



UNIVERSITAS
INDONESIA

Veritas, Probatum, Justitia

FAKULTAS

ILMU
KOMPUTER

Slide 11

Database Design Theory & Normalization (part 2)

CSF2600700 - BASIS DATA

SEMESTER GENAP 2019/2020





Tujuan Pemelajaran

- Mengevaluasi model basis data relasional yang telah dibuat berdasarkan teori desain database relasional sehingga dapat menghasilkan skema database yang “baik” dalam level conceptual maupun physical.
- *Implicit goals: information preservation and minimum redundancy.*





Outline

1. Panduan Informal dalam Merancang Basis Data Relasional



2. *Functional Dependency*



3. Normalisasi Berdasarkan *Primary Key*

4. *General Normal Form*

5. *Boyce Codd Normal Form*







Normalisasi Berdasarkan *Primary Key*





Normalisasi

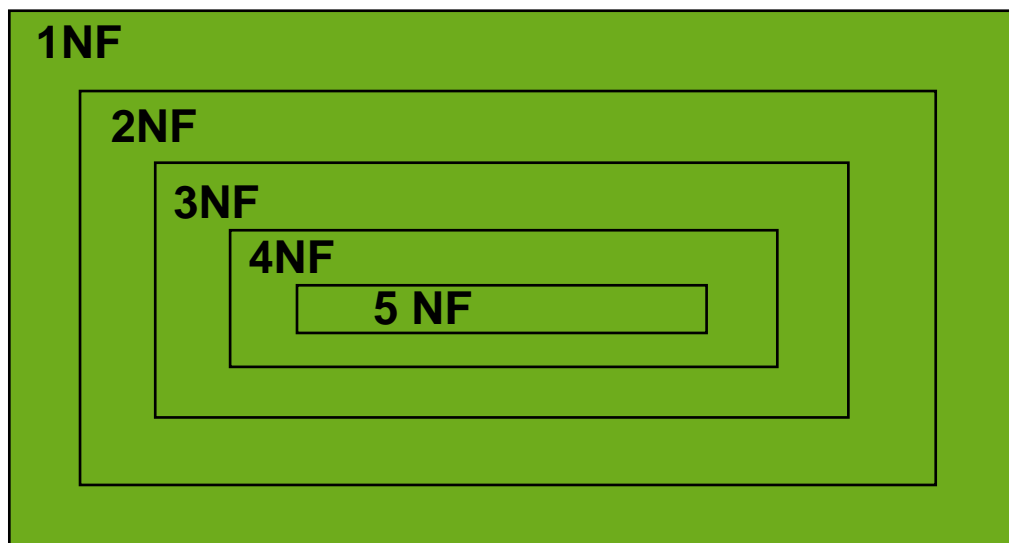
- Normalisasi (Normalization):
 - Proses dekomposisi relasi yang masih “buruk” dengan memecah atribut-atributnya untuk membentuk beberapa relasi
- Bentuk Normal (Normal Form)
 - Kondisi (dengan menggunakan FD dan key) yang menentukan apakah suatu skema relasi memenuhi kriteria tertentu





Bentuk Normal

- Ada beberapa bentuk normal berdasarkan sejumlah kriteria:
 - *Primary keys* (1NF, 2NF, 3NF)
 - *All Candidate Keys* (2NF, 3NF, BCNF)
 - *Multivalued Dependencies* (4NF)
 - *Join Dependencies* (5NF)





Penggunaan Bentuk Normal

- Normalisasi perlu dilakukan agar rancangan basis data yang dihasilkan berkualitas baik dan memenuhi sifat yang diinginkan
- Normalisasi dalam praktiknya sulit dilakukan jika batasan-batasan yang menjadi dasar normalisasi sulit dimengerti atau sulit dideteksi
- Perancang basis data tidak perlu melakukan normalisasi sampai bentuk tertinggi
 - Normalisasi biasanya dilakukan sampai mencapai 3NF atau BCNF
- Denormalisasi: kebalikan dari proses normalisasi, yakni menggabungkan beberapa relasi, membawa ke bentuk normal yang lebih rendah





Key Attributes

- **Superkey** dari relasi **R**: himpunan attribute dari R yang dapat membedakan satu tuple dengan tuple lainnya pada R.
- **Key K**: Superkey minimal, sedemikian hingga penghapusan salah satu attribute dari K akan menyebabkan K tidak lagi menjadi key
- Jika relasi punya beberapa key, masing-masing disebut **candidate key**. Salah satu dari candidate key dipilih menjadi **primary key**.
- **Prime attribute**: Attribute yang menjadi anggota dari candidate key
- **Non prime attribute**: Attribute yang bukan merupakan anggota candidate key manapun.





Bentuk Normal Pertama (1NF)

- Bentuk normal pertama (1NF) merupakan bagian dari definisi relasi.
- Bentuk normal pertama tidak mengizinkan:
 - *Composite attributes,*
 - *Multivalued attributes,*
 - *Nested relations.*





Normalisasi dari Multivalued ke 1NF

(a)

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations

(b)

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations
Research	5	333445555	{Bellaire, Sugarland, Houston}
Administration	4	987654321	{Stafford}
Headquarters	1	888665555	{Houston}

(c)

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocation
Research	5	333445555	Bellaire
Research	5	333445555	Sugarland
Research	5	333445555	Houston
Administration	4	987654321	Stafford
Headquarters	1	888665555	Houston

Figure 10.8


Normalization into 1NF.

(a) A relation schema that is not in 1NF. (b) Example state of relation DEPARTMENT. (c) 1NF version of the same relation with redundancy.



Sample Solution for Example 1

DEPARTMENT

Dname	<u>Dnumber</u>	Dmgr_ssn	Dlocations
			

1st Option

Dname	<u>Dnumber</u>	Dmgr_ssn	<u>Dlocation</u>
-------	----------------	----------	------------------

2nd Option

DEPT_1

Dname	<u>Dnumber</u>	Dmgr_ssn
-------	----------------	----------

DEPT_2

<u>Dnumber</u>	<u>Dlocation</u>
----------------	------------------

3rd Option

Dname	<u>Dnumber</u>	Dmgr_ssn	<u>Dlocation1</u>	<u>Dlocation2</u>	<u>Dlocation3</u>
-------	----------------	----------	-------------------	-------------------	-------------------





Normalisasi dari Nested Relation ke 1NF

(a)

EMP_PROJ		Projs	
Ssn	Ename	Pnumber	Hours

(b)

EMP_PROJ			
Ssn	Ename	Pnumber	Hours
123456789	Smith, John B.	1	32.5
		2	7.5
666884444	Narayan, Ramesh K.	3	40.0
453453453	English, Joyce A.	1	20.0
		2	20.0
333445555	Wong, Franklin T.	2	10.0
		3	10.0
		10	10.0
		20	10.0
999887777	Zelaya, Alicia J.	30	30.0
		10	10.0
987987987	Jabbar, Ahmad V.	10	35.0
		30	5.0
987654321	Wallace, Jennifer S.	30	20.0
		20	15.0
888665555	Borg, James E.	20	NULL

(c)

EMP_PROJ1	
Ssn	Ename

EMP_PROJ2		
Ssn	Pnumber	Hours

Figure 10.9

Normalizing nested relations into 1NF. (a) Schema of the EMP_PROJ relation with a *nested relation* attribute PROJS. (b) Example extension of the EMP_PROJ relation showing nested relations within each tuple. (c) Decomposition of EMP_PROJ into relations EMP_PROJ1 and EMP_PROJ2 by propagating the primary key.





Bentuk Normal Kedua (2NF)

- Menggunakan konsep FD dan primary key
- Definisi:
 - *Full functional dependency*: suatu FD $X \rightarrow Y$ sedemikian hingga jika salah satu attribute dari X dibuang, maka FD tersebut tidak ada lagi.
- Contoh:
 - $\{SSN, PNUMBER\} \rightarrow HOURS$ merupakan full FD, karena tidak ada $SSN \rightarrow HOURS$ dan $PNUMBER \rightarrow HOURS$
 - $\{SSN, PNUMBER\} \rightarrow ENAME$ merupakan *partial dependency*, bukan full FD karena terdapat dependency $SSN \rightarrow ENAME$





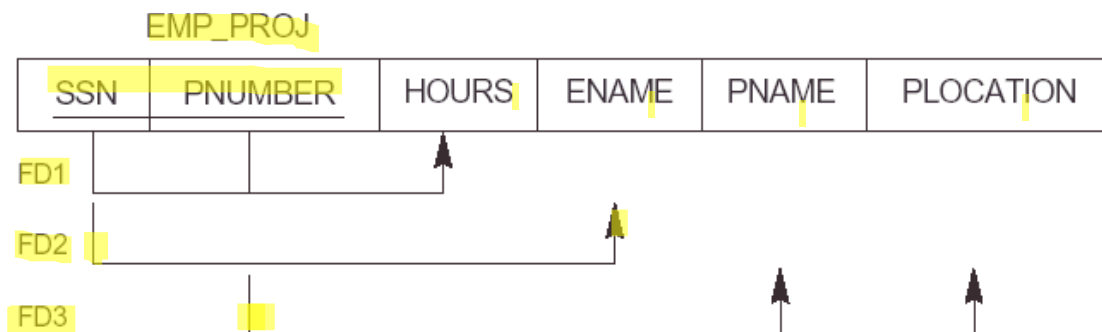
Bentuk Normal Kedua (2NF)

- Suatu relasi R berada dalam bentuk normal kedua (2NF) jika R berada dalam 1NF dan setiap nonprime attribute A dalam relasi R bersifat full functional dependent terhadap primary key.
- Suatu relasi R dapat didekomposisi ke relasi-relasi yang memenuhi 2NF melalui proses normalisasi 2NF



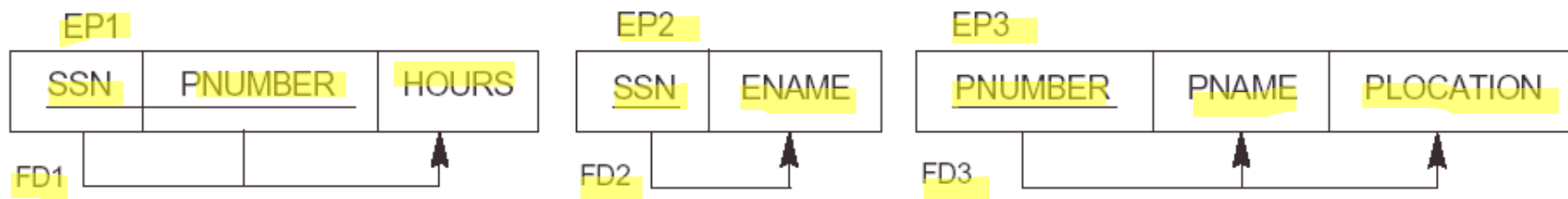


Contoh Normalisasi 2NF



- **ENAME** fully dependent on **SSN**
- **PNAME**, **PLOCATION** fully dependent on **PNUMBER**

Normalisasi 2NF





Bentuk Normal Ketiga (3NF)

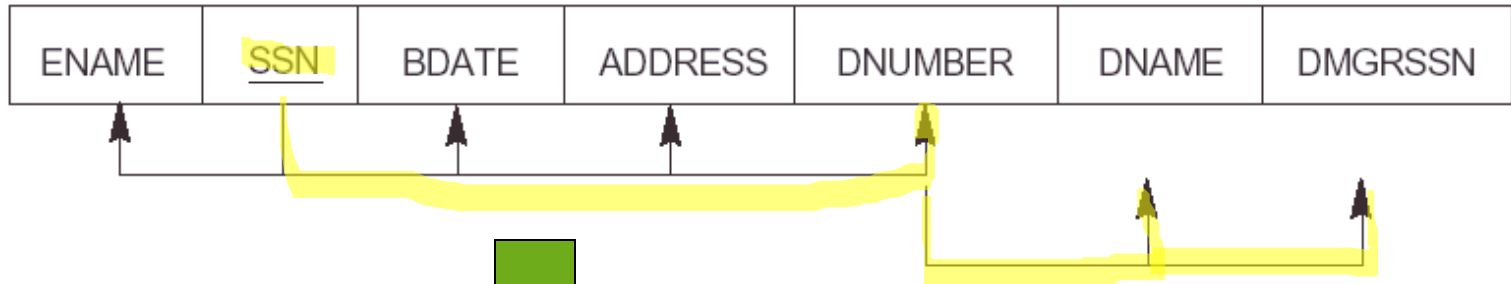
- Suatu relasi R berada dalam bentuk normal ketiga (3NF) jika R berada dalam 2NF dan tidak ada nonprime attribute A di R yang memiliki dependensi transitif terhadap primary key.
- Suatu relasi R dapat didekomposisi ke relasi-relasi yang memenuhi 3NF melalui proses normalisasi 3NF
- Catatan:
 - Pada FD $X \rightarrow Y$ dan $Y \rightarrow Z$, dengan X sebagai primary key, kita mempertimbangkan ini sebagai sebuah problem hanya jika Y bukan sebuah candidate key.
 - Jika Y merupakan candidate key, tidak ada masalah dengan dependensi transitif ini
 - Contoh: Relasi EMP (SSN, Emp#, Salary) tidak melanggar 3NF karena meskipun terdapat FD $SSN \rightarrow Emp\# \rightarrow Salary$, Emp# merupakan candidate key



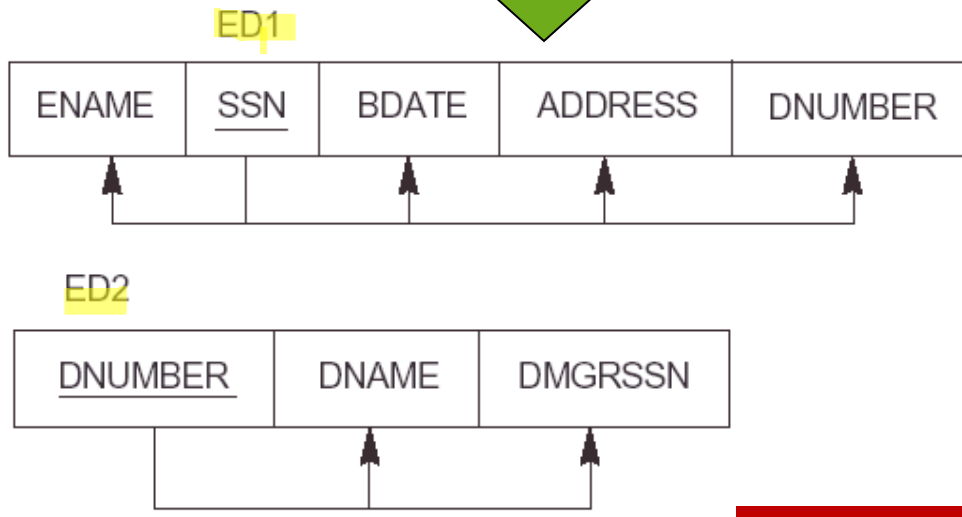
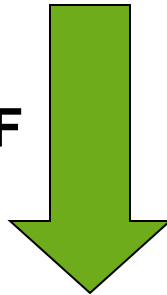


Contoh Normalisasi 3NF

EMP_DEPT



Normalisasi 3NF





Bentuk Normal Didefinisikan secara Informal

- *1st normal form*
 - *All attributes depend on the key*
- *2nd normal form*
 - *All attributes depend on the whole key*
- *3rd normal form*
 - *All attributes depend on nothing but the key*





Rangkuman Bentuk Normal Berdasarkan Primary Key

Table 10.1

Summary of Normal Forms Based on Primary Keys and Corresponding Normalization

Normal Form	Test	Remedy (Normalization)
First (1NF)	Relation should have no multivalued attributes or nested relations.	Form new relations for each multivalued attribute or nested relation.
Second (2NF)	For relations where primary key contains multiple attributes, no nonkey attribute should be functionally dependent on a part of the primary key.	Decompose and set up a new relation for each partial key with its dependent attribute(s). Make sure to keep a relation with the original primary key and any attributes that are fully functionally dependent on it.
Third (3NF)	Relation should not have a nonkey attribute functionally determined by another nonkey attribute (or by a set of nonkey attributes). That is, there should be no transitive dependency of a nonkey attribute on the primary key.	Decompose and set up a relation that includes the nonkey attribute(s) that functionally determine(s) other nonkey attribute(s).



General Normal Form





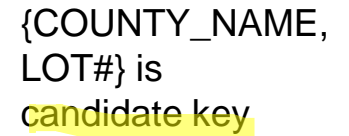
Definisi dalam General Normal Form

- Definisi-definisi sebelumnya hanya mempertimbangkan primary key
- Definisi selanjutnya mempertimbangkan semua candidate key pada suatu relasi
- Skema relasi R berada pada 2NF (general definition) jika setiap *nonprime attribute A* pada R bersifat *full functional dependent* pada *setiap key* pada R





LOTS



Normalisasi 2NF



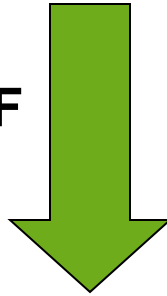


Definisi dalam General Normal Form

- Definisi:
 - Superkey pada relasi R: himpunan attributes S dari R yang berisi key dari R
 - Skema relasi R berada dalam 3NF dengan syarat jika terdapat FD $X \rightarrow Y$ maka:
 - (a) X merupakan superkey dari R atau
 - (b) Y merupakan prime attribute dari R
- BCNF tidak membolehkan kondisi (b) di atas



LOTS1





Contoh Normalisasi 3NF

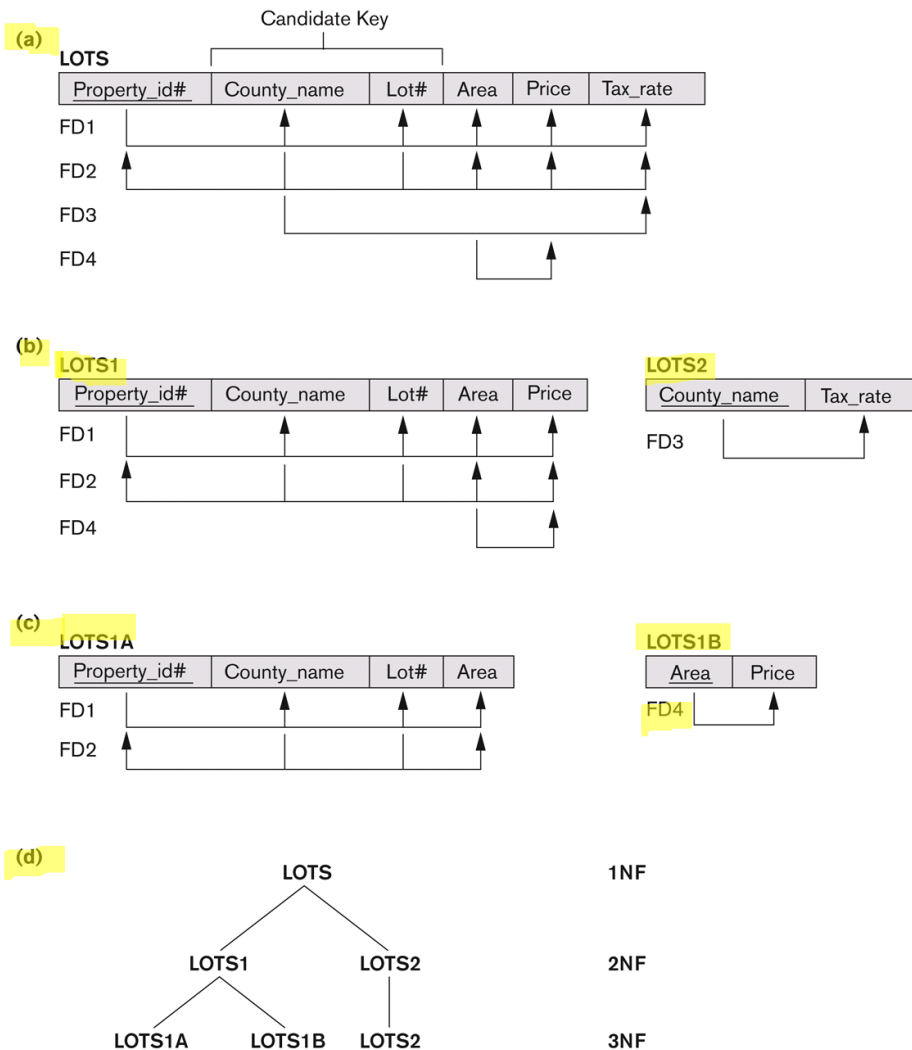


Figure 10.11

Normalization into 2NF and 3NF. (a) The LOTS relation with its functional dependencies FD1 through FD4. (b) Decomposing into the 2NF relations LOTS1 and LOTS2. (c) Decomposing LOTS1 into the 3NF relations LOTS1A and LOTS1B. (d) Summary of the progressive normalization of LOTS.



Boyce Codd Normal Form





Boyce Codd Normal Form (BCNF)

- Suatu skema relasi R berada pada BCNF dengan syarat jika terdapat FD $X \rightarrow Y$ pada R, maka X merupakan superkey dari R.
- Setiap BCNF pasti memenuhi 3NF
- Ada relasi yang berada pada 3NF namun tidak BCNF
- Tujuan normalisasi umumnya untuk mencapai 3NF atau BCNF





Contoh BCNF

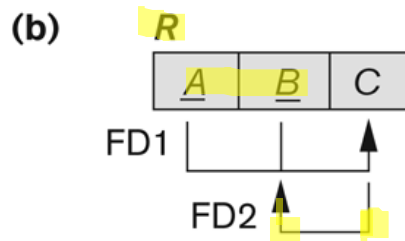
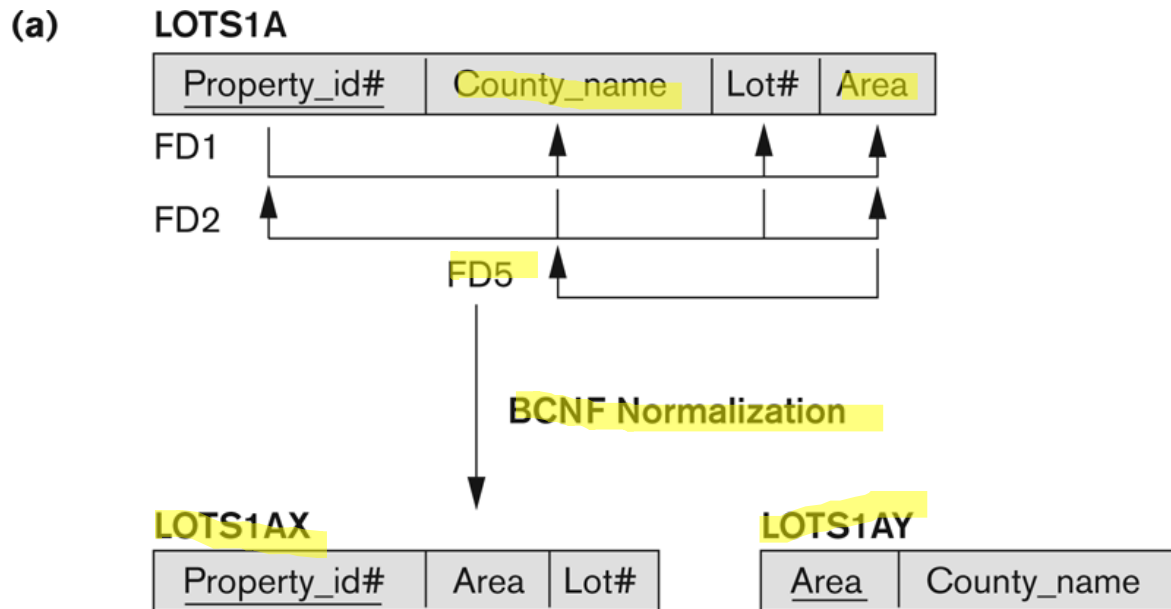


Figure 10.12

Boyce-Codd normal form. (a) BCNF normalization of LOTS1A with the functional dependency **FD2** being lost in the decomposition. (b) A schematic relation with FDs; it is in 3NF, but not in BCNF.





Contoh: Relasi pada 3NF Namun tidak BCNF

TEACH

Student	Course	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar

Figure 10.13

A relation TEACH that
is in 3NF but not
BCNF.





Normalisasi BCNF

Ada 2 FD pada relasi
TEACH:

- fd1: { student, course} -> instructor
- fd2: instructor -> course

{student, course}
merupakan candidate key
untuk teach

- Relasi ini berada pada 3NF
namun tidak pada BCNF

Relasi yang belum BCNF
dapat didekomposisi
untuk mencapai BCNF,
namun kadang-kadang
dapat menghilangkan FD
yang semula telah ada

TEACH

Student	Course	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar





Normalisasi BCNF

Ada 3 dekomposisi yang mungkin untuk relasi TEACH:

- {student, instructor} and {student, course}
- {course, instructor} and {course, student}
- {instructor, course} and {instructor, student}

Semua dekomposisi tersebut kehilangan FD1 : { **student, course** }
-> **instructor**

- Semua FD yang ada pada relasi sedapat mungkin dijaga, namun dapat dimaklumi jika hilang pada normalisasi BCNF
- Namun sifat lossless dan non-additivity property setelah dekomposisi harus tetap dijaga

TEACH

Student	Course	Instructor
Narayan	Database	Mark
Smith	Database	Navathe
Smith	Operating Systems	Ammar
Smith	Theory	Schulman
Wallace	Database	Mark
Wallace	Operating Systems	Ahamad
Wong	Database	Omiecinski
Zelaya	Database	Navathe
Narayan	Operating Systems	Ammar





Referensi

- Elmasri & Navathe, Fundamental of Database Systems, 7th Edition, Chapter 15, 2015

