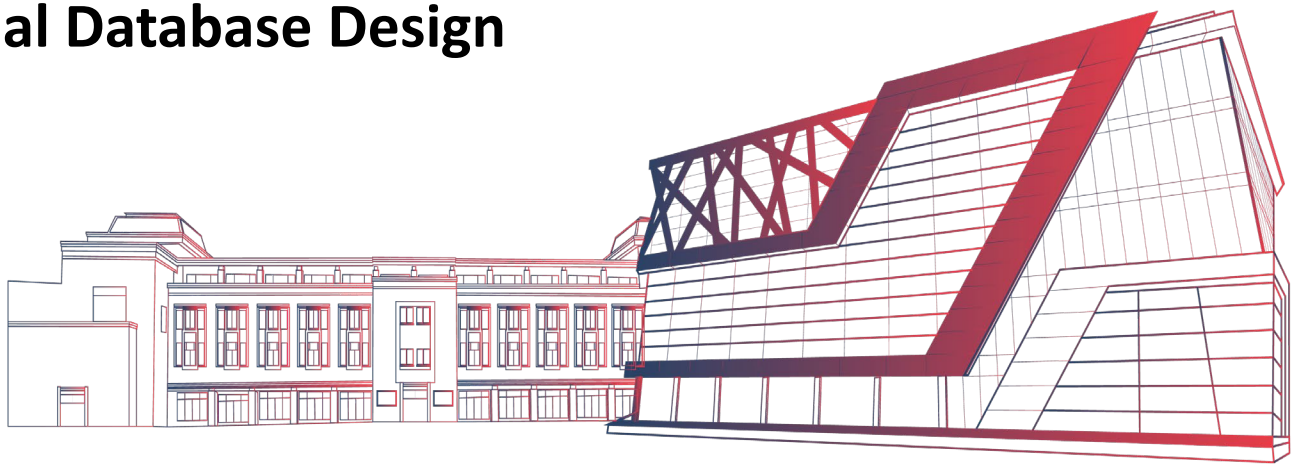
THE NORTHCAP UNIVERSITY

UNIT III

Relational Database Design



30 YEARS OF EXCELLENCE

NAAC Accredited

Lecture 13

THE NORTHCAP UNIVERSITY

Relational Database Design

30 YEARS OF EXCELLENCE

NAAC Accredited

What is Normalization?

THE NORTHCAP UNIVERSITY

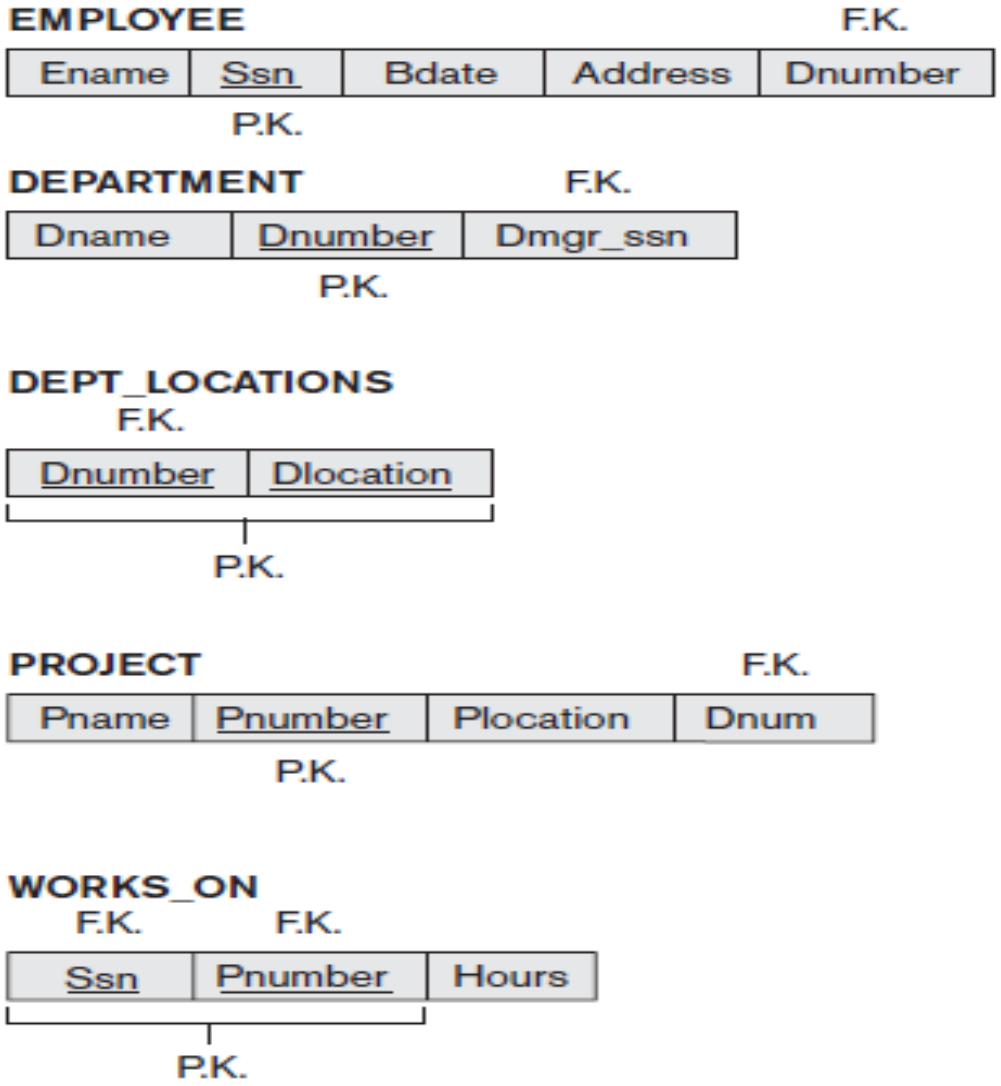
- Eliminating or reducing redundancy in relations.

- Why is redundancy created in DB?

  - When 2 or more independent relations are stored in a single relation

  - **Eg. if EMPLOYEE and DEPARTMENT relations are stored as a single relation**

COMPANY Database



Why Redundancy?

DEPARTMENT

Dname	Dnumber	Dmgr_ssn
Research	5	333445555
Administration	4	987654321
Headquarters	1	888665555

EMPLOYEE

Ename	Ssn	Bdate	Address	Dnumber
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4
Wallace, Jennifer S.	987654321	1941-06-20	291Berry, Bellaire, TX	4
Narayan, Ramesh K.	666884444	1962-09-15	975 Fire Oak, Humble, TX	5
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1

2 independent relations combined

Redundancy

EMP_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research	333445555
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research	333445555
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration	987654321
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration	987654321
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research	333445555
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research	333445555
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration	987654321
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters	888665555

Problems Encountered due to Redundancy

- **Update Anomaly:**

EMPLOYEE

Ename	Ssn	Bdate	Address	Dnumber
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4
Wallace, Jennifer S.	987654321	1941-06-20	291Berry, Bellaire, TX	4
Narayan, Ramesh K.	666884444	1962-09-15	975 Fire Oak, Humble, TX	5
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1

PROJECT

Pname	Pnumber	Plocation	Dnum
ProductX	1	Bellaire	5
ProductY	2	Sugarland	5
ProductZ	3	Houston	5
Computerization	10	Stafford	4
Reorganization	20	Houston	1
Newbenefits	30	Stafford	4

Update Anomaly (Cont.)

2 independent relations combined

Redundancy

Redundancy

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
123456789	1	32.5	Smith, John B.	ProductX	Bellaire
123456789	2	7.5	Smith, John B.	ProductY	Sugarland
666884444	3	40.0	Narayan, Ramesh K.	ProductZ	Houston
453453453	1	20.0	English, Joyce A.	ProductX	Bellaire
453453453	2	20.0	English, Joyce A.	ProductY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProductY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProductZ	Houston
333445555	10	10.0	Wong, Franklin T.	Computerization	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganization	Houston
999887777	30	30.0	Zelaya, Alicia J.	Newbenefits	Stafford
999887777	10	10.0	Zelaya, Alicia J.	Computerization	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Computerization	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Newbenefits	Stafford
987654321	30	20.0	Wallace, Jennifer S.	Newbenefits	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganization	Houston
888665555	20	Null	Borg, James E.	Reorganization	Houston

Changing the name of project number “30” from “Newbenefits” to “Customer-Accounting” may cause this update to be made for all employees working on project “30”.

Problems Encountered due to Redundancy

- **Insertion Anomaly:**

- *Cannot insert a project unless an employee is assigned to it.*
- *If we still want to insert a project, the employee details will be filled by dummy data which causes inconsistency*

Redundancy Redundancy

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
123456789	1	32.5	Smith, John B.	ProductX	Bellaire
123456789	2	7.5	Smith, John B.	ProductY	Sugarland
666884444	3	40.0	Narayan, Ramesh K.	ProductZ	Houston
453453453	1	20.0	English, Joyce A.	ProductX	Bellaire
453453453	2	20.0	English, Joyce A.	ProductY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProductY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProductZ	Houston
333445555	10	10.0	Wong, Franklin T.	Computerization	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganization	Houston
999887777	30	30.0	Zelaya, Alicia J.	Newbenefits	Stafford
999887777	10	10.0	Zelaya, Alicia J.	Computerization	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Computerization	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Newbenefits	Stafford
987654321	30	20.0	Wallace, Jennifer S.	Newbenefits	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganization	Houston
888665555	20	Null	Borg, James E.	Reorganization	Houston

Problems Encountered due to Redundancy

- **Deletion Anomaly:**

If a project is deleted, consequently all the employees who work on that project will be deleted. Thus, loosing employee information.

Redundancy | Redundancy

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
123456789	1	32.5	Smith, John B.	ProductX	Bellaire
123456789	2	7.5	Smith, John B.	ProductY	Sugarland
666884444	3	40.0	Narayan, Ramesh K.	ProductZ	Houston
453453453	1	20.0	English, Joyce A.	ProductX	Bellaire
453453453	2	20.0	English, Joyce A.	ProductY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProductY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProductZ	Houston
333445555	10	10.0	Wong, Franklin T.	Computerization	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganization	Houston
999887777	30	30.0	Zelaya, Alicia J.	Newbenefits	Stafford
999887777	10	10.0	Zelaya, Alicia J.	Computerization	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Computerization	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Newbenefits	Stafford
987654321	30	20.0	Wallace, Jennifer S.	Newbenefits	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganization	Houston
888665555	20	Null	Borg, James E.	Reorganization	Houston

Guidelines for Normalized DB

THE NORTHCAP UNIVERSITY

- *Attributes of different entities should not be mixed in the same relation and only foreign keys should be used to refer to other entities.*
- *Design a schema that does not suffer from the insertion, deletion and update anomalies.*
- *Relations should be designed such that their tuples will have as few NULL values as possible Attributes that are NULL frequently could be placed in separate relations.*
- *The relations should be designed to satisfy the lossless join condition. No spurious tuples should be generated by doing a natural-join of any relations.*

NAAC Accredited

Solution of Redundancy

EMP_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname	Dmgr_ssn

EMPLOYEE

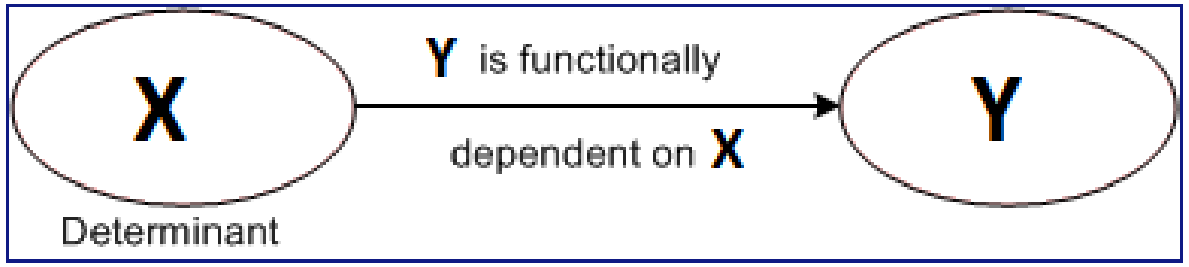
Ename	Ssn	Bdate	Address	Dnumber

DEPARTMENT

Dname	Dnumber	Dmgr_ssn

Functional Dependency

- If one set of attributes in a table determines another set of attributes in the table, then the second set of attributes is said to be functionally dependent on the first set of attributes.



Functional dependency between X and Y

- FDs are **constraints** that are derived from the *meaning* and *interrelationships* of the data attributes

Functional Dependencies (Cont.)

- A set of attributes X *functionally determines* a set of attributes Y if the value of X determines a unique value for Y.

- X → Y holds if whenever two tuples have the same value for X, they *must have* the same value for Y.

- For any two tuples t1 and t2 in any relation:

If **t1[X] = t2[X],** ***then*** **t1[Y] = t2[Y]**

- X → Y in R specifies a *constraint* on all relation instances r(R)

Examples of FD

THE NORTHCAP UNIVERSITY

- social security number determines employee name

SSN → ENAME

- project number determines project name and location

PNUMBER → {PNAME, PLOCATION}

- employee ssn and project number determines the hours per week that the employee works on the project

{SSN, PNUMBER} → HOURS

30 YEARS OF EXCELLENCE | NAAC Accredited

Functional Dependencies (Cont.)

- If K is a key of R,

  then K functionally determines all attributes in R

  since, two distinct tuples with t1[K]=t2[K] not possible

EMPLOYEE

Ename	Ssn	Bdate	Address	Dnumber

SSN → ENAME

SSN → BDATE

SSN → ADDRESS

SSN → DNUMBER

Examples of FD

PubId	PubName	PubPhone
1	Big House	123-456-7890
2	Small House	999-456-7890
3	Alpha Press	111-111-7890

AuId	AuName	AuPhone
1	Sleepy	999-456-9999
2	Snoopy	999-456-1111
3	Grumpy	111-111-7890
4	Jones	222-222-2222
5	Smith	333-333-3333
6	Joyce	444-444-4444
7	Snoopy	555-555-5555

Table Scheme: {PubID, PubName, PubPhone}

Functional Dependencies:

{PubId} → {PubPhone}

{PubId} → {PubName}

{PubName, PubPhone} → {PubID}

Table Scheme: {AuID, AuName, AuPhone}

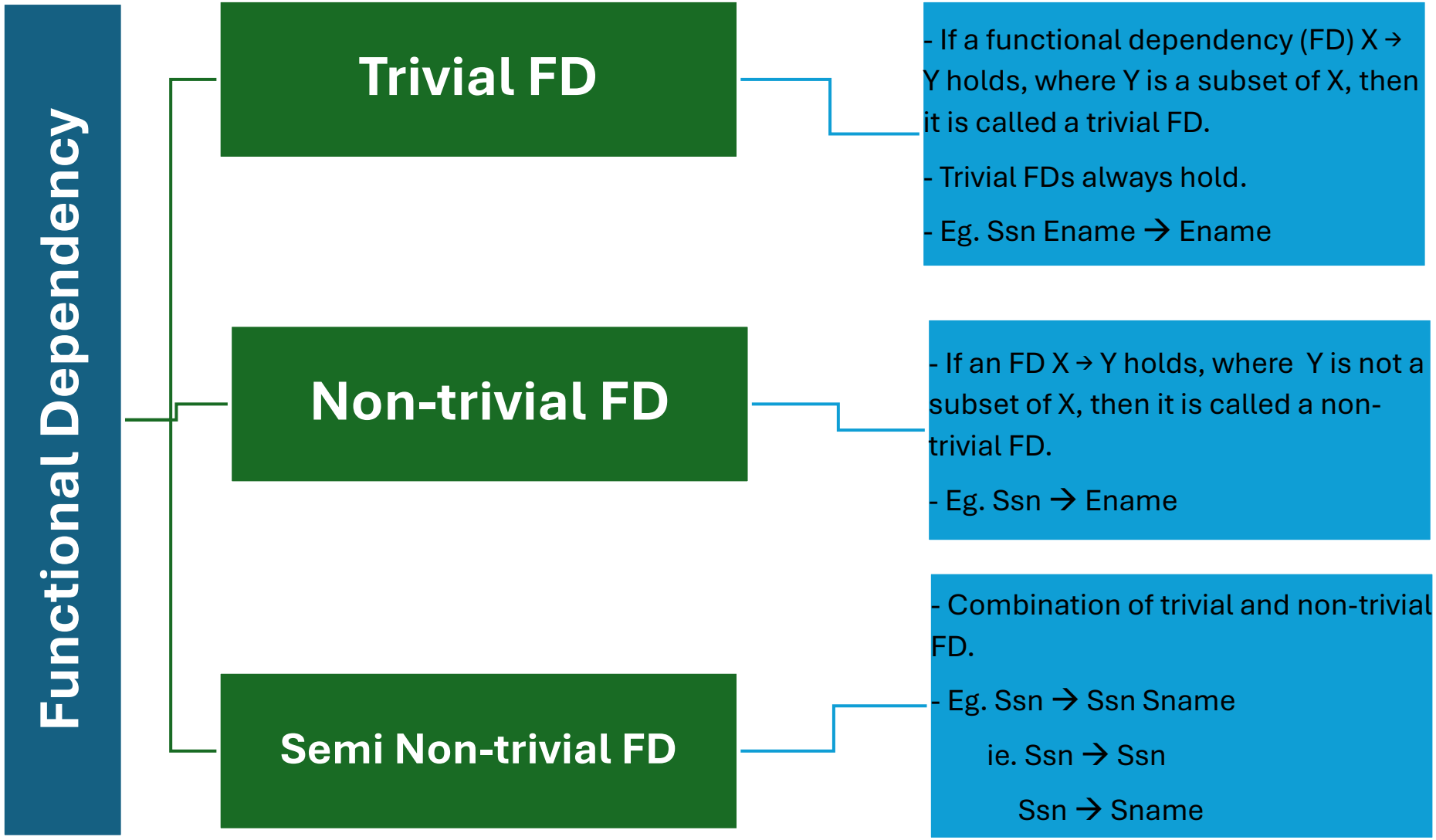
Functional Dependencies:

{AuId} → {AuPhone}

{AuId} → {AuName}

{AuName, AuPhone} → {AuID}

FD Categorization



FD Important Note

- Trivial FD are always implied but non-trivial FD can be implied/not implied.
- A relation with no non-trivial FD is also possible.

Example:

A	B	C
3	4	2
3	4	3
3	6	2
3	6	3
4	4	2

A → B B → A AB → C
A → C B → C BC → A
A → BC B → AC AC → B

None of the above non-trivial FDs are implied. Thus, the candidate key for this relation is ABC

- A FD is implied/not implied is decided on the basis on data.
- Trivial/non-trivial FD is decided on the basis of attributes.

Inference Rules for FDs

- **Armstrong's inference rules:**

IR1. Reflexive: If Y *subset-of* X, then X → Y

IR2. Augmentation: If X → Y, then XZ → YZ

IR3. Transitive: If X → Y and Y → Z, then X → Z

Inference Rules for FDs (Cont.)

- **Additional Inference Rules:**

IR4. Decomposition If X → YZ, then X → Y and X → Z

IR5. Union: If X → Y and X → Z, then X → YZ

IR6. Psuedotransitivity: If X → Y and WY → Z,

then WX → Z

THE NORTHCAP UNIVERSITY

Thanks!!

30 YEARS OF EXCELLENCE

NAAC Accredited