

Pertemuan 6

Normalization

Nur Iksan

Pendidikan Teknik Informatika dan Komputer UNNES

normalisasi

Normalisasi : Teknik/pendekatan yang digunakan dalam membangun disain logik database relasional melalui organisasi himpunan data dengan tingkat ketergantungan fungsional dan keterkaitan yang tinggi sedemikian sehingga menghasilkan struktur tabel yang normal.

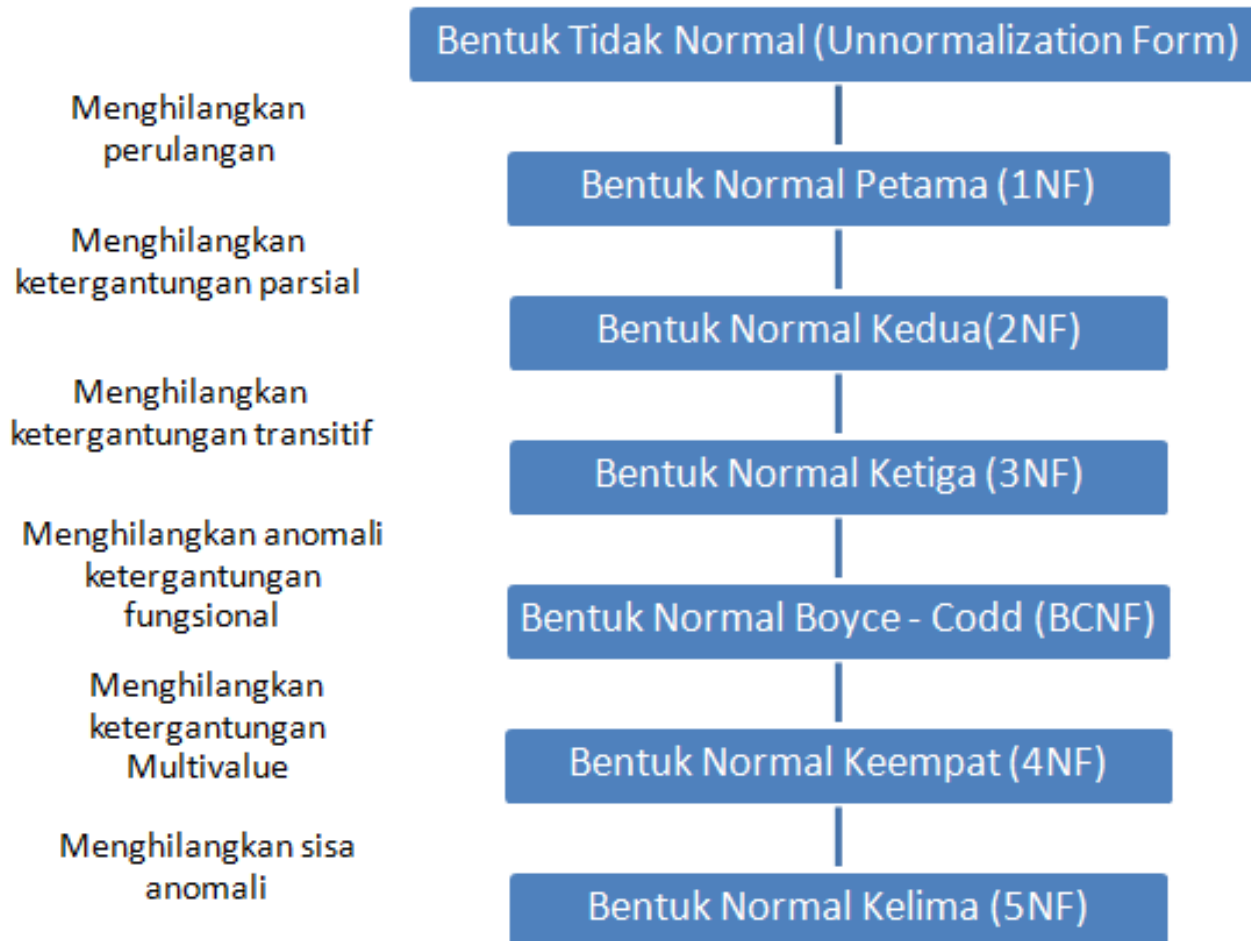
Tujuan :

- **Minimalisasi redundansi (pengulangan data)**
- **Memudahkan identifikasi entitas**
- **Mencegah terjadinya anomali**

Beberapa bentuk normal (normal forms, NF) :

- **1NF, 2NF, 3NF, BCNF**
based on keys and functional dependencies
- **4NF, 5NF**
based on keys and multi-valued dependencies

Tahapan Proses Normalisasi



Boyce Codd Normal Form (BCNF)

- Suatu relasi disebut memenuhi BCNF jika dan hanya jika setiap determinan yang ada pada relasi tersebut adalah candidate key.
Definisi yang lain :
Suatu relasi disebut memenuhi BCNF jika untuk setiap FD nontrivial : $X \rightarrow A$ atribut X adalah superkey.
- Untuk normalisasi ke bentuk BCNF, maka tabel 3NF didekomposisi menjadi beberapa tabel yang masing-masing memenuhi BCNF.
- Tujuan membentuk BCNF :
 - :: semantik multiple candidate key menjadi lebih eksplisit (FD hanya pada candidate key).
 - :: menghindari update anomali yang masih mungkin terjadi pada 3NF.

Dari definisi 3NF dan BCNF, maka apabila suatu relasi memenuhi BCNF pasti memenuhi 3NF, tetapi belum tentu sebaliknya.

Boyce Codd Normal Form (BCNF)

Contoh :

Diketahui tabel $R=(A,B,C)$

dengan FD : $A \rightarrow B$ dan $B \rightarrow C$ maka R bukan BCNF, sebab :

- A superkey ?

$A \rightarrow B$ (diketahui)

$A \rightarrow B$ dan $B \rightarrow C$ maka $A \rightarrow C$ (transitif)

$A \rightarrow A$ (refleksif)

Sehingga $A \rightarrow (A,B,C)$ atau $A \rightarrow R$. Jadi A superkey.

- B superkey ?

$B \rightarrow C$ (diketahui)

$B \rightarrow B$ (refleksif)

Tapi $B \not\rightarrow A$. Sehingga $B \not\rightarrow A,B,C$ atau B bukan superkey.

Agar R memenuhi BCNF maka didekomposisi menjadi :

$R_1=(A,B)$; FD : $A \rightarrow B$ dan

$R_2=(B,C)$; FD : $B \rightarrow C$.

sehingga R_1 dan R_2 masing-masing memenuhi BCNF. Sebab A dan B dua-duanya sekarang menjadi superkey.

Boyce Codd Normal Form (BCNF)

Contoh :

Diketahui tabel $R=(A,B,C)$

dengan FD : $AB \rightarrow C$ dan $C \rightarrow B$. Apakah :

- 3NF ?
- BCNF ?

- R memenuhi 3NF karena :
 $AB \rightarrow C$; maka $AB \rightarrow ABC$, atau $A \rightarrow R$. Jadi AB superkey dari R
 $C \rightarrow B$; maka $AC \rightarrow AB$, atau $AC \rightarrow ABC$ dan $AC \rightarrow R$.
Jadi AC juga superkey (sekalius juga candidate key) dari R
Karena AB superkey dan C subset candidate key maka R memenuhi 3NF
- R bukan BCNF karena :
AB superkey tetapi C bukan superkey.

Boyce Codd Normal Form (BCNF)

Students

sid	name	age
53666	Jones	18
53668	Smith	18
53669	Melissa	17
53670	Hilden	19

Students=(sid, name, age)

FD : sid → name, age

- **BCNF, sebab sid superkey**

Books

bid	title	year
B001	MySQL	2002
B002	Algorithm	2003
B003	Visual Foxpro 6.0	2003
B004	Visual basic 6.0	2005

Books=(bid, title, year)

FD : bid → title, year

- **BCNF, sebab bid superkey**

Pinjam

idpinjam	sid	bid	date
P-01	53666	B002	10/11/2005
P-02	53668	B001	10/11/2005
P-03	53668	B004	11/12/2005
P-04	53670	B002	14/11/2005

Pinjam=(idpinjam, sid, bid, date)

FD : idpinjam → bid, date

- **Bukan BCNF, sebab idpinjam bukan superkey**

idpinjam ↗ sid

Boyce Codd Normal Form (BCNF)

Pinjam

idpinjam	sid	bid	date
P-01	53666	B002	10/11/2005
P-02	53668	B001	10/11/2005
P-03	53668	B004	11/12/2005
P-04	53670	B002	14/11/2005

Didekomposisi menjadi :

Pinjam1

idpinjam	sid
P-01	53666
P-02	53668
P-03	53668
P-04	53670

FD trivial
→ BCNF

Pinjam2

idpinjam	bid	date
P-01	B002	10/11/2005
P-02	B001	10/11/2005
P-03	B004	11/12/2005
P-04	B002	14/11/2005

idpinjam → bid, date
idpinjam superkey
→ BCNF

Comparison of BCNF And 3NF

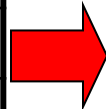
- **It is always possible to decompose a relation into relations in 3NF and**
 - **the decomposition is lossless**
 - **the dependencies are preserved**
- **It is always possible to decompose a relation into relations in BCNF and**
 - **the decomposition is lossless**
 - **it may not be possible to preserve dependencies.**

Comparison of BCNF And 3NF

- Contoh kasus redundansi pada 3NF
Jadwal = (Nim, Modul, Dosen)
FD = {Dosen → Modul}
Relasi ini memenuhi 3NF, karena tidak ada ketergantungan transitif.
Tetapi tidak memenuhi BCNF karena dari Dosen → Modul maka Dosen bukan candidate key.
Alternatif yang dilakukan adalah dekomposisi tabel menjadi :

<u>NIM</u>	<u>Modul</u>	Dosen
P11.2004.0129	VB.Net	Ajib
P11.2004.0130	Prolog	Aris
P11.2004.0129	VB Net	Ajib
P11.2004.0201	VB Net	Budi
P11.2004.0250	Prolog	Jono
P11.2004.0260	VB.Net	Budi

NOT BCNF



<u>NIM</u>	Dosen
P11.2004.0129	Ajib
P11.2004.0130	Aris
P11.2004.0129	Ajib
P11.2004.0201	Budi
P11.2004.0250	Jono
P11.2004.0260	Budi

BCNF

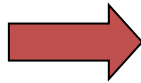
Dosen	<u>Modul</u>
Ajib	VB.Net
Aris	Prolog
Jono	Prolog
Budi	VB.Net

Design Goals

- **Goal for a relational database design is:**
 - **BCNF.**
 - **Lossless join.**
 - **Dependency preservation.**
- **If we cannot achieve this, we accept one of**
 - **Lack of dependency preservation**
 - **Redundancy due to use of 3NF**

Design Steps

(doesn't meet the definition of a relation)



Entity Set

Remove multivalued & repeating attributes. Meet definition of relation

First Normal Form

Remove partial dependencies

Second Normal Form

Remove transitive dependencies

Third Normal Form

Select relation where all determinants are candidate key

Boyce-Codd Normal Form

Praktisi database kebanyakan menganggap bahwa tingkatan normalisasi hingga BCNF atau 3NF dianggap sudah cukup untuk meminimalisasi masalah dalam desain database (redundansi, lossless, dependency preservation)

Additional Material (Normal Form Other)

Multivalued Dependency (MVD)

Beberapa bentuk normal (normal forms, NF) :

- **1NF, 2NF, 3NF, BCNF**
based on keys and functional dependencies
- **4NF, 5NF**
based on keys and multi-valued dependencies)

Multivalued Dependencies (MVD)

Misal R adalah skema relasi. A, B dan C adalah subset atribut dari R maka A disebut multidependensi pada B, ditulis $A \twoheadrightarrow B$, jika untuk setiap nilai A terdapat sekumpulan nilai B dan sekumpulan nilai C, tetapi nilai B dan nilai C independen.

Multivalued Dependency (MVD)

- Relasi MVD melibatkan minimal 3 atribut relasi, misal A-B-C
- Untuk setiap nilai A, terdapat sekumpulan nilai B dan sekumpulan nilai C.
- Sekumpulan nilai B independen dengan sekumpulan nilai C, semikian juga sebaliknya.

Penawaran Mata Kuliah

Mata Kuliah	Instruktur	Pustaka
Basis Data	Himawan	CJ Date
	Purwanto	HF Korth
	Aris Marjuni	
Matematika	M Sidiq	C Liu MT Lang



Penawaran Mata Kuliah

Mata Kuliah	Instruktur	Pustaka
Basis Data	Himawan	CJ Date
Basis Data	Purwanto	CJ Date
Basis Data	Aris Marjuni	CJ Date
Basis Data	Himawan	HF Korth
Basis Data	Purwanto	HF Korth
Basis Data	Aris Marjuni	HF Korth
Matematika	M Sidiq	C Liu
Matematika	M Sidiq	MT Lang

MVD : Mata Kuliah →→ Instruktur
 Mata Kuliah →→ Pustaka

Multivalued Dependency (MVD)

Refleksif : $x \twoheadrightarrow x$

Augmentation : Jika $x \twoheadrightarrow y$ maka $xz \twoheadrightarrow yz$

Transitif : Jika $x \twoheadrightarrow y$ dan $y \twoheadrightarrow z$ maka $x \twoheadrightarrow z-y$

Pseudotransitif : Jika $x \twoheadrightarrow y$ dan $yw \twoheadrightarrow z$ maka $xw \twoheadrightarrow z-yw$

Union : Jika $x \twoheadrightarrow y$ dan $x \twoheadrightarrow z$ maka $x \twoheadrightarrow yz$

Dekomposisi : Jika $x \twoheadrightarrow y$ dan $x \twoheadrightarrow z$ maka $x \twoheadrightarrow yz$ dan $x \twoheadrightarrow z-y$

Komplemen : Jika $x \twoheadrightarrow y$ dan $z = R - x - y$ maka $x \twoheadrightarrow z$

Fourth Normal Form (4NF)

Skema relasi R disebut 4NF jika dan hanya jika

- BCNF dan
- Tidak terdapat MVD.
- Apabila terdapat MVD dalam R, misal $X \twoheadrightarrow Y$ maka harus memenuhi salah satu :
 - MVD adalah trivial atau
 - X adalah superkey

Fourth Normal Form (4NF)

Misal diketahui $R = (A, B, C, D, E, F)$

MVD : $A \twoheadrightarrow B$ dan $CD \twoheadrightarrow EF$. Bila diasumsikan R memenuhi BCNF, apakah memenuhi 4NF ? Jika belum upayakan menjadi 4NF.

- Bukan 4NF sebab terdapat MVD dan nontrivial.
- Dekomposisi R :
Dari $A \twoheadrightarrow B$, dibentuk $R1=(A,B)$
Karena $A \twoheadrightarrow B$ maka $A \twoheadrightarrow C,D,E,F$ dan dibentuk $R2=(A,C,D,E,F)$
- $R1=(A,B)$
memenuhi 4NF, sebab MVD : $A \twoheadrightarrow B$ trivial
- $R2=(A,C,D,E,F)$
Bukan 4NF, karena MVD : $CD \twoheadrightarrow EF$ nontrivial
- Dekomposisi $R2$:
Dari $CD \twoheadrightarrow EF$ dibentuk $R21=(C,D,E,F)$
Karena $CD \twoheadrightarrow EF$ dan $CD \twoheadrightarrow A$, dibentuk $R22=(C,D,A)$
- $R21=(C,D,E,F)$
4NF, karena MVD : $CD \twoheadrightarrow EF$ trivial
- $R22=(C,D,A)$
4NF, karena tidak ada MVD.

$R1=(A,B)$

$R21=(C,D,E,F)$

$R22=(C,D,A)$

Memenuhi 4NF

Fourth Normal Form (4NF)

Misal diketahui $R = (A, B, C, G, H, I)$

MVD : $A \twoheadrightarrow B$, $B \twoheadrightarrow HI$ dan $CG \twoheadrightarrow H$. Bila diasumsikan R memenuhi BCNF, apakah memenuhi 4NF ? Jika belum upayakan menjadi 4NF.

- Bukan 4NF sebab terdapat MVD, nontrivial dan bukan superkey.
- Dekomposisi R :
Dari $A \twoheadrightarrow B$, dibentuk $R1=(A,B)$
Karena $A \twoheadrightarrow B$ maka $A \twoheadrightarrow C,G,H,I$ dan dibentuk $R2=(A,C,G,H,I)$
- $R1=(A,B)$
memenuhi 4NF, sebab MVD : $A \twoheadrightarrow B$ trivial
- $R2=(A,C,G,H,I)$
Bukan 4NF, karena MVD : $CG \twoheadrightarrow H$ nontrivial
- Dekomposisi $R2$:
Dari $CG \twoheadrightarrow H$ dibentuk $R21=(C,G,H)$
Karena $CG \twoheadrightarrow H$ maka $CG \twoheadrightarrow A,I$ dan dibentuk $R22=(C,G,A,I)$
- $R21=(C,G,H)$ 4NF, karena MVD : $CG \twoheadrightarrow H$ trivial, tetapi $R22=(C,G,A,I)$ bukan 4NF, karena ada MVD : $A \twoheadrightarrow I$ (dari $A \twoheadrightarrow B$, $B \twoheadrightarrow HI$ maka $A \twoheadrightarrow HI$ dan $A \twoheadrightarrow I$).
- Dekomposisi $R22$ menjadi $R221=(A,I)$ dan $R222=(A,C,G)$ dan masing-masing merupakan bentuk 4NF.

$R1=(A,B)$
 $R21=(C,G,H)$
 $R221=(A,I)$
 $R222=(A,C,G)$
Memenuhi 4NF

Fourth Normal Form (4NF)

Penawaran Mata Kuliah

Mata Kuliah	Instruktur	Pustaka
Basis Data	Himawan	CJ Date
Basis Data	Purwanto	CJ Date
Basis Data	Aris Marjuni	CJ Date
Basis Data	Himawan	HF Korth
Basis Data	Purwanto	HF Korth
Basis Data	Aris Marjuni	HF Korth
Matematika	M Sidiq	C Liu
Matematika	M Sidiq	MT Lang

▪ BCNF, sebab :

- 1NF, Yes
- 2NF, Yes. Sebab semua atribut mrpk atribut kunci.
- 3NF, Yes. Sebab tidak ada ketergantungan transitif.
- BCNF, Yes. Tidak ada FD pada atribut bukan kunci.
- 4NF, No. Sebab terdapat MVD yang nontrivial.

Dekomposisi

MVD : Mata Kuliah →→ Instruktur
Mata Kuliah →→ Pustaka

Fourth Normal Form (4NF)

Penawaran Mata Kuliah

Mata Kuliah	Instruktur	Pustaka
Basis Data	Himawan	CJ Date
Basis Data	Purwanto	CJ Date
Basis Data	Aris Marjuni	CJ Date
Basis Data	Himawan	HF Korth
Basis Data	Purwanto	HF Korth
Basis Data	Aris Marjuni	HF Korth
Matematika	M Sidiq	C Liu
Matematika	M Sidiq	MT Lang

Inst

Mata Kuliah	Instruktur	▪ 4NF
Basis Data	Himawan	
Basis Data	Purwanto	
Basis Data	Aris Marjuni	
Matematika	M Sidiq	

Mata Kuliah → Instruktur

Pust

Mata Kuliah	Pustaka	▪ 4NF
Basis Data	CJ Date	
Basis Data	HF Korth	
Matematika	C Liu	
Matematika	MT Lang	

Mata Kuliah → Pustaka

Fourth Normal Form (4NF)

R1

X	Y	Z
1	1	1
1	1	2
1	2	1
1	2	2
2	2	1
2	2	2
3	3	3

4NF ? No.

- MVD nontrivial
- Bukan Superkey

R2

X	Y	Z
1	1	1
1	1	2
1	2	1
1	2	2
2	2	1
2	1	2

4NF ? Yes

R3

X	Y	Z
1	1	1
1	1	2
1	2	1
2	2	1
2	2	2

4NF ? Yes

Fifth Normal Form (5NF)

Misal SPJ berisi data-data Supplier (S) yang memasok barang (P) untuk suatu Project (J). Satu supplier dapat memasok banyak barang, supplier juga Dapat berpartisipasi pada banyak proyek.

SPJ

S	P	J
S1	P1	J2
S1	P2	J1
S2	P1	J1
S1	P1	J1

Key : S, P, J
MVD : $S \twoheadrightarrow P$
 $S \twoheadrightarrow J$
4NF ? No

SP

S	P
S1	P1
S1	P2
S2	P1

$S \twoheadrightarrow P$
4NF ? Yes

PJ

P	J
P1	J2
P2	J1
P1	J1

$S \twoheadrightarrow J$
4NF ? Yes

Jika SP dan PJ di-project join-kan kembali on P akan diperoleh :

SP join PJ

S	P	J
S1	P1	J2
S1	P1	J1
S2	P1	J2
S2	P1	J1
S1	P2	J1

← Tupel asing

Ternyata projection-join dari dekomposisinya tidak mengembalikan Ke relasi asal.

Fifth Normal Form (5NF)

Relasi hasil project-join tsb akan kembali ke bentuk relasi asal SPJ setelah di project-join
Kan dengan SJ pada S dan J

SP join PJ

S	P	J
S1	P1	J2
S1	P1	J1
S2	P1	J2
S2	P1	J1
S1	P2	J1

Project-join
on (S,J)

SJ

S	J
S1	J2
S1	J1
S2	J1



SPJ (Original)

S	P	J
S1	P1	J2
S1	P1	J1
S2	P1	J1
S1	P2	J1

SP

S	P
S1	P1
S1	P2
S2	P1

PJ

P	J
P1	J2
P2	J1
P1	J1

SJ

S	J
S1	J2
S1	J1
S2	J1

Dekomposisi Akhir

Fifth Normal Form (5NF)

- Disebut juga Projection-Join Normal Form (PJ/NF)
- Skema relasi R disebut memenuhi 5NF jika tidak dapat dibuat menjadi beberapa tabel kecil yang lossless melalui operasi projection-join atau
Skema relasi R disebut memenuhi 5NF jika setiap join dependency (JD) nontrivial pada R diimplikasikan oleh Candidate key dari R
- Join Dependency
Batasan dekomposisi yang lossless pada sejumlah operasi project-join

Fifth Normal Form (5NF)

SPJ

S	P	J
S1	P1	J2
S1	P2	J1
S2	P1	J1
S1	P1	J1

Key : S, P, J

MVD : $S \twoheadrightarrow P$

$S \twoheadrightarrow J$

4NF ? No

SP

S	P
S1	P1
S1	P2
S2	P1

$S \twoheadrightarrow P$

4NF ? Yes

5NF ? Yes

PJ

P	J
P1	J2
P2	J1
P1	J1

$S \twoheadrightarrow J$

4NF ? Yes

5NF ? Yes

SJ

S	J
S1	J2
S1	J1
S2	J1

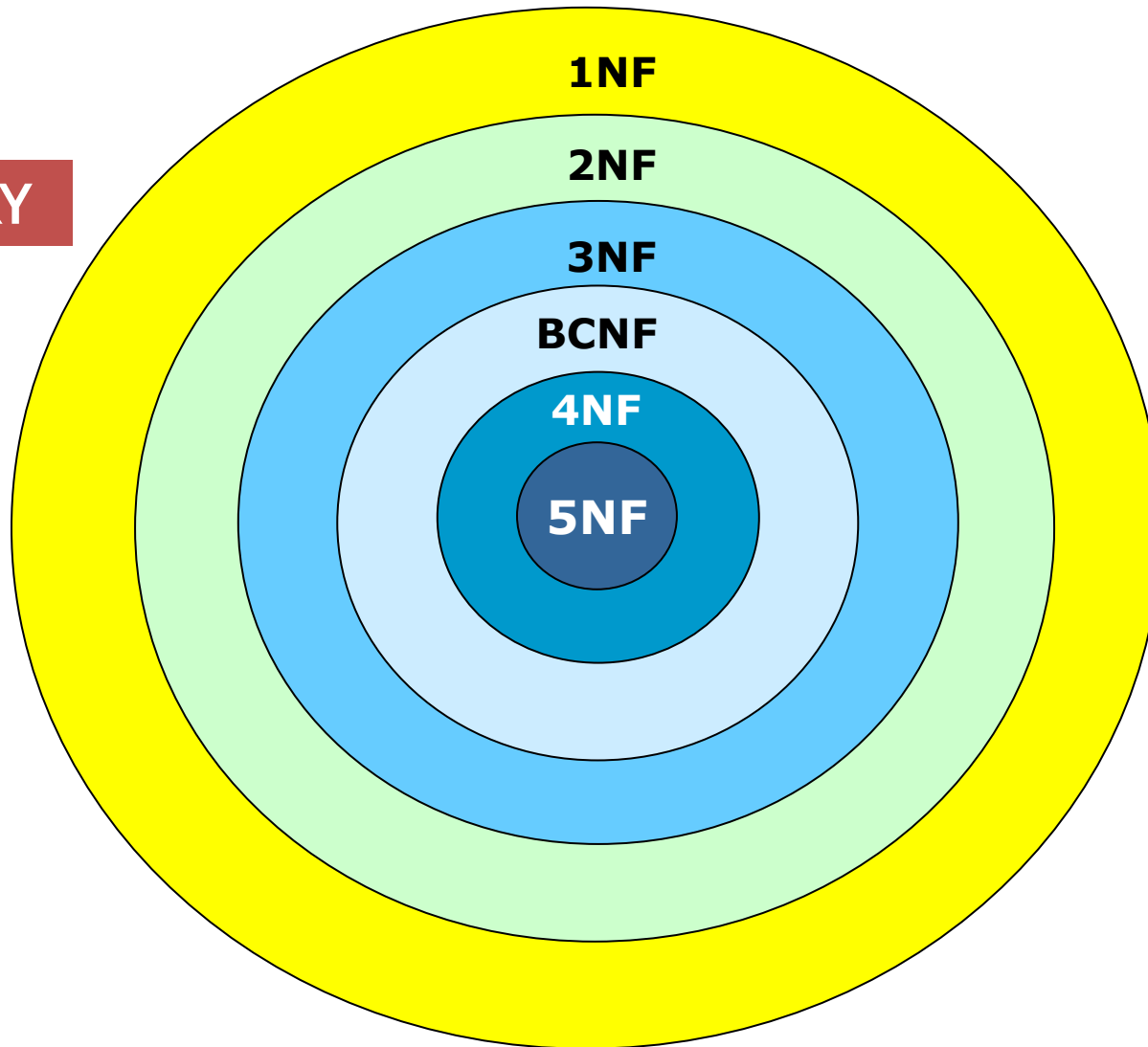
4NF ? Yes, tidak ada MVD

5NF ? Yes

5NF ?

**Yes, tidak dapat didekomposisi lagi
Trivial pada candidate key**

SUMMARY



LEVEL NORMALISASI

SUMMARY

It is best to find a database design that meet the 3-criteria :

- **NF (Normal-Form)^{note}**
- **Dependency Preservation**
- **Lossless-Join**

note

If we only have functional dependencies (FD), the first criteria is just BCNF.

If we only have multivalued dependency (MVD), the first criteria is just 4NF.

If we only have join dependency (JD), the first criteria is just 5NF.

If we cannot meet all 3-criteria, we compromise on 4NF, and accept BCNF, or even 3NF if necessary, to ensure dependency preservation.