Normal Form Identification

✓ 1NF		BCNF X
✓ 2NF		3NF X
✓ 3NF		2NF X
✓ BCNF		1NF

R(ABCDEFGH) CK: ABC

{ABC → DE, E → FG, H → G, G → H, ABC → EF}

ABC → DE	E → FG	H → G	G → H	ABC → EF	
✓	X	X	X	✓	BCNF
✓	X	X	X	✓	3NF
✓	✓	✓	✓	✓	2NF
✓	✓	✓	✓	✓	1NF

2NF

R(ABCDE) CK: {AB, C, D, BE}

{AB → C, C → D, D → E, E → A, D → B}

AB → C	C → D	D → E	E → A	D → B	
✓	✓	✓	X	✓	BCNF
✓	✓	✓	✓	✓	3NF
✓	✓	✓	✓	✓	2NF
✓	✓	✓	✓	✓	1NF

not in BCNF But in 3NF

R(ABCD) CK: AB

$\{AB \to C, \; BC \to D\}$

2NF

R(ABCDEF) CK: $\{AB, FB, EB, CB\}$

$\{AB \to C, \; C \to DE, \; E \to F, \; F \to A\}$

1NF

R(ABCDEFGH) CK: $\{A, H, F, C\}$

$\{AB \to CD, \; D \to EG, \; F \to H, \; C \to EF, \; H \to A$

$G \to B, \; A \to B\}$

2NF

R(ABCDEFG) CK: $\{ABG, CG, BEG, FG, DG\}$

R(ABCDEFG) CK: {ABG, CG, BEG, FG, DG}

{AB → CDEF, C → ADE, D → EBF, F → DA, BE → AF}

3NF But not in BCNF

Design Goals- Redundancy, Lossless join decomposition, Dependency preserving decomposition

Red   LJD   D.P.

When we cannot meet all three design criteria, we abandon BCNF and accept a weaker form called **third normal form (3NF)**.
It is always possible to find a dependency-preserving lossless-join decomposition that is in 3NF.

(high Red)	1NF <	2NF <	$\boxed{3NF}$ <	BCNF	(0%) Red
	✓	✓	✓	✓	LJD
	✓	✓	✓	×	DP

R(ABCD)   FD: {AB → CD , D → A}

3NF But not in BCNF

R ABCD

R_1 DBC

~~DBC: CK~~

CK: DB

DB → C

R_2 AD

D: CK

D → A

Which one of the following statements about normal forms is FALSE?

(a) BCNF is stricter than 3NF

(b) Lossless, dependency-preserving decomposition into 3NF is always possible

(c) Lossless, dependency-preserving decomposition into BCNF is always possible

(d) Any relation with two attributes is in BCNF

R(AB) CK

1. R(AB) $A \to B$ — A BCNF
2. R(AB) $B \to A$ — B BCNF
3. R(AB) $A \to B$, $B \to A$ — A, B BCNF
4. R(AB) { } — AB BCNF

8

Multivalued Dependency and Fourth Normal Form

R

<u>X</u>	Y	Z
a	b/c	d/e
f	g/h	i/j

Not in 1 NF

Multivalued attribute

X →→ Z

R

<u>X</u>	<u>Y</u>	<u>Z</u>	
a	b.	d	t1
a	c.	e	t3
a	b	e	t2
a	c	d	t4
f	g	i	
f	h	j	
f	g	j	
f	h	i	

PK (<u>XYZ</u>)

Now in 1 NF
Have MVD

Multivalued dependencies are a consequence of first normal form (1NF), which disallows an attribute in a tuple to have a set of values, and the accompanying process of converting an unnormalized relation into 1NF. If we have **two or more multivalued independent attributes** in the same relation schema, we get into a problem of having to repeat every value of one of the attributes with every value of the other attribute to keep the relation state consistent and to maintain the independence among the attributes involved. This constraint is specified by a multivalued dependency.

Informally, whenever **two independent** 1:N relationships X:Y and X:Z are mixed in the same relation, R(X, Y, Z), an MVD (Multivalued dependency) may arise.

Definition: A multivalued dependency X →→ Y specified on relation schema R, where X and Y are both subsets of R, specifies the following constraint on any relation state r of R: If two tuples t1 and t2 exist in r such that t1[X] = t2[X], then two tuples t3 and t4 should also exist in r with the following properties, where we use Z to denote (R − (X ∪ Y)):

- t3[X] = t4[X] = t1[X] = t2[X]
- t3[Y] = t1[Y] and t4[Y] = t2[Y]
- t3[Z] = t2[Z] and t4[Z] = t1[Z]

Note: The tuples t1, t2, t3, and t4 are not necessarily distinct.

Whenever X →→ Y holds, we say that X **multidetermines** Y. Because of the symmetry in the definition, whenever X →→ Y holds in R, so does X →→ Z. Hence, X →→ Y implies X →→ Z and therefore it is sometimes written as X →→ Y|Z.

X →→ Y   X →→ Z

EMP

<u>Ename</u>	<u>Pname</u>	<u>Dname</u>
Smith	X	John
Smith	Y	Anna
Smith	X	Anna
Smith	Y	John

The EMP relation with two MVDs: Ename →→ Pname and Ename →→ Dname

PK: <u>Ename Pname Dname</u>

A tuple in this EMP relation represents the fact that an employee whose name is Ename works on the project whose name is Pname and has a dependent whose name is Dname. An employee may work on several projects and may have several dependents, and the employee's projects and dependents are independent of one another.

To keep the relation state consistent and to avoid any spurious relationship between the two independent attributes, we must have a separate tuple to represent every combination of an employee's dependent and an employee's project. In the relation state shown above, the employee with Ename Smith works on two projects 'X' and 'Y' and has two dependents 'John' and 'Anna', and therefore there are four tuples to represent these facts together.

The relation EMP is an **all-key relation (with key made up of all attributes)** and therefore has no f.d.'s and as such qualifies to be a BCNF relation. We can see that there is an obvious redundancy in the relation EMP—the dependent information is repeated for every project and the project information is repeated for every dependent.

Decomposing the EMP relation into two 4NF relations EMP_PROJECTS and EMP_DEPENDENTS

EMP_PROJECTS

<u>Ename</u>	<u>Pname</u>
Smith	X
Smith	Y

EMP_DEPENDENTS

<u>Ename</u>	<u>Dname</u>
Smith	John
Smith	Anna

An MVD $X \twoheadrightarrow Y$ in R is called a trivial MVD if
(a) Y is a subset of X, or
(b) $X \cup Y = R$.

An MVD that satisfies neither (a) nor (b) is called a nontrivial MVD.

For example, the relation EMP_PROJECTS has the trivial MVD Ename →→ Pname and the relation EMP_DEPENDENTS has the trivial MVD Ename →→ Dname.

4 NF

A relation schema R is in 4NF with respect to a set of dependencies F (that includes functional dependencies and multivalued dependencies) if, for every nontrivial multivalued dependency $X \twoheadrightarrow Y$ in F^+, X is a superkey for R.

We can state the following points:

- An all-key relation is always in BCNF since it has no FDs.
- An all-key relation such as the EMP relation, which has no FDs but has the MVD Ename →→ Pname | Dname, is not in 4NF.
- A relation that is not in 4NF due to a nontrivial MVD must be decomposed to convert it into a set of relations in 4NF.
- The decomposition removes the redundancy caused by the MVD.

Pizza Delivery Permutations

Restaurant	Pizza Variety	Delivery Area
A1 Pizza	Thick Crust	Springfield
A1 Pizza	Thick Crust	Shelbyville
A1 Pizza	Thick Crust	Capital City
A1 Pizza	Stuffed Crust	Springfield
A1 Pizza	Stuffed Crust	Shelbyville
A1 Pizza	Stuffed Crust	Capital City
Elite Pizza	Thin Crust	Capital City
Elite Pizza	Stuffed Crust	Capital City
Vincenzo's Pizza	Thick Crust	Springfield
Vincenzo's Pizza	Thick Crust	Shelbyville
Vincenzo's Pizza	Thin Crust	Springfield
Vincenzo's Pizza	Thin Crust	Shelbyville

not in 4 NF

Varieties By Restaurant

Restaurant	Pizza Variety
A1 Pizza	Thick Crust
A1 Pizza	Stuffed Crust
Elite Pizza	Thin Crust
Elite Pizza	Stuffed Crust
Vincenzo's Pizza	Thick Crust
Vincenzo's Pizza	Thin Crust

in 4NF

Delivery Areas By Restaurant

Restaurant	Delivery Area
A1 Pizza	Springfield
A1 Pizza	Shelbyville
A1 Pizza	Capital City
Elite Pizza	Capital City
Vincenzo's Pizza	Springfield
Vincenzo's Pizza	Shelbyville

in 4 NF