

Database Design:

**Further Normalization: 1NF, 2NF, 3NF, BCNF
4NF, 5NF**

Surat Kointarangkul

Functional Dependencies

- Let R be a relation, and let X and Y be arbitrary subsets of the set of attributes of R . Then we say that Y is functionally dependent on X -- in symbols,

$$X \rightarrow Y$$

if and only if each X -value in R has associated with it precisely one Y -value in R .

Functional Dependencies

S

<u>S#</u>	<u>CITY</u>
-----------	-------------

S1	London
----	--------

S2	Paris
----	-------

S3	Paris
----	-------

S4	London
----	--------

$S\# \rightarrow CITY$

Normal Forms

Universe of relations (normalized and unnormalized)

1NF relations (normalized relation)

2NF relations

3NF relations

BCNF relations

4NF relations

5NF relations

Nonloss Decomposition

S	<u>S#</u>	<u>STATUS</u>	<u>CITY</u>
	S3	30	Paris
	S5	30	Athens

Nonloss Decomposition

S	<u>S#</u>	<u>STATUS</u>	<u>CITY</u>
	S3	30	Paris
	S5	30	Athens

(a)	SST	<u>S#</u>	<u>STATUS</u>
		S3	30
		S5	30

SC	<u>S#</u>	<u>CITY</u>
	S3	Paris
	S5	Athens

(b)	SST	<u>S#</u>	<u>STATUS</u>
		S3	30
		S5	30

STC	<u>STATUS</u>	<u>CITY</u>
	30	Paris
	30	Athens

First Normal Form

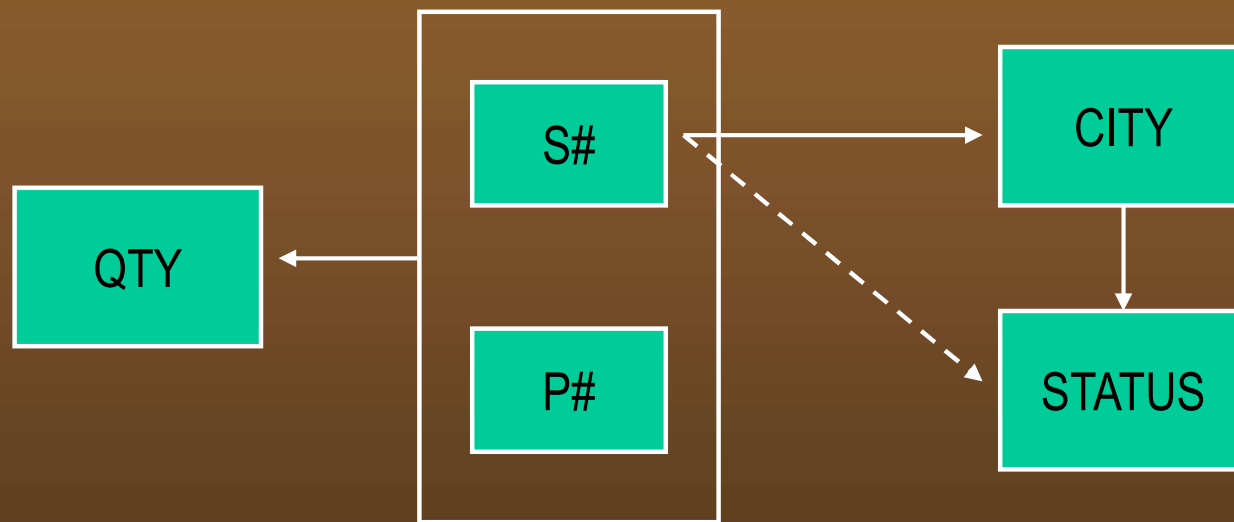
- A relation is in 1NF if and only if all underlying domains contain scalar values only.

First Normal Form

FIRST

S#	STATUS	CITY	P#	QTY
S1	20	London	P1	300
S1	20	London	P2	200
S1	20	London	P3	400
S1	20	London	P4	200
S1	20	London	P5	100
S1	20	London	P6	100
S2	10	Paris	P1	300
S2	10	Paris	P2	400
S3	10	Paris	P2	200
S4	20	London	P2	200
S4	20	London	P4	300
S4	20	London	P5	400

First Normal Form



Update Anomalies

- Add
- Delete
- Update

Second Normal Form

- (definition assuming only one candidate key, which is thus the primary key): A relation is in 2 NF if and only if it is in 1NF and every nonkey attribute is irreducibly dependent on the primary key.

Second Normal Form

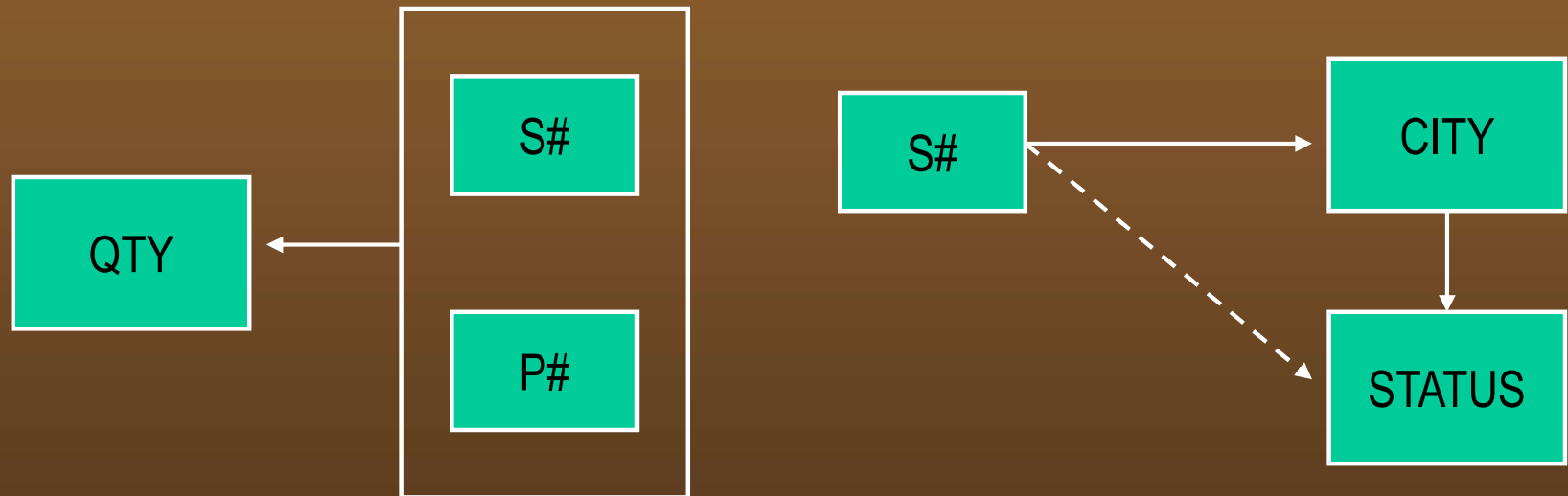
SECOND

<u>S#</u>	<u>STATUS</u>	<u>CITY</u>
S1	20	London
S2	10	Paris
S3	10	Paris
S4	20	London
S5	30	Athens

FIRST

<u>S#</u>	<u>P#</u>	<u>QTY</u>
S1	P1	300
S1	P2	200
S1	P3	400
S1	P4	200
S1	P5	100
S1	P6	100
S2	P1	300
S2	P2	400
S3	P2	200
S4	P2	200
S4	P4	300
S4	P5	400

Second Normal Form



Third Normal Form

- (definition assuming only one candidate key, which is thus the primary key): A relation is in 3 NF if and only if it is in 2NF and every nonkey attribute is nontransitively dependent on the primary key.

Third Normal Form

SC

<u>S#</u>	<u>CITY</u>
S1	London
S2	Paris
S3	Paris
S4	London
S5	Athens

CS

<u>CITY</u>	<u>STATUS</u>
Athens	30
London	20
Paris	10
Rome	50

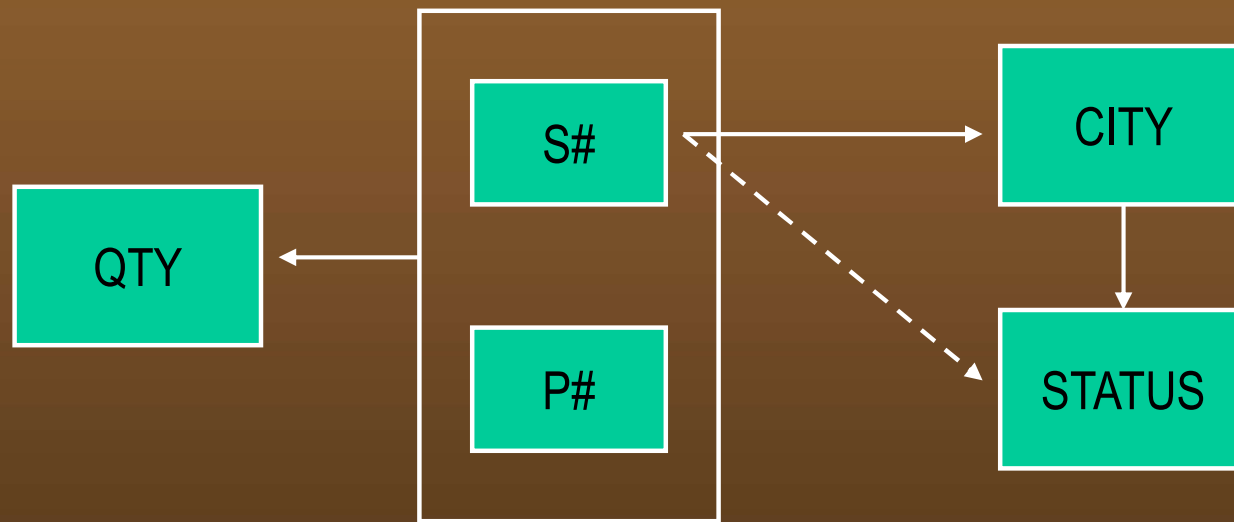
FIRST

<u>S#</u>	<u>P#</u>	<u>QTY</u>
S1	P1	300
S1	P2	200
S1	P3	400
S1	P4	200
S1	P5	100
S1	P6	100
S2	P1	300
S2	P2	400
S3	P2	200
S4	P2	200
S4	P4	300
S4	P5	400

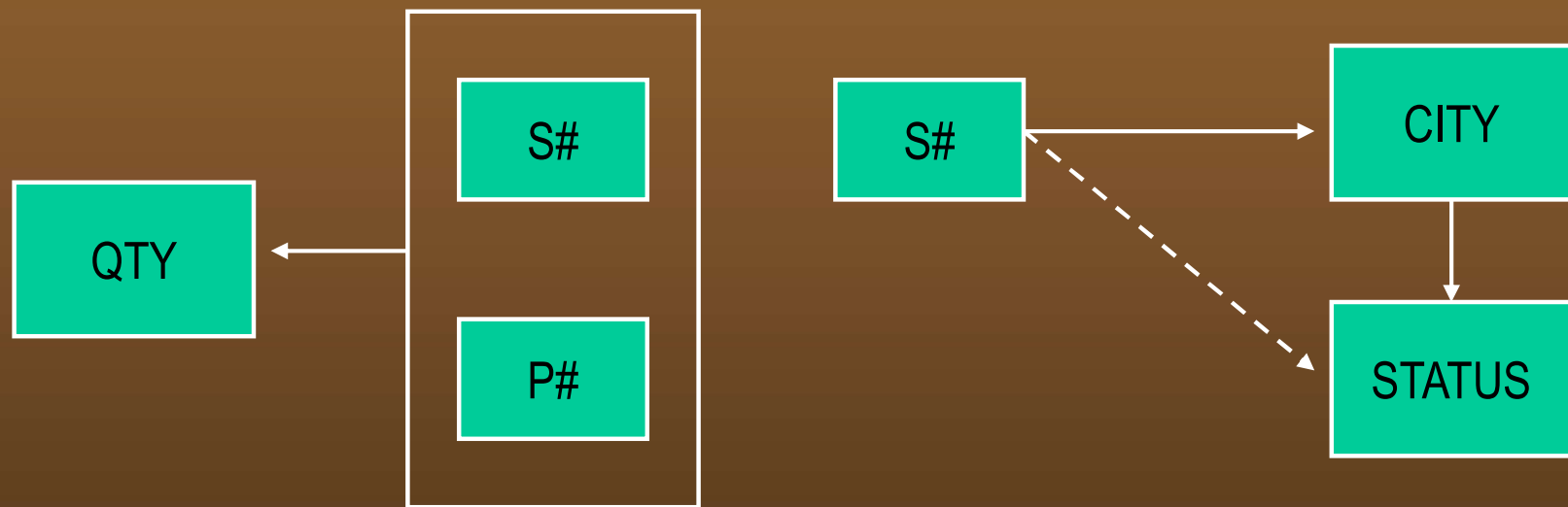
Third Normal Form



Comparision: First Normal Form



Comparison: Second Normal Form



Comparison: Third Normal Form



Exercise 1

Student

id	name	advisor	office
101	Mary	Tom	426
102	David	Tom	426
103	David	Norman	511
104	Philips	Norman	208
105	Rose	Norman	208

- คำถาม
1. เขียน FD diagram ของ relation ดังกล่าว
 2. Relation ดังกล่าวอยู่ใน Normal Form ระดับใด?
 3. หากไม่อยู่ในระดับ3NF ให้นักศึกษา Normalize ให้อยู่ในระดับ3NF

Exercise 2

Register

Id	Name	Code	Desc	Grade
101	Mary	AC201	Accounting	A
102	David	IS201	Information Systems	B
103	David	AC201	Accounting	C+
104	Philips	MK201	Marketing	A
101	Mary	IS201	Information Systems	B
104	Philips	IS201	Information Systems	A

ต้องการ

1. เขียน FD diagram ของ relation ดังกล่าว
2. Relation ดังกล่าวอยู่ใน normal form ระดับใด
3. ให้นักศึกษา normalize relation ดังกล่าวให้อยู่ในระดับ 3NF

Exercise 3

Library

id	name	call	copy	title	out	author
101	Mary	102	1	Physics	12/7/00	Mathew
102	David	146	8	Math	14/7/00	Jane
103	Steve	187	4	Art	14/7/00	Paul
101	Mary	102	4	Physics	14/7/00	Mathew
104	Rose	146	8	Math	16/7/00	Jane

Exercise 3

กรณีที่ 1.

- หนังสือหนึ่งเล่ม ผู้ยืมจะต้องยืมอย่างน้อยหนึ่งวัน
- หนึ่งคนยืมได้หลายเล่มในหนึ่งวัน

ต้องการ

1. เขียน FD diagram ของ relation ดังกล่าว
2. Relation ดังกล่าวอยู่ใน Normal Form ระดับใด?
3. หากไม่อยู่ในระดับ3NF ให้นักศึกษา Normalize ให้อยู่ในระดับ3NF

Exercise 3

กรณี 2.

หนังสือหนึ่งเล่ม ผู้ยืมสามารถยืมและคืนวันเดียวกัน แต่ผู้ยืมจะไม่ยืมซ้ำเล่มเดิม (call เดิม copy เดิม) ในวันเดียวกัน
หนึ่งคนยืมได้หลายเล่มในวัน

ต้องการ

1. เขียน FD diagram ของ relation ดังกล่าว
2. Relation ดังกล่าวอยู่ใน Normal Form ระดับใด?
3. หากไม่อยู่ในระดับ3NF ให้นักศึกษา Normalize ให้อยู่ในระดับ3NF

Exercise 3

กรณีที่ 3.

หนังสือหนึ่งเล่ม ผู้ยืมสามารถยืมและคืนวันเดียวกัน และผู้ยืมสามารถยืมซ้ำเล่มเดิม (call เดิม copy เดิม) ในวันเดียวกัน
หนึ่งคนยืมได้หลายเล่มในวัน

ต้องการ

1. เขียน FD diagram ของ relation ดังกล่าว
2. Relation ดังกล่าวอยู่ใน Normal Form ระดับใด?
3. หากไม่อยู่ในระดับ 3NF ให้นักศึกษา Normalize ให้อยู่ในระดับ 3NF

Exercise 3

กรณีี่ 4.

ในโลกแห่งความเป็นจริง หากหนังสือเล่มนั้นมี copy เดียว ทางห้องสมุดจะไม่มีไม่
หมายเลข COPY ให้หนังสือ ระบบ database จะทำอย่างไร?

Independent Relations

SECOND (S#, CITY, STATUS)

A SC (S#, CITY) and CS (CITY, STATUS)

B SC (S#, CITY) and SS (S#, STATUS)

Independent Relations

Projections $R1$ and $R2$ of a relation R are independent if and only if

- Every FD in R is a logical consequence of those in $R1$ and $R2$, and
- The common attributes of $R1$ and $R2$ form a candidate key for at least one of the pair.

Boyce/Codd Normal Form

Boyce/Codd normal form will be used when a relation

- had two (or more) candidate keys, such that
- the two candidate keys were composite, and
- they overlapped

Boyce/Codd Normal Form

A relation is in BCNF if and only if the only determinants are candidate keys.

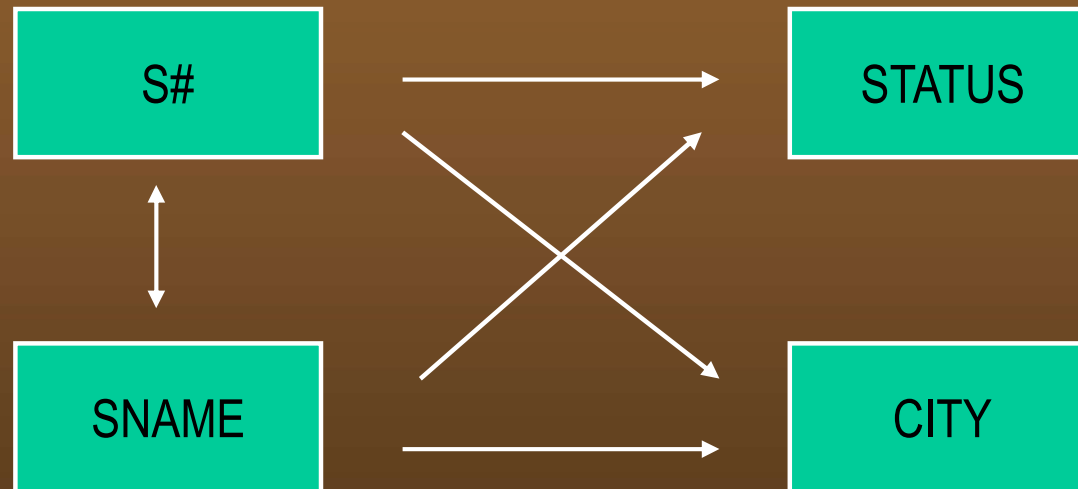
Boyce/Codd Normal Form

S (S#, SNAME, STATUS, CITY)

CANDIDATE KEY (S#)

CANDIDATE KEY (SNAME)

Boyce/Codd Normal Form



Boyce/Codd Normal Form

S (S#, SNAME, STATUS, CITY)

CANDIDATE KEY (S#)

CANDIDATE KEY (SNAME)

1NF



2NF



3NF



BCNF



Boyce/Codd Normal Form

SSP	S#	SNAME	P#	QTY
	S1	Smith	P1	300
	S1	Smith	P2	200
	S1	Smith	P3	400
	S1	Smith	P4	200

1NF ?

2NF ?

3NF ?

BCNF ?

Boyce/Codd Normal Form

SSP	S#	SNAME	P#	QTY
	S ₁	Smith	P ₁	300
	S ₁	Smith	P ₂	200
	S ₁	Smith	P ₃	400
	S ₁	Smith	P ₄	200

1NF



2NF



3NF



BCNF

Boyce/Codd Normal Form

SSP (S#, SNAME, P#, QTY)

A SS (S#, SNAME) and SP (S#, P#, QTY)

OR

B SC (S#, SNAME) and SP (SNAME, P#, QTY)

Boyce/Codd Normal Form

SJT (S, J, T)

Student **S** is taught subject **J** by teacher **T**

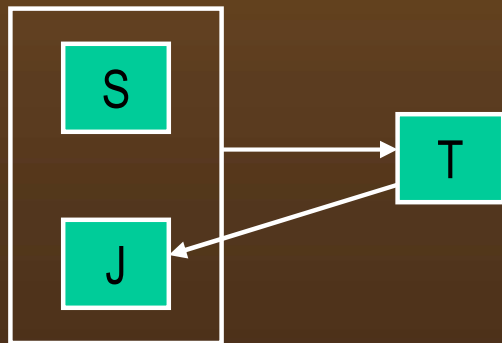
Additional constraint:

- For each subject, each student of that subject is taught by only one teacher.
- Each teacher teaches only one subject (but each subject is taught by several teachers).

Boyce/Codd Normal Form

SJT

S	J	T
Smith	Math	Prof. White
Smith	Physics	Prof. Green
Jones	Math	Prof. White
Jones	Physics	Prof. Brown



Candidate Keys: (S, J) (S, T)

Boyce/Codd Normal Form

SJT

S	J	T
Smith	Math	Prof. White
Smith	Physics	Prof. Green
Jones	Math	Prof. White
Jones	Physics	Prof. Brown



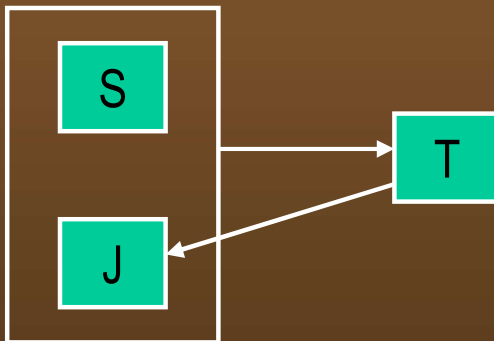
ST (S, T)



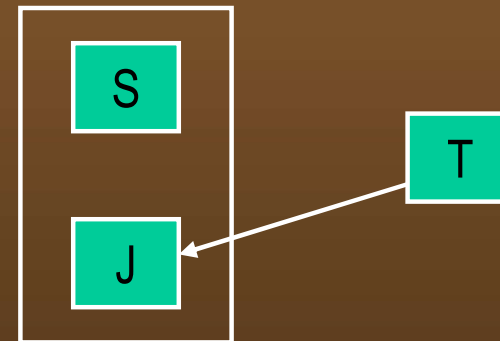
TJ (T, J)

Functional Dependency Diagram

Before
Decomposition



After
Decomposition



Boyce/Codd Normal Form

In sum

- (a) decompose a relation into BCNF
- (b) decompose it into independent

it is not always possible to satisfy both objectives simultaneously.

Boyce/Codd Normal Form

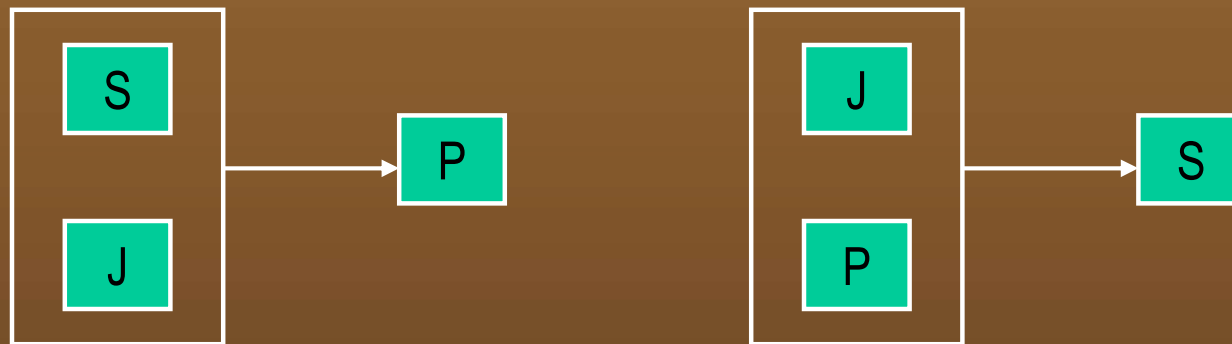
SJP (S, J, P)

Student **S** was examined in subject **J** and achieved position **P** in the class list.

Additional constraint:

- There are no ties; that is, no two students obtained the same position in the same subject.

Boyce/Codd Normal Form



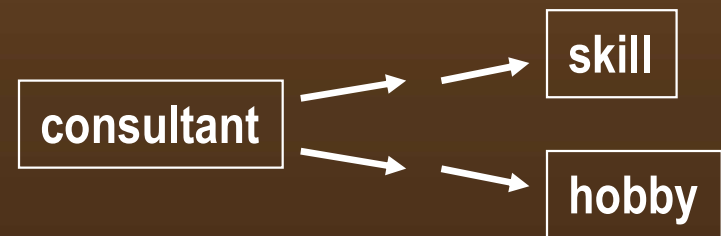
CANDIDATE KEYS : (S, J) ; (J, P)

Q1. Do we have to test for BCNF?

Q2. Is SPJ in BCNF?

Multivalued Dependence

<i>Consultant</i>	<i>Skill</i>	<i>Hobby</i>
Blake	SQL *Forms	Golf
Blake	SQL *Forms	Chess
Blake	Cobol	Golf
Blake	Cobol	Chess
Blake	SQL *Plus	Golf
Blake	SQL *Plus	Chess



[Go to Normalized Relation](#)

Fourth Normal Form

A relation is in 4th normal form if, and only if, it does not contain independent multivalued dependencies

Fourth Normal Form

<i>Consultant</i>	<i>Skill</i>
Blake	SQL *Forms
Blake	Cobol
Blake	SQL *Plus

SKILLS

<i>Consultant</i>	<i>Hobby</i>
Blake	Golf
Blake	Chess

HOBBIES

[Back to Unnormalized Relation](#)

Fifth Normal Form

AGENT	COMPANY	PRODUCT
Smith	Ford	car
Smith	Ford	truck
Smith	GM	car
Smith	GM	truck
Jones	Ford	car

AGENT	COMPANY
Smith	Ford
Smith	GM
Jones	Ford

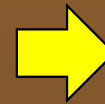
COMPANY	PRODUCT
Ford	car
Ford	truck
GM	car
GM	truck

AGENT	PRODUCT
Smith	car
Smith	truck
Jones	car

Agent ขาย Product อะไร
หาก Company ผลิต Product นั้น
Agent ก็จะขาย Product ของ
Company นั้น

Fifth Normal Form

AGENT	COMPANY	PRODUCT
Smith	Ford	car
Smith	Ford	truck
Smith	GM	car
Smith	GM	truck
Jones	Ford	car



AGENT	COMPANY
Smith	Ford
Smith	GM
Jones	Ford

COMPANY	PRODUCT
Ford	car
Ford	truck
GM	car
GM	truck

AGENT	PRODUCT
Smith	car
Smith	truck
Jones	car

Agent ขาย Product อะไร
หาก Company ผลิต Product นั้น
Agent ก็จะขาย Product ของ
Company นั้น

Fifth Normal Form (Project / join NF)

<i>Cons_id</i>	<i>Skill</i>	<i>Project</i>
1024	SQL *Plus	NATPOW22
1024	SQL *Forms	BT533
1326	SQL *Forms	NATPOW22

A consultant is assigned to a specific project because of a given skill requirement for that project.

<i>Cons_id</i>	<i>Skill</i>	<i>Project</i>
1024	SQL *Plus	NATPOW22
1024	SQL *Forms	BT533
1326	SQL *Forms	NATPOW22



ACHIEVEMENT

<i>Cons_id</i>	<i>Skill</i>
1024	SQL *Plus
1024	SQL *Forms
1326	SQL *Forms

REQUIREMENT

<i>Skill</i>	<i>Project</i>
SQL *Plus	NATPOW22
SQL *Forms	BT533
SQL *Forms	NATPOW22

ASSIGNMENT

<i>Cons_id</i>	<i>Project</i>
1024	NATPOW22
1024	BT533
1326	NATPOW22

<i>Cons_id</i>	<i>Skill</i>	<i>Project</i>
1024	SQL *Plus	NATPOW22
1024	SQL *Forms	BT533
1326	SQL *Forms	NATPOW22
1024	SQL *Forms	NATPOW22



Join Dependencies

ACHIEVEMENT

<i>Cons_id</i>	<i>Skill</i>
1024	SQL *Plus
1024	SQL *Forms
1326	SQL *Forms

REQUIREMENT

<i>Skill</i>	<i>Project</i>
SQL *Plus	NATPOW22
SQL *Forms	BT533
SQL *Forms	NATPOW22

ASSIGNMENT

<i>Cons_id</i>	<i>Project</i>
1024	NATPOW22
1024	BT533
1326	NATPOW22

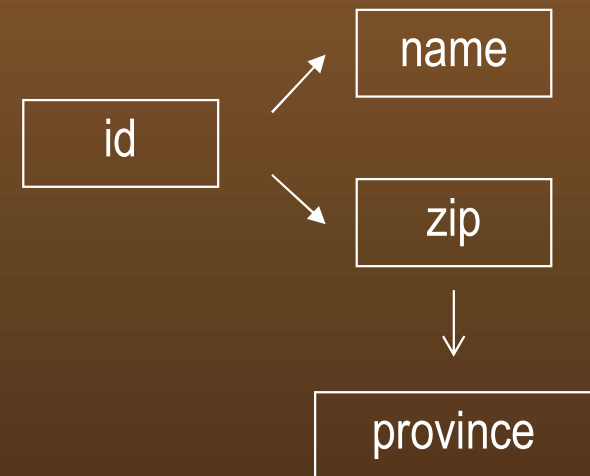
Fifth Normal Form

A relation is in 5th NF if and only if it is in 4NF and does not have a join dependency.

Denormalization

Student (id, name, province, zip)

Id	name	province	zip
101	Somsak	Bangkok	10400
102	Sunee	Bangkok	10280
103	Vinai	Bangkok	10540
104	Somsri	Chiangmai	52000
105	Chai	Chiangmai	52100



Answer 1

2. อยู่ในระดับ 2 NF

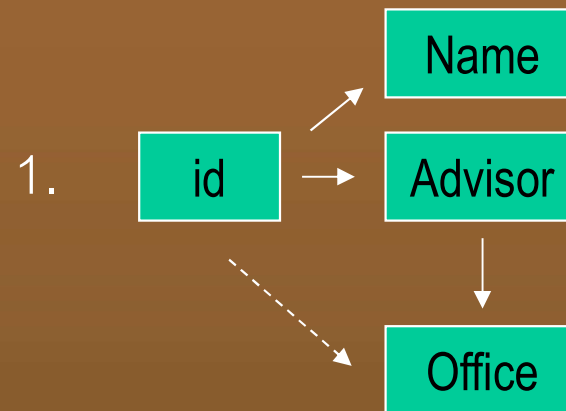
3. Normalize ได้สอง relations ดังนี้

Student

id	name	advisor
101	MaryTom	Tom
102	David	Tom
103	Steve	Tom
104	Philips	Norman
105	Rose	Norman

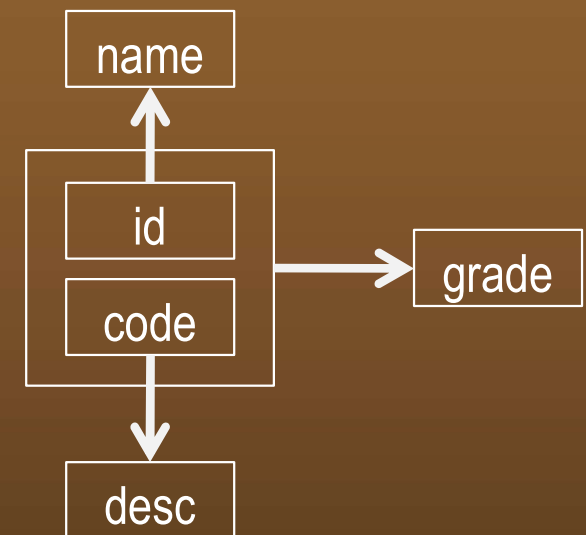
Advisor

advior	office
426	
Norman	208



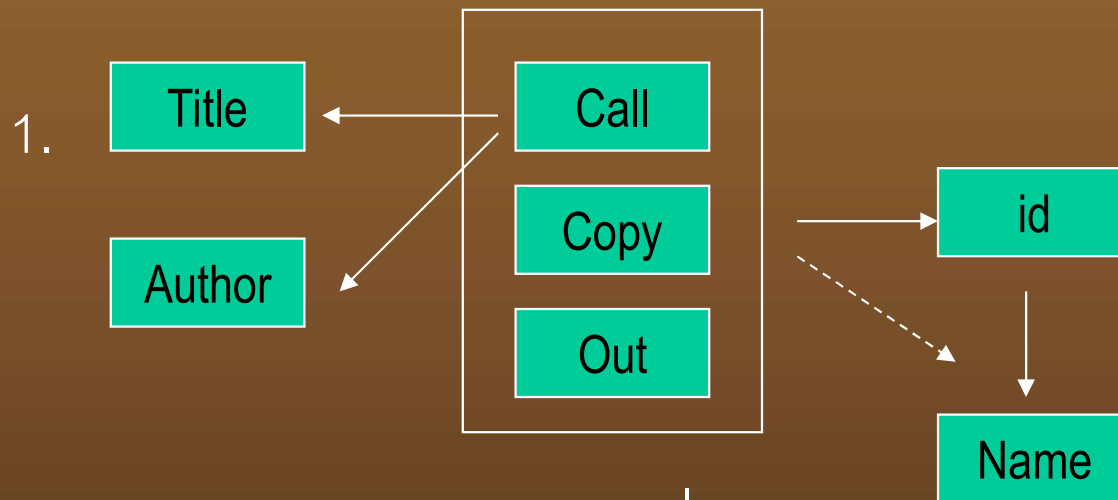
Answer 2

- อยู่ระดับ 1 NF
- Normalize ได้ตารางดังนี้
 - Student (id, name)
 - Course (code, desc)
 - Register (id, code, grade)



Answer 3

กรณีนี้ที่ 1



2. อยู่ใน Normal Form ระดับที่ 1

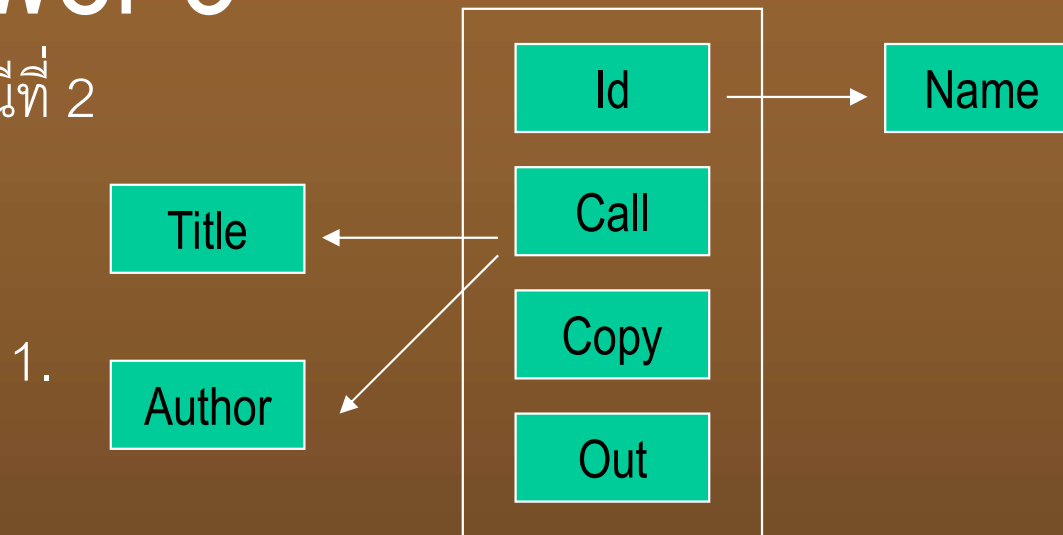
3. Book (CALL, title, author)

Borrow (CALL, COPY, OUT, id)

Student (ID, name)

Answer 3

กรณีนี้ที่ 2



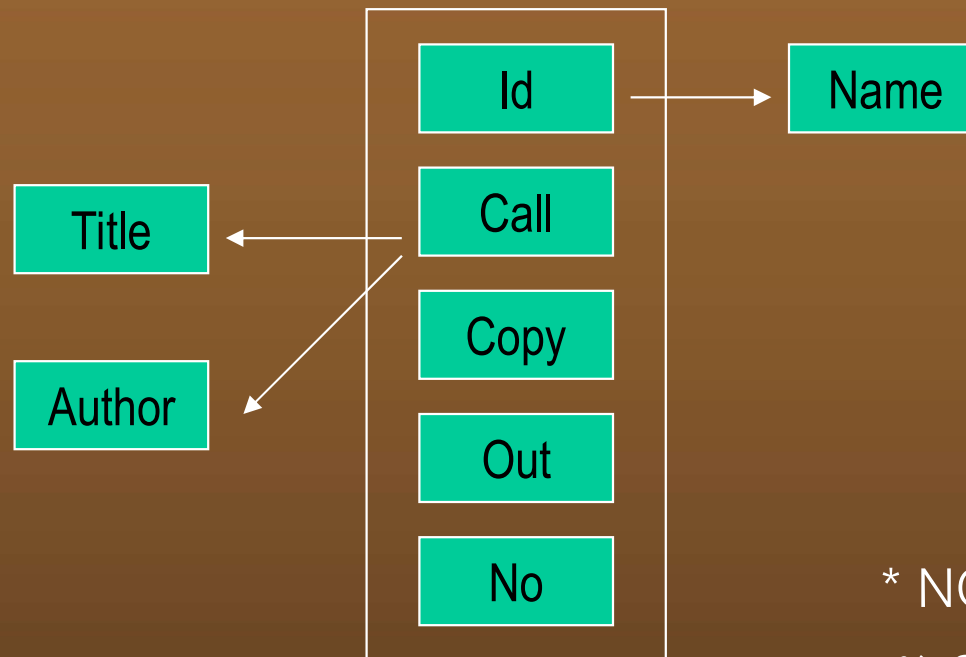
2. อยู่ใน Normal Form ระดับที่ 1

3. Book (CALL, title, author)
Borrow (ID, CALL, COPY, OUT)
Student (ID, name)

Answer 3

กรณีนี้ที่ 3

1.



2. อยู่ใน Normal Form ระดับที่ 1

3. Book (CALL, title, author)

Borrow (ID, CALL, COPY, OUT, NO)

Student (ID, name)

* NO คือครั้งที่
นักศึกษายืมหนังสือ
เล่มนั้นในวันนั้น