Data Modeling continued.
Extended E-R Diagrams and Business Rules.

1. ERDs

Guidelines to identifying entities and relationships.
Start with writing a scenario that describes the objectives and processes of the desired system.

**Discovering Entities:**

look for **nouns** that are:
- Tangible things (e.g. Airplane, Book, Classroom, Course)
- Roles, people (e.g. Customer, Vendor, Student, Patient)
- Events (e.g. Accident, Flight, Service Call, Appointment)
- Interactions (e.g. Marriage, Sale, Purchase, Enrollment)

If you need to save data about many **properties** of a thing, then that thing is likely to be an entity.
If something has many **instances**, then that thing is probably an entity.

**Naming rules:**
- use common (business) names
- use singular nouns for entities and attributes

**Identifying Relationships**

In the scenario description, a **verb** that denotes an association or a state is likely to represent a relationship.
Such a verb is usually a relationship if both sides of the verb are nouns

- Note: more than one relationship can exist between two entities.
  - For example, a house can be owned by a person, and can be occupied by a person.

**Naming rules:**
- use a verb or a verb - phrase

Checking the E-R diagram for **completeness**:
Entities: must be specified and appropriately named
  All relevant attributes are specified
  The composition of all composite attributes is indicated
  An identifier attribute(s) (Primary Key) is(are) indicated (possible exceptions: Associative Entities)
Relationships: must be specified and appropriately named
  Cardinality labels are specified completely (both minimum and maximum cardinality)
Practice problem: create an E-R model for the following scenario:

The H.I. Topi School of Business operates international business programs in 10 locations throughout Europe. The School had its first class of 900 graduates in 1965. The School keeps track of each graduate’s student number, name when a student, country of birth, current country of citizenship, current name, current address, and the name of each major the student completed (each student has one or two majors). In order to maintain strong ties to its alumni, the School holds various events around the world. Events have a title, date, location, and type (for example, reception, dinner, or seminar). The School needs to keep track of which graduates have attended which events. For an attendance by a graduate at an event, a comment is recorded about information School officials learned from the graduate at that event. The School also keeps in contact with graduates by mail, e-mail, telephone, and fax interactions. As with events, the School records information learned from the graduate from each of these contacts. When a School official knows that he or she will be meeting or talking to a graduate, a report is produced showing the latest information about that graduate and the information learned during the past two years from that graduate from all contacts and events the graduate attended.
2. Enhanced E-R Model introduces the supertype/subtype relationship otherwise known as (generalization/specialization).

Relation between subclass and superclass can be described by “is a” relation, e.g.
- Each Cat is an Animal, Dog is an Animal, etc…
- Each Student is a Person, Each Employee is a Person

Supertype - a general entity type, can be further specialized into Subtypes.
Subtype - an entity that “is a kind of” the Entity Supertype.

Subtype Entity inherits all attributes and relationships of its supertype, but provides more specific details about its own characteristics, that are not properties of the Supertype:

- attributes of its own and/or:
- relationships with other entities distinct from those of other subtypes

Practice Problems.
1. Establish supertype/subtype structure of the following entity types:
   (a) Game, Bridge, Chess, Go, Board Game, Card Game, Poker
   (b) Polygon, Circle, Triangle, Rectangle, Geometric Figure.

2. Identify entities in the following description of a campus map. Then establish supertype/subtype relationships between them.

A college campus consists of residential, administrative and academic buildings, library, parking lots, student centers and various fields. Each campus objects has a unique identification number and is depicted on the campus map. Residential buildings are undergraduate or graduate dorms, and the President’s residence. Academic buildings house classrooms, labs, and auditoriums, sports facilities, and academic departments. Administrative buildings house various offices. Parking lots are restricted for student use only, staff only and mixed (students, staff, and visitors).
Notation:

- Attributes shared by all entities

Example:

[Diagram showing relationships between Patient, Outpatient, Resident Patient, and Bed with attributes such as Patient_ID, Admit_Date, Physician_ID, Responsible Physician, Outpatient, Resident Patient, Bed_ID, Date_Discharged, etc.]
Practice questions:

True or false? Answer the following questions based on the diagram above:

1. Every OUTPATIENT has an Admit_Date ___
2. Every PATIENT has a Date_Discharged ___
3. Every RESIDENT PATIENT is cared for by a RESPONSIBLE PHYSICIAN ___
4. Every PATIENT is cared for by a RESPONSIBLE PHYSICIAN ___
5. Every PATIENT is assigned a BED ___

3. Supertype/Subtype Constraints

Completeness constraint: total specialization vs. partial specialization

Completeness constraint specifies whether or not each instance of an entity supertype must also be an instance of at least one subtype. In other words, there are no “pure” supertype instances

- Total specialization (double line notation) - each instance of a supertype entity must also be an instance of at least one subtype. In other words, there are no supertype instances in the business that are not an instance of some subtype.

  E.g. A PATIENT must be either an OUTPATIENT or a RESIDENT PATIENT

- Partial specialization (single line notation) – a supertype entity instance may not be an instance of any subtypes.
  E.g. A BUILDING can be a RESIDENTIAL BUILDING or a COMMERCIAL BUILDING or neither (e.g. an Army Warehouse).

Disjointness constraint: whether or not a supertype instance may simultaneously be a member of two (or more) subtypes

- Disjoint Rule (disjoint subtypes) - letter “d” notation
  An entity instance can be a member of one and only one subtype
  E.g. all PERSONS are either MALE or FEMALE

- Overlap Rule (overlapping subtypes) - letter “o” notation
  An entity instance can simultaneously be a member of more than one subtypes
  E.g. an ATHLETE can be both a RUNNER and a JUMPER

Practice: Problems 6, 7 (p. 158) from Chapter 4.

Business Rule Classification:
- Derivation – describe derivations from data
- Structural Assertion – define structure of data – entities, attributes, relationships
- Action Assertion – constrain actions (insert, delete, update record) to be taken on a database.

Some assertions can be encoded in SQL - decoupled from the application code.

Action Assertions
Specifying Restrictions (R-assertion)

Business Rule 1: For a faculty member to be assigned to teach a section of a course, the faculty member must be qualified to teach the course for which that section is scheduled

This assertion states that for an association Is_Assigned between Faculty and Section to be created, there must already exist
a. an association Is_Qualified between that Faculty and a Course, and
b. an association between that Course and the Section

Specifying a Limit (LIM-assertion)

Business Rule 2: For a faculty member to be assigned to teach a section of a course, the faculty member must not be assigned to teach a total of more than three course sections

This assertion states that there is an Upper Limit on the number of Is_Assigned associations between each Faculty and a Section.
Practice problems: Interpret the following assertions:

1. Patient_ID  Admit_Date  Physician_ID
   PATIENT  Admits  RESPONSIBLE PHYSICIAN
   OUTPATIENT  RESIDENT PATIENT  BED
   Checkbox_Date  Date_Discharge

2. May_attend  R
   Has_completed  R
   CONCERT  HOMEWORK
   STUDENT