# Third normal form

From Wikipedia, the free encyclopedia

The **third normal form** (**3NF**) is a normal form used in database normalization. 3NF was originally defined by E.F. Codd[1] in 1971. Codd's definition states that a table is in 3NF if and only if both of the following conditions hold:

- The table is in second normal form (**2NF**)
- No non-prime attribute of the table is transitively dependent on a candidate key

## Contents



A **non-prime attribute** is an attribute that does not belong to any candidate key. A **transitive dependency** is a functional dependency $X \rightarrow Z$ in which $Z$ is not immediately dependent on $X$, but rather on a third set of attributes $Y$ which depends on $X$. That is, $X \rightarrow Z$ by virtue of $X \rightarrow Y$ and $Y \rightarrow Z$.

An alternative formulation of Codd's definition, given by Carlo Zaniolo[2] in 1982, is this: a table is in 3NF if and only if, for each of its functional dependencies $X \rightarrow A$, **at least one** of the following conditions holds:

- $X$ contains $A$, or
- $X$ is a superkey, or
- $A$ is a **prime attribute** (i.e., $A$ is contained within a candidate key)

Zaniolo's definition has the advantage of giving a clear sense of the difference between 3NF and the more stringent Boyce-Codd normal form (**BCNF**). BCNF simply eliminates the third alternative ("$A$ is a prime attribute").

## Example

An example of a 2NF table that fails to meet the requirements of 3NF is:

**Tournament Winners**

| Tournament | Year | Winner | Winner Date of Birth |
|---|---|---|---|
| Indiana Invitational | 1998 | Al Fredrickson | 21 July 1975 |
| Cleveland Open | 1999 | Bob Albertson | 28 September 1968 |
| Des Moines Masters | 1999 | Al Fredrickson | 21 July 1975 |
| Indiana Invitational | 1999 | Chip Masterson | 14 March 1977 |

The only candidate key is {Tournament, Year}.

The breach of 3NF occurs because the non-prime attribute Winner Date of Birth is transitively dependent on {Tournament, Year} via the non-prime attribute Winner. The fact that Winner Date of Birth is functionally dependent on Winner makes the table vulnerable to logical inconsistencies, as there is nothing to stop the same person from being shown with different dates of birth on different records.

In order to express the same facts without violating 3NF, it is necessary to split the table into two:

**Tournament Winners**

| Tournament | Year | Winner |
|---|---|---|
| Indiana Invitational | 1998 | Al Fredrickson |
| Cleveland Open | 1999 | Bob Albertson |
| Des Moines Masters | 1999 | Al Fredrickson |
| Indiana Invitational | 1999 | Chip Masterson |

**Player Dates of Birth**

| Player | Date of Birth |
|---|---|
| Chip Masterson | 14 March 1977 |
| Al Fredrickson | 21 July 1975 |
| Bob Albertson | 28 September 1968 |

Update anomalies cannot occur in these tables, which are both in 3NF.

# Normalization beyond 3NF

Most 3NF tables are free of update, insertion, and deletion anomalies. Certain types of 3NF tables, rarely met with in practice, are affected by such anomalies; these are tables which either fall short of Boyce-Codd normal form (BCNF) or, if they meet BCNF, fall short of the higher normal forms 4NF or 5NF.

# References

1. ^ Codd, E.F. "Further Normalization of the Data Base Relational Model." (Presented at Courant Computer Science Symposia Series 6, "Data Base Systems," New York City, May 24th-25th, 1971.) IBM Research Report RJ909 (August 31st, 1971). Republished in Randall J. Rustin (ed.), *Data Base Systems: Courant Computer Science Symposia Series 6*. Prentice-Hall, 1972.
2. ^ Zaniolo, Carlo. "A New Normal Form for the Design of Relational Database Schemata." *ACM Transactions on Database Systems* 7(3), September 1982.

# Further reading

- Date, C. J. (1999), *An Introduction to Database Systems (http://www.aw-bc.com/catalog/academic/product/0,1144,0321197844,00.html)* (8th ed.). Addison-Wesley Longman. ISBN 0-321-19784-4.
- Kent, W. (1983) *A Simple Guide to Five Normal Forms in Relational Database Theory (http://www.bkent.net/Doc/simple5.htm)*, Communications of the ACM, vol. 26, pp. 120-125

# External links

- Litt's Tips: Normalization (http://www.troubleshooters.com/littstip/ltnorm.html)
- Rules Of Data Normalization (http://www.datamodel.org/NormalizationRules.html)
- Database Normalization Basics (http://databases.about.com/od/specificproducts/a/normalization.htm) by Mike Chapple (About.com)

- An Introduction to Database Normalization (http://dev.mysql.com/tech-resources/articles/intro-to-normalization.html) by Mike Hillyer.
- Normalization (http://www.utexas.edu/its/windows/database/datamodeling/rm/rm7.html) by ITS, University of Texas.
- A tutorial on the first 3 normal forms (http://phlonx.com/resources/nf3/) by Fred Coulson
- Description of the database normalization basics (http://support.microsoft.com/kb/283878) by Microsoft
- Database Debunkings (http://www.dbdebunk.com/): Fabian Pascal, Chris Date, and Hugh Darwen

**Topics in Database normalization**

First normal form | Second normal form | **Third normal form**
Boyce-Codd normal form | Fourth normal form | Fifth normal form | Domain/key normal form | Sixth normal form
Denormalization

Retrieved from "http://en.wikipedia.org/wiki/Third_normal_form"

Category: Database normalization