

Konsep Normalisasi dan Anomali Tabel

Daftar Isi

Konsep Normalisasi dan Anomali Tabel

Pengantar

11.1. Pengantar Konsep Normalisasi

 Tahapan Normalisasi

 Bentuk Normal Kesatu

 Bentuk Normal Kedua

 Bentuk Normal Ketiga

 Boyce Codd Normal Form

11.2 Konsep Normalisasi

11.3 Modification Anomaly

 Deletion Anomaly

 Insertion Anomaly

 Update Anomaly

11.4. Keuntungan dan Kerugian Normalisasi

11.5. Latihan Soal

11. Konsep Normalisasi dan Anomali Tabel

Pengantar

Modul ini berisi tentang konsep normalisasi dan anomali tabel. Normalisasi adalah suatu teknik untuk menghasilkan sekumpulan relasi/tabel yang memiliki karakteristik tertentu, untuk memenuhi kebutuhan organisasi. Proses normalisasi diperlukan dalam membentuk tabel-tabel yang normal. Tabel-tabel yang normal dalam sebuah database hanya dapat dibentuk jika tabel-tabel tersebut bebas dari anomali. Terdapat beberapa jenis anomali, antara lain insert anomali, delete anomali, dan update anomali. Diperlukan kemampuan untuk mengidentifikasi ketiga jenis anomali tersebut. Proses normalisasi diharapkan dapat menghilangkan ke-3 jenis anomali ini. Selain itu proses normalisasi juga dapat menghilangkan kerangkapan data dan memudahkan dalam modifikasi data.

Penulis berharap, setelah mempelajari modul ini mahasiswa mampu menggunakan konsep normalisasi dan anomali tabel sehingga dapat membentuk relasi/tabel yang normal dalam sebuah database untuk memenuhi kebutuhan data organisasi.

Kompetensi

Setelah mempelajari materi pokok bahasan disini, mahasiswa diharapkan:

- Mampu menjelaskan konsep normalisasi.
- Mampu memahami konsep anomali.
- Mampu mengidentifikasi anomali dalam sebuah relasi.

11.1 Pengantar Konsep Normalisasi

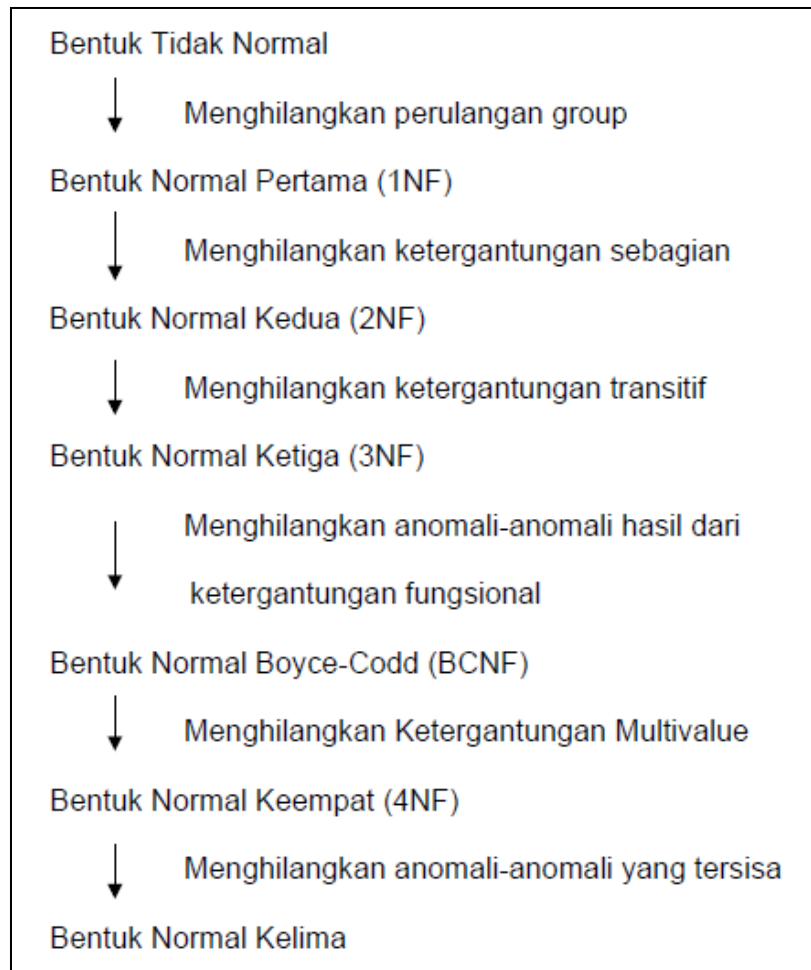
Ketika kita akan merancang sebuah database untuk sistem database relasional maka tujuan utama dalam mengembangkan model data logikalnya adalah untuk menciptakan representasi data yang tepat bagi hubungan maupun batasannya (*constraints*). Untuk mencapai hal tersebut kita harus mengidentifikasi sekumpulan relasi yang tepat. Teknik yang digunakan untuk membantu mengidentifikasi relasi tersebut disebut dengan *normalisasi*.

Proses normalisasi pertama kali diperkenalkan oleh E.F. Codd pada tahun 1972. Normalisasi sering dilakukan sebagai serangkaian tes pada relasi untuk menentukan apakah suatu relasi sudah memenuhi atau masih melanggar persyaratan bentuk normal tertentu. Pada awalnya terdapat 3 jenis bentuk normal yang diusulkan, yaitu bentuk normal ke satu (1NF), bentuk normal kedua (2NF), dan bentuk normal ketiga (3NF). Setelah itu R. Boyce dan E.F. Codd memperkenalkan *Boyce Codd Normal Form (BCNF)*, bentuk normal yang lebih tinggi dari bentuk normal ketiga pada tahun 1974. Pada perkembangan selanjutnya muncul pula bentuk normal ke-4 dan ke-5.

Tahapan Normalisasi

Berikut ini adalah tahapan normalisasi. Penjelasan lebih rinci akan diberikan pada modul 12.

Bentuk normal kedua (2NF) adalah lebih baik dari bentuk normal kesatu (1NF); bentuk normal ketiga (3NF) adalah lebih baik dari Bentuk normal kedua 2NF. Untuk kepentingan rancangan database bisnis, bentuk normal ketiga (3NF) adalah bentuk terbaik dalam proses normalisasi (sudah mencukupi). Normalisasi dengan level paling tinggi tidak selalu diharapkan. Jadi normalisasi dilakukan, sepanjang dirasa sudah cukup normal (dengan mengikuti prasyarat normalisasi diatas).



Gambar 11.1 Tahapan Normalisasi

Bentuk Normal Kesatu (1NF)

Suatu relasi dikatakan sudah memenuhi bentuk normal kesatu bila setiap data bersifat atomik yaitu setiap irisan baris dan kolom hanya mempunyai satu nilai data

Tujuan 1NF adalah:

- Membuang adanya pengulangan (Redudansi) data,
- Menghindari adanya pencatatan Null Value, dan
- Menjaga setiap entry data dari relasi (perpotongan bariskolom) memiliki maksimal satu nilai tunggal.

Bentuk Normal Kedua (2NF)

Suatu relasi dikatakan sudah memenuhi bentuk normal kedua bila relasi tersebut sudah memenuhi bentuk normal kesatu, dan atribut yang bukan key sudah tergantung penuh terhadap keynya.

Bentuk Normal Ketiga (3NF)

Suatu relasi dikatakan sudah memenuhi bentuk normal ketiga bila relasi tersebut sudah memenuhi bentuk normal kedua dan atribut yang bukan key tidak tergantung transitif terhadap keynya.

Boyce Codd Normal Form (BCNF)

Suatu relasi R dikatakan dalam bentuk BCNF jika dan hanya jika setiap atribut kunci (Key) pada suatu relasi adalah kunci kandidat (candidate key).

11.2 Konsep Normalisasi

Suatu rancangan database disebut buruk jika terdapat beberapa fakta berikut ini: data yang sama tersimpan di beberapa tempat yang berbeda, tidak mampu untuk menghasilkan informasi tertentu, kehilangan informasi, terjadi duplikasi data (pengulangan) yang menyebabkan pemborosan ruang penyimpanan serta timbulnya null value. Fakta ini disebabkan oleh adanya anomali (penyimpangan).

Sebaliknya, tabel-tabel dalam sebuah database yang baik harus memenuhi aturan normalisasi, yaitu bebas dari ketergantungan struktural atau anomali yang disebabkan oleh modifikasi data. Ini disebut dengan *modification anomaly*. *Modification anomaly* dibagi menjadi 3, yaitu: *deletion anomaly*, *insertion anomaly*, dan *update anomaly*. Dengan demikian tujuan dari normalisasi adalah untuk menghilangkan duplikasi/kerangkapan data, mengurangi kompleksitas data, dan mempermudah modifikasi data.

Pada bentuk normal, setiap baris dalam suatu tabel harus unik, setidaknya pada satu atribut yang disebut sebagai primary key. Tabel-tabel pada database dihubungkan dengan menanamkan primary key dari suatu tabel ke tabel yang berhubungan sebagai foreign key. Nilai atribut di setiap kolom harus dari kelas atau tipe data yang sama. Dan setiap kolom pada tabel harus memiliki nama yang unik.

11.3 Modification Anomaly

Pada bagian ini akan dijelaskan tentang 3 jenis modification anomaly, yaitu *deletion anomaly*, *insertion anomaly*, dan *update anomaly*

Deletion Anomaly

Perhatikan tabel EQUIPMENT_REPAIR

ItemNumber	Type	AcquisitionCost	RepairNumber	RepairDate	RepairAmount
100	Drill Press	3500	2000	5/5/2004	375
200	Lathe	4750	2100	5/7/2004	255
100	Drill Press	3500	2200	5/19/2004	178
300	Mill	27300	2300	5/19/2004	1785
100	Drill Press	3500	2400	5/11/2004	0
100	Drill Press	3500	2600	6/1/2004	275

Gambar 11.2 Tabel EQUIPMENT_REPAIR

Misalkan kita menghapus data dengan RepairNumber 2100. Data yang terhapus bukan hanya data tentang repair, tetapi data tentang mesin juga akan terhapus (informasi tentang mesin Lathe dengan AcquisitionPrice 4750 akan hilang).

Dari kasus di atas, ketika kita menghapus satu baris pada tabel tersebut, kita akan kehilangan 2 fakta sekaligus (machine dan repair) karena struktur tabel tersebut memaksa demikian. Kondisi tersebut dinamakan **deletion anomaly**.

Insertion Anomaly

Pada tabel EQUIPMENT_REPAIR, misalkan kita ingin memasukan data *repair*. Untuk memasukan data tersebut selain diperlukan:

- Data *Repair*, yaitu: RepairNumber, RepairDate, dan RepairAmount.
- Juga perlu data *Machine*, yaitu: ItemNumber, Type, dan AcquisitionCost.

Artinya ketika kita hanya perlu memasukan data tentang satu entity (*repair*), kita diharuskan memasukan data tentang entity lain (*machine*) karena kondisi struktur tabelnya mengharuskan demikian. Kondisi ini dinamakan **insertion anomaly**

Update Anomaly

Pada tabel EQUIPMENT_REPAIR, misalkan kita akan mengubah data (update). Misalkan kita akan mengubah data baris terakhir menjadi

ItemNumber	Type	AcquisitionCost	RepairNumber	RepairDate	RepairAmount
100	Drill Press	3500	2000	5/5/2004	375
200	Lathe	4750	2100	5/7/2004	255
100	Drill Press	3500	2200	5/19/2004	178
300	Mill	27300	2300	5/19/2004	1785
100	Drill Press	3500	2400	5/11/2004	0
100	Drill Press	5500	2600	6/1/2004	275

Gambar 11.3 Tabel EQUIPMENT_REPAIR (update anomaly)

Terlihat bahwa “Drill Press” memiliki dua nilai AcquisitionCost yang berbeda. Sehingga terjadi *data inconsistency*. Kondisi demikian dinamakan **update anomaly**

Ringkasan teori normalisasi dapat digambarkan dengan tabel berikut ini.

Source of Anomaly	Normal Forms	Design Principles
Functional dependency	1NF, 2NF, 3NF, BCNF (Boyce-Codd Normal Form)	BCNF: Design tables so that every determinant is a candidat key
Multivalued dependency	4NF	4NF: Move each multivalued dependency to table of its own
Data constraints and oddities	5NF, DK/NF	DK/NF: Make every constraint a logical consequence of candidate keys and domains

11.4 Keuntungan dan Kerugian Normalisasi

Keuntungan normalisasi

- Mengeliminasi *modification anomalies*
- Mengurangi duplikasi data
 - ✓ Mengeliminasi masalah *data integrity*
 - ✓ *Space* penyimpanan data lebih hemat

Kerugian

- SQL query akan lebih rumit, terutama untuk mengakses data dari banya tabel
- DBMS akan berjalan lebih lambat karena perlu kerja yang ekstra

11.5. Latihan Soal

1. Apa yang dimaksud dengan bentuk normal?
2. Apa yang dimaksud dengan modification anomaly? Ada berapa jenis modification anomaly?
3. Beri contoh modification anomaly (contoh tidak boleh sama dengan modul).
4. Jelaskan mengapa data yang rangkap dapat menyebabkan masalah integritas data?
5. Apakah yang disebut dengan 1NF, 2NF, 2NF, 3NF, dan BCNF?

Daftar Pustaka

1. Connolly, Thomas M., Begg, Carolyn E., and Strachan, Anne D., Database System. A practical pproach to Design, Implementation, and Management, Addison Wesley Company, 1996.
2. David M. Kroenke, *Basis data Processing, Fundamentals, Design and Implementation*, 12nd Edition, Prentice-Hall Int'l Edition, 2012
3. Hall, James A., *Accounting Information Systems*, 6th edition, South Western Cengage Learning, 2008

