1.264 Lecture 8

Data modeling

This class: Hand in exercises on paper after class
Next class: No reading. Exercises due <u>after</u> class

Data models

- **Data model is representation of**
  - **Things (or entities or objects) of importance to a business or a system**
  - **How the things relate to each other**
- **It is built and modified until it represents the business well enough to write a system.**
- **Data models are extended to become class diagrams in the Unified Modeling Language [UML] by adding the behaviors of each entity to the model**
- **Data models are sometimes built during requirements, and other times during design phase**
  - **The earlier the better. I always build them during requirements.**

Logical data modeling

- **Method to discover the data, relationships and rules of a business, collectively called the business rules**
- **Logical data models are the basis of:**
  - **Physical data models, or actual databases**
  - **Applications, parts of which can be automatically generated from the data model**
- **Small model for broker of transportation services**
  - **Small, but says a lot about broker**
  - **Gives good picture of what database should look like**
  - **Also gives good picture of underlying business rules of broker**
  - **Useful in requirements analysis and scrubbing**

Transportation Broker Data Model
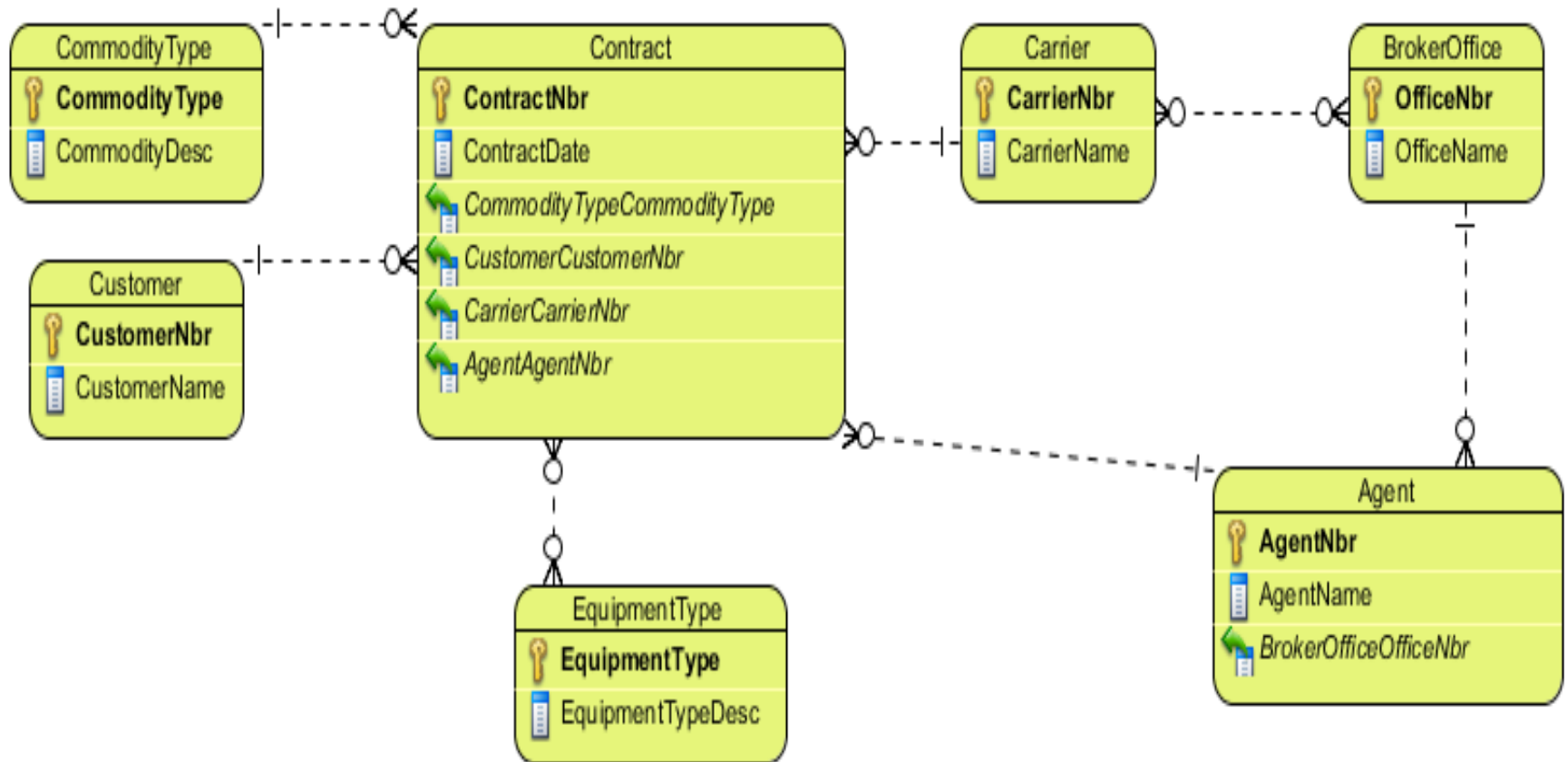
Broker Business Rules

- A carrier can be associated with many offices
- An office can be associated with many carriers
- A carrier can issue many contracts
- A contract is issued by one carrier
- An office can employ many agents
- An agent is employed by one office
- An agent can sell many contracts
- A contract is serviced by only one agent
- A contract can serve to carry only one commodity type
- A commodity type can be carried under many contracts
- A contract can be associated with many equipment types
- An equipment type can be associated with many contracts
- A customer can be served by many contracts
- A contract covers one customer

Data model purpose

- **Business needs to build logical data model so users and developers both understand business rules of company**
  - **Models enable users and developers to have single view of system**
  - **Sometimes users note this is first time they understood business rules!**
- **Converting logical to physical data model (database) is very straightforward these days.**
  - **Little need for separate physical model for online databases**
  - **Create integer system-generated keys instead of strings and composite keys for performance**
  - **We still create separate physical models for data warehouses, read-only databases and some other special cases**

Data modeling concepts

- **Entities (objects, tables)**
- **Attributes (properties)**
- **Keys (primary and foreign)**
- **Relationships**
- **Referential integrity**

Entity type and entity occurrence

Entity type

Department

DeptNbr
DeptName
DeptType
DeptStatus

Table, class

Entity occurrence

Department			
DeptNbr	**DeptName**	**DeptType**	**DeptStatus**
930	Receiving	Mfg	Active
378	Assembly	Mfg	Active
372	Finance	Adm	Active
923	Planning	Adm	Active
483	Construction	Plant	Inactive

Row, object

Entities

- **"Department" is an entity type**
  - **In a software program, "department" is a class**
  - **In a database, "department" is a table**
- **"Department 101" is an occurrence of entity type "Department"**
  - **In a software program, "department 101" is an object, which is an instance of class "department"**
  - **In a database, "department 101" is a row in a table**
- **Entities are things, often physical, that have facts associated with them.**
- **Processes are almost never entities. For example:**
  - **Order entry is not an entity**
  - **Orders and customers are entities**
  - **Reports are not entities**
- **Entity type descriptions should be as extensive as possible in developing a model.**

Entity type description

- **Poor description (I've seen lots of these)**
  - **Vendor: Someone we buy products from.**
- **Good description (I've never seen one like this in real life)**
  - **Vendor: A US corporation we have reviewed with respect to their qualifications for providing products to our company. Vendors are rated based on price, quality, delivery performance and financial stability. Each vendor is classified by one vendor status: approval pending, approved, rejected or inactive. This approval decision is made in a weekly meeting among purchasing, manufacturing and finance. Purchasing requests that rejected vendors be kept in the database for future reference. Purchasing expects 500 vendors will be maintained at any one time. Of these, 200 will be active, 25 pending, 75 inactive and 100 rejected. Contact Joan Smith in Purchasing for more information.**

Attributes

- **Attributes are a data item or property associated with an entity type**
    - **They are typically nouns (quantity, type, color, …)**
    - **Example: Employee**
        - **ID**
        - **Name**
        - **Social security number**
        - **Address**
        - **Phone**

Entity type/attribute exercise

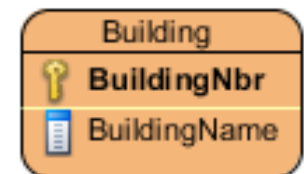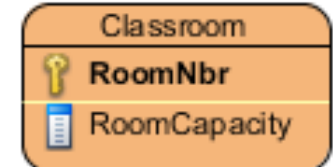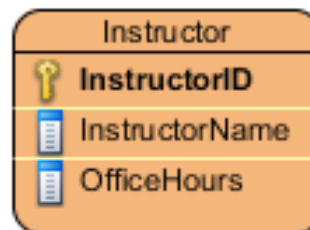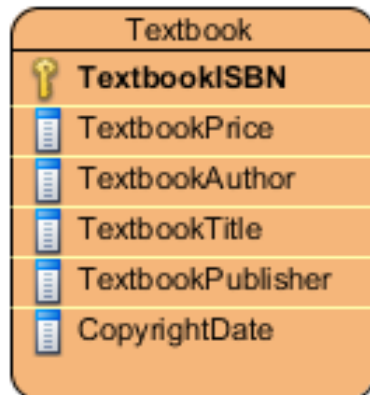## 1. Identify which are types and which are attributes:

- **Instructor**
- **Teaching assistant (TA)**
- **Course section number**
- **Building name**
- **Course number**
- **Textbook price**
- **Teaching asst (TA) name**
- **Instructor ID**
- **Textbook author**
- **Course title**
- **Textbook**
- **Classroom**
- **Textbook ISBN**
- **Section days**
- **Instructor office hours**
- **Textbook title**
- **Classroom number**
- **TA student ID**
- **Instructor name**
- **Textbook publisher**
- **Section capacity**
- **Course objective**
- **Copyright date**
- **Building number**
- **Course section**
- **Course**
- **Building**
- **Section time**
- **Classroom capacity**

Entity type/attribute exercise

2. Draw an entity type box and its attributes for each:

- Instructor
- Teaching assistant (TA)
- Course section number
- Building name
- Course number
- Textbook price
- Teaching asst (TA) name
- Instructor ID
- Textbook author
- Course title
- Textbook
- Classroom
- Textbook ISBN
- Section days

- Instructor office hours
- Textbook title
- Classroom number
- TA student ID
- Instructor name
- Textbook publisher
- Section capacity
- Course objective
- Copyright date
- Building number
- Course section
- Course
- Building
- Section time
- Classroom capacity

Solution

Course
- 🔑 **CourseNbr**
- 📄 CourseTitle
- 📄 CourseObjective

CourseSection
- 🔑 **CourseSectionNbr**
- 📄 SectionDate
- 📄 SectionTime
- 📄 SectionCapacity

TeachingAsst
- 🔑 **TAStudentID**
- 📄 TAName

Classroom
- 🔑 **RoomNbr**
- 📄 RoomCapacity

Textbook
- 🔑 **TextbookISBN**
- 📄 TextbookPrice
- 📄 TextbookAuthor
- 📄 TextbookTitle
- 📄 TextbookPublisher
- 📄 CopyrightDate

Instructor
- 🔑 **InstructorID**
- 📄 InstructorName
- 📄 OfficeHours

Building
- 🔑 **BuildingNbr**
- 📄 BuildingName

Domain entity type

- **Also called pick list, validation list, etc.**
- **Department name example**

Domain entity type

Department			
DeptNbr	**DeptName**	**DeptType**	**DeptStatus**
930	Receiving	Mfg	Active
378	Assembly	Mfg	Active
372	Finance	Adm	Active
923	Planning	Adm	Active
483	Constructi	Plant	Inactive

| ValidDeptType |
|---|
| **DeptType** |
| Mfg |
| Adm |
| Plant |
| Sales |
| Operations |

Relationships

- **Entities are drawn as boxes, as in the broker diagram**
- **Relationships are lines between boxes**
- **Cardinality is the expected number of related occurrences between the two entities in the relationship**
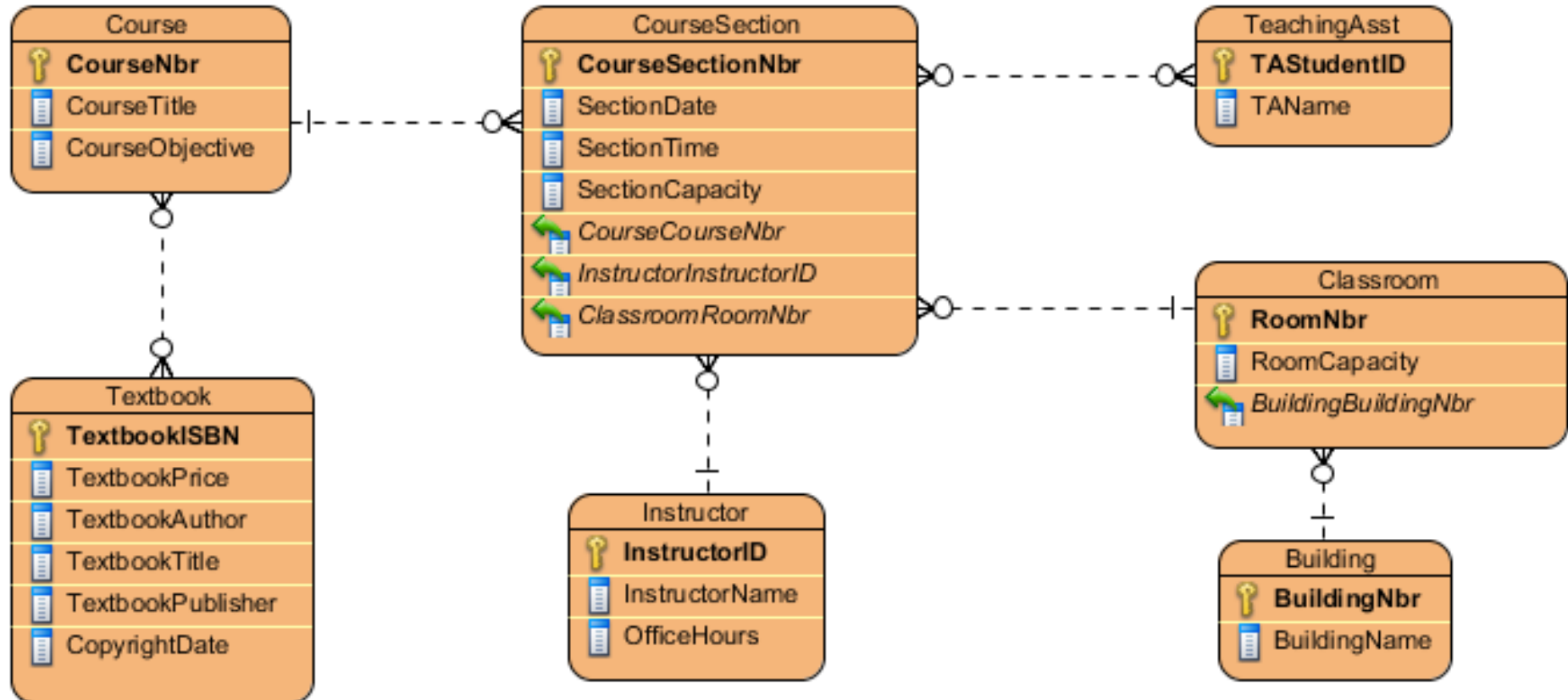- **Relationships + cardinality = business rules**

One   **Zero or many**

(Instructor)   **(Course section)**
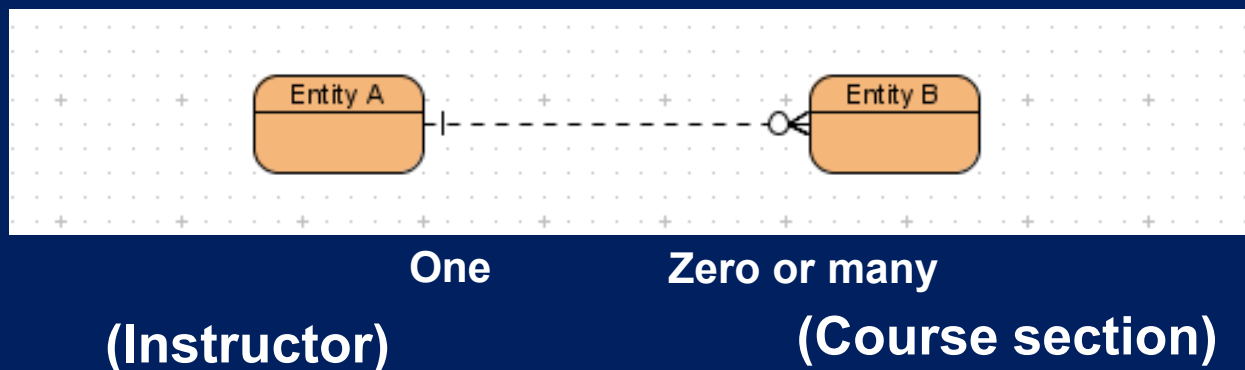
Relationships and Cardinality

Exercise: Draw the relationships among these entities
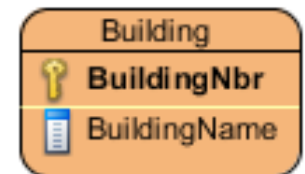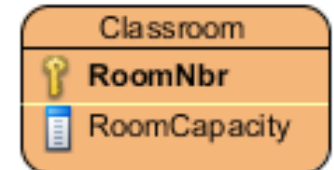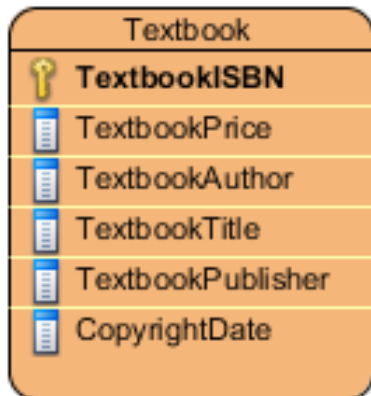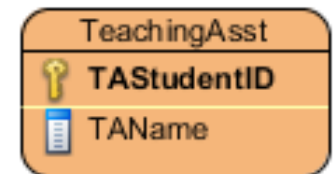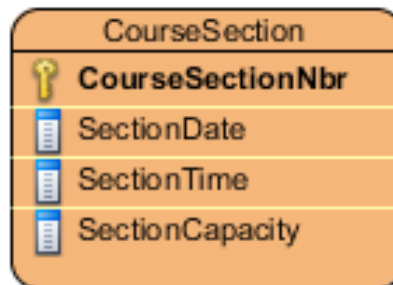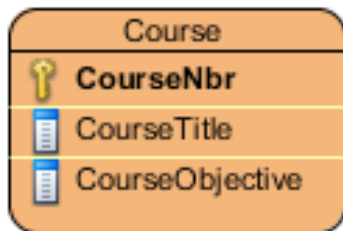
Solution

We're getting there: we've defined entities, attributes and relationships. We still have to add keys and more entities

Solution: Course example

- Course may be offered in many (0,1 or more) sections

- Course section must be associated with a course

- Course section may be taught by many (0,1 or more) TAs

- TA may teach many (0, 1 or more) course sections

- Course section must be taught by 1 instructor (??)

- Instructor may teach many sections

- Course may use many textbooks (all sections use same)

- Textbook may be used in many courses

- Building may contain many rooms

- A room is in only one building

- A course section may use a room

- A room may be used by many course sections (not at same time)

MIT OpenCourseWare

1.264J / ESD.264J Database, Internet, and Systems Integration Technologies
Fall 2013