Unit 6

View Design and View Integration

本單元目的

- ## What is view?
  - View 指某一使用者或一群使用者所看到的資料庫內容或應用程式的資料需求.

- ## View design
  - 抓住各使用者的需求物件, 把這些物件用 ER model 表現出整個系統的 conceptual schema.

- ## View integration
  - 把數個 conceptual schema 整合為一個 global conceptual schema, 以代表整個應用系統的需求.

Outline

- View Design From Natural Language Requirements

- View Design Starting From Forms

- View Integration in the Large

- Conflict Analysis and Resolution

- Merging of Views

View Design From Natural Language Requirements

- Step 1   Requirement analysis
- Step 2   Initial design
- Step 3   Schema design

View Design From Natural Language Requirements

Step 1. Requirement analysis

1.1. Analyze requirements and filter ambiguities

搞清楚敘述內容, 找出不明確的用語, 重新定義

1.2. Partition sentences into homogeneous sets

把敘述同一概念的句子放在同一組

Step 2. Initial design

2.1. Build a global skeleton schema

根據已分群的句子, 畫出 schema 初步架構, 原則上一群句子描述一個 entity.

Step 3. Schema design 對每個 concept, 反覆使用

Top-down primitives or bottom-up primitives or inside-out primitives 直到需求一一表達在 concept schema 之內.

An example of natural language description of requirement

Step 1: Requirement analysis

Line 1: In a university database, we represent data about
Line 2: students and professors. {For students, we
Line 3: represent last name, age, sex, city and state of
Line 4: birth, city and state of residence of their
Line 5: families, places and states where they lived before
Line 6: (with the period they lived there), courses that
Line 7: they have passed, with name, code, professor,
Line 8: grade, and date.} We also represent courses they
Line 9: are presently attending, and for each day, places

cities

In the
current year

number
of years

classroom

Example(cont.)

course

Line 10: and hours where classes are held (each course

Line 11: meets at most once in one day). For graduate

Line 12: students, we represent the name of the advisor

Line 13: and the total number of credits in the last year.

Line 14: For Ph.D. students, we represent the title and the

research area

professor

Line 15: of their thesis. For teachers, we represent last

Line 16: name, age, place and state of birth, name of the

Line 17: department they belong to, telephone number

Line 18: title, status, and topics of their research.

research area

Example(cont.)

Filtering the ambiguities

Line	Term	New Term	Reason for the Correction
5	Places	Cities	Place is a generic word
6	Period	Number of years	Period is a generic word
9	Presently	In the current year	Presently is ambiguous
9	Day	Day of the week	More specific
9	Places	Rooms	Homonym for places in line 5
10	Classes	Courses	Synonym for courses in line 8
15	Teacher	Professor	Synonym for professor in line 2
16	Place	City	Same as in line 5
17	Telephone	Telephone of the department	More specific
18	Status	Marital status	Status is ambiguous
18	Topic	Research area	Synonym for research area at line 14

Partitioning of sentences into homogeneous groups

In a university database, we represent data about students and professors.

General sentences

For students, we represent last name, age, sex, city and state of birth, city and state of residence of their families, cities and states where they lived before (with the number of years they lived there), courses that they have passed, with name, code,professor, grade, and date.

Sentences on students

We also represent courses they are attending in the current year, and for each day of the week, rooms and hours where courses are held (each course meets at most once in one day).

Sentences on courses

For graduate students, we represent the name of the advisor and the total number of credits in the last year. For Ph.D.students, we represent the title and the research area of their thesis.

Sentences on specific types of students

For professors, we represent their last name, age, city and state of birth, name of the department they belong to, telephone number of the department, title, marital status, and research area.

Sentences on professors

Example(cont.)

Step 2: Initial design

Find the first skeleton schema
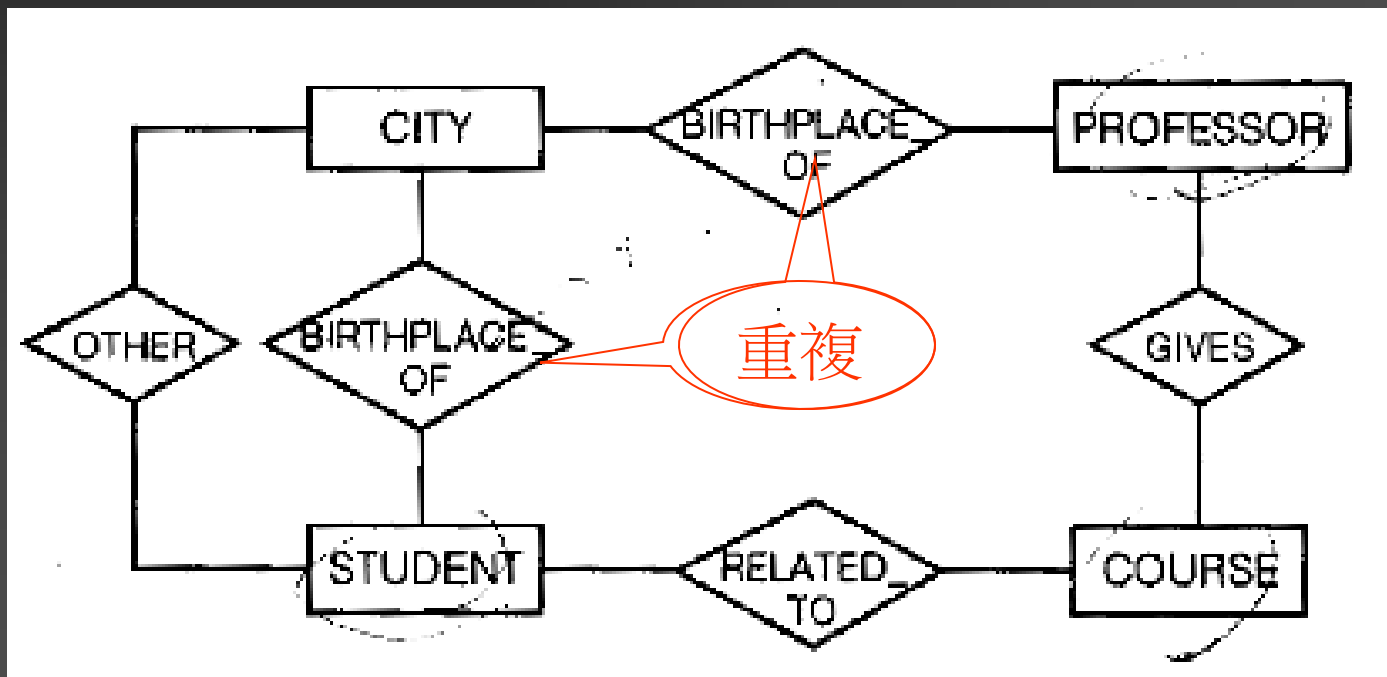
Example(cont.)
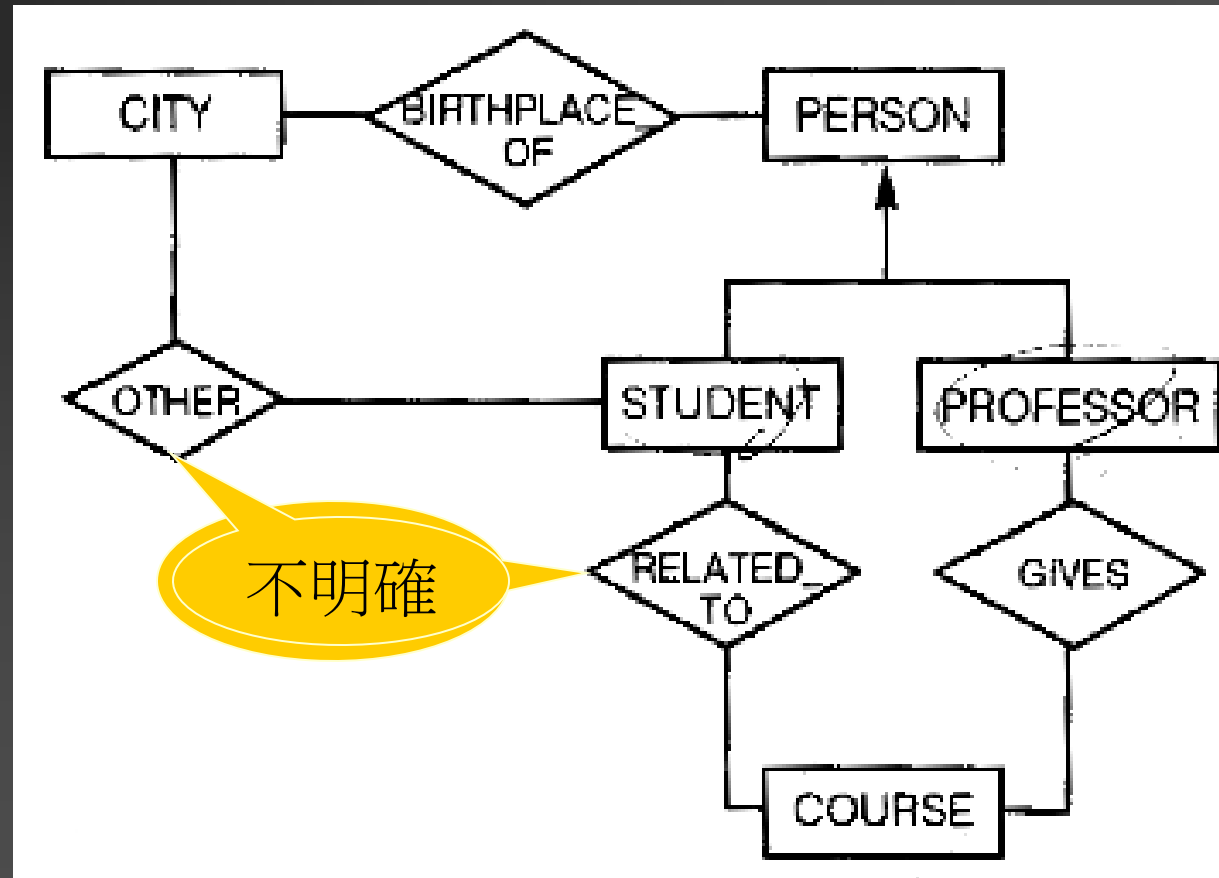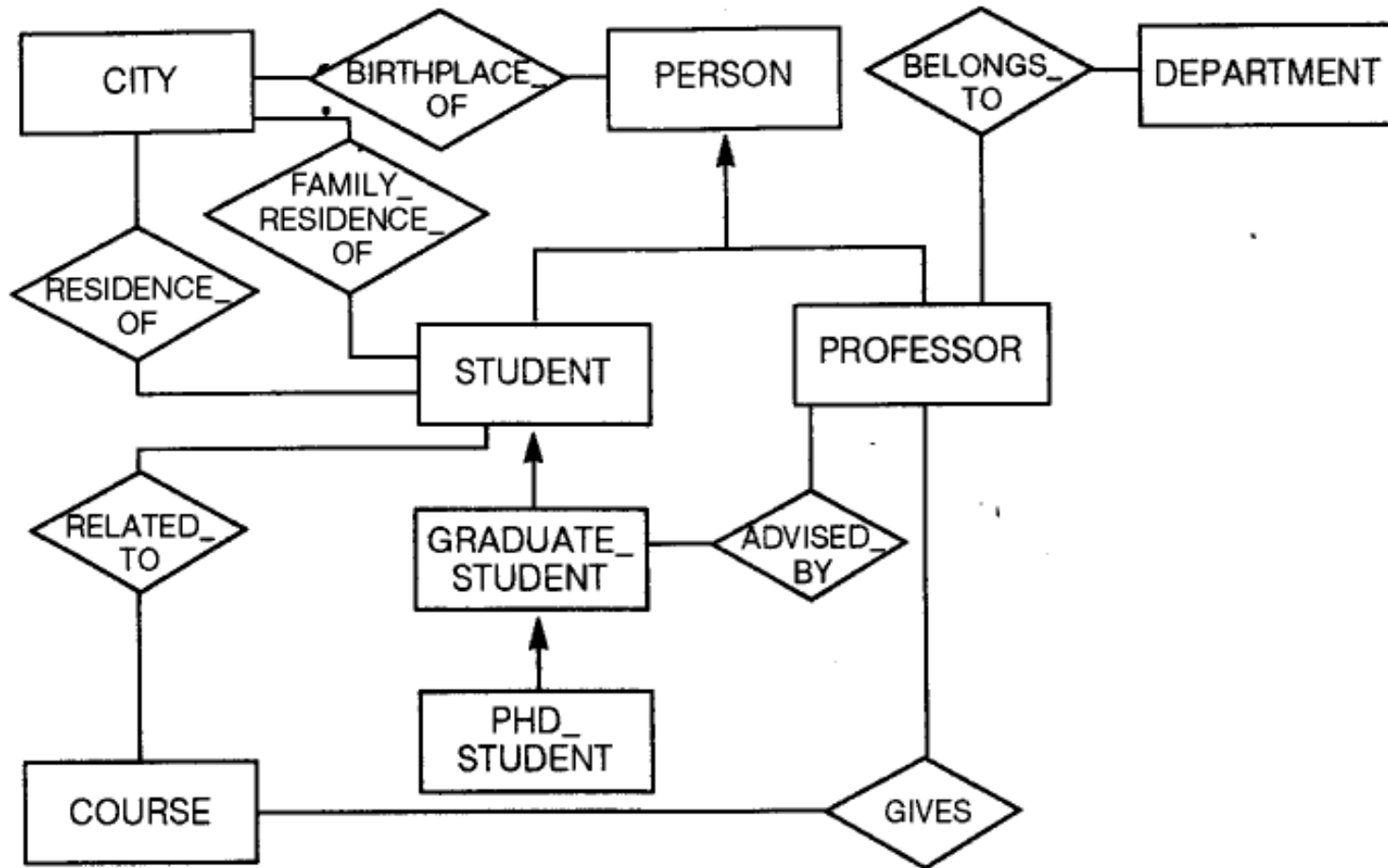
Example(cont.)

Figure 4.7 Refinement of the skeleton schema

Example(cont.)

Final schema

加上全部的
attribute 和
cardinality
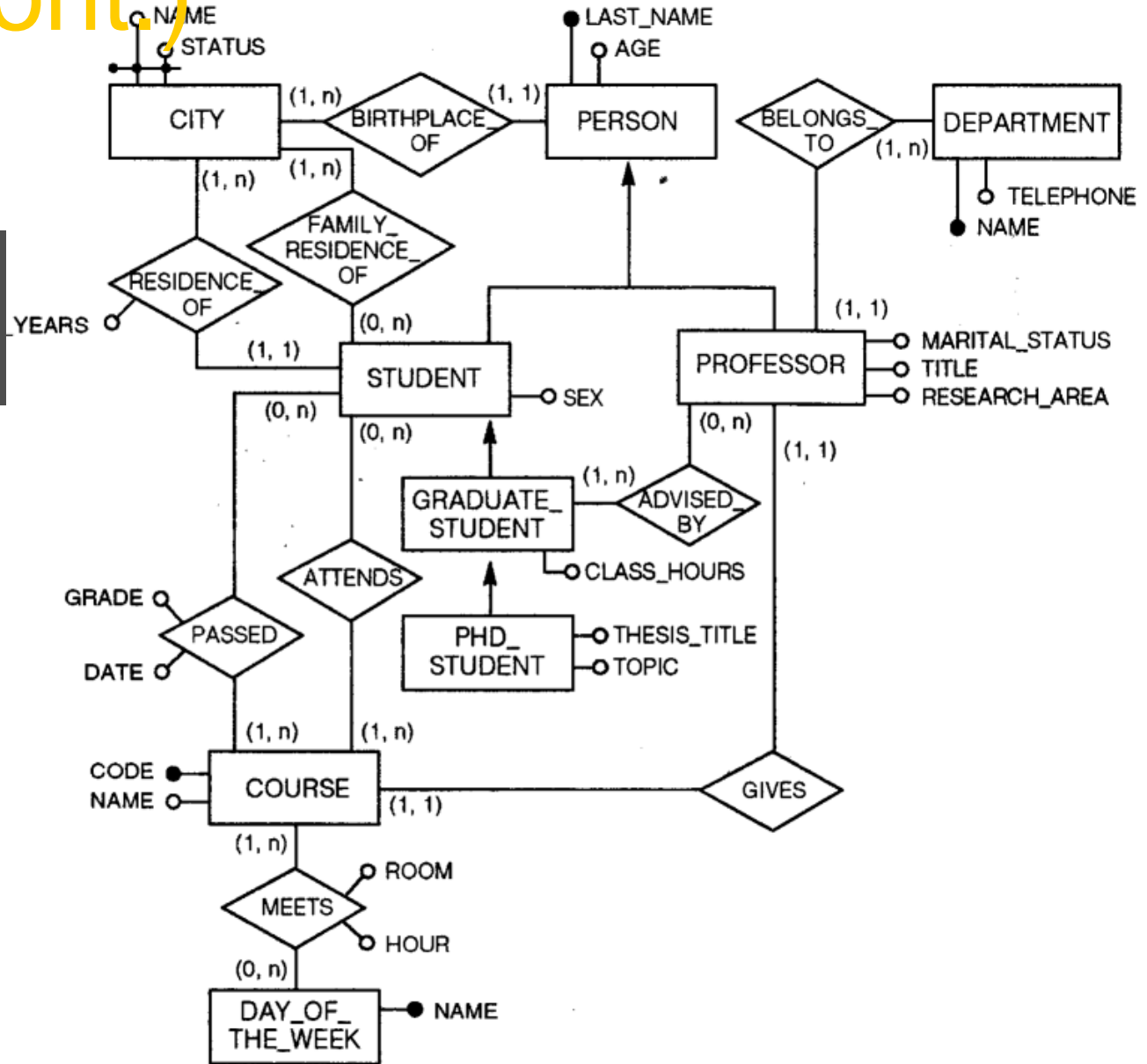
Figure 4.8   Final schema

View Design Starting From Forms

Step 1: Requirements analysis

1.1 Distinguish extensional, intensional and descriptive parts of the form

找出表格中的各部分

1.2 Select areas and subareas

選出表格中相關資料群

- 一張表格可分成四個部分
  - 簽名認證部分
  - 使用者填入內容(extension part)
  - 欄位名稱(intension part)
  - 說明指示部分(descriptive part)

View Design Starting From Forms

Step 2: Initial design

    2.1 Build a global skeleton schema

    建立初步架構圖

Step 3: Schema design for each area

    3.1 Build the area schema

    3.2 Merge the area schema

Example

Example (cont.)

Step 1: Find the areas and sub-areas.
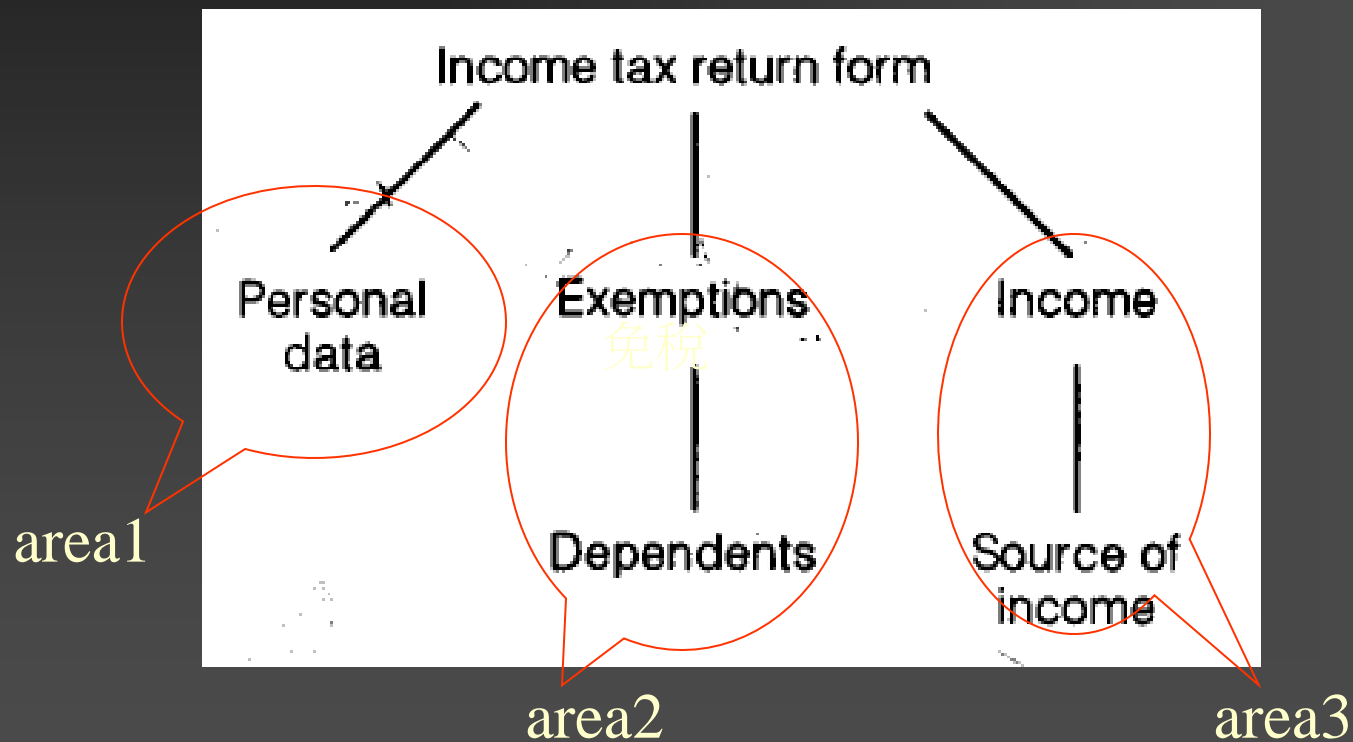
Example (cont.)

Step 2: Build a global skeleton schema
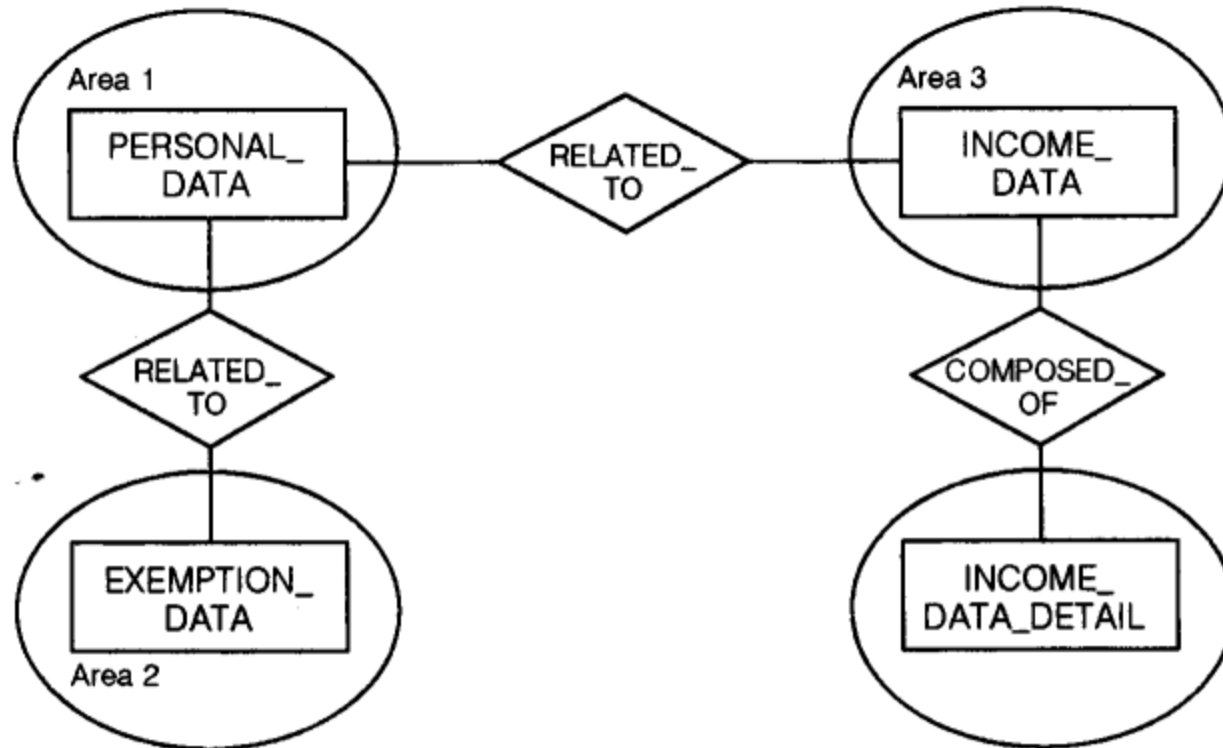將同一階層的概念抓下來當作 skeleton schema 的 entities.

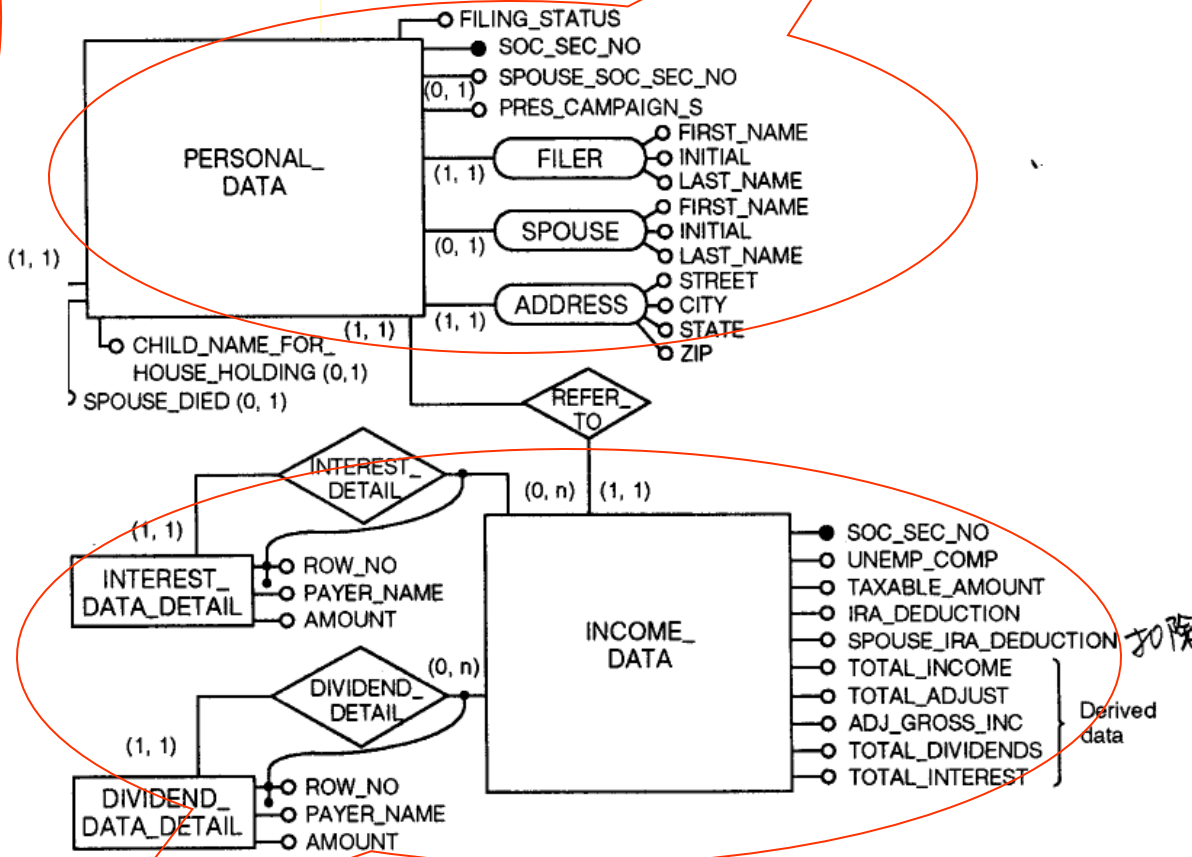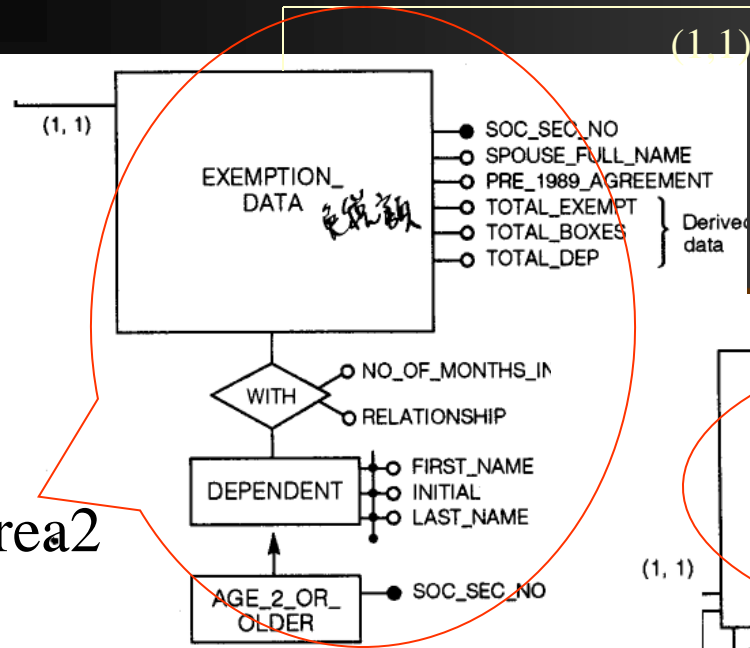Figure 4.13   Skeleton schema of the database for the Income Tax Return Form
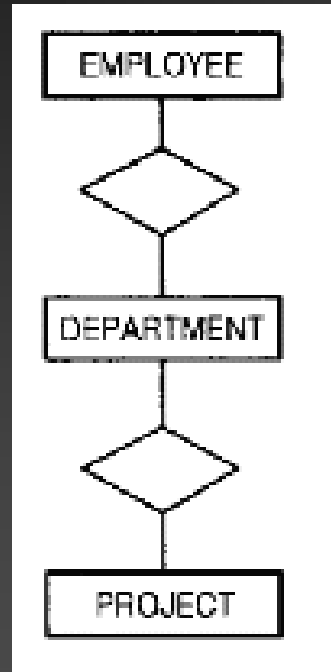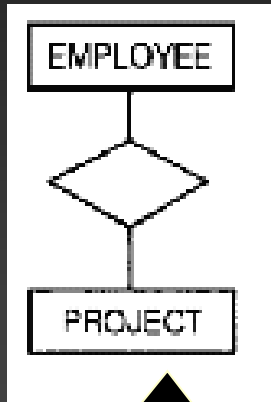
Example (cont.)

area1

area2

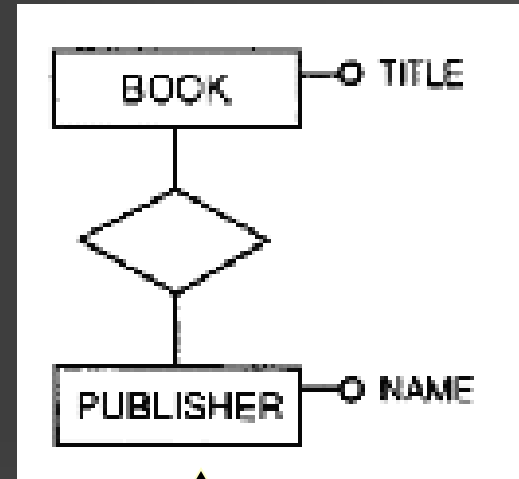area3

Final conceptual schema for Income Tax Return Form

View Integration

- 把數個 conceptual schema 整合為一個 global conceptual schema, 以代表整個應用系統的需求.
- 常用在大型資訊系統或分散式資料庫
- 整合時發生同樣觀念不同表示法或是同一事件但解釋不同的現象稱為衝突 conflict.
- 整合時必須先把衝突的概念一一列出, 將互相衝突的概念一致化, 再行整合.
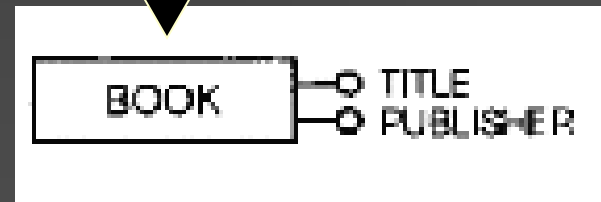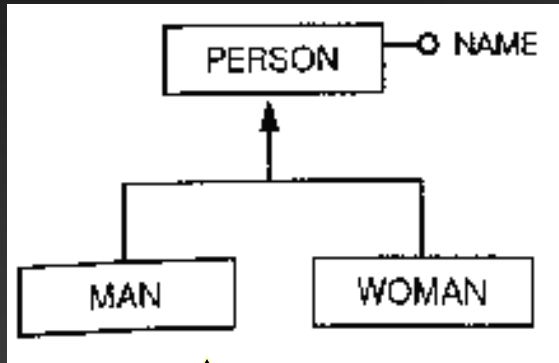- 大型系統整合時,
  - 同時整合數個 schema 或是
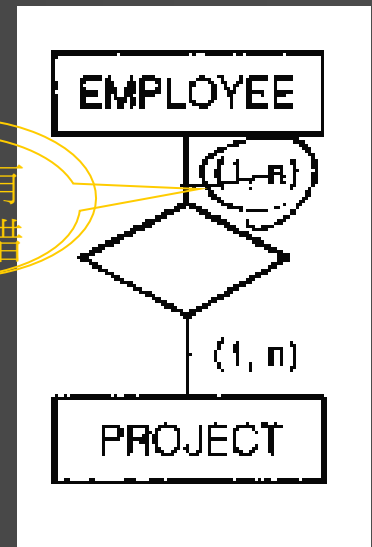 - 一次加入一個 schema
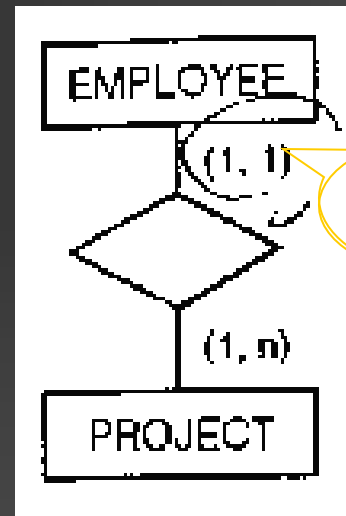
Conflicts

不同觀念

不同圖表達相同的概念

Conflicts

不同的圖表現相同的概念

一定有
一個錯

不相容的設計規格

Approach to View Integration

Schema 1   Schema 2

Conflict analysis

分別列出
schema1 和
schema2 概念
衝突的部分

Conflict resolution

分別將schema1 和
schema2 概念衝突
的部分一致化

Schema merging

整合結果
的 schema
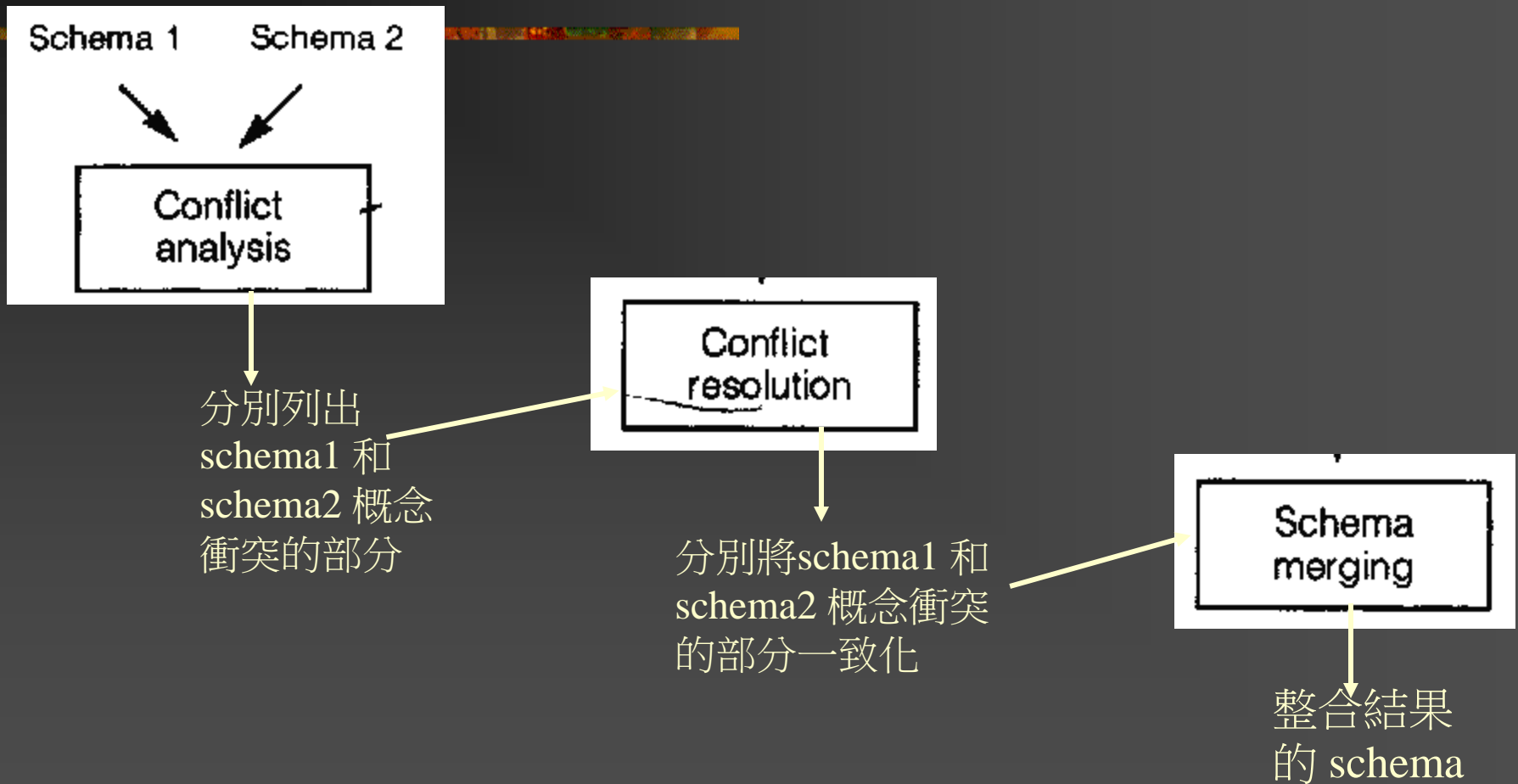
The most general approach to view integration

Initial schema

Intermediate design steps
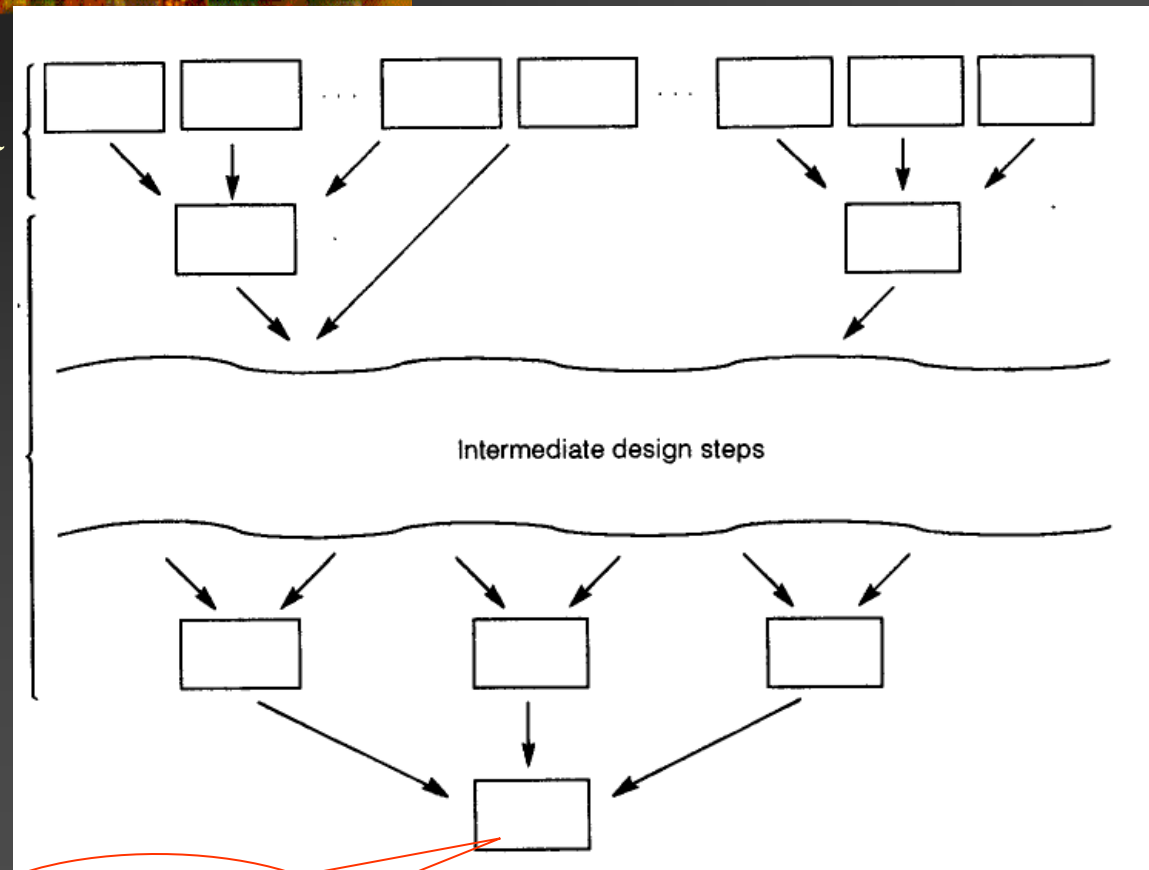
Final schema
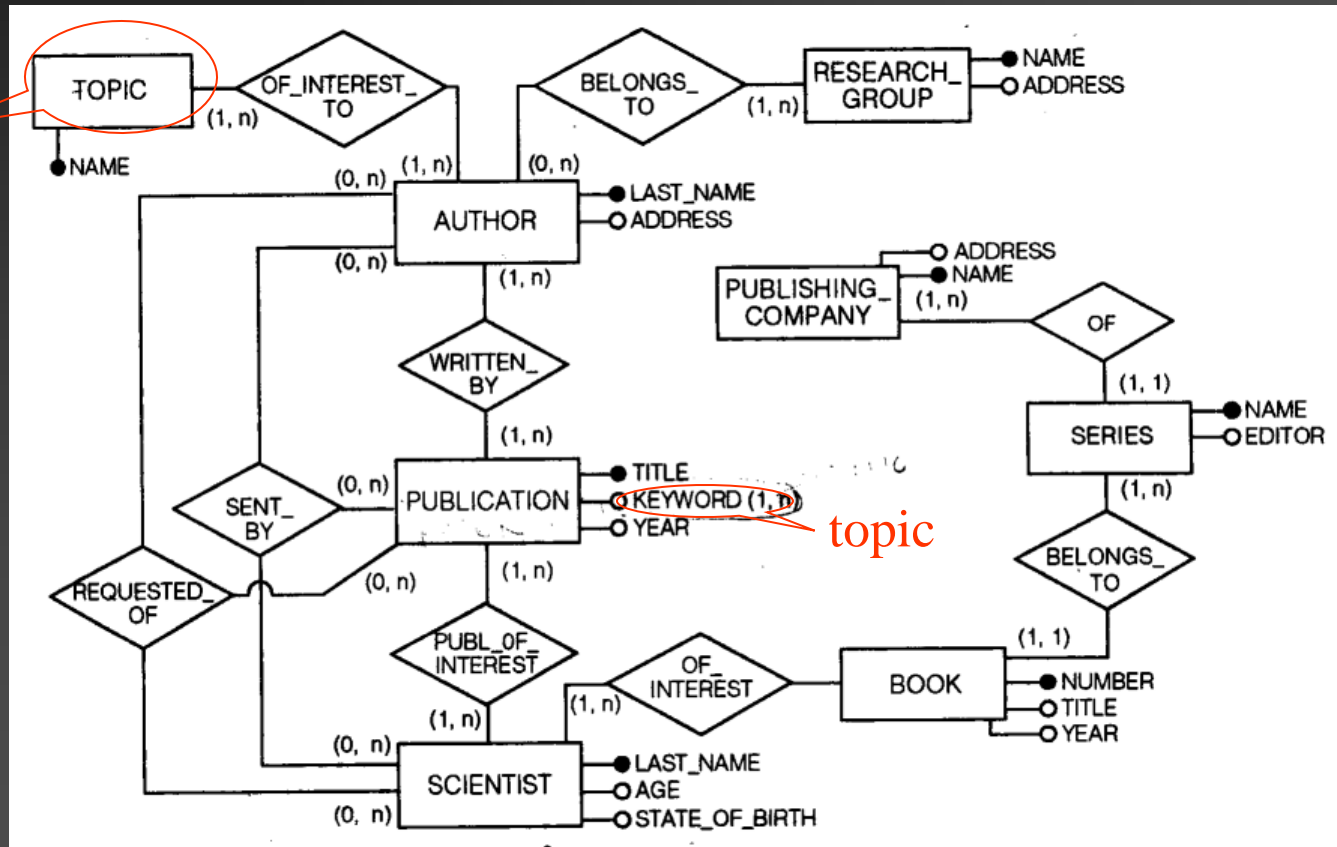
The suggested sequence of view-integration activities
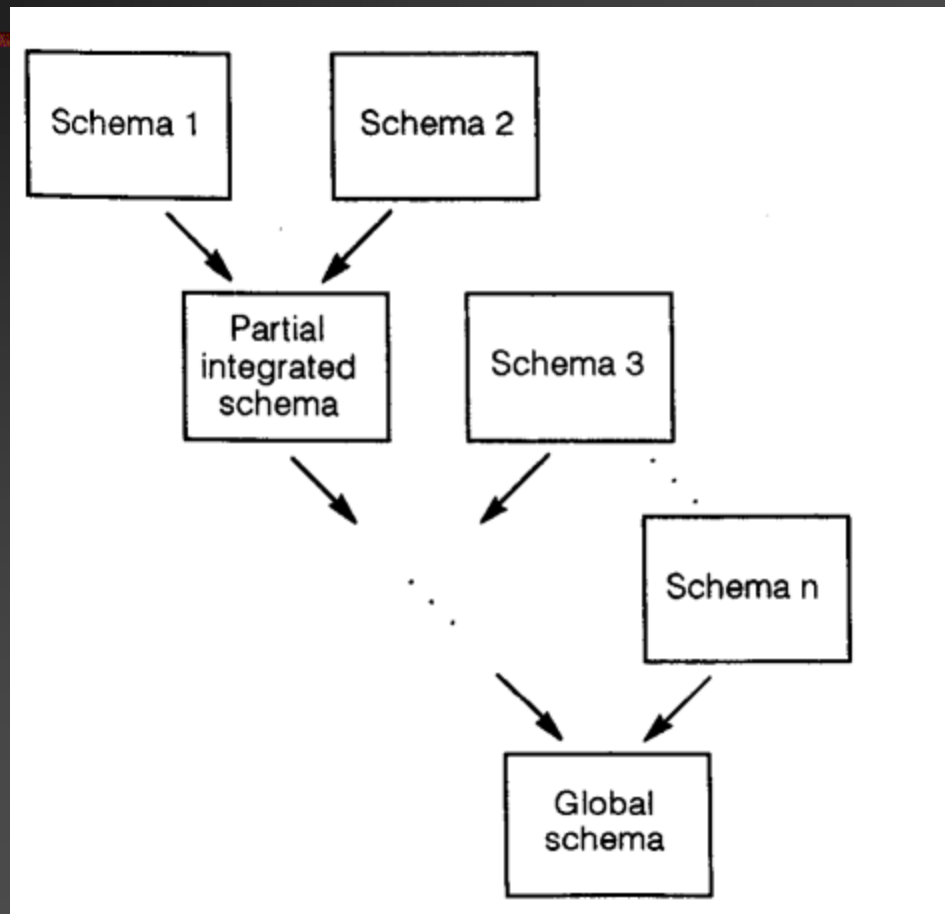
一次加入一個

Example of integrating two schemas

Scientist schema (Schema 1)

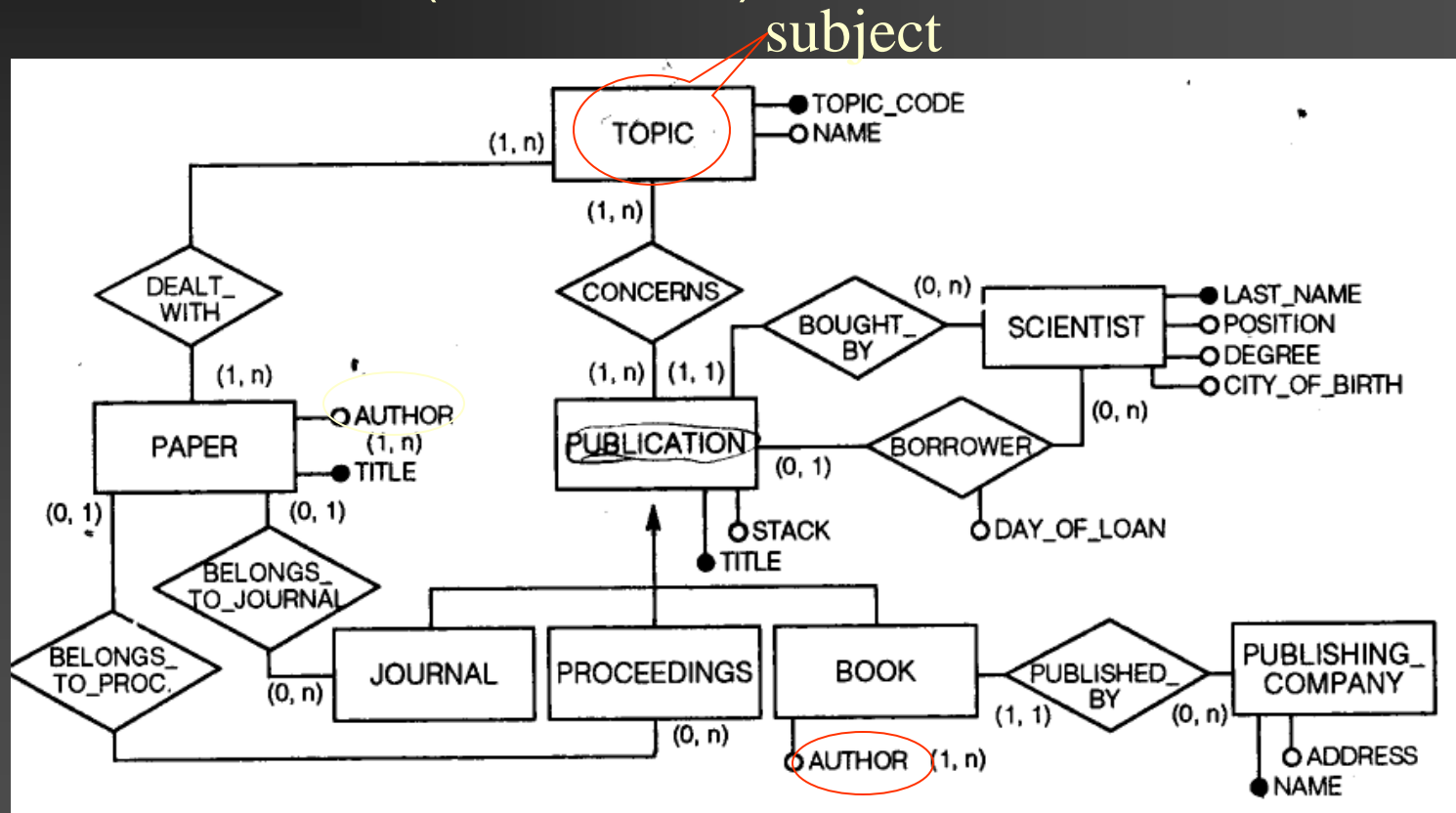# Example of integrating two schemas

Librarian schema (Schema 2)

Table 5.1   Concepts in Schema 1 (Scientist Schema)

Name	Description
Author	Authors of publications of interest to scientists
Publication	Publications kept by scientists in their private cabinets; they are usually directly obtained by scientists from authors
Topic	Research areas of interest to authors
Requested of	Connects papers that have been requested by some scientist to the author of whom the request has been made
Sent by	Connects papers that have been sent by authors to scientists who have requested them

Table 5.2   Concepts in Schema 2 (Librarian Schema)

Name	Description
Publication	Publications presently kept in the library
Paper	Papers published in journals or proceedings kept in the library
Topic	Topics of papers
Bought by	Indicates which scientist is responsible for the grant used to purchase the publication

Conflict analysis and resolution
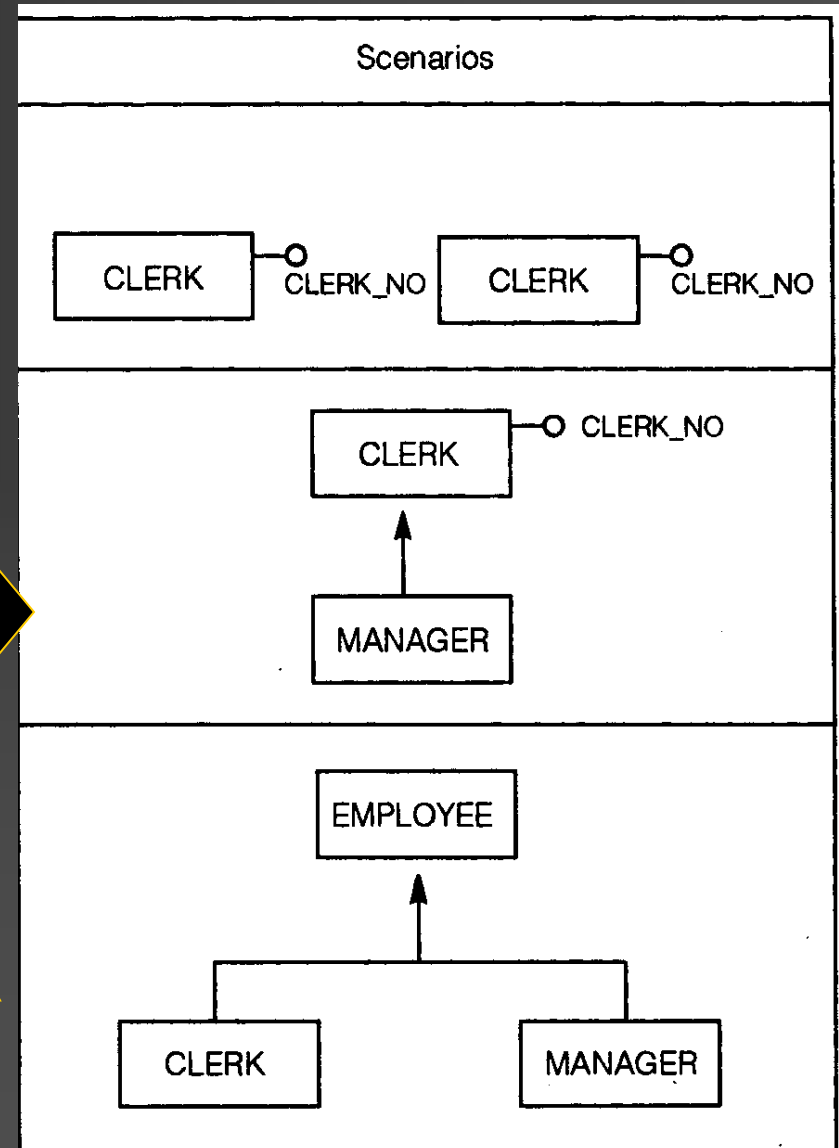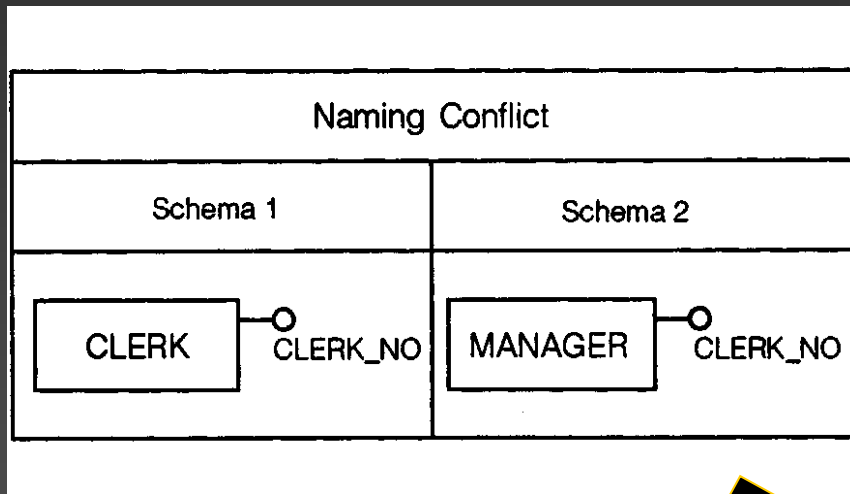
- **Name conflict analysis**
  - 找出同名異義或異名同義者
  - 找出在不同的 schema 中不同概念間有互相限制的關係存在 (inter-schema property)
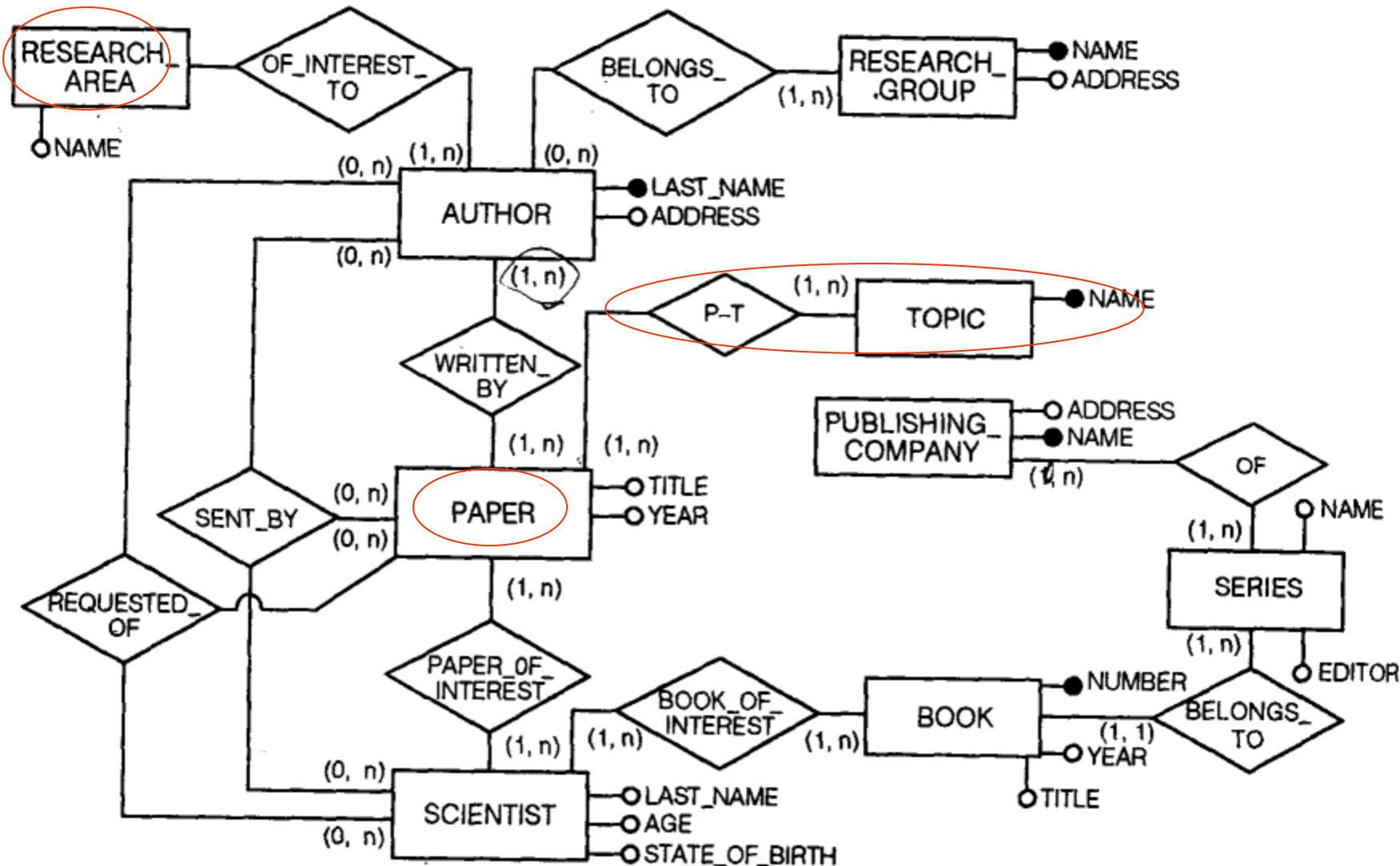- **Structural conflict analysis**
  - 相同的概念, 用不同的schema 表達
    - 例: AUTHOR 在 schema 1 是 entity
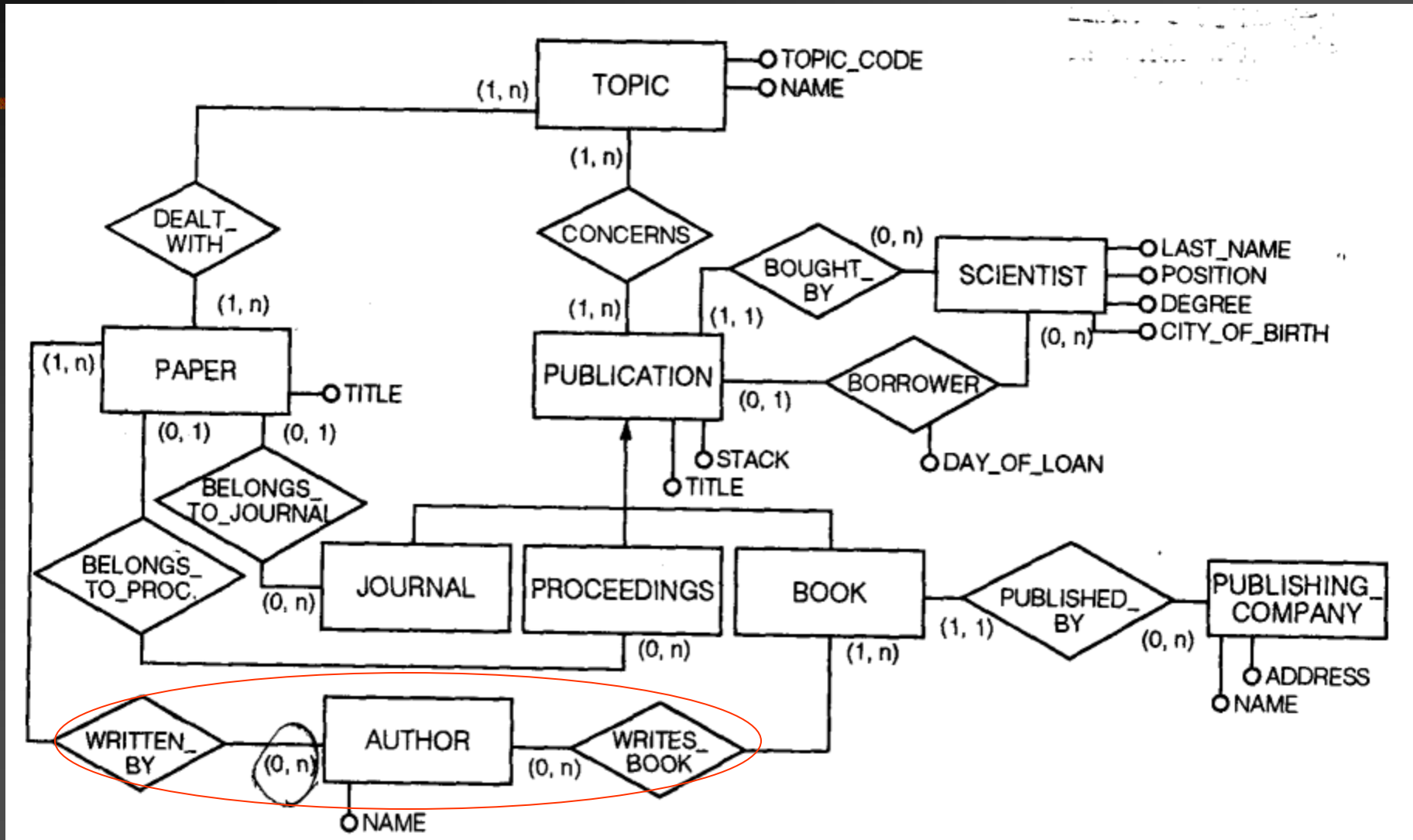   - 在 schema 2 是 attribute

Conflict analysis and resolution

Three possible modification for a name conflict.

Modified Scientist Schema

Modified Librarian Schema

Merging of views

- 把完全相同的 entity直接放入 result schema.
    - 例: PUBLISHING-COMPANY
- 相同的 entity, 但不同的 attributes, 將之聯集起來, 找出 identifier.
    - 例: SCIENTIST
- 若有 generalization 表現出 inter-schema properties, 加入 result schema.
    - 例: TOPIC and RESEARCH_AREA 用 CONCERNS 相連
- 當全部的 entities 皆放入 result schema 之後, 加入 relationships & generalization.
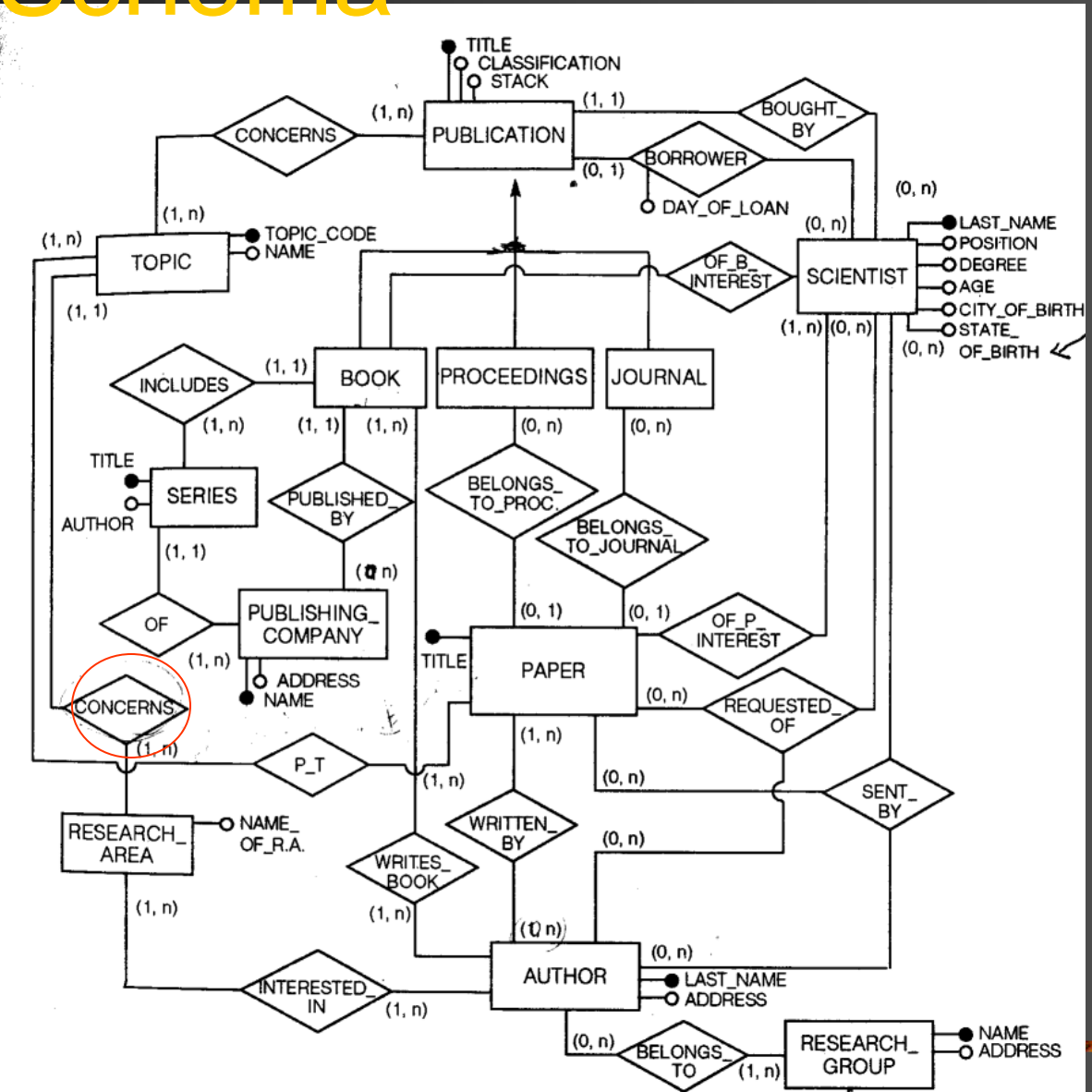
Resulting Schema

Figure 5.8 Global schema after merging

Exercise 1

Build a conceptual schema for the following natural language description.

Design a database system for managing information about routes supported by a bus company. Each route served by the company has a starting place and an ending place, but it can go through several intermediate stops. The company is distributed over several branches. Not all the cities where the buses stop have a branch; however, each branch must be at a city located along the bus routes. There can be multiple branches in the same city and also multiple stops in the same city. One bus is assigned by the company to one route; some routes can have multiple buses. Each bus has a driver and an assistant, who are assigned to the bus for the day.

Exercise 2

Build a conceptual schema for the following natural language description.

Design the database for the administration and reservation office of a bus company. Each passenger can book a seat on a given portion of the routes served by each bus; routes have a starting place, an ending place, and several intermediate places. Passengers can specify whether they want to be in the smoking or nonsmoking section. Some passengers can get in the bus even if they do not have a reservation, when some seats are left empty. With each reservation, the last name, initials, and telephone number of the passenger is stored. Sometimes, trips are not made because of bad weather conditions; in this case, passengers holding reservations are notified. At the end of the trip, the driver's assistant reports to the company the total amount of tickets purchased on the bus by passengers and reports this amount to the administrative office of the branch at the route's destination.

Exercise 3

Build a conceptual schema that contains all data mentioned in the following form.
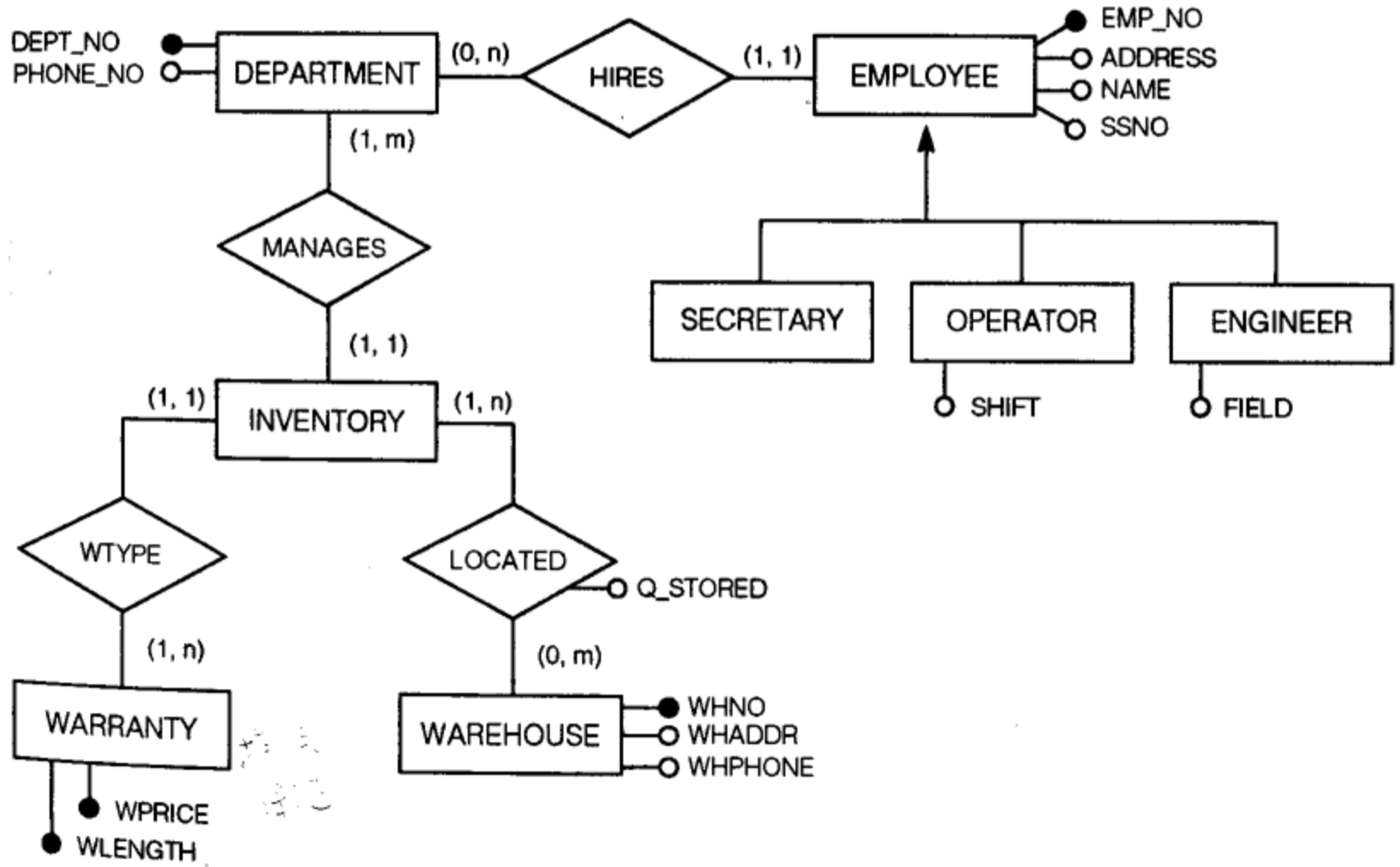
Exercise 4

4. Integrate the two schemas, which represent sales in a company and the structure of its departments and personnel, producing a single schema.
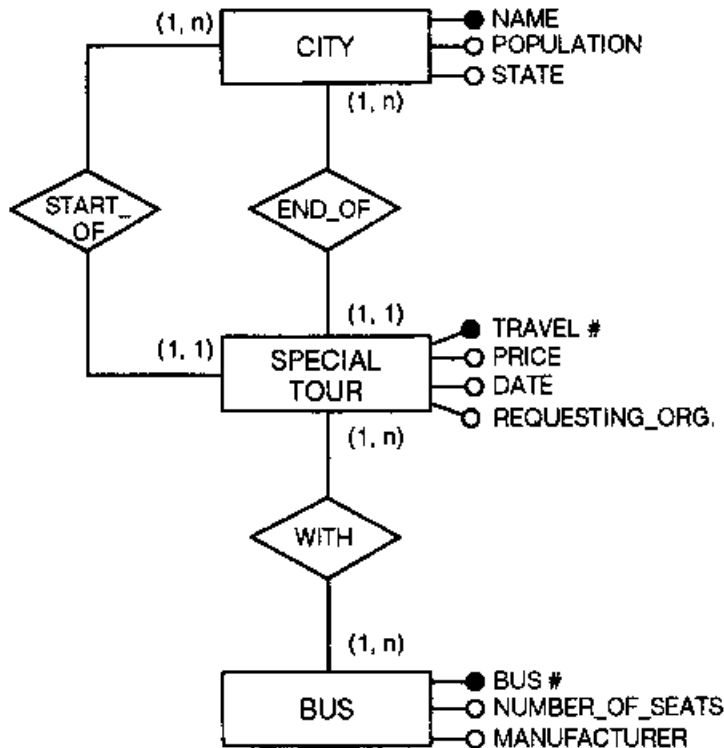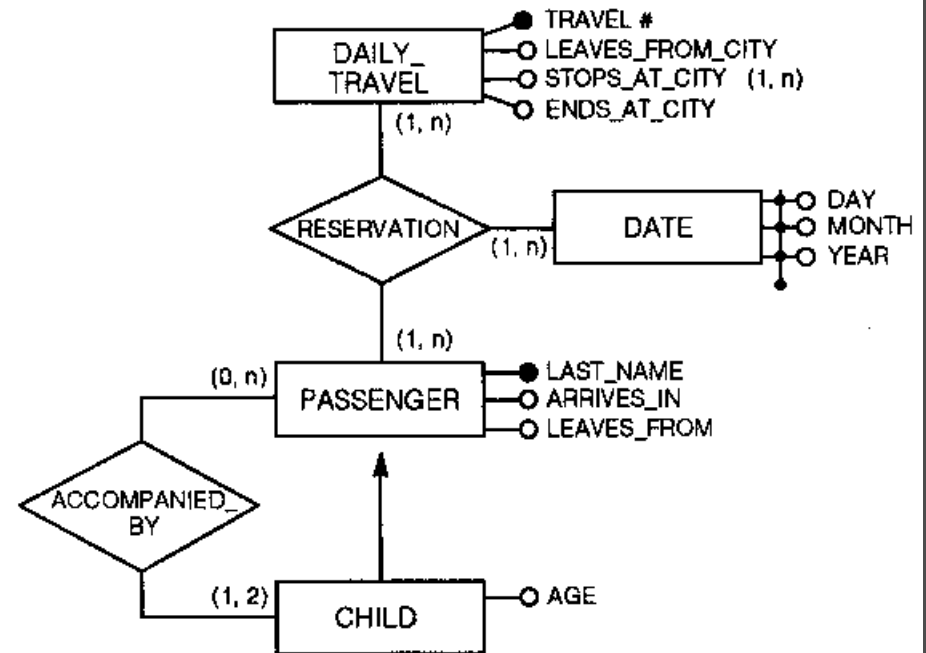

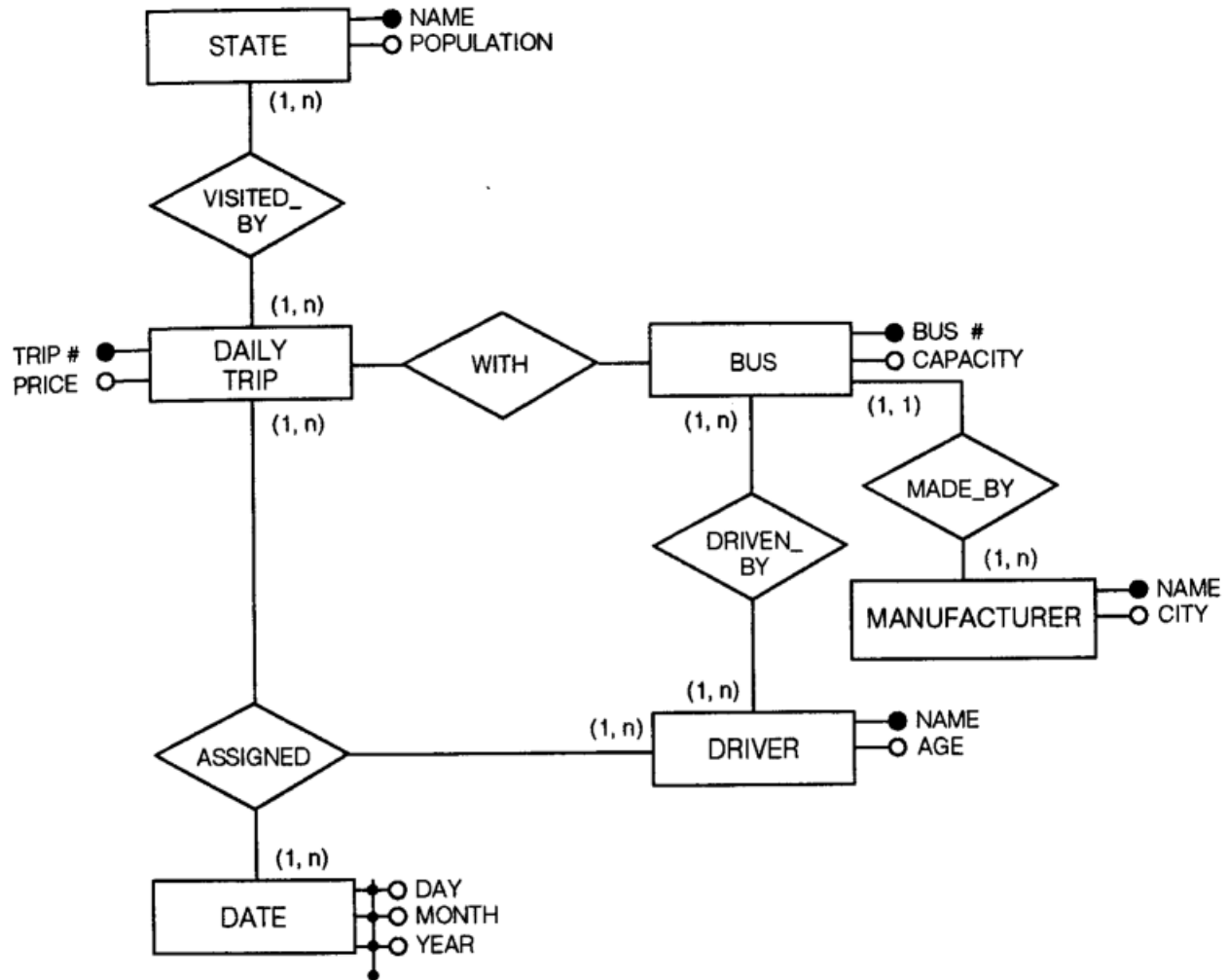(a) First schema

(b) Second schema

Unit 6
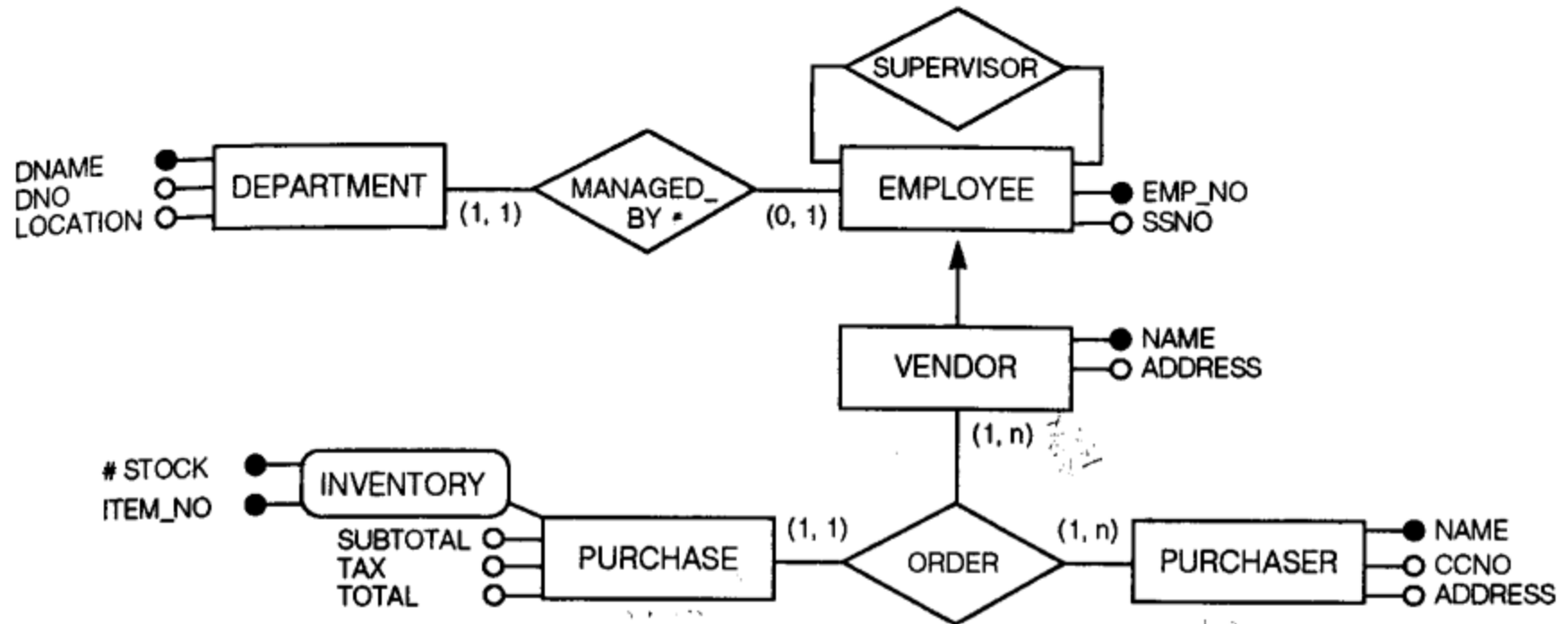
Exercise 5

Integrate the following three schemas which represent special tours, daily travels, and reservations for daily travels, producing a single schema.

(a) First schema

(b) Second schema

(c) Third schema

Figure 5.10 (cont'd)   Trip database