**MASSACHUSETTS INSTITUTE OF TECHNOLOGY**
SLOAN SCHOOL OF MANAGEMENT

15.565 Integrating Information Systems:
Technology, Strategy, and Organizational Factors

15.578 Global Information Systems:
Communications & Connectivity Among Information Systems

Spring 2002

Lecture 16

SCHEMA INTEGRATION
AMONG DATABASES

SCHEMA INTEGRATION

SCHEMA (VIEW)   INTEGRATED SCHEMA

USA

SCHEMA (VIEW)

Europe

Global
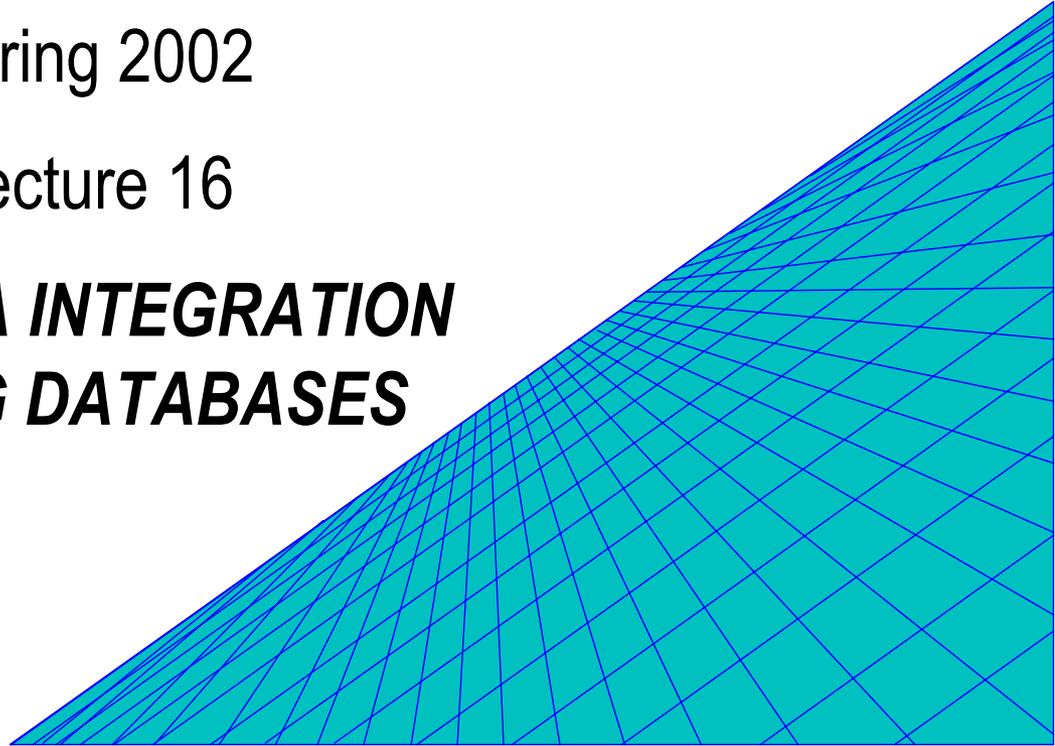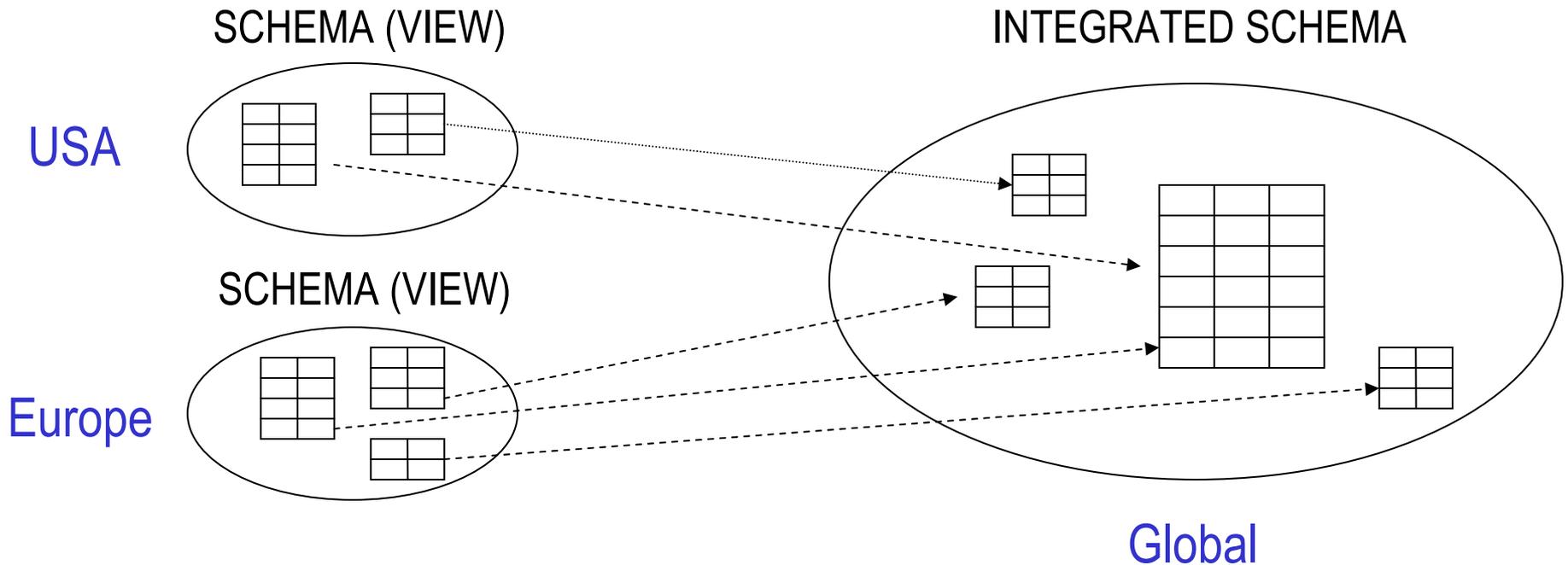
- **PURPOSES FOR SCHEMA / VIEW INTEGRATION**

  – **NEW DATABASE DESIGN**: MERGE MULTIPLE PEOPLE'S VIEWS

  – **DATABASE REDESIGN**: MERGE MULTIPLE EXISTING DB'S TO NEW DB

  – **DATA WAREHOUSE**: CREATE INTEGRATED SCHEMA / DATABASE

  – **INTEGRATED VIEW FOR DISTRIBUTED HETEROGENEOUS DBMS**

EVOLUTION OF CONCEPTUAL DATABASE DESIGN

- **TRADITIONAL DATA MODELS**

HIERARCHICAL   **NETWORK**   **RELATIONAL**

```
                        PRODUCT
    DIVISION                          DEPARTMENT
                  DEPT    CUSTOMER       Table
  DEPARTMENT
                        EMPLOYEE       EMPLOYEE
                                         Table
EMPLOYEES  PRODUCTS
                        PROJECT        PRODUCT
                                         Table
```
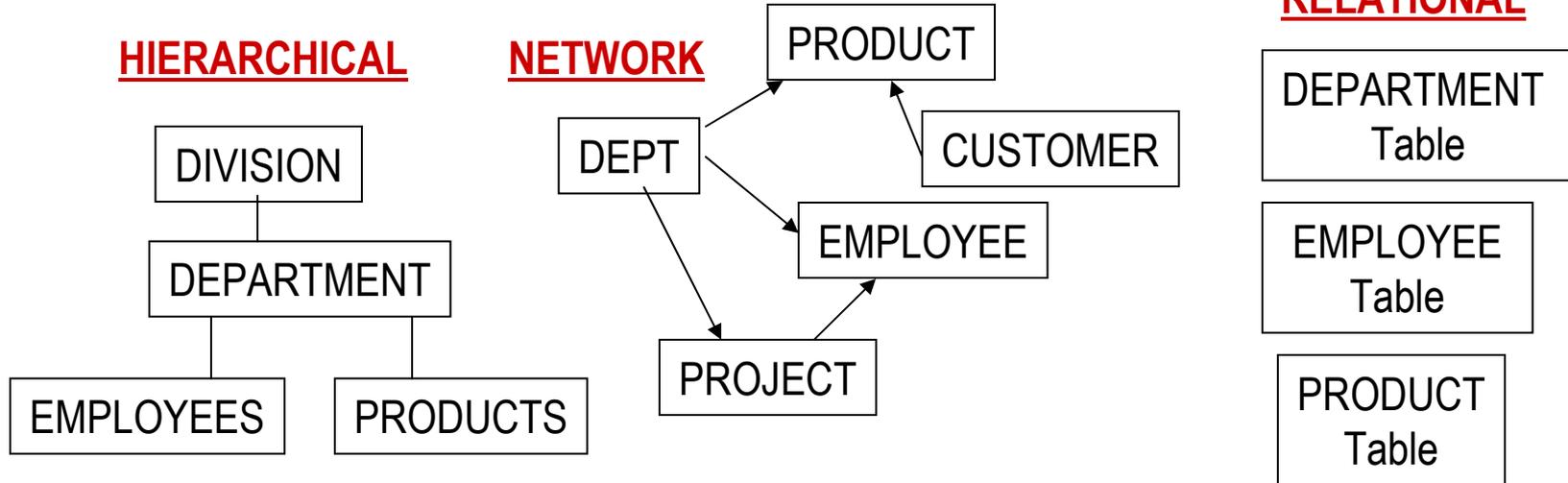
- **SEMANTICALLY -- RICHER DATA MODELS**
  - **ENTITY-RELATIONSHIP (ER) MODEL**
  - SEMANTIC DATA MODEL
  - OBJECT-ORIENTED (OO) DATA MODEL
  - OBJECT-RELATIONAL (OR) DATA MODEL

- **"CONCEPTUAL" DATABASE DESIGN**
  - LEADS TO "CONCEPTUAL SCHEMA"
  - AIDS INTEGRATION ("SCHEMA INTEGRATION")

3

PROCESS OF SCHEMA INTEGRATION

- DATA -- AN AUTONOMOUS RESOURCE
- NEED TO CAPTURE "MEANING" OF DATA
- DESIRE "ENTERPRISE-WIDE" VIEW OF DATA

TERMINOLOGY

- **SCHEMA INTEGRATION**
  - COMPONENT SCHEMA
  - INTEGRATED SCHEMA
- **VIEW INTEGRATION**
  - USER VIEW
  - CONCEPTUAL VIEW
- **DATABASE INTEGRATION**
 - LOCAL SCHEMA
  - GLOBAL SCHEMA

EXAMPLE

- TWO SCHEMAS DEVELOPED BY TWO GROUPS
 - BOTH FOCUS ON BOOKS/PUBLICATIONS

ENTITY-RELATION (ER) DIAGRAM – For Book Distributor

- ENTITIES
- **RELATIONSHIPS**
- *ATTRIBUTES*

Book Distributor

Name — Publisher — Address

Publisher ◇ Book — Title

Book ◇ University — Name, State

Book ◇ Topics — Name

University Library

Title, Code, Publisher

◇ Keyword — Title, Code, Research Area

Step 1. Original schemas

Entity and Attribute Equivalence

Name — Publisher — Address

Publisher ◇ Book — Title

Book ◇ University — Name, State

Book ◇ Topics — **Title**

Publication — Title, Code, Publisher

◇ **Topics** — Title, Code, Research Area
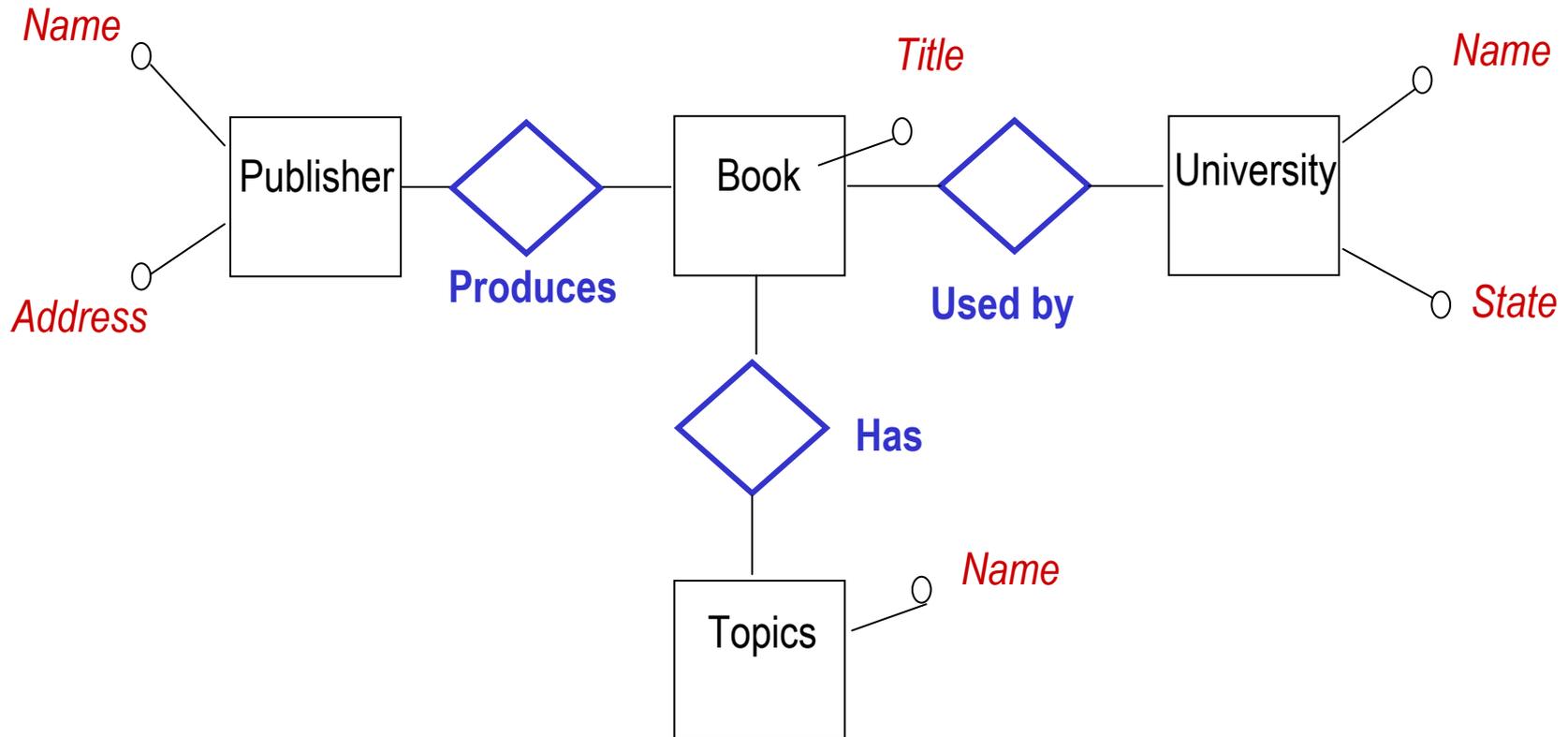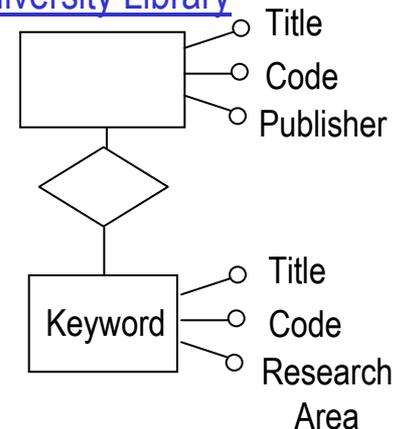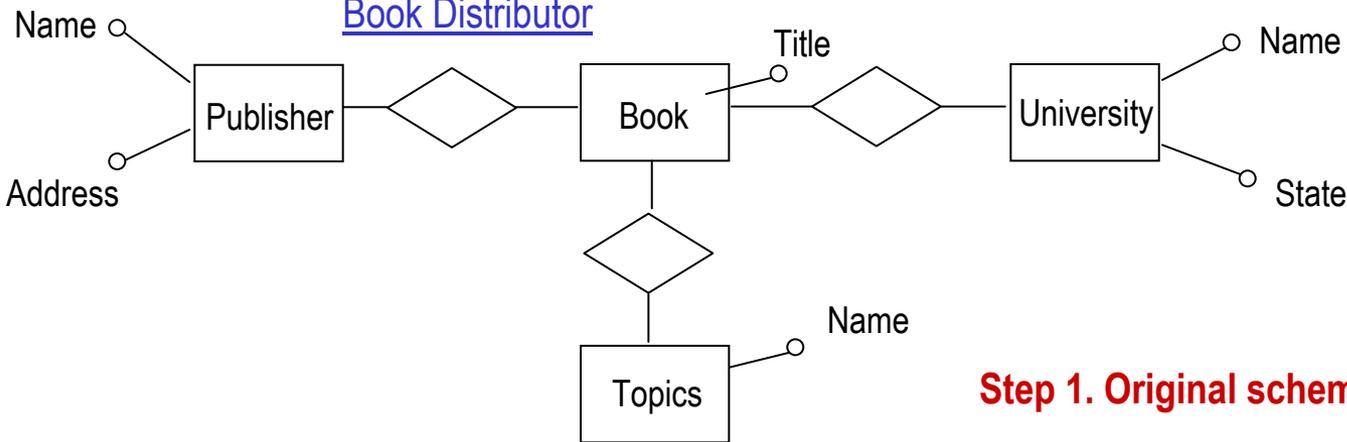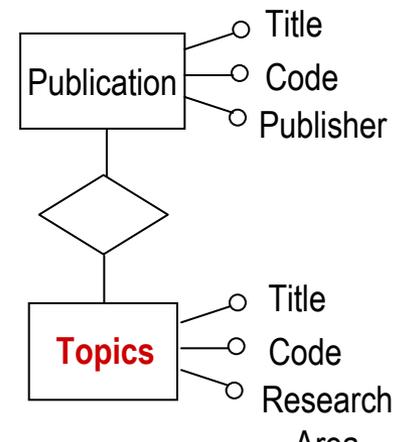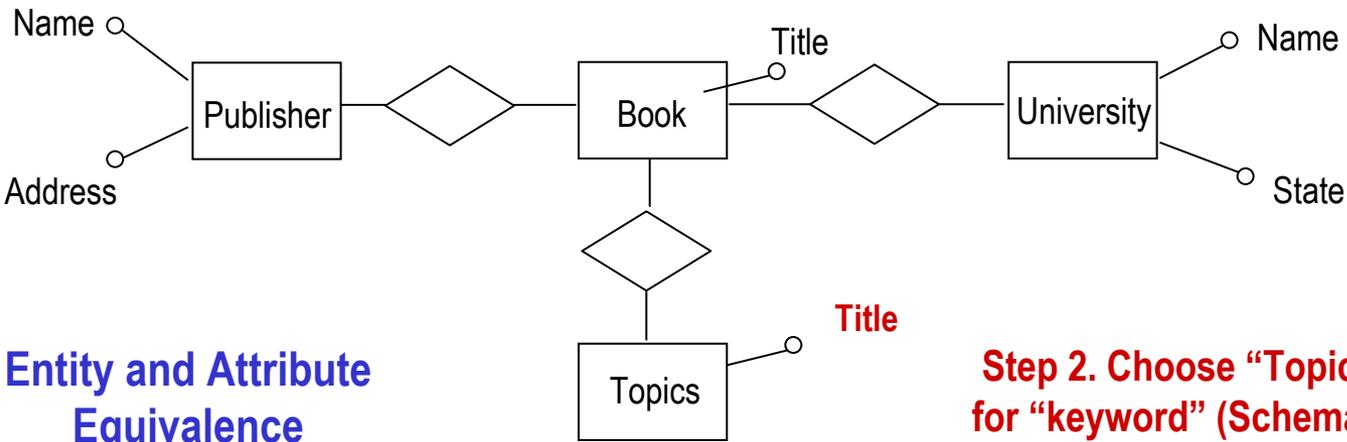
Step 2. Choose "Topics" for "keyword" (Schema 2).

Attribute vs Entity

Name — Publisher — Address

Publisher ◇ Book — Title

Book ◇ University — Name, State

Book ◇ Topics — Title

Publisher ◇ Publication — Name, Title, Code

Publication ◇ Topics — Title, Code, Research Area
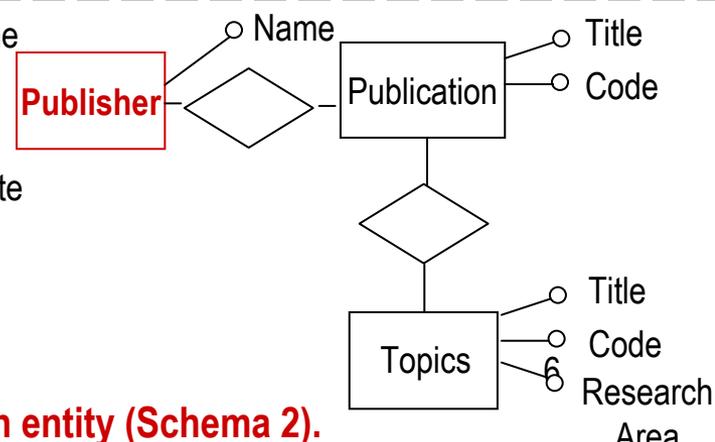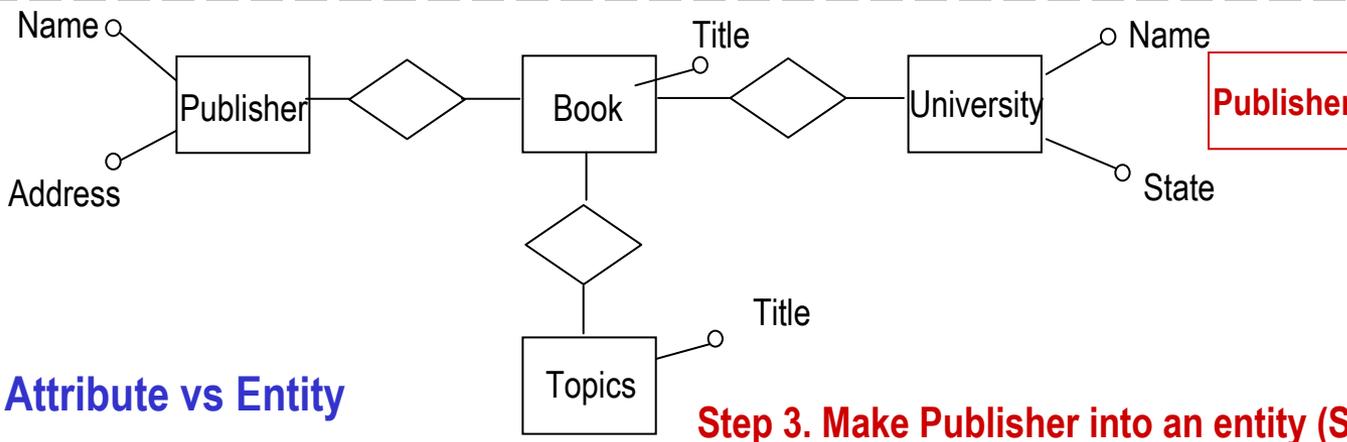
Step 3. Make Publisher into an entity (Schema 2).

Step 4. Superimposition of schema

Publisher — Name, Address

Publication — Title, Code

Topics — Title, Code, Research Area

Book — Title

University — Name, State

Generalization & Specialization

Step 5. Creation of subset relationship

Publisher — Name, Address

Publication — Title, Code

Topics — Title, Code, Research Area

Book — Title

University — Name, State

7

Step 6. Drop the properties of Book common to Publication

CAUSES OF DIFFERENCES

- DIFFERENT PERSPECTIVES
  - DIFFERENT NAME FOR SAME CONCEPT
  - INTERVENING STEPS
    - EMPLOYEE-DEPARTMENT
    - EMPLOYEE-PROJECT-DEPARTMENT

- EQUIVALENCE AMONG CONSTRUCTS
  - MODEL AS ATTRIBUTE OR ENTITY (E.G., PUBLISHER)

- INCOMPATIBLE DESIGN SPECIFICATIONS
 - RELATIONSHIP DIFFERENCES
  (E.G., 1:N VS. N:M)

SOME EXAMPLES

- NAMING CONFLICTS
    - **HOMONYMS**: SAME NAME, DIFFERENT CONCEPT

        E.G., EQUIPMENT

        E.G., "SIZE" OF DRESS = 1 NUMBER, "SIZE" OF PANTS = 2 NUMBERS)

    - **SYNONYMS**: SAME CONCEPT, DIFFERENT NAMES

        E.G., CLIENT, CUSTOMER, PATIENT, ...

- STRUCTURAL CONFLICTS
    - **TYPE**: ENTITY VS. ATTRIBUTE   (E.G., PUBLISHER)
    - **DEPENDENCY**: 1.1 VS. N:M   (E.G., HUSBAND: WIFE)
    - **KEY**: ALTERNATE WAYS TO IDENTIFY ENTITY (E.G., SS# VS. EMP#)
    - **BEHAVIORAL**: DIFFERENT INSERTION/DELETION POLICIES

        (E.G., CAN YOU HAVE DEPARTMENT WITH NO EMPLOYEES?)

EXAMPLE SCHEMAS TO BE INTEGRATED:

Departmental database
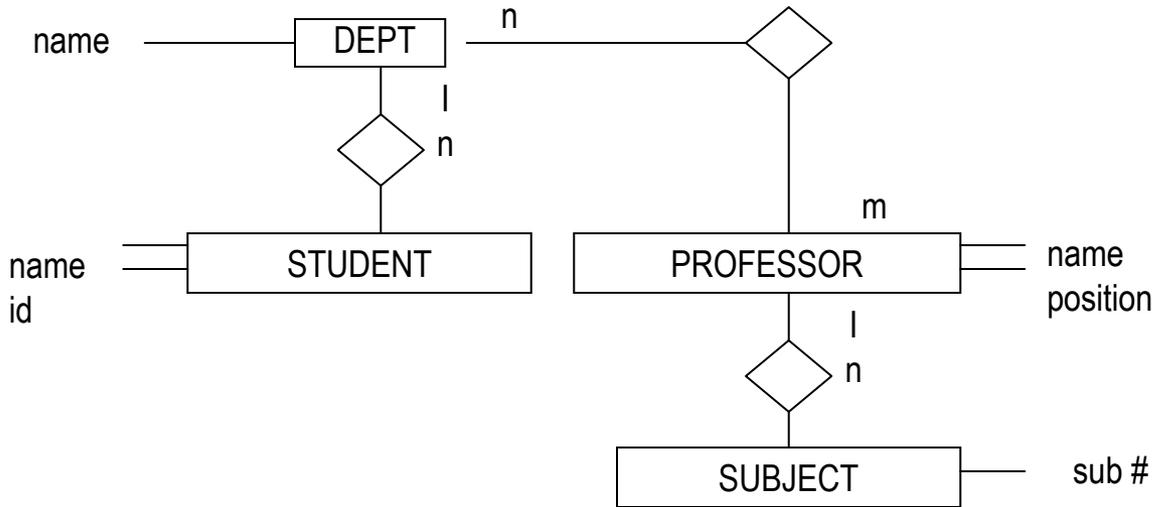
Registrar's database
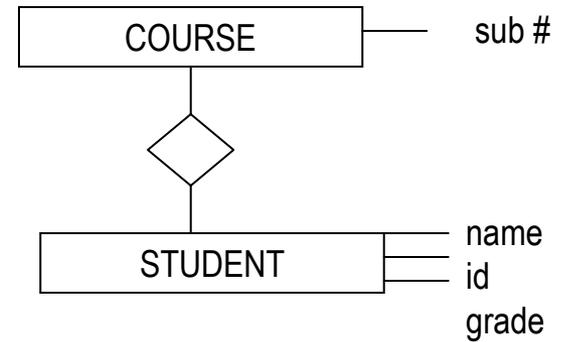
Staff Telephone directory
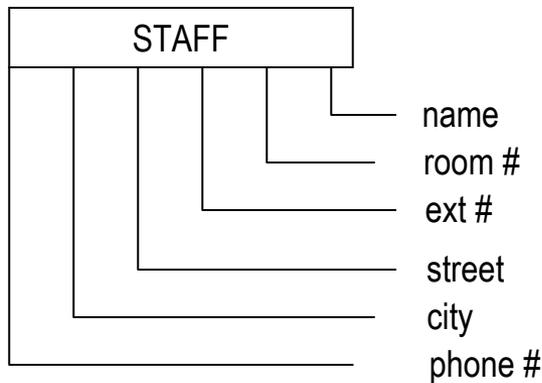
Student Telephone directory
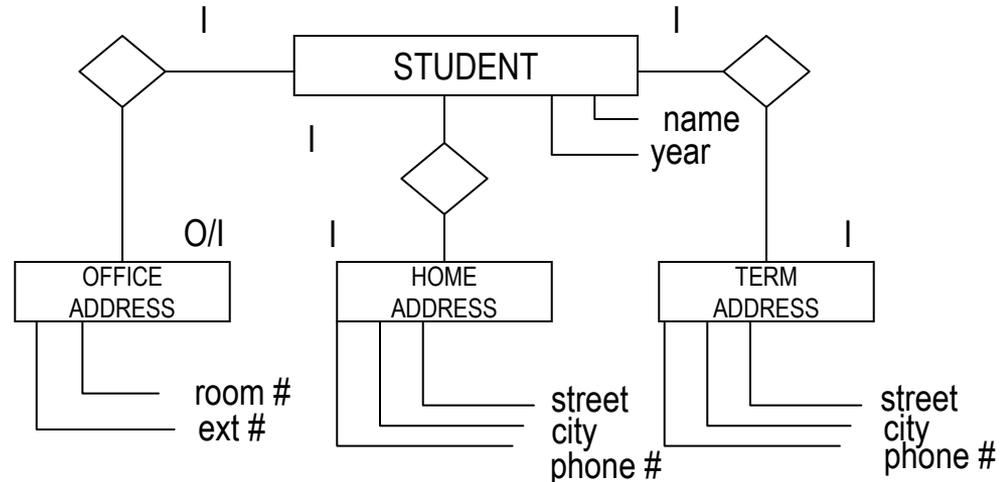
INTEGRATED SCHEMA: (an exercise for the reader)

SUMMARY

- INCREASING NEED TO PROVIDE AN INTEGRATED GLOBAL VIEW OF AN ORGANIZATION'S INFORMATION
  - AND SOMETIMES RELATED ORGANIZATIONS (CUSTOMERS / SUPPLIER)

- AN IMPORTANT STEP IS THE CREATION OF A GLOBAL SCHEMA
  - INTEGRATES SEPARATE SCHEMAS
  - CONTAINS ALL THE CRITICAL INFORMATION NEEDED