $\rightarrow\rightarrow$  FUND, where (COMPANY, PROJECT) represents a combination of an COMPANY with PROJECT. A record including (COMPANY, PROJECT, WORK, FUND) would violate Fourth Normal Form (4NF). Two records, containing C11(COMPANY, PROJECT, WORK) and C21(COMPANY, PROJECT, FUND), respectively, would satisfy Fourth Normal Form (4NF). So, we observed that, Fourth Normal Form (4NF) uses multivalued dependencies without loss of information.<sup>[6]</sup>

### 3. MULTIVALUED DEPENDENCIES AND FIFTH NORMAL FORM (5NF)

Fifth Normal Form (5NF) reconstruct information to smaller pieces of information that can be maintained with less redundancy. Consider two records, C11(COMPANY, PROJECT, WORK) and C21(COMPANY, PROJECT, FUND), which satisfy Fourth Normal Form (5NF).

Now, suppose that certain rule was in effect :

- i. COMPANY have PROJECTS
- ii. PROJECTS have number of WORK
- iii. For each WORK there was FUND

In this case, it turn out that we can reconstruct all the true facts from a normalized form consisting to three separate record types, each containing two fields. C1(COMPAY, PROJECT), C2(PROJECT, WORK), and C3(WORK, FUND). These three record types are in the Fifth Normal Form (5NF), whereas the corresponding two records previously shown is not. Also, these three records in Fifth Normal Form (5NF) is also in 4NF, 3NF, 2NF, and 1NF. So, we observed that, Fifth Normal Form (5NF) is achieved using projection and join without redundancy.

For, these three records which are in Fifth Normal Form (5NF), we observe that :

- i. COMPANY may have several PROJECT, but each PROJECT may have not have same WORK.
- ii. PROJECT may have several WORKS, but each WORK may not have same FUND.

#### 4. CONCLUSION

Fourth Normal Form (4NF) and Fifth Normal Form (5NF) both supports multivalued dependencies. Multivalued dependency extent the understanding of the logical design of relational databases which leads to Fourth Normal Form (4NF). Fifth Normal Form (5NF) is ultimate normal form when projection and joins are allowed. Though, Fourth Normal Form (4NF) and Fifth Normal Form (5NF) deals with multivalued dependencies, constraint of functional dependency on field differentiate both. A record in Fifth Normal Form (5NF) violates Fourth Normal Form (4NF) by decomposing into two records if constraints on fields are not independent.

#### REFERENCES

1. Edgar Codd. "A Relational Model of Data for Large Shared Data Banks". Communications of the ACM. Volume 13, Number 6. June 1970. Pages 377-387.
2. Edgar Codd. "Further Normalization of the Data Base Relational Model". Courant Computer Science Symposia Series 6, "Data Base Systems", New York City. May 24-25, 1971.
3. Ronald Fagin, "Multivalued Dependencies and a New Normal Form for Relational Database". ACM Transactions on Database Systems. Volume 2, Number 3. September 1977. Pages 262-278.
4. Ronald Fagin, "Normal Forms and Relational Database Operators". ACM SIGMOD International Conference on Management of Data, Boston, Mass. May 31 – June 1, 1979.
5. William Kent, "A Simple Guide to Five Normal Forms in Relational Database Theory", Communications of the ACM. Volume 26, Number 2. February 1982. Pages 120-125.
6. Wu, M.S. (1992), The Practical Need for Fourth Normal Form. In Proceedings of the 23rd ACM SIGCSE technical symposium on Computer Science education, Kansas City, Missouri, (pp. 19-23)
7. Yu Liping, Pan Yunto, Wu Yishan. "Research on Data Normalization Methods in Multi-Attribute Evaluation". IEEE – Computational Intelligence and Software Engineering, 2009. International Conference on 11-13 Dec. 2009. pp : 1-5.
8. Date, C.J. Database in Depth: Relational Theory for Practitioners. O'Reilly (2005), p. 142.
9. C.J. Date, Hugh Darwen, Nikos Lorentzos. Temporal Data and the Relational Model. Morgan Kaufmann (2002), p. 176.