Understanding Entity Relationship Diagrams

Dr. Chawalit Jeenanunta
Objective

- Business rule representation
- Diagram rules
- Alternative notations
Business Rules

- Enforce organizational policies
- Promote efficient communication
- Formal representation in ERD
- Informal representation in documentation associated with an ERD
- Use rules language to formally represent in relational database after conversion
Formal Representation

- **Primary key** constraints: entity identification
- **Named** relationships: direct connections among business entities
- **Identification** dependency: knowledge of other entities for identification
- **Cardinalities**: restrict number of related entities in a business situation
- **Generalization hierarchies**: classification of business entities and organizational policies
Informal Representation

- Specify as documentation associated elements of an ERD
- **Candidate key constraints**: alternate ways to identify business entities
- **Reasonable values**: fixed collection of values or consistent with another attribute
- **Null value constraints**: data collection completeness
- **Default values**: simplify data entry and provide value when unknown
Diagram Rules

- Ensure that ERD notation is correctly used
- Similar to syntax rules for a computer language
- Completeness rules
  - no missing symbols or specifications
- Consistency rules:
  - no conflicts among symbols or specifications
Diagram Rules

- Apply these rules when completing an ERD to ensure that there are no notation errors in your ERD.

- Similar to syntax rules for a computer language:
  - Ensures proper language structure, not correct meaning
  - Diagram rules ensure proper structure among symbols
  - Do not ensure that you have considered multiple alternatives, correctly represented user requirements, and properly documented your design
Completeness Rules

1. **Primary Key Rule:**
   - all entity types have a PK (direct, indirect, or inherited)

2. **Naming Rule:**
   - all entity types, relationships, and attributes have a name

3. **Cardinality Rule:**
   - cardinality is specified in both directions for each relationship

4. **Entity Participation Rule:**
   - all entity types participate in at least one relationship except for entity types in a generalization hierarchy

5. **Generalization Hierarchy Participation Rule:**
   - at least one entity type in a generalization hierarchy participates in a relationship
Completeness Rules

- The first three rules are mandatory. A finished ERD should not violate the PK, Naming, and Cardinality rules.

- PK rule:
  - Direct: Table contains the primary key attribute(s)
  - Indirect: Table borrows (id dependent) for part or all of PK
  - Inherited: Table inherits PK from a supertype

- The next two rules are optional. Most ERDs will not violate the Entity Participation and Generalization Hierarchy Participation rules.

- Rule 5 applies to an entire generalization hierarchy, not to each entity type in a generalization hierarchy. In other words, at least one entity type in a generalization hierarchy should be connected to at least one entity type not in the generalization hierarchy. In many generalization hierarchies, multiple entity types participate in relationships. Generalization hierarchies permit subtypes to participate in relationships thus constraining relationship participation.
Primary Key Rule Issue

- Primary key rule is simple in most cases
- For some weak entities, the PK rule is subtle
  - Weak entity with only one 1-M identifying relationship
  - Weak entity must have a local key to augment the borrowed PK from the parent entity type
  - Borrowed PK alone cannot identify weak entity instances because there can be many weak entity instances related to the same parent entity
  - Violation of PK rule if local key is missing
  - Associative entity types do not need to provide a local key although they can if needed
PK Rule Violation Example

PK rule violation
A single 1-M identifying relationship
Room does not have a local key.

Building
BldgID
BldgName
BldgLocation

Contains

Room
RoomNo
RoomCapacity
Naming Consistency Rules

- Consistency rules: no conflicting specifications
- Naming rules: no conflict among names

- **Entity Name Rule:**
  - entity type names must be unique

- **Attribute Name Rule:**
  - attribute names must be unique within each entity type and relationship

- **Inherited Attribute Rule:**
  - attribute names in a subtype do not match inherited (direct or indirect) attribute names.
Relationship Names

- No uniqueness requirement
- Participating entities provide a context for relationship names
- Use unique names as much as possible to distinguish relationships
- Must provide unique names for multiple relationships between the same entity types
Connection Consistency Rules

- no conflicts or redundancies among relationships

**Relationship/Entity Connection Rule:**

- relationships connect two entity types (not necessarily distinct)
- Do not connect relationships directly (connect through entity types)
- Self referencing relationships: connect the same entity type two times
Connection Consistency Rules

- **Redundant Foreign Key Rule:**
  - Foreign keys are redundant with 1-M relationships
  - Use FKs in the relational model, not in ERDs
  - Violation of this rule is common:
    - confusion between the ERDs and relational table design
  - Conversion replaces relationships with foreign keys
Identification Dependency Rules

- Common source of diagram errors
- no conflicts among components of identification dependency (weak entity, identifying relationships, cardinality specification)

- Weak entity rule:
  - weak entities have at least one identifying relationship

- Identifying relationship rule:
  - at least one participating entity type must be weak for each identifying relationship

- Identification dependency cardinality rule:
  - Minimum and maximum cardinality must (1,1) for a weak entity in all identifying relationships
  - (1,1) cardinality should appear near the parent entity type
  - Common source of diagram errors
Example of Diagram Errors

- Weak entity rule violation:
  - Faculty is a weak entity but it is not involved in any identifying relationships
  - Resolution: remove weak entity symbols
  - In some cases, resolution involves changing a relationship to identifying
Example of Diagram Errors

- Identifying relationship rule violation:
  - Has is an identifying relationship but neither Offering nor Course is a weak entity
  - Resolution: make Has a regular (non-identifying) relationship
  - Sometimes resolution involves making an entity type weak
Example of Diagram Errors

- **Identification Dependency**
  - Cardinality rule:
    - The min/max cardinality of the Registers relationship should be (1,1) near Student
  - **Resolution**: reverse the cardinalities on the Registers relationship
  - Sometimes the resolution does not involve reversing the cardinality but just changing one cardinality specification
Example of Diagram Errors

- Redundant foreign key rule:
  - CourseNo in Offering is redundant with the Has relationship
  - Resolution: remove the CourseNo attribute in Offering
  - This rule can be violated even if the FK attribute has a different name than associated PK. It is more difficult to detect a redundancy if the FK has a different name.
Corrected ERD

- **UnivPerson**
  - SSN
  - Name
  - City
  - State
  - Zip

- **Student**
  - StdClass
  - StdMajor
  - StdGPA

- **Offering**
  - OfferNo
  - OffLocation
  - OffTime

- **Faculty**
  - FacSalary
  - FacRank
  - FacHireDate

- **Course**
  - CourseNo
  - CrsDesc
  - CrsUnits

- **Enrollment**
  - EnrGrade

- **Supervises**

- **Teaches**

- **Registers**

- **Has**
ERD Variations

- No standard ERD notation
  - Many notations: too many; source of confusion
  - Many variations of a given notation: many variations of the Crow’s Foot notation
  - Crow’s foot notation is widely used but it has many variations
- Placement of cardinality symbols
- Rule variations
- Be prepared to adjust to the ERD notation in use by each employer

- Symbol variations
  - Different symbols for entity types, relationships, and attributes
  - Different symbols for identification dependency and generalization hierarchies
  - Different definitions for commonly used terms: existent dependent is defined incorrectly by some authors to mean identification dependency
ERD Rule Variations

- Lack of ERD standards
- M-way relationships
- M-N relationships
- Relationships with attributes
- Self-referencing relationships
- Relationships connected to other relationships
- Adapt to notations in work environments
Chen ERD Notation

- Original ERD notation
- Widely known and used
- Variations:
  - Different relationship symbols
  - Reversed cardinality positions
  - Attributes are sometimes shown in ovals attached to entity types and relationships
  - M-way relationships supported
  - Different symbol for associative entity types and weak entities
Chen ERD Notation

**Course**
- **CourseNo**
- **CrsDesc**
- **CrsUnits**

Minimum cardinality for Course: (0:N)

**Has**

Maximum cardinality for Offering: (1:1)

**Offering**
- **OfferNo**
- **OffLocation**
- **OffTime**
- ...

Minimum cardinality for Offering
Unified Modeling Language

- Standard notation for object-oriented modeling
  - Objects
  - Object features
  - Interactions among objects
- UML supports class diagrams, interface diagrams, and interaction diagrams
- More complex than ERD notation
Simple Class Diagram

Object name

Attributes

Operations

Cardinality

Role name

Offering

OfferNo : Long
OffTerm : String
OffYear : Integer
OffLocation : String

EnrollmentCount() : Integer
OfferingFull() : Boolean

Faculty

FacSSN : String
FacFirstName : String
FacLastName : String
FacDOB : Date

FacAge() : Integer

Teaches

0..1

TaughtBy

0..n
Association Class

- Association class: similar to a M-N relationship with attributes

```
Association class

Offering
OfferNo : Long
OffTerm : String
OffYear : Integer
OffLocation : String
EnrollmentCount() : Integer
OfferingFull() : Boolean

Enrollment
EnrGrade : Numeric

Student
StdSSN : String
StdFirstName : String
StdLastName : String
StdDOB : Date
StdAge() : Integer

Takes 0..n
0..n Enrolls
```
Generalization Relationship

- Generalization and inheritance are original features of the UML, not added later as generalization hierarchies are for ERDs
Composition Relationship

- Composition is similar to identification dependency
- Dark diamond for a composition class

Composition symbol (dark diamond)
Summary

- Data modeling is an important skill
- Crow’s Foot ERD notation is widely used
- Use notation precisely
- Use the diagram rules to ensure structural consistency and completeness
- Understanding the ERD notation is a prerequisite to applying the notation on business problems