Vietnam and Japan
Joint ICT HRD Program

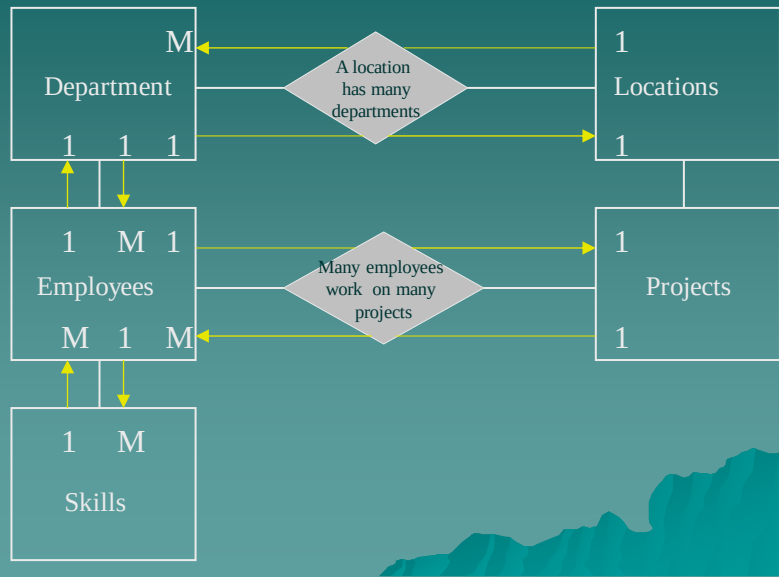
ITSS Software Development

Chapter 8. Normalization of the data structure using E-R

Nguyen Thi Thu Trang
trangntt-fit@mail.hut.edu.vn

1

Example: Entity Relationship Diagram (ERD)



Design with Normalization

- **Normalization**: the process of steps that will identify, for elimination, redundancies in a database design.

Normalization Forms

- First Normal Form: eliminate repeating groups → *repeating items within an entity are brought out as independent entity.*
- Second Normal Form: eliminate columns (attributes) that depend only on parts of the primary key → *the items depending on only some parts of composite key are brought out as independent entity*
- Third Normal Form: eliminate columns that don't depend on the primary key at all → the items depending on non-primary key or items are brought out as independent entity.

4

First Normal Form: eliminate repeating groups

Denormalized EMPL

EMP NO	LAST NAME	WORK DEPT	DEPT NAME	SKILL 1	SKILL2	SKILLN
000030	KWAN	GRE	OPERATIONS	141		
000250	SMITH	BLU	PLANNING	002	011	067
000270	PEREZ	RED	MARKETING	415	447	
000300	SMITH	BLU	PLANNING	011	032	

Normalized to 1NF - EMPL

EMP NO	LAST NAME	WORK DEPT	DEPT NAME
000030	KWAN	GRE	OPERATIONS
000250	SMITH	BLU	PLANNING
000270	PEREZ	RED	MARKETING
00030	SMITH	BLU	PLANNING

EMPL_SKILL TABLE

EMP NO	SKILL	SKILL DESC
000030	141	RESEARCH
000250	002	BID PREP
000250	011	NEGOTIATION
000250	067	PROD SPEC
000270	415	BENEFITS ANL
000270	447	TESTING
000300	011	NEGOTIATION
000300	032	INV CONTROL

Second Normal Form: eliminate the columns that depend only on part of the key

Normalized – 1NF

EMPL_SKILL TABLE

EMP NO	SKILL	SKILL DESC
000030	141	RESEARCH
000250	002	BID PREP
000250	011	NEGOTIATION
000250	067	PROD SPEC
000270	415	BENEFITS ANL
000270	447	TESTING
000300	011	NEGOTIATION
000300	032	INV CONTROL

Normalized – 2NF

EMPL_SKILL TABLE

EMP NO	SKILL	DATE CERT
000030	141	
000250	002	
000250	011	
000250	067	
000270	415	
000270	447	
000300	011	
000300	032	

SKILL DESC TABLE

SKILL	SKILL DESC
141	RESEARCH
002	BID PREP
011	NEGOTIATION
067	PROD SPEC
415	BENEFITS ANL
447	TESTING
011	NEGOTIATION
032	INV CONTROL

Third Normal Form: eliminate columns that don't depend on the key at all

Denormalized EMPL

EMP NO	LAST NAME	WORK DEPT	DEPT NAME	MGRNO	...
000030	KWAN	GRE	OPERATIONS	000080	
000250	SMITH	BLU	PLANNING	000010	
000270	PEREZ	RED	MARKETING	000020	
000300	SMITH	BLU	PLANNING	000010	

Normalized to 3NF

EMPL

EMP NO	LAST NAME	WORK DEPT	...
000030	KWAN	GRE	
000250	SMITH	BLU	
000270	PEREZ	RED	
00030	SMITH	BLU	

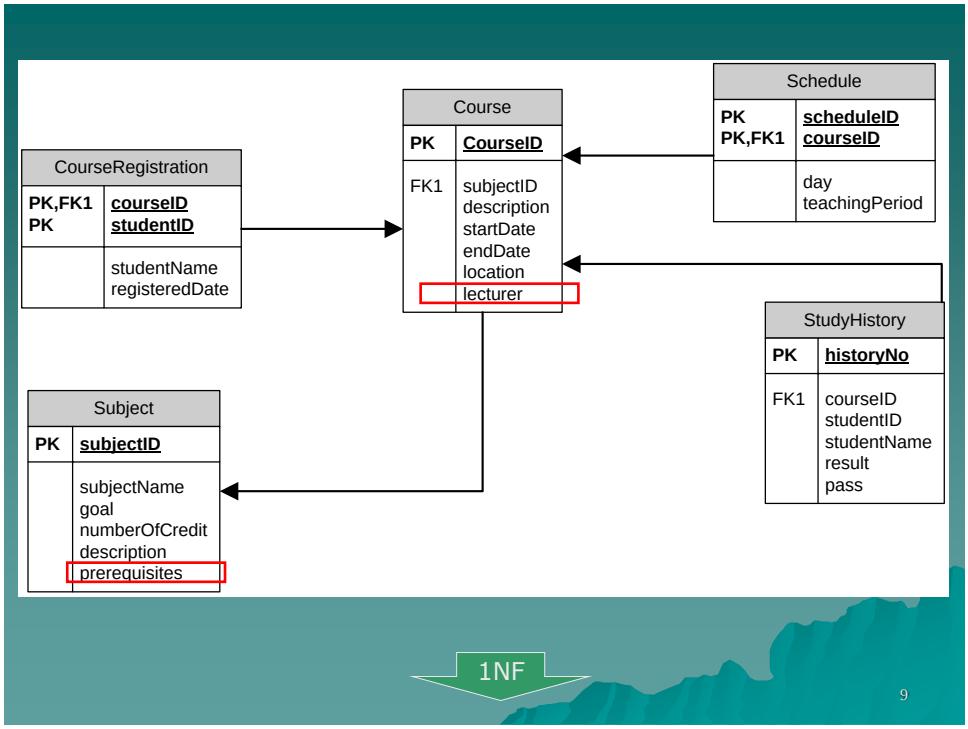
DEPT

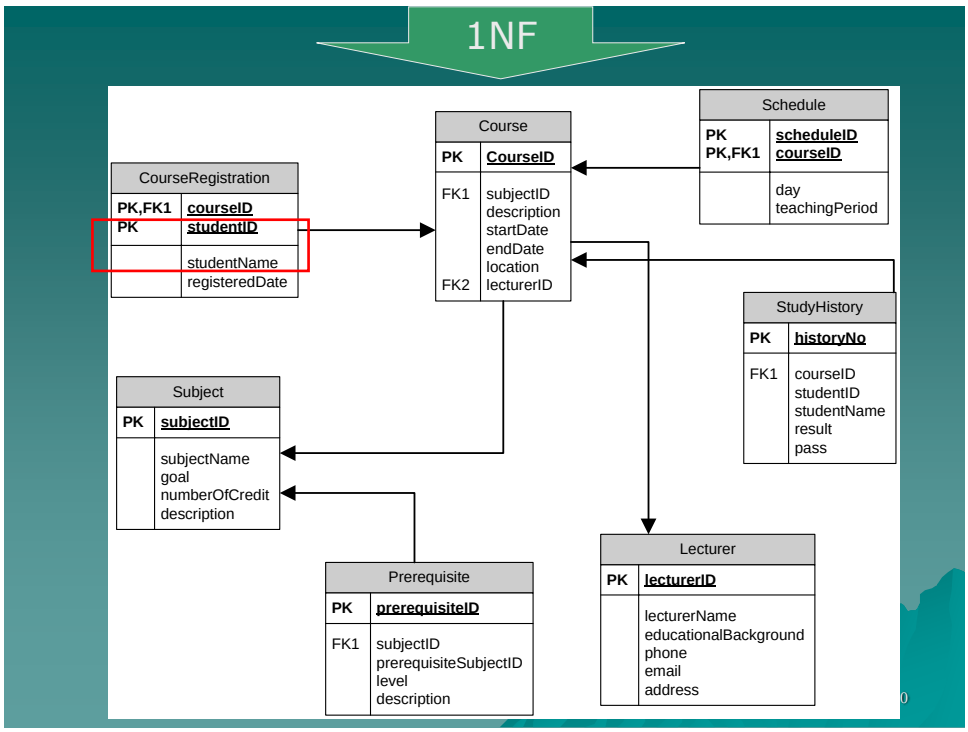
WORK DEPT	DEPT NAME	MGRNO
GRE	OPERATIONS	000080
BLU	PLANNING	000010
RED	MARKETING	000020
BLU	PLANNING	000010

"Register for course" use case

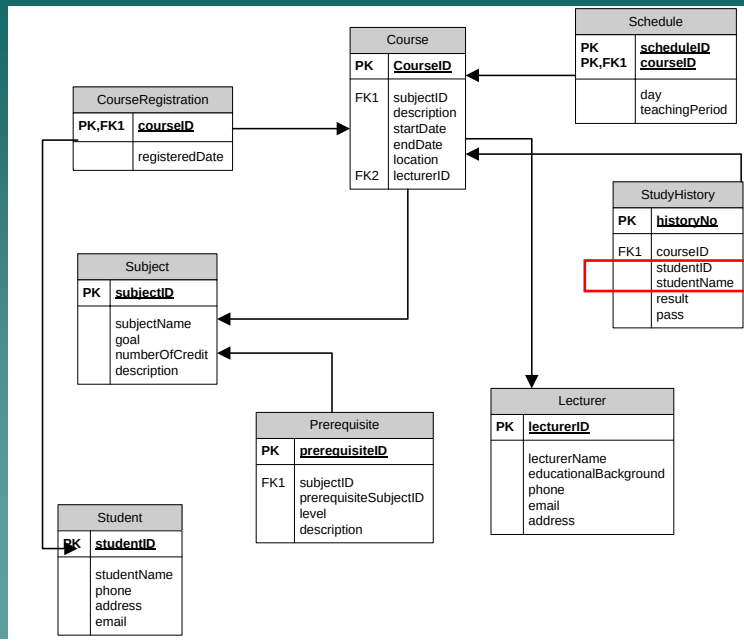
- Make the E-R diagram from the previous step for "Register for course" use case to become:
  - The first normal form
  - The second normal form
  - The third normal form

8


Schedule
PK
PK,FK1
scheduleID
courseID
day
teachingPeriod
Course
PK
CourseID
FK1
subjectID
description
startDate
endDate
location
lecturer
CourseRegistration
PK,FK1
PK
courseID
studentID
studentName
registeredDate
StudyHistory
PK
historyNo
FK1
courseID
studentID
studentName
result
pass
Subject
PK
subjectID
subjectName
goal
numberOfCredit
description
prerequisites
1NF
9

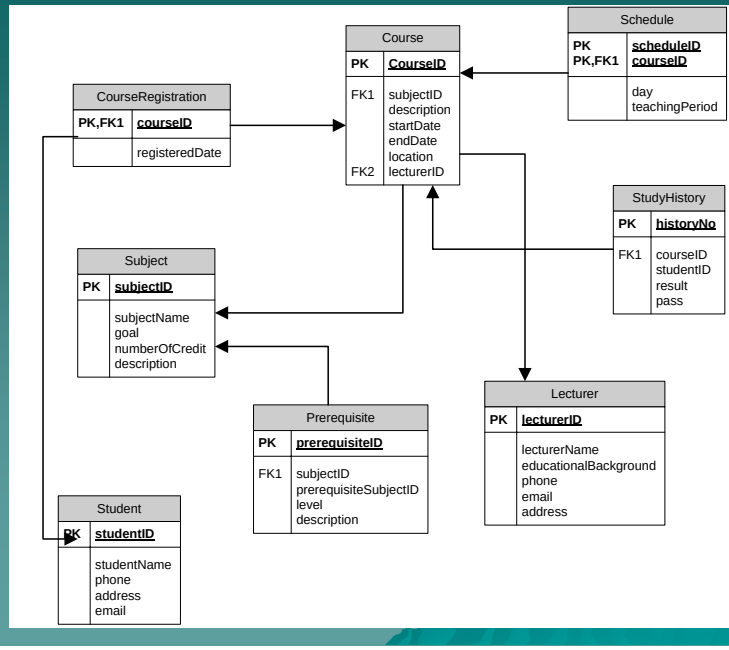

1NF
Schedule
PK
PK,FK1
scheduleID
courseID
day
teachingPeriod
Course
PK
CourseID
FK1
subjectID
description
startDate
endDate
location
FK2
lecturerID
CourseRegistration
PK,FK1
PK
courseID
studentID
studentName
registeredDate
StudyHistory
PK
historyNo
FK1
courseID
studentID
studentName
result
pass
Subject
PK
subjectID
subjectName
goal
numberOfCredit
description
Prerequisite
PK
prerequisiteID
FK1
subjectID
prerequisiteSubjectID
level
description
Lecturer
PK
lecturerID
lecturerName
educationalBackground
phone
email
address
10

2NF


Course
PK CourseID
FK1 subjectID
description
startDate
endDate
location
FK2 lecturerID
Schedule
PK scheduleID
PK,FK1 courseID
day
teachingPeriod
CourseRegistration
PK,FK1 courseID
registeredDate
StudyHistory
PK historyNo
FK1 courseID
studentID
studentName
result
pass
Subject
PK subjectID
subjectName
goal
numberOfCredit
description
Prerequisite
PK prerequisiteID
FK1 subjectID
prerequisiteSubjectID
level
description
Lecturer
PK lecturerID
lecturerName
educationalBackground
phone
email
address
Student
PK studentID
studentName
phone
address
email

11

3NF


Course
PK CourseID
FK1 subjectID
description
startDate
endDate
location
FK2 lecturerID
Schedule
PK scheduleID
PK,FK1 courseID
day
teachingPeriod
CourseRegistration
PK,FK1 courseID
registeredDate
StudyHistory
PK historyNo
FK1 courseID
studentID
result
pass
Subject
PK subjectID
subjectName
goal
numberOfCredit
description
Prerequisite
PK prerequisiteID
FK1 subjectID
prerequisiteSubjectID
level
description
Lecturer
PK lecturerID
lecturerName
educationalBackground
phone
email
address
Student
PK studentID
studentName
phone
address
email

12

Checkpoints: Normalization

- Have you

Question?