Entity Relationship (E-R) Modeling

In this chapter, you will learn:

- What a conceptual model is and what its purpose is
- The difference between internal and external models
- How internal and external models serve the database design process
- How relationships between entities are defined and refined, and how such relationships are incorporated into the database design process
- How ERD components affect database design and implementation

Basic Modeling Concepts

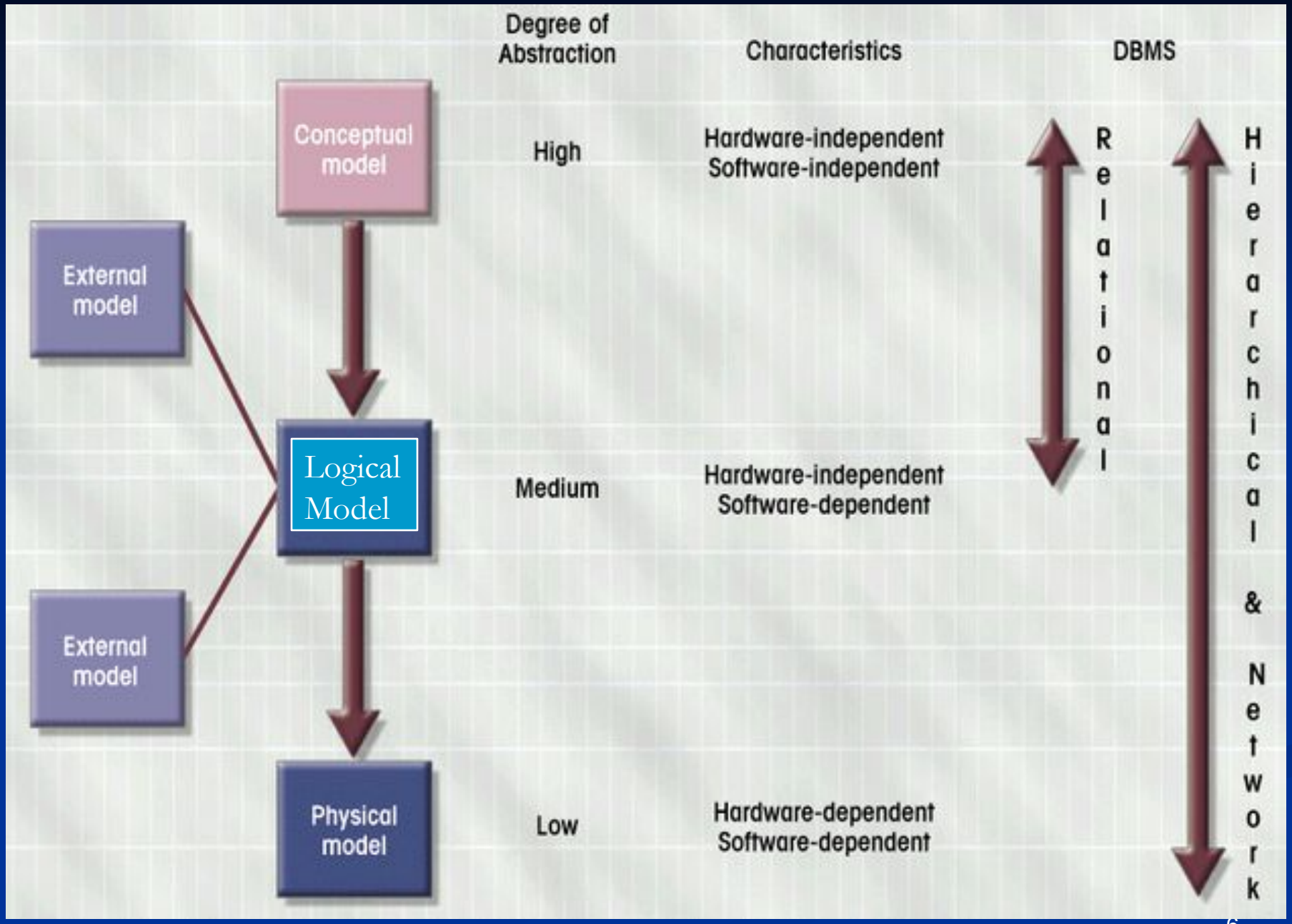
- Art and science
- Good judgment coupled with powerful design tools
- Models
  - "Description or analogy used to visualize something that cannot be directly observed" *Webster's Dictionary*

Basic Modeling Concepts

- Data Model
  - Relatively simple representation of complex real-world data structures
  - Basic tools for database design
  - Communication
  - Good database design starts with a good design of data model.

Data Models: Degrees of Data Abstraction

- Three different models (according to the degree of abstract)
  - **Conceptual model**
    - Global view of data
    - Basis for identification and description of main data items
  - **Logical model**
    - Representation of database as seen by DBMS
    - Adapts conceptual model to specific DBMS
  - **Physical model**
    - Lowest level of abstraction
    - Describe the ways data being stored


Degree of Abstraction
Characteristics
DBMS
Conceptual model
High
Hardware-independent
Software-independent
External model
Logical Model
Medium
Hardware-independent
Software-dependent
External model
Physical model
Low
Hardware-dependent
Software-dependent
Relational
Hierarchical & Network

The Entity Relationship (E-R) Model

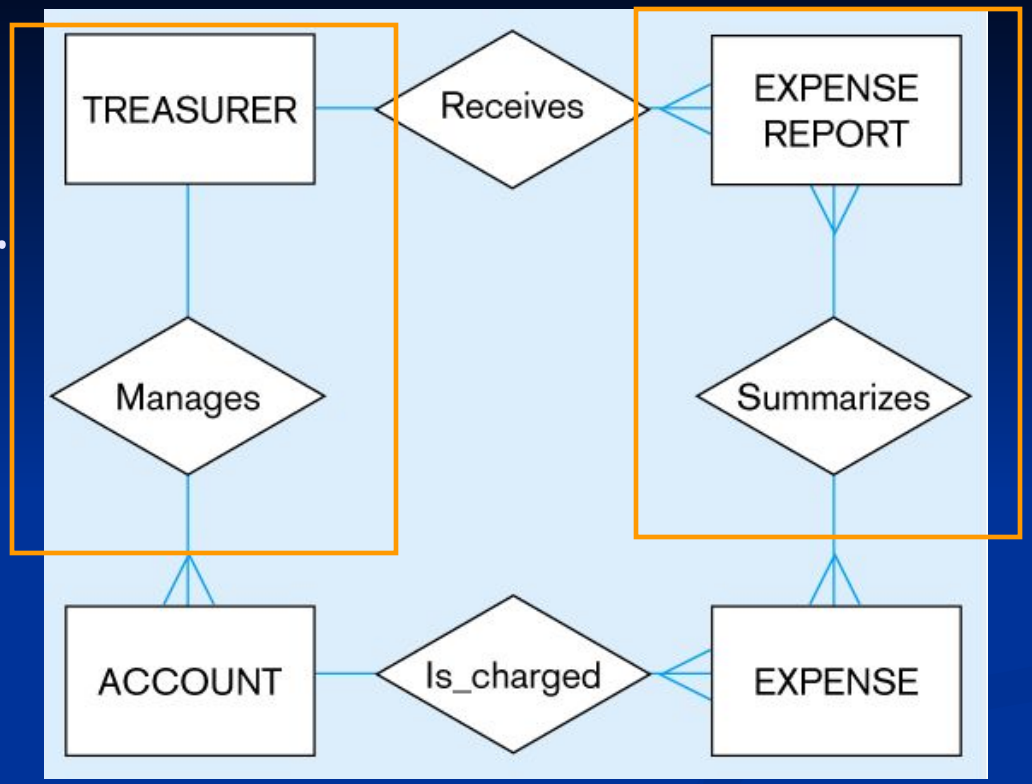
- Represents conceptual view
- Main Components
  - Entities
    - Corresponds to entire table, not row
    - Represented by rectangle
  - Attributes
  - Relationships

What Should an Entity Be?

- SHOULD BE:
  - An object that will have many instances in the database
  - An object that will be composed of multiple attributes
  - An object that we are trying to model
- SHOULD NOT BE:
  - A user of the database system
  - An output of the database system (e.g. a report)

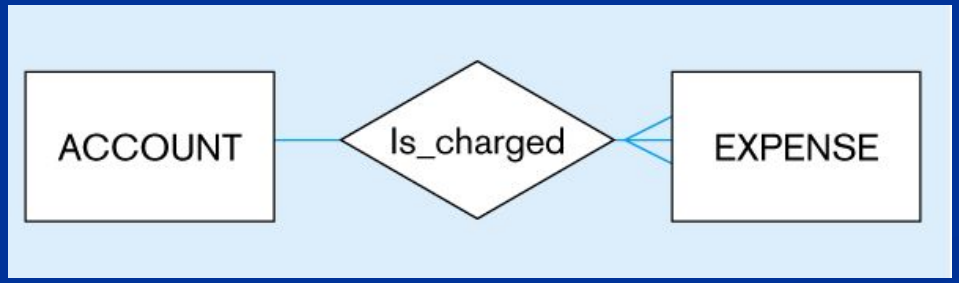
Figure 3-4

Inappropriate entities



System user

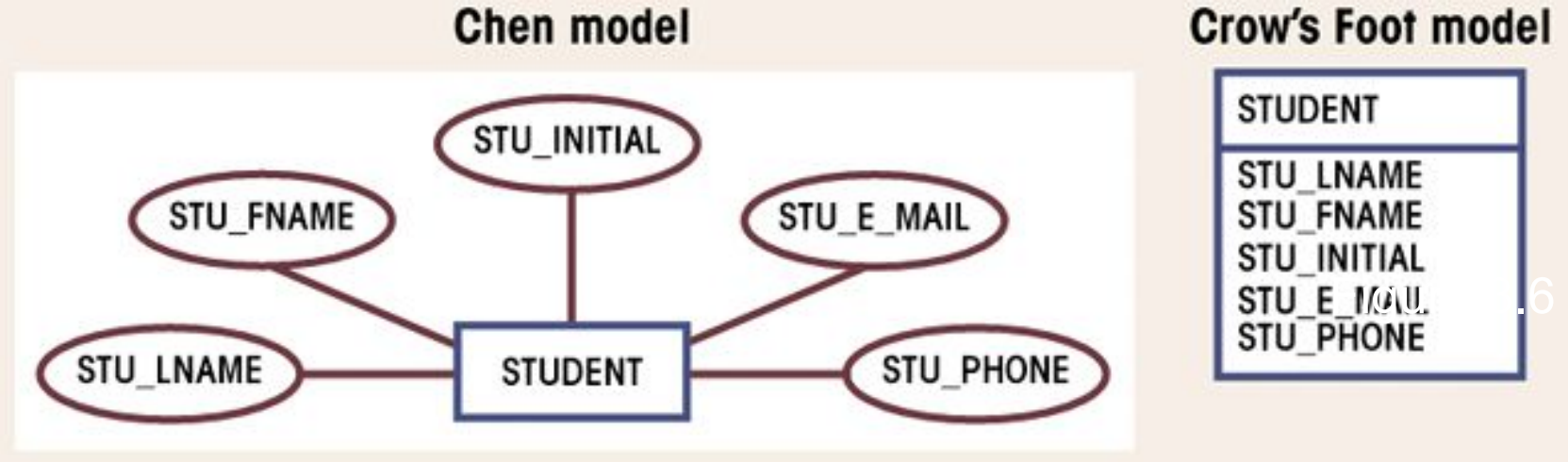
System output



Appropriate entities

Attributes

- Characteristics of entities
- Domain is set of possible values
- Primary keys underlined



	CLASS_CODE	CRS_CODE	CLASS_SECTION	CLASS_TIME	ROOM_CODE	PROF_NUM
▶	10012	ACCT-211	1	MWF 8:00-8:50 a.m.	BUS311	105
	10013	ACCT-211	2	MWF 9:00-9:50 a.m.	BUS200	105
	10014	ACCT-211	3	TTh 2:30-3:45 p.m.	BUS252	342
	10015	ACCT-212	1	MWF 10:00-10:50 a.m.	BUS311	301
	10016	ACCT-212	2	Th 6:00-8:40 p.m.	BUS252	301
	10017	CIS-220	1	MWF 9:00-9:50 a.m.	KLR209	228
	10018	CIS-220	2	MWF 9:00-9:50 a.m.	KLR211	114
	10019	CIS-220	3	MWF 10:00-10:50 a.m.	KLR209	228
	10020	CIS-420	1	W 6:00-8:40 p.m.	KLR209	162
	10021	QM-261	1	MWF 8:00-8:50 a.m.	KLR200	114
	10022	QM-261	2	TTh 1:00-2:15 p.m.	KLR200	114
	10023	QM-362	1	MWF 11:00-11:50 a.m.	KLR200	162
	10024	QM-362	2	TTh 2:30-3:45 p.m.	KLR200	162

FIGURE 3.7 THE CLASS TABLE (ENTITY) COMPONENTS AND CONTENTS

- Simple
  - Cannot be subdivided
  - Age, sex, marital status
- Composite (avoided)
  - Can be subdivided into additional attributes
  - Address into street, city, zip
- Single-valued
  - Can have only a single value
  - Person has one social security number
- Multi-valued (avoid) in RDBMS
  - Can have many values
  - Person may have several college degrees
  - How to avoid? (see next slide)
- Derived
  - Can be derived with algorithm (Age can be derived from date of birth)
  - Need to be stored

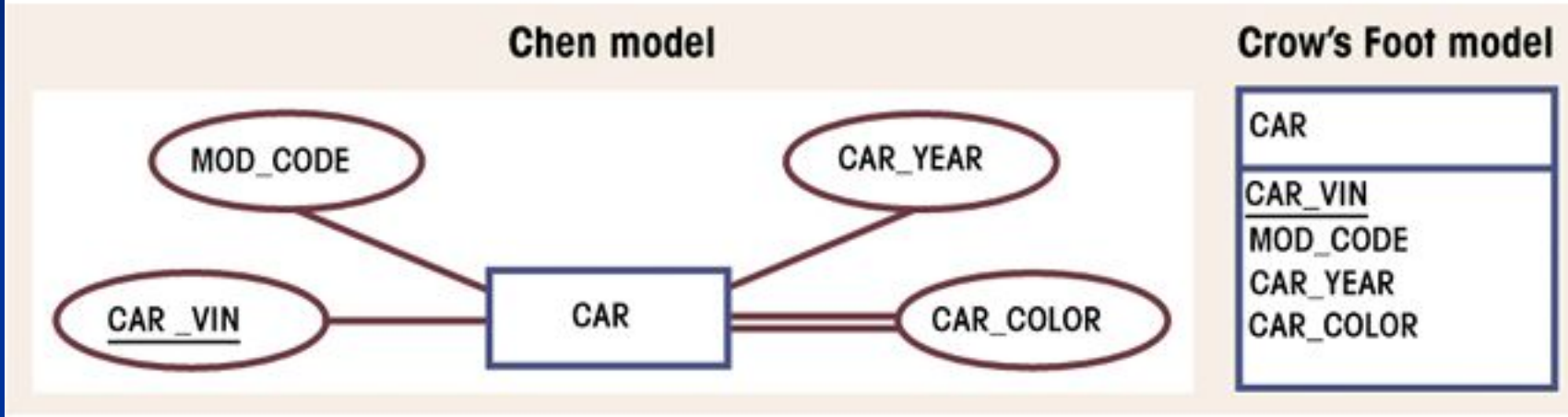


FIGURE 3.8 A MULTIVALUED ATTRIBUTE IN AN ENTITY

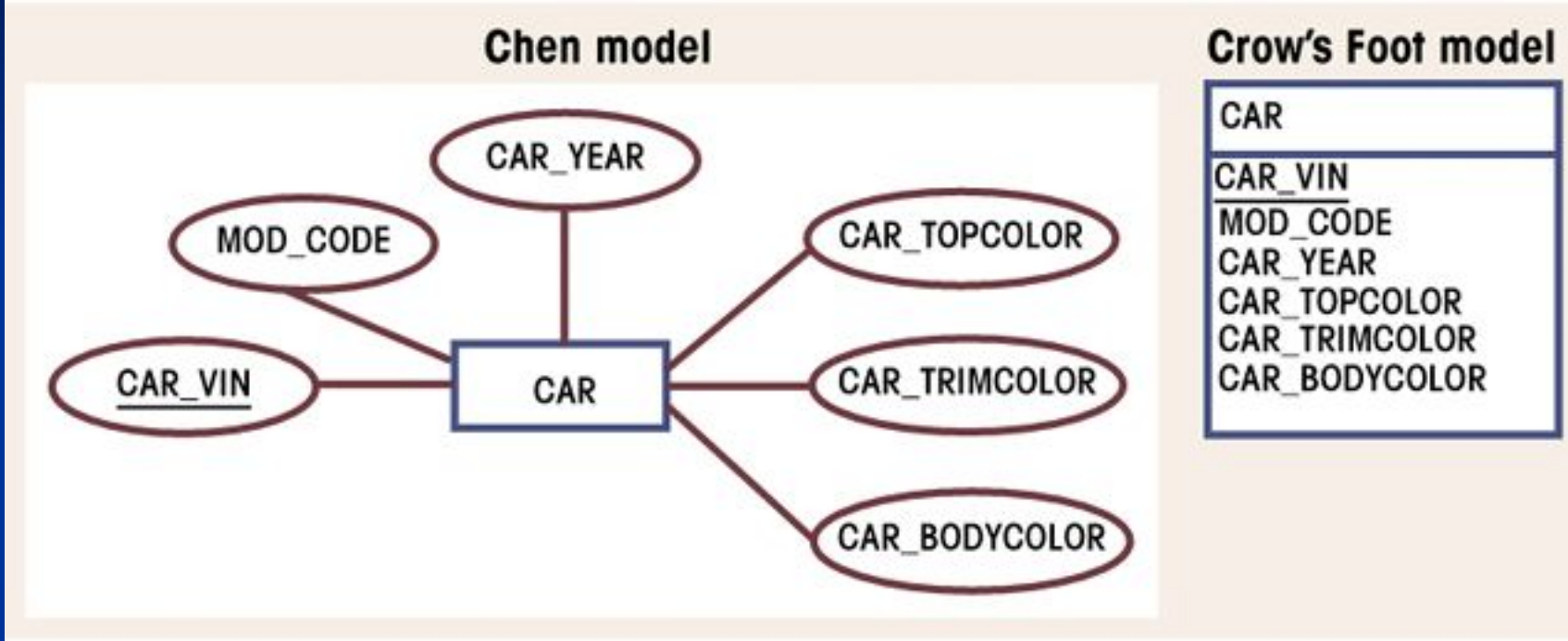


FIGURE 3.9 SPLITTING THE MULTIVALUED ATTRIBUTE INTO NEW ATTRIBUTES

TABLE 3.1 COMPONENTS OF THE MULTIVALUED ATTRIBUTE

SECTION	COLOR
Top	White
Body	Blue
Trim	Green
Interior	Blue

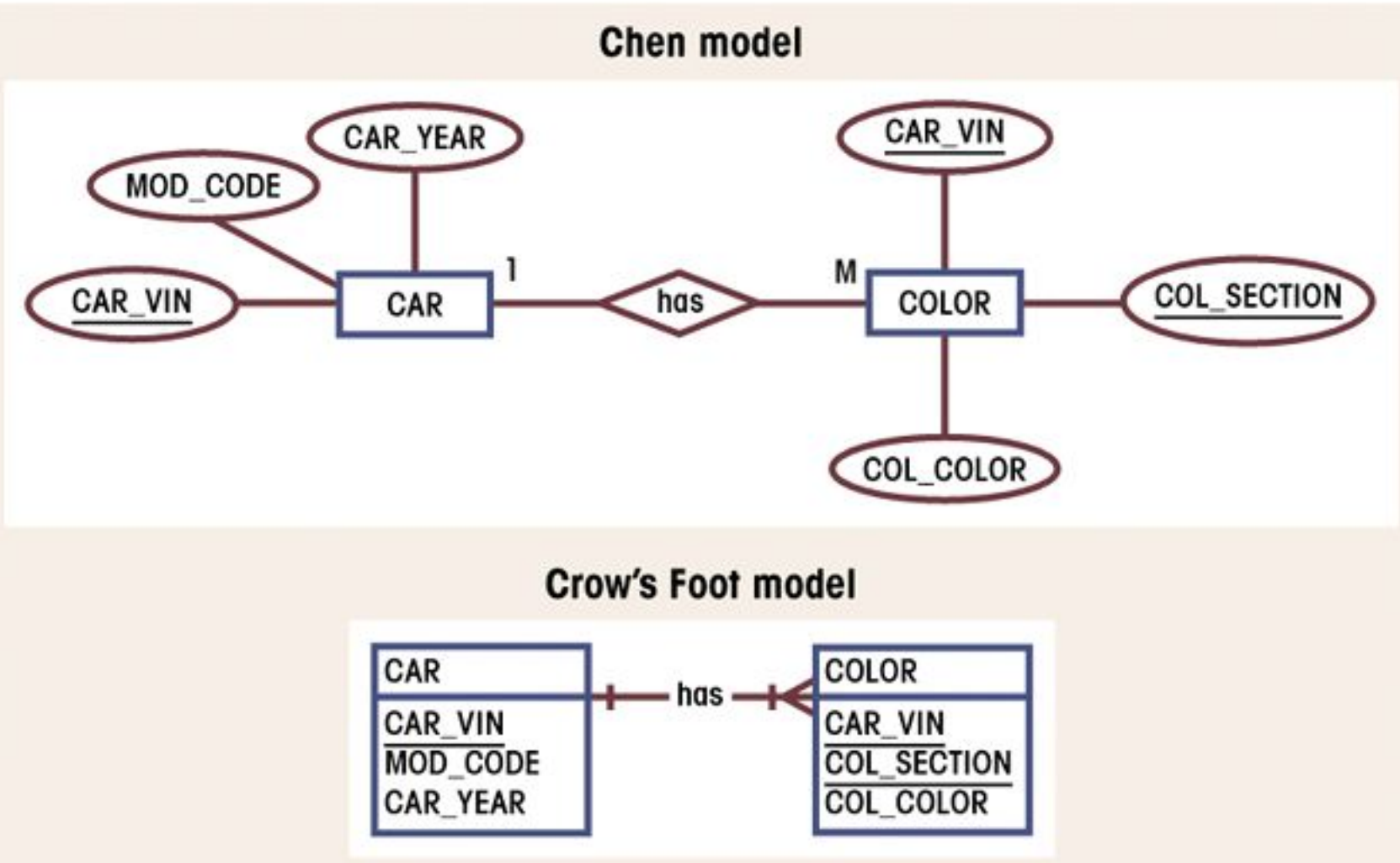


FIGURE 3.10 A NEW ENTITY SET COMPOSED OF A MULTIVALUED ATTRIBUTE'S COMPONENTS

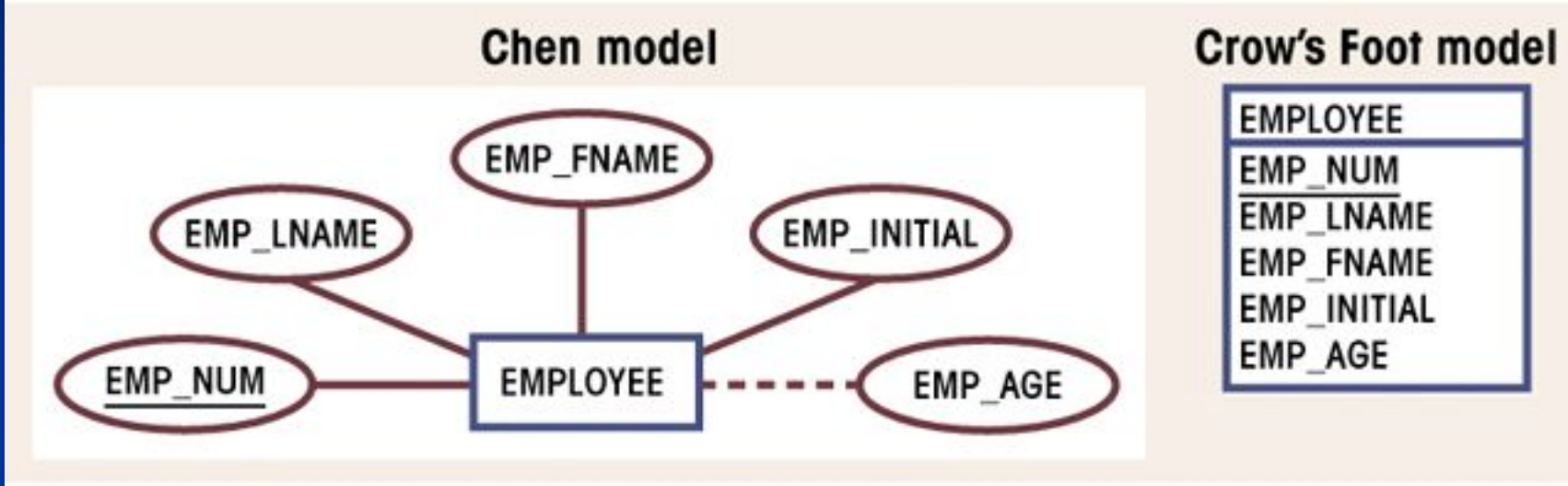
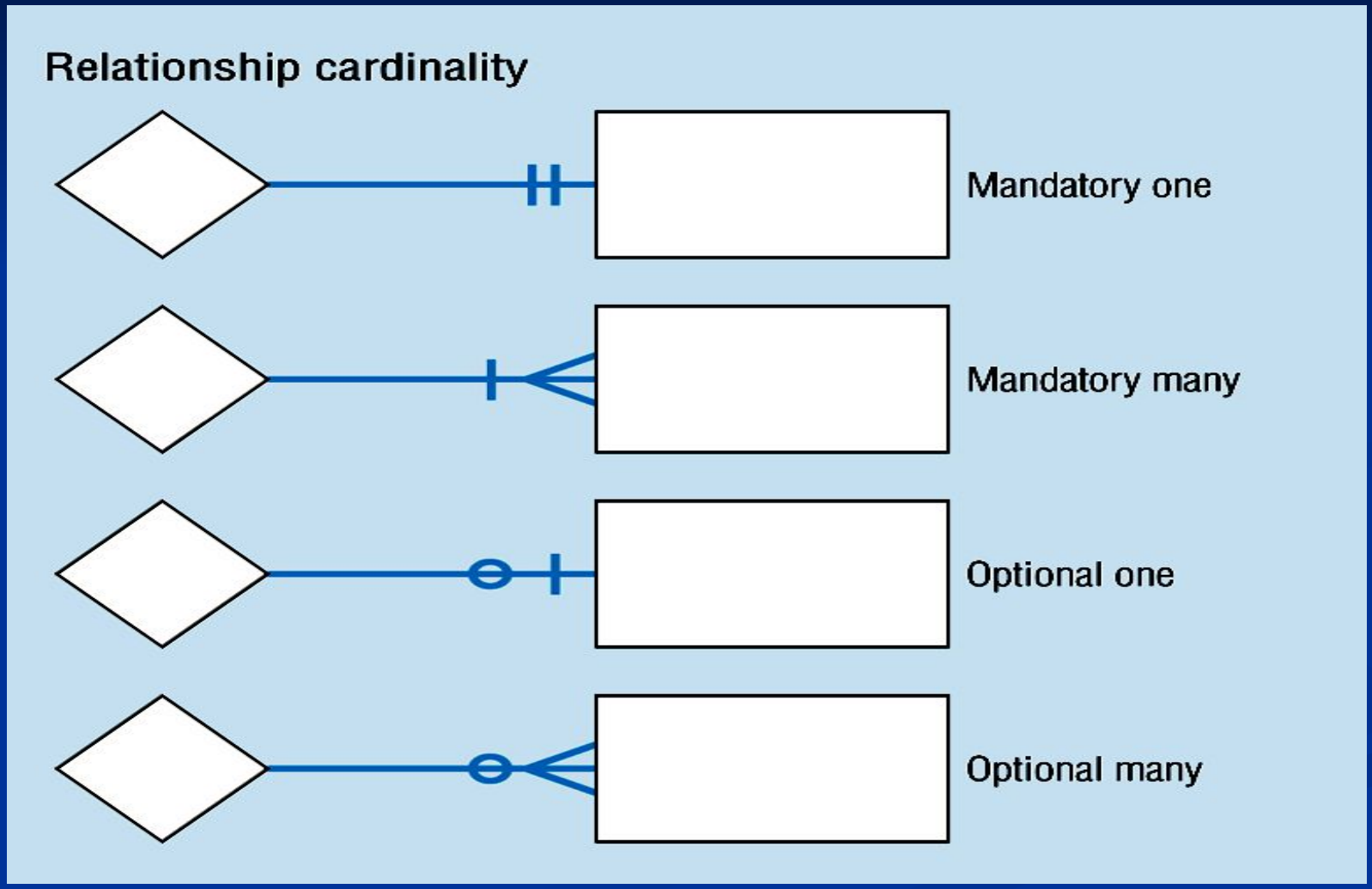


FIGURE 3.11 DEPICTION OF A DERIVED ATTRIBUTE

Relationships

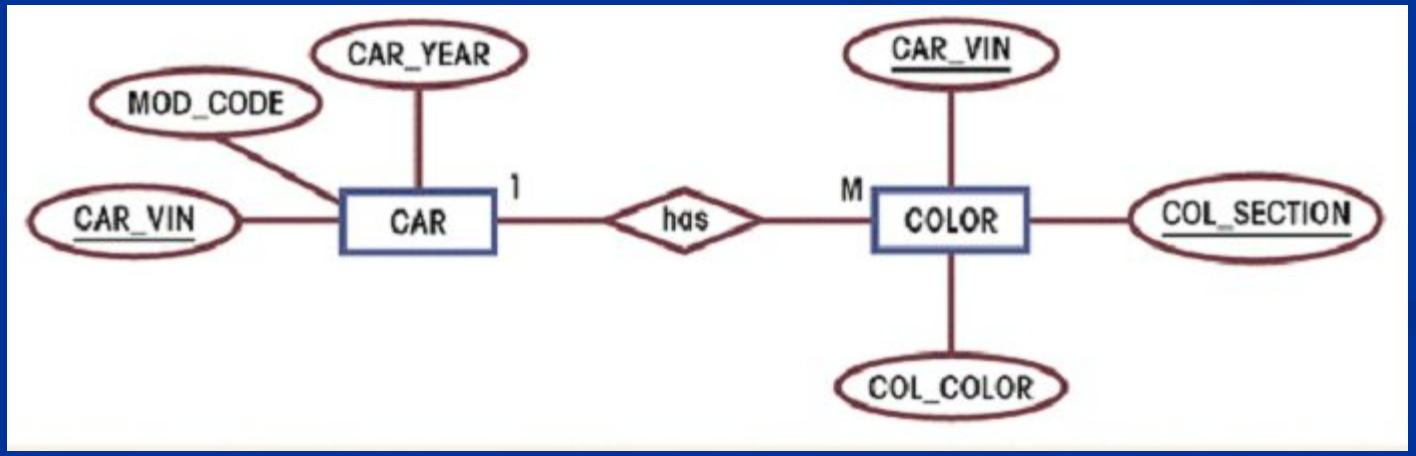
- Association between entities
- Connected entities are called participants
- Operate in both directions
- Connectivity describes relationship classification
  - 1:1, 1:M, M:N
- Cardinality
  - Expresses number of entity occurrences associated with one occurrence of related entity

Relationship Participation



Relationship Strength

- Existence dependence
  - Entity's existence depends on existence of related entities
  - Existence-independent entities can exist apart from related entities
  - Example:



Relationship Strength

- Weak (non-identifying)
- Strong (identifying)

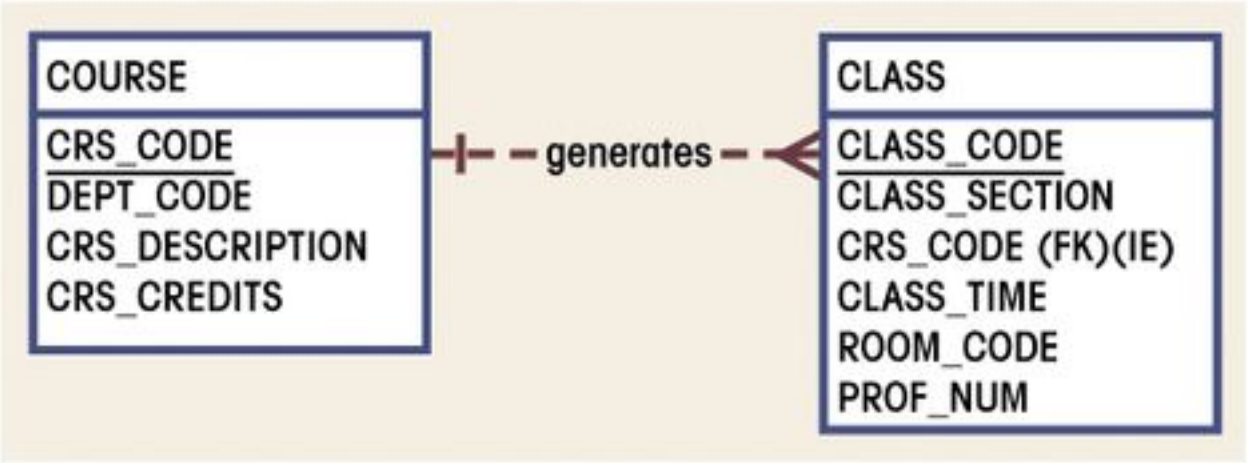


FIGURE 3.13 A WEAK (NON-IDENTIFYING) RELATIONSHIP BETWEEN COURSE AND CLASS

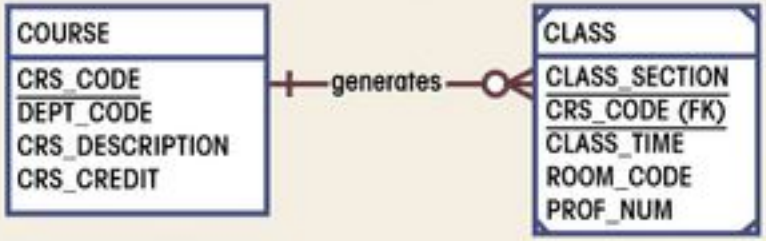


Table name: COURSE

Database: CH3_TEXT

CRS_CODE	DEPT_CODE	CRS_DESCRIPTION	CRS_CREDIT
ACCT-211	ACCT	Accounting I	3
ACCT-212	ACCT	Accounting II	3
CIS-220	CIS	Intro. to Microcomputing	3
CIS-420	CIS	Database Design and Implementation	4
MATH-243	MATH	Mathematics for Managers	3
QM-261	CIS	Intro. to Statistics	3
QM-362	CIS	Statistical Applications	4

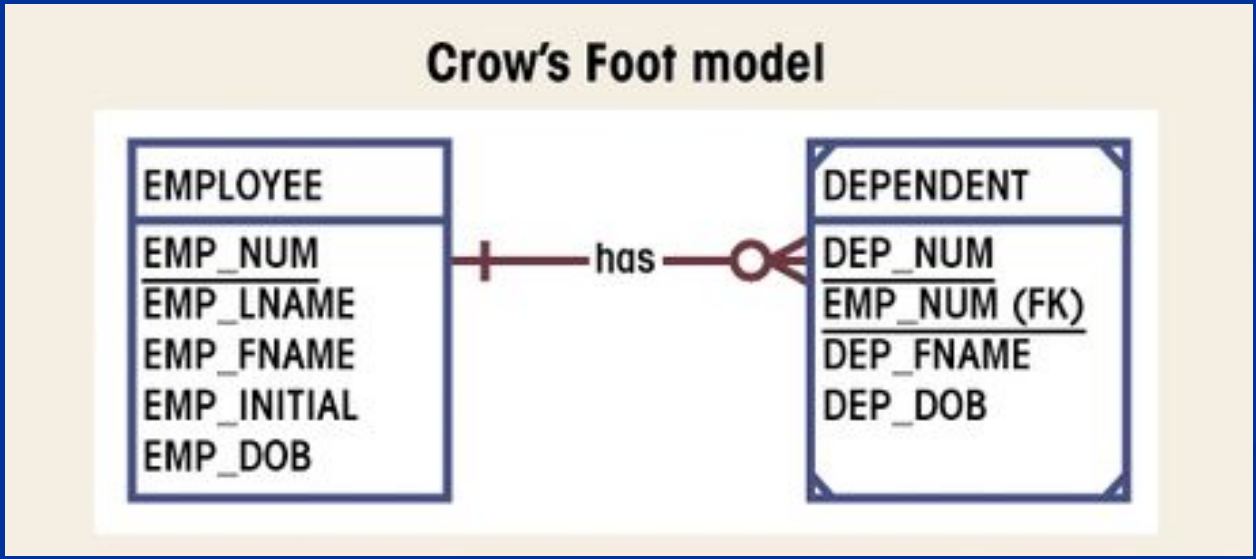
Table name: CLASS_FIG3_15

CRS_CODE	CLASS_SECTION	CLASS_TIME	ROOM_CODE	PROF_NUM
ACCT-211	1	MWF 8:00-8:50 a.m.	BUS311	105
ACCT-211	2	MWF 9:00-9:50 a.m.	BUS200	105
ACCT-211	3	TTh 2:30-3:45 p.m.	BUS252	342
ACCT-212	1	MWF 10:00-10:50 a.m.	BUS311	301
ACCT-212	2	Th 6:00-8:40 p.m.	BUS252	301
CIS-220	1	MWF 9:00-9:50 a.m.	KLR209	228
CIS-220	2	MWF 9:00-9:50 a.m.	KLR211	114
CIS-220	3	MWF 10:00-10:50 a.m.	KLR209	228
CIS-420	1	W 6:00-8:40 p.m.	KLR209	162
QM-261	1	MWF 8:00-8:50 a.m.	KLR200	114
QM-261	2	TTh 1:00-2:15 p.m.	KLR200	114
QM-362	1	MWF 11:00-11:50 a.m.	KLR200	162
QM-362	2	TTh 2:30-3:45 p.m.	KLR200	162

FIGURE 3.15 A STRONG (IDENTIFYING) RELATIONSHIP BETWEEN COURSE AND CLASS

Weak Entity

- Existence-dependent on another entity
- Has primary key that is partially or totally derived from parent entity



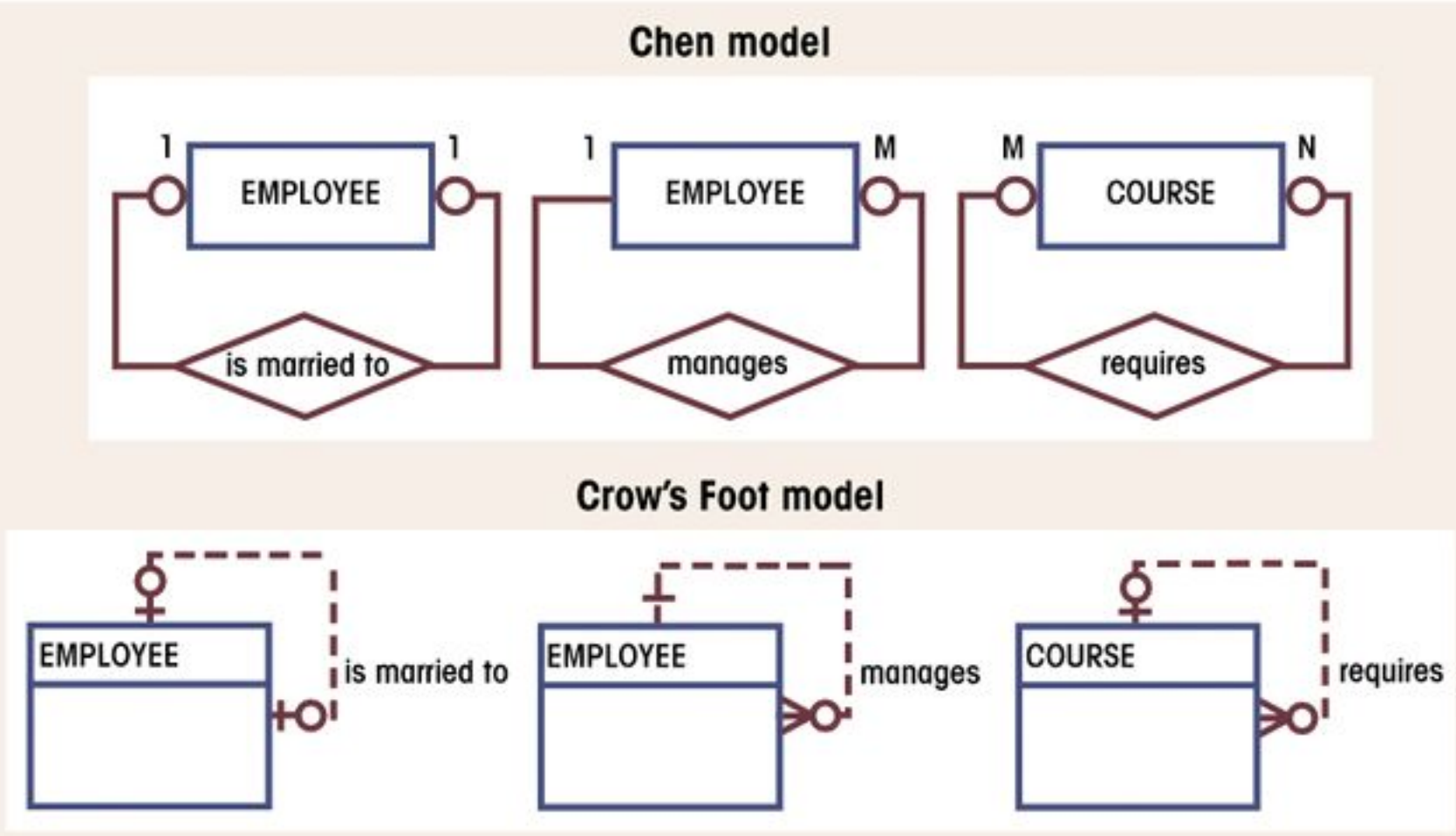
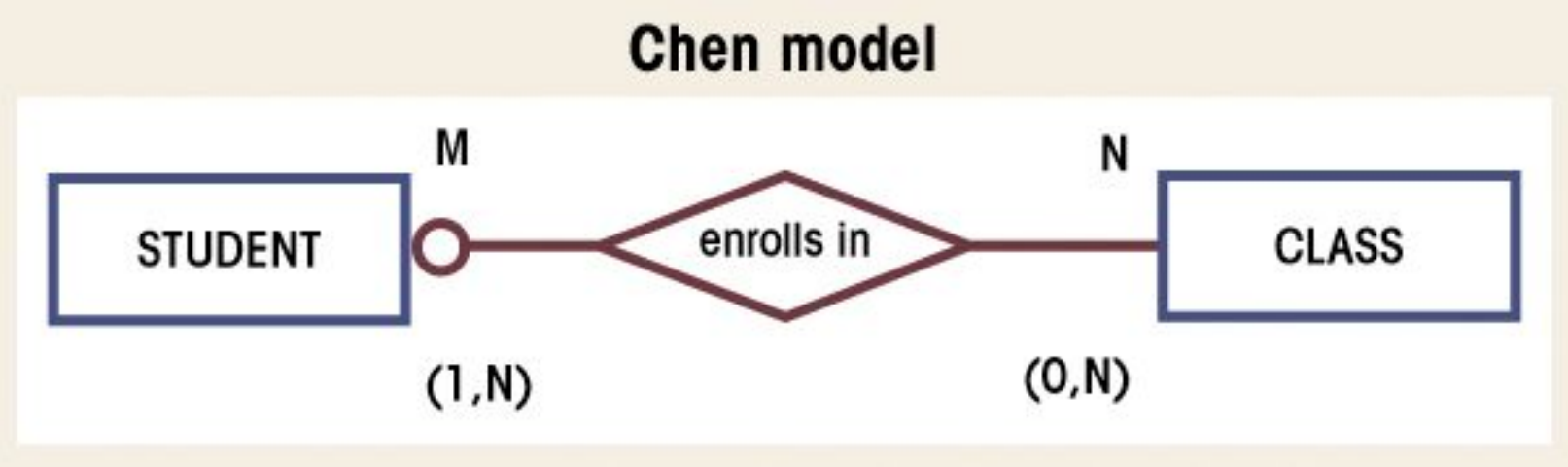


FIGURE 3.23 AN E-R REPRESENTATION OF RECURSIVE RELATIONSHIPS

Composite Entities

- Used to 'bridge' between M:N relationships
- Bridge entities composed of primary keys of each entity needing connection



Composite Entities (con't.)

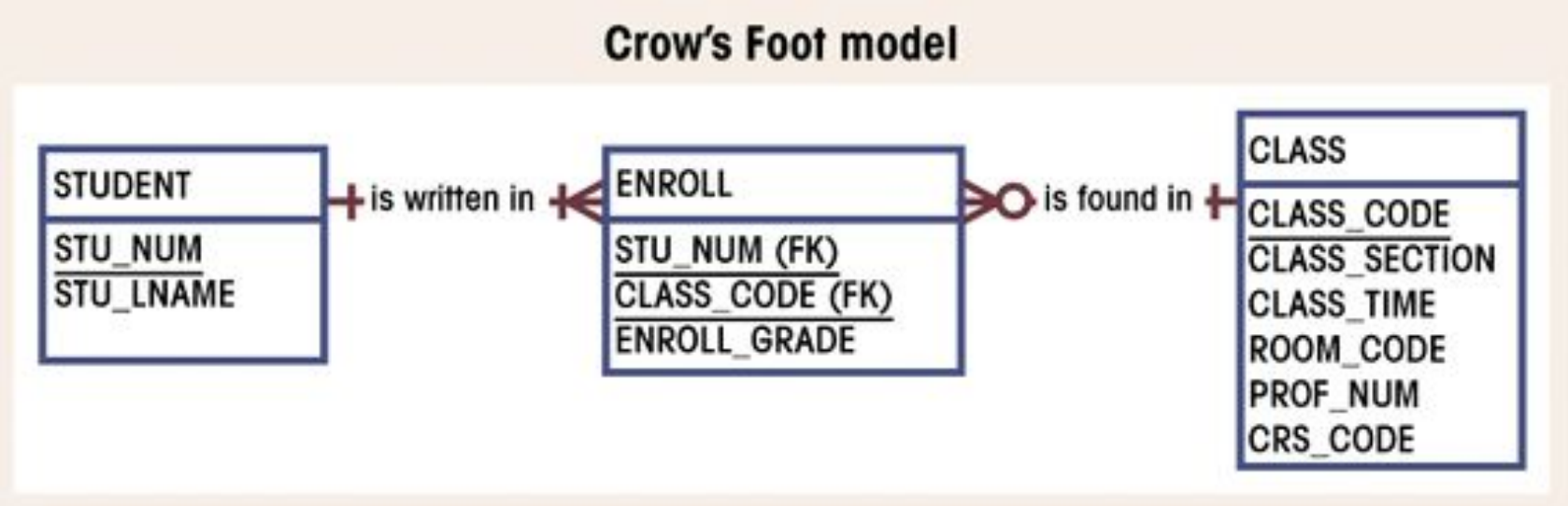
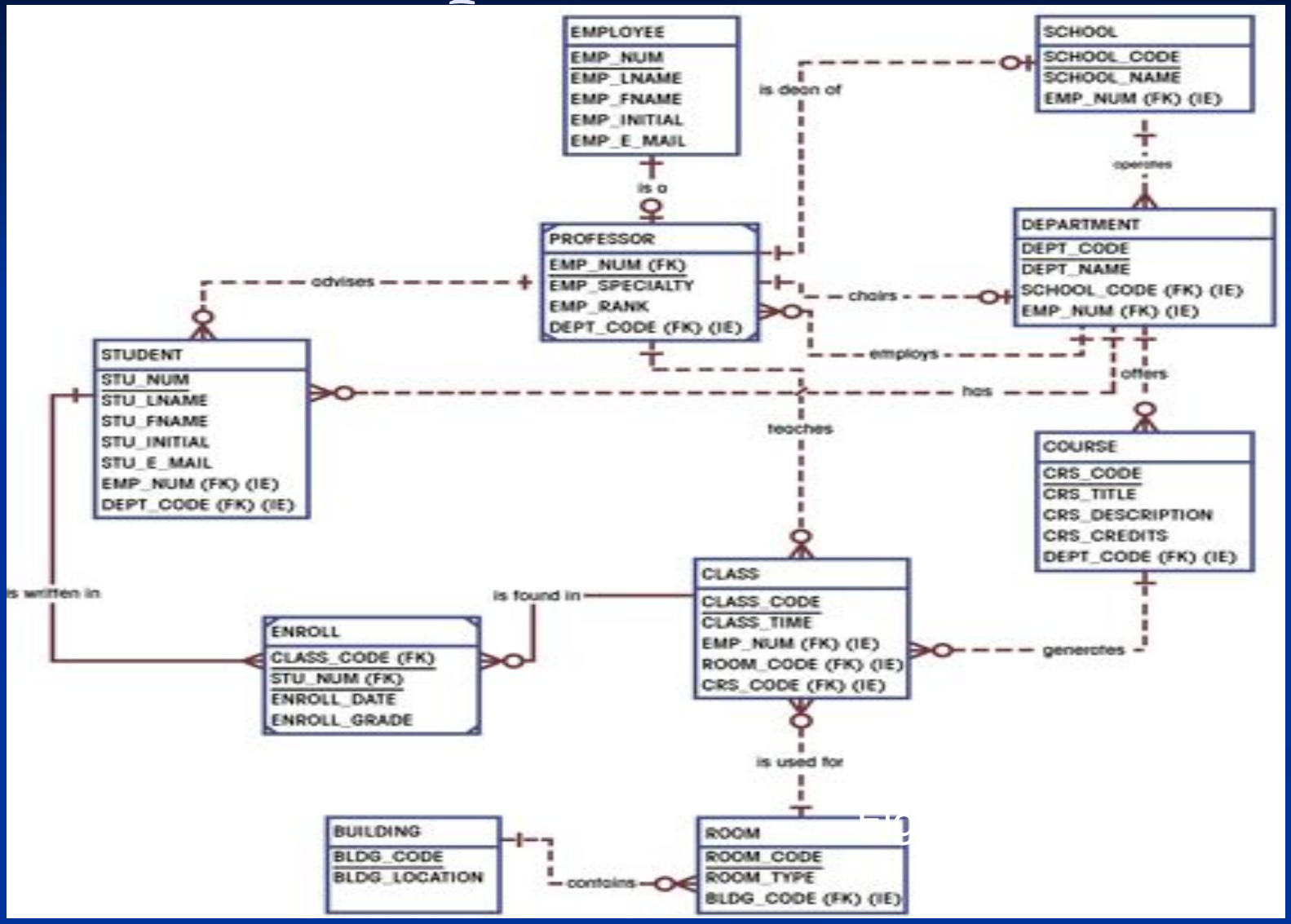


Figure 3.31

Developing an E-R Diagram

(With some repetition)

Completed ERD



Introduction

- Entity Relationship Modeling (ERM)
  - a technique used to analyze & model the data in organizations using an Entity Relationship (E-R) diagram.

Why use ER Diagrams ?

- provides a global quick reference to an organization's data structures.
- can be used individually to design an Information System's (IS) data structure
- can be used with Data Flow Diagrams to provide a more comprehensive IS logical design.

ERD Development Process

- Identify the entities
- Determine the attributes for each entity
- Select the primary key for each entity
- Establish the relationships between the entities
- Draw an entity model
- Test the relationships and the keys

A Simple Example

- STUDENTs attend DEGREEs that consist of many COURSESs.
- A single COURSE (i.e. English) can be studied in many different DEGREE.
- Each STUDENT may only attend one DEGREE.

Identify the entities

- Any entity can be classified in one of the following categories:
- Regular :
  - any physical object, event, or abstract concept that we can record facts about.
- Weak :
  - any entity that depends on another entity for its existence.

Determine the Attributes

- Every Entity has attributes.
- Attributes are characteristics that allow us to classify/describe an entity
- e.g., entity STUDENT has the attributes:
  - student number
  - name
  - date of birth
  - course number

Key Attributes

- Certain attributes identify particular facts within an entity, these are known as KEY attributes.

- The different types of KEY attribute are:
  - Primary Key
    - Composite Primary Key
  - Foreign Key

Key Definitions

- Primary Key:
  - One attribute whose value can uniquely identify a complete record (one row of data) within an entity. Irreducible
- Composite Primary Key
  - A primary key that consists of two or more attribute within an entity. Irreducible
- Foreign Key
  - A copy of a primary key that exists in another entity for the purpose of forming a relationship between the entities involved.

ER Diagram Components

- Every entity diagram consists of the following components:

Entity (labelled box)

Relationship line

Degrees of a Relationship

One-to-one (1:1)

One-to-many (1:n)

Many-to-many (n:m)

NOTE: Every many to many relationship consists of two one to many relationships working in opposite directions

Degrees of relationship, alternative representation

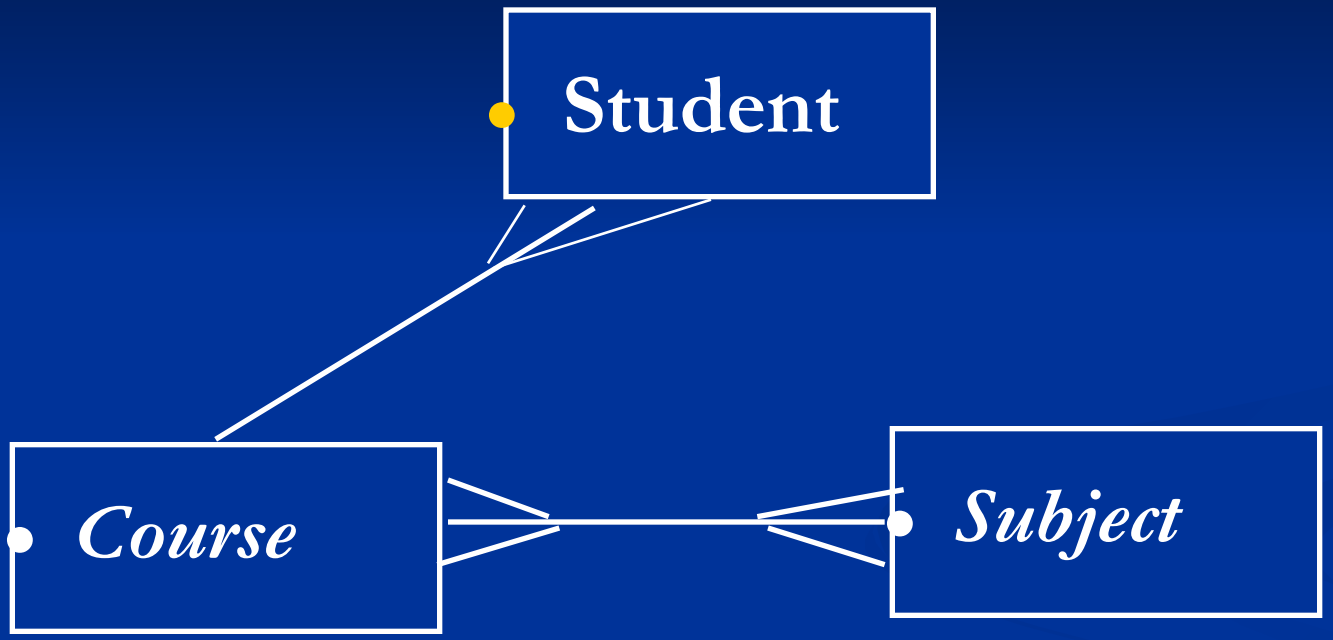
One-to-one (1:1)

One-to-many (1:n)

Many-to-many (n:m)

NOTE: Every many to many relationship consists of two one to many relationships working in opposite directions

A Sample ER Diagram



A Student Record Entity Diagram

Library Case Study

- When a library first receives a book from a publisher it is sent, together with the accompanying delivery note, to the library desk. Here the delivery note is checked against a file of books ordered.
- If no order can be found to match the note, a letter of enquiry is sent to the publishers. If a matching order is found, a catalogue note is prepared from the details on the validated delivery note.
- The catalogue note, together with the book, is sent to the registration department. The validated delivery note is sent to the accounts department where it is stored.
- On receipt of an invoice from the publisher, the accounts department checks its store of delivery notes. If the corresponding delivery note is found then an instruction to pay the publishers is made, and subsequently a cheque is sent. If no corresponding delivery note is found, the invoice is stored in a pending file.

Summary

In today's session we have learned to:

- Identify the entities
- Determine the attributes for each entity
- Select the primary key for each entity
- Establish the relationships between the entities
- Draw an entity model

A Case Study

Conference centre booking system

A conference centre takes bookings from clients who wish to hold courses or conferences at the centre. When clients make bookings they specify how many people are included in the booking, and of these, how many will be resident during the booking, and how many will require catered or non-catered accommodation at the centre.
The centre contains a number of facilities which may be required by clients making bookings as follows:

A. There are 400 bedrooms for clients who will be resident during the Course or conference.
B. A maximum of 250 catered people can be handled at any one time.
C. Six main lecture theatres providing seating for 200 people.
D. Twenty seminar rooms each able to accommodate 25 people.
E. Video conference facilities. The video conference facilities consist of four separate video conference networks. Each video conference network has a large screen based in one of the main lecture theatres, along with 3 satellite screens each of which is based in one of the seminar rooms.

Draw an entity relationship diagram for the case, stating any assumptions you deem necessary.

Shipping company example

- The London and Ireland Shipping Company PLC (LISC) was founded in 1852 and owns a fleet of cargo ships. The company had historically run passenger liners, but recent policy decisions involved the sale of all passenger-carrying vessels. The company currently has 14 vessels, including one oil tanker and one tugboat operating out of Liverpool. Most of the vessels are registered in Liberia for tax reasons.
- Each ship has one or more holds divided into spaces. The holds are defined by steel bulkheads and the spaces are defined by shelf racks or other physical dividers. Sister ships, built by the same shipbuilders and to the same designs have similar names, such as *Pride of Ireland, Queen of Ireland, Song of Ireland* and *Warrior of Ireland*. Sister ships also have identical cargo storage facilities.
- LISC issues contracts to agents for one or more manifests (lists of cargo items to be shipped). LISC's charges for cargo carried are based on the number of spaces the cargo requires for storage. The types of cargo typically carried by LISC include grain, coal and ores (carried only in ships equipped with bulk cargo holds). They also transport sacked grain, heavy cases, containers (which may be carried on deck), pallets and so on.
- Cargo items may take up less than one space in a hold, or one or more spaces, depending on the size of the item. A space may therefore contain several small cargo items.
- The ships owned by LISC are kept as busy and as full as possible, in order to maximise the profits that each vessel makes and minimise running & operating costs. LISC's ships ply most of the seas of the world, but tend to operate mainly in the Mediterranean, the North and Mid Atlantic and the Indian Ocean. Different ships require different crew complements.
- LISC intends to create a computer based information system that will be able to perform the following tasks:
- • record the voyages of each ship with the start and end ports.
- • record the cargo held by a ship on each voyage
- • keep records of their employees and the ships they are assigned to
- • producing invoices for agents and customers
- • keep a record of customers' payments on invoices
- • analyse the efficiency of use of cargo space and of percentage wasted cargo space for ships voyages