

Databases and Advanced Data Techniques

Tarapong Sreenuch

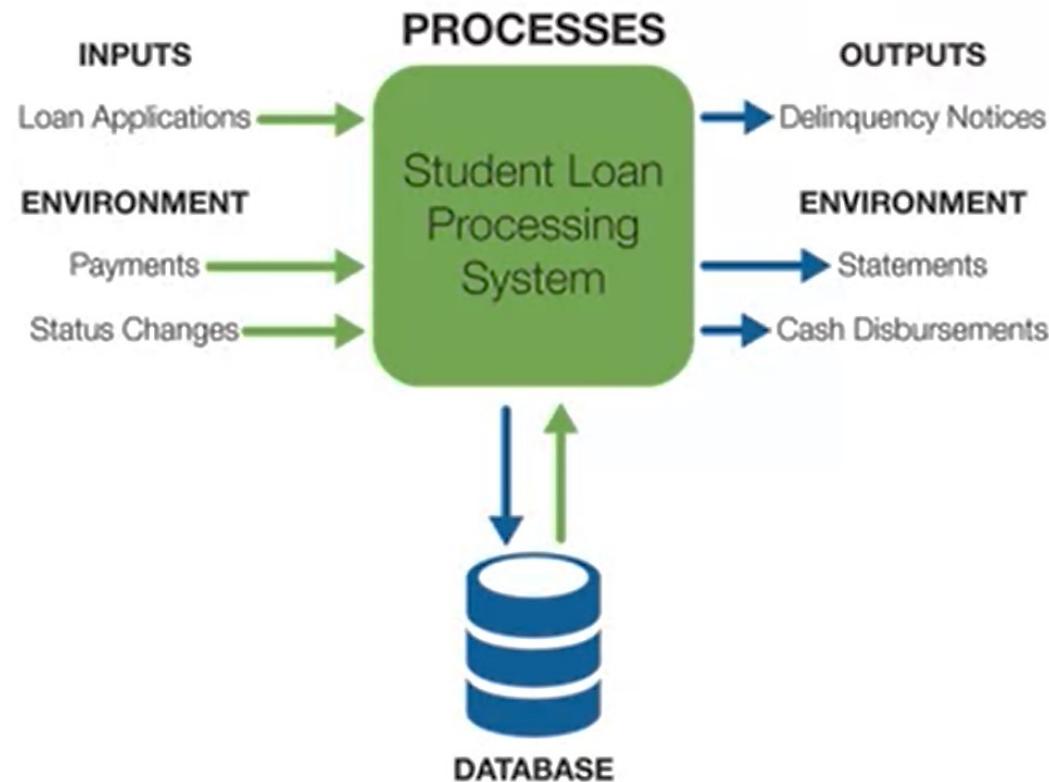
Database Design and Entity-Relationship Diagram

Lecture Outline

- Database Development Process
- ERD Notation and Rules
- Relationship Variations
- Conceptual Data Modelling Goals and Challenges
- Data Modelling Examples
- Wrapping Up



Information System

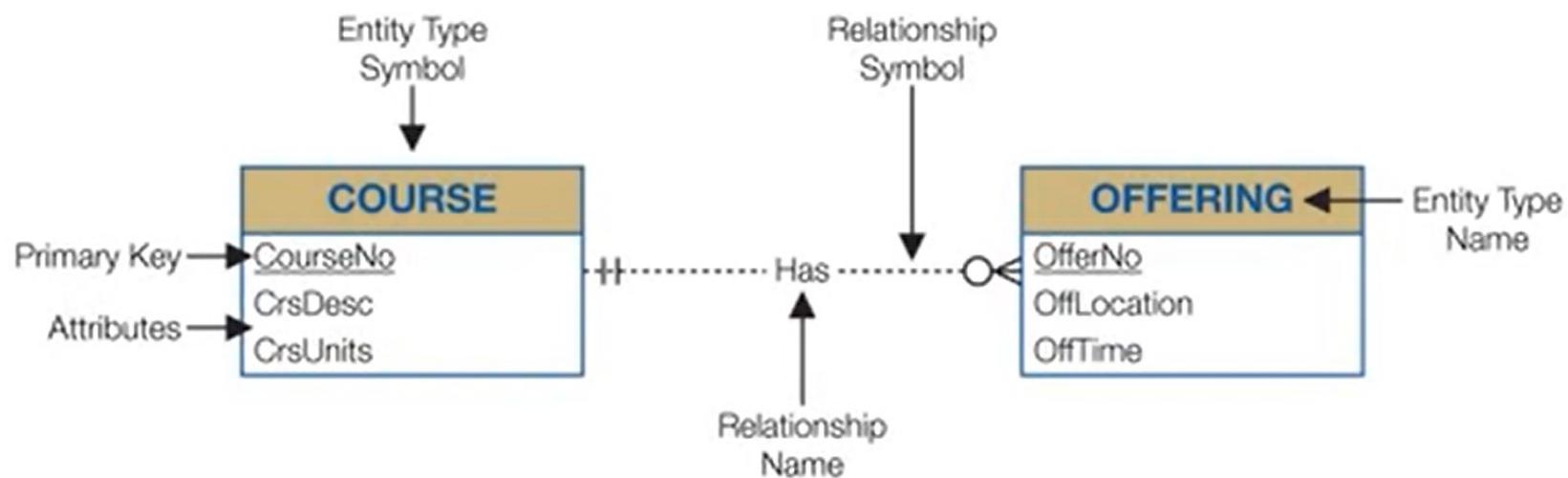


Broad Goals of Database Development

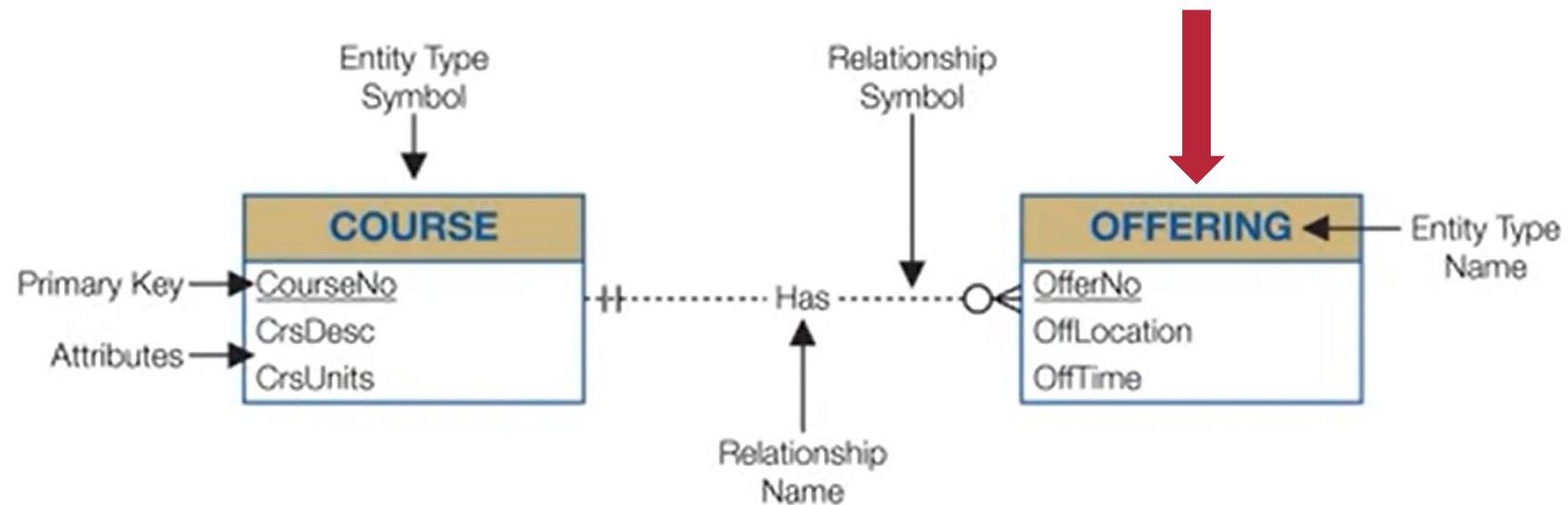
- Develop a Common Vocabulary
- Define Business Rules
- Ensure Data Quality
- Provide Efficient Implementation



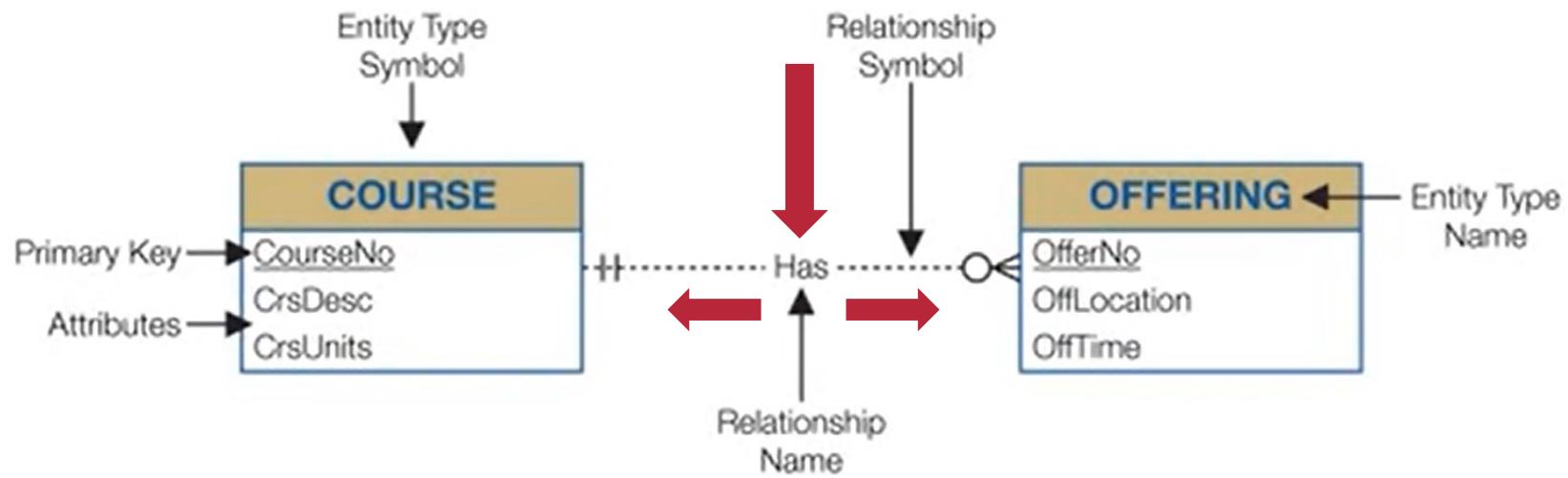
ERD Diagram: Basic Symbols



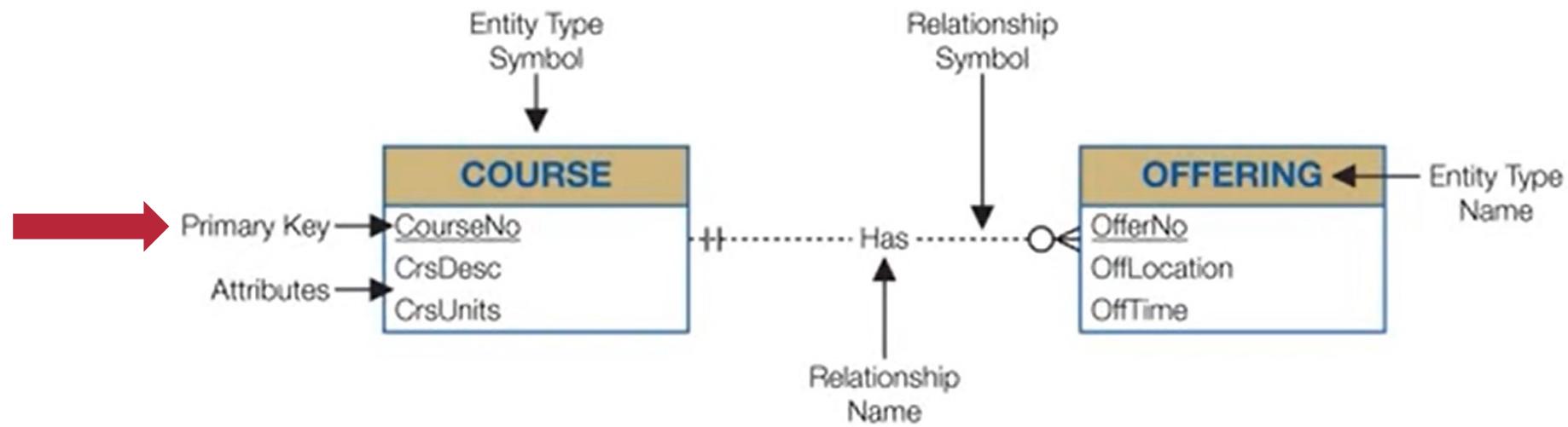
ERD Diagram: Basic Symbols



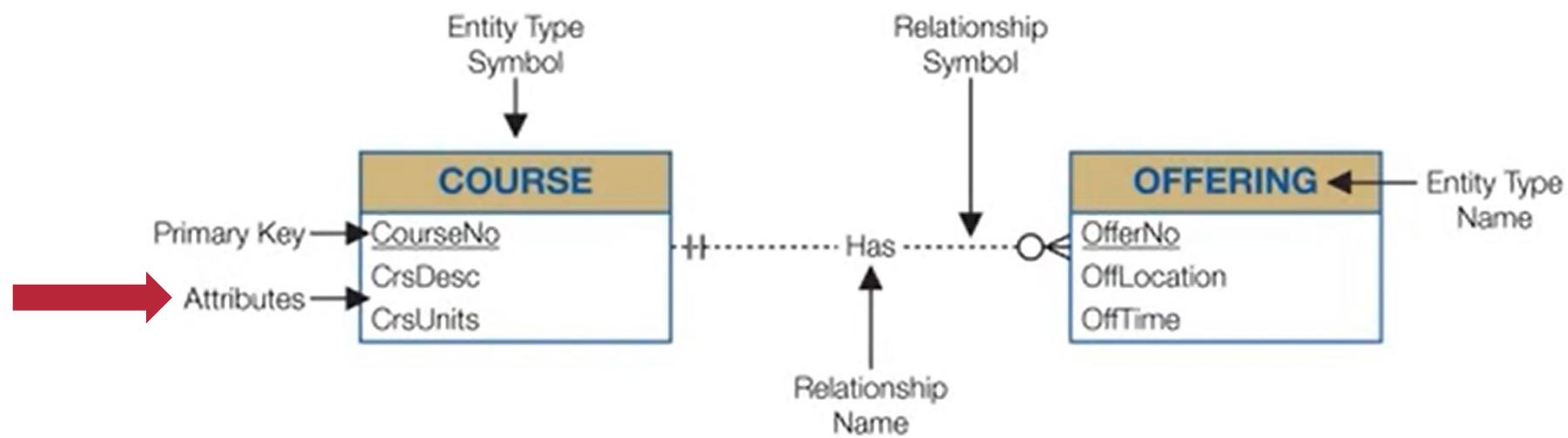
ERD Diagram: Basic Symbols



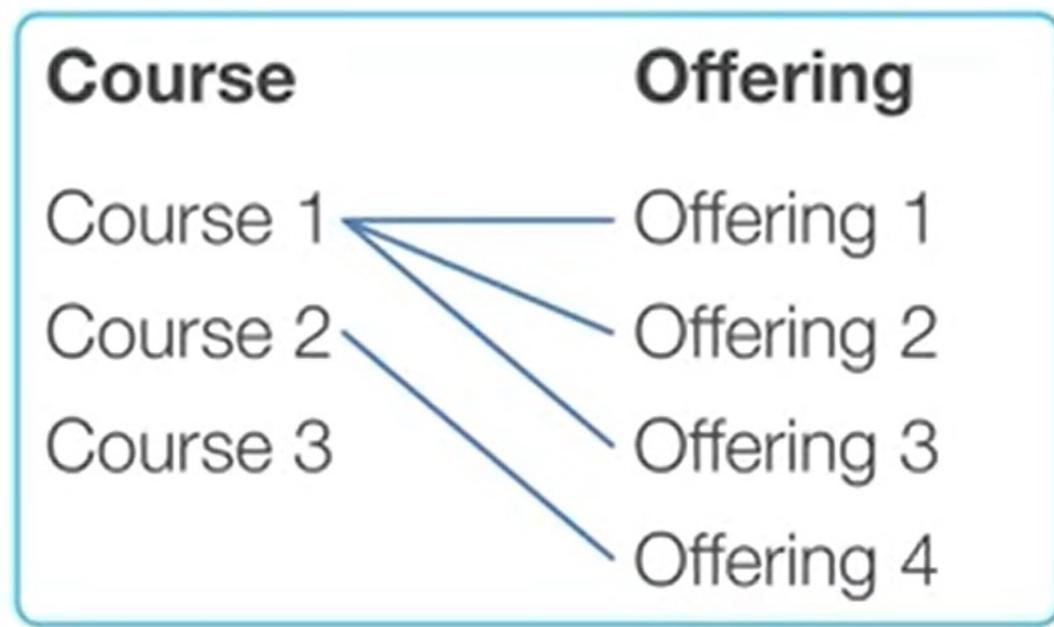
ERD Diagram: Basic Symbols



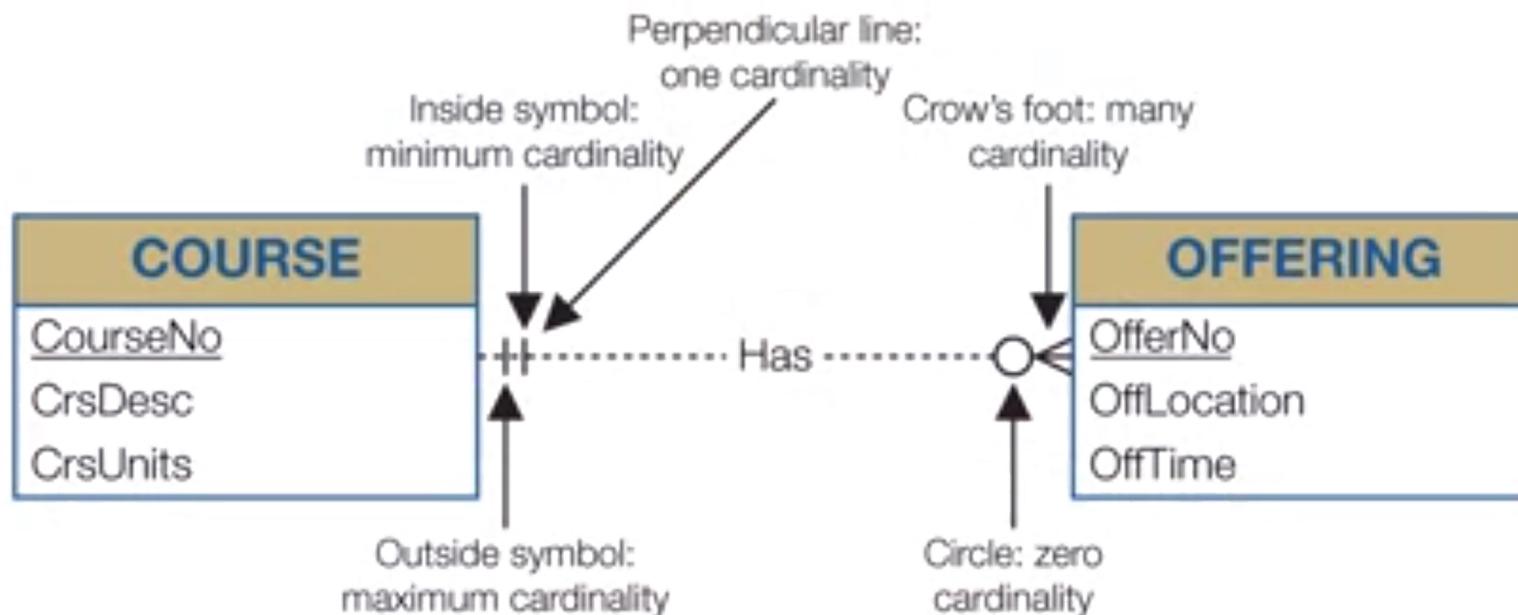
ERD Diagram: Basic Symbols



Cardinalities



Cardinality Notation



Important Cardinalities

Classification	Cardinality Restrictions
Mandatory	Minimum cardinality ≥ 1
Optional	Minimum cardinality = 0
Functional or single-valued	Minimum cardinality = 1
1-M	Maximum cardinality = 1 in one direction; maximum cardinality > 1 in the other direction
M-N	Maximum cardinality > 1 in both directions
1-1	Maximum cardinality = 1 in both directions



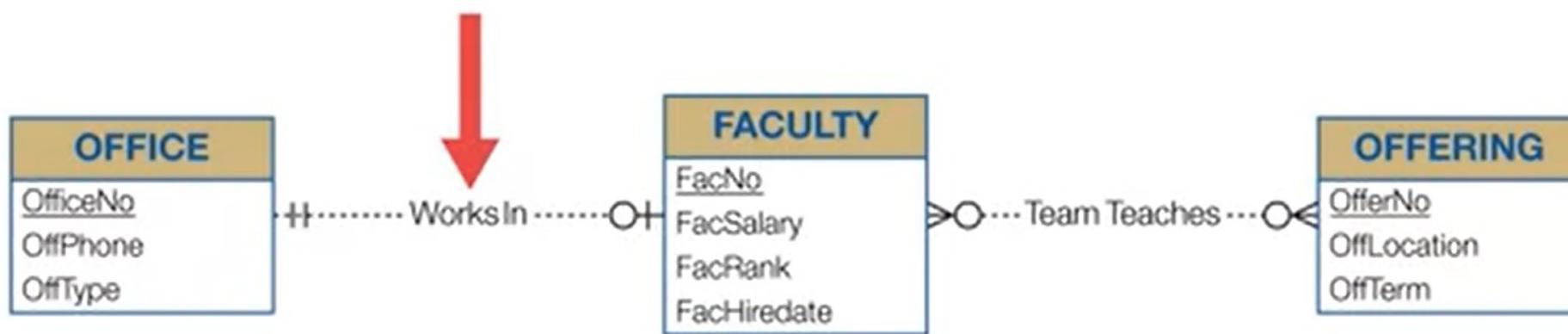
More Relationship Examples

Many-To-Many



More Relationship Examples

One-To-One



Comparison To DB Vendor-Specific Notation

ERD (Crow's Foot)

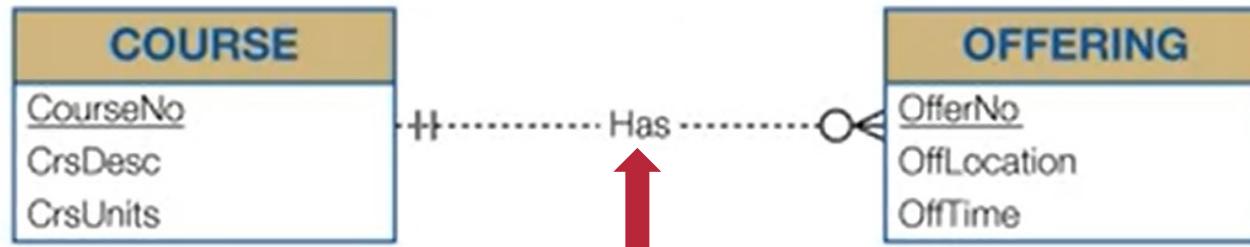


Relational Model Diagram



Comparison To DB Vendor-Specific Notation

ERD (Crow's Foot)



Relational Model Diagram



Comparison To DB Vendor-Specific Notation

ERD (Crow's Foot)



Relational Model Diagram



Comparison To DB Vendor-Specific Notation

ERD (Crow's Foot)



Relational Model Diagram

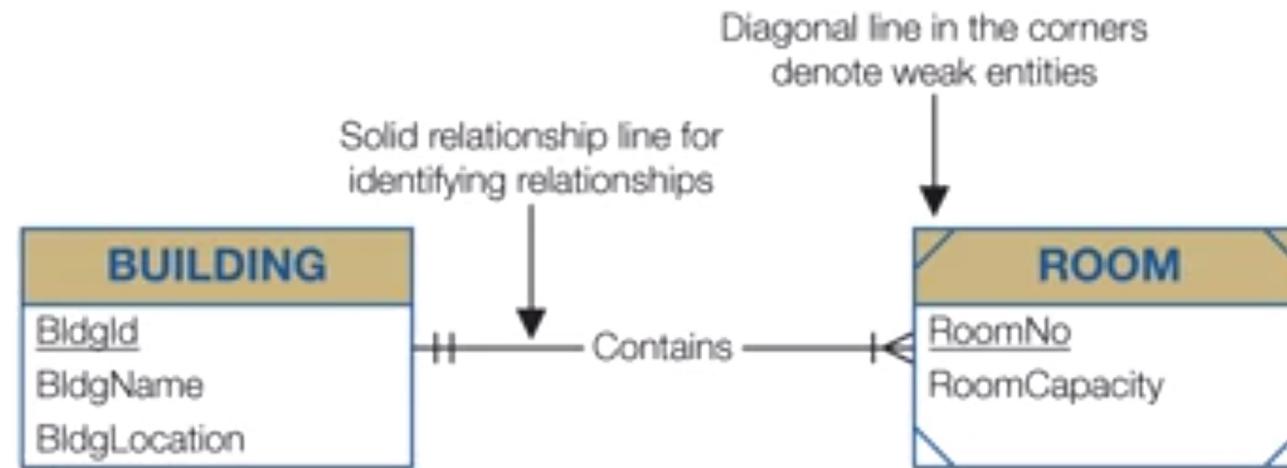


Identification Dependency

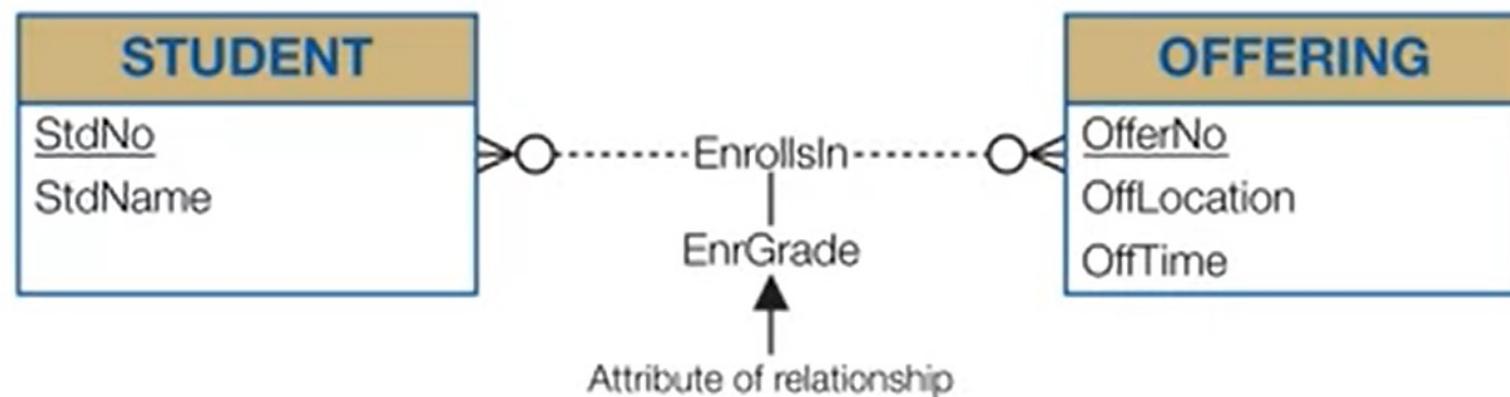
- '*Weak Entity Type*' borrows part or entire primary key.
- '*Identifying Relationship*' provides a component of a primary key for a weak entity type.
- '*Identification Dependency*' is a weak entity type and one or more identifying relationships.



Identification Dependency Symbols



M-N Relationships With Attributes



M-N Relationships With Attributes (cont.)



M-N Relationships With Attributes (cont.)

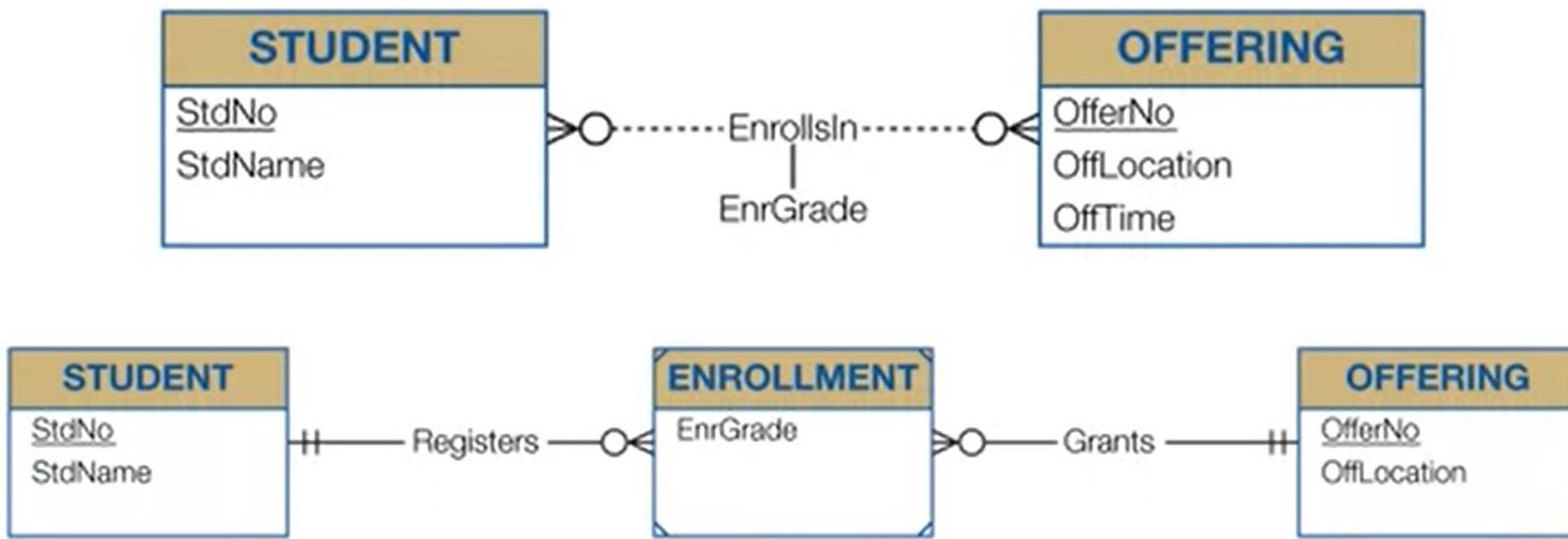


M-N Relationship Equivalency Rule

- Replace M-N relationship
 - Associative entity type
 - Two identifying 1-M relationships



M-N Relationships With Attributes



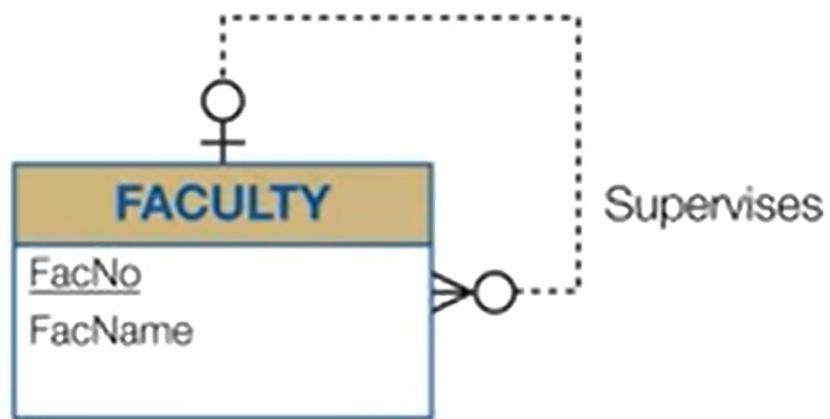
M-N Relationship Equivalency Rule

- Replace M-N relationship
 - Associative entity type
 - Two identifying 1-M relationships
- M-N relationship versus associative entity type
 - Largely preference
 - Associative entity type is more flexible in some situations

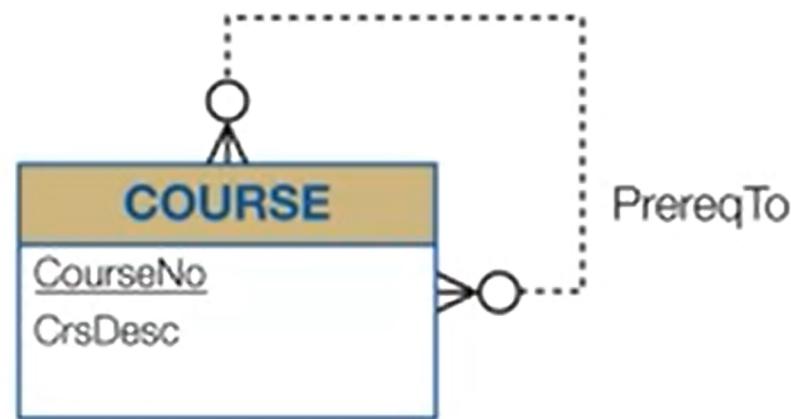


ERD Notation For Self-Referencing Relationships

Manager - Subordinate

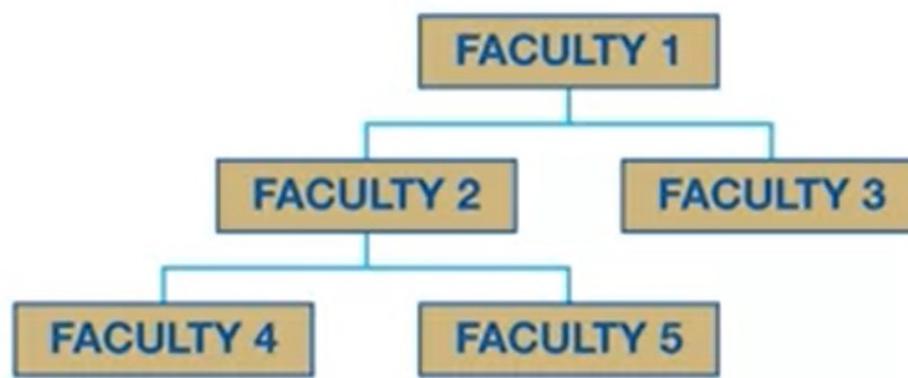


Course Prerequisites

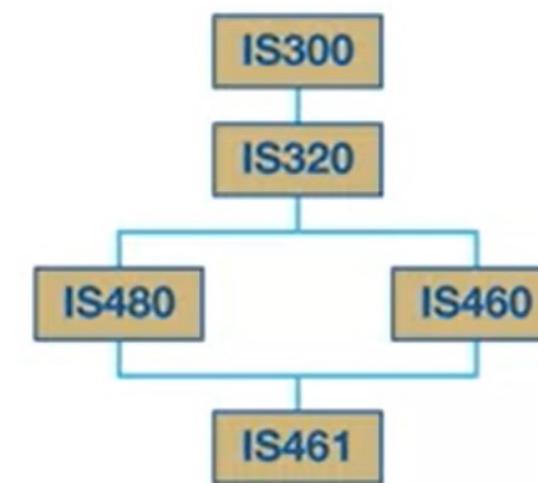


Instance Diagrams For Self-Referencing Relationships

Supervises



PreReqTo



Associative Entity Types For M-Way Relationships

- '*M-Way Relationship*' is an association of more than 2 entity types.
- '*Indirect Representation*' is an associative entity type and a collection of 1-M relationships.
- It is very rare for M greater than 3.



Associative Entity Types For M-Way Relationships

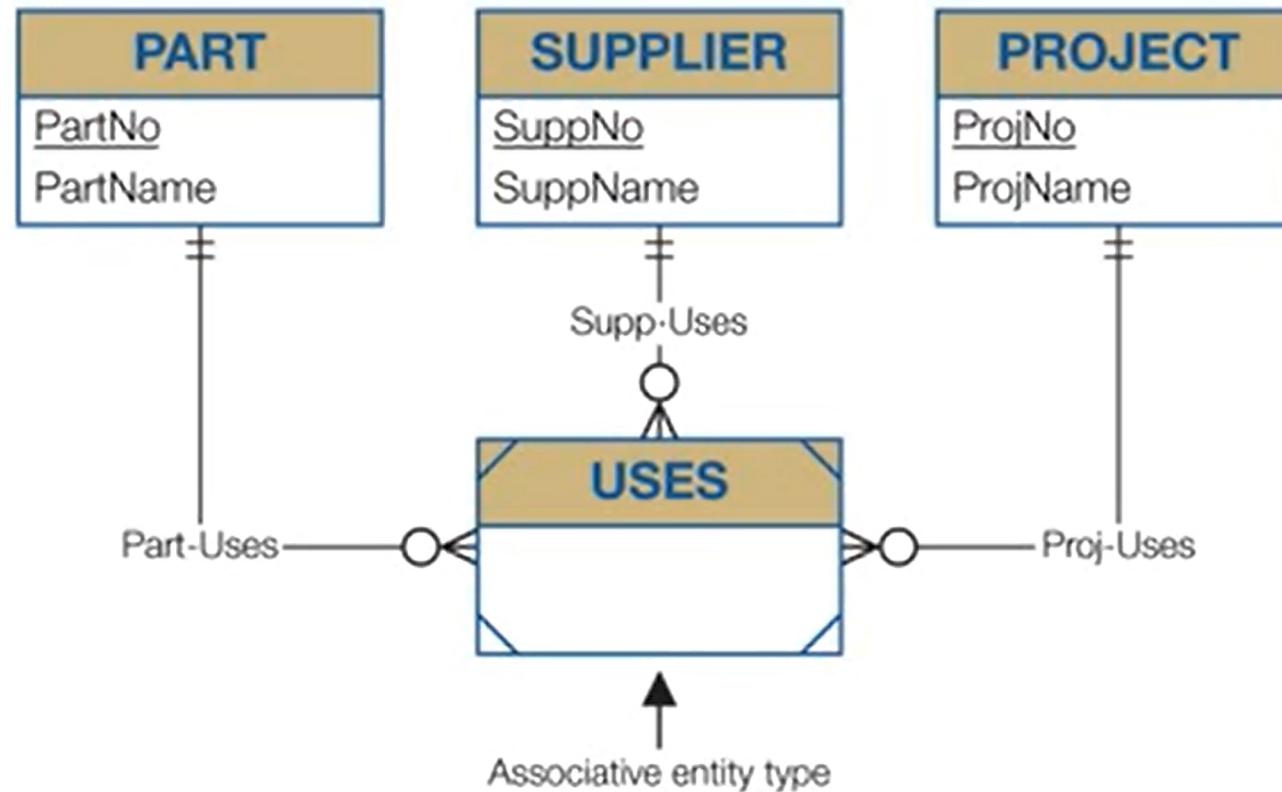


Diagram Rules

- They ensure that ERD notations is correctly used.
- They are similar to syntax rules for a computer language.
- '*Completeness Rules*' ensure no missing specifications.



Primary Key Rule

All entity types have a PK (direct or indirect).



Naming Rule

All entity types, relationships and attributes have a name.



Cardinality Rule

Cardinality is specified in both directions for each relationship.

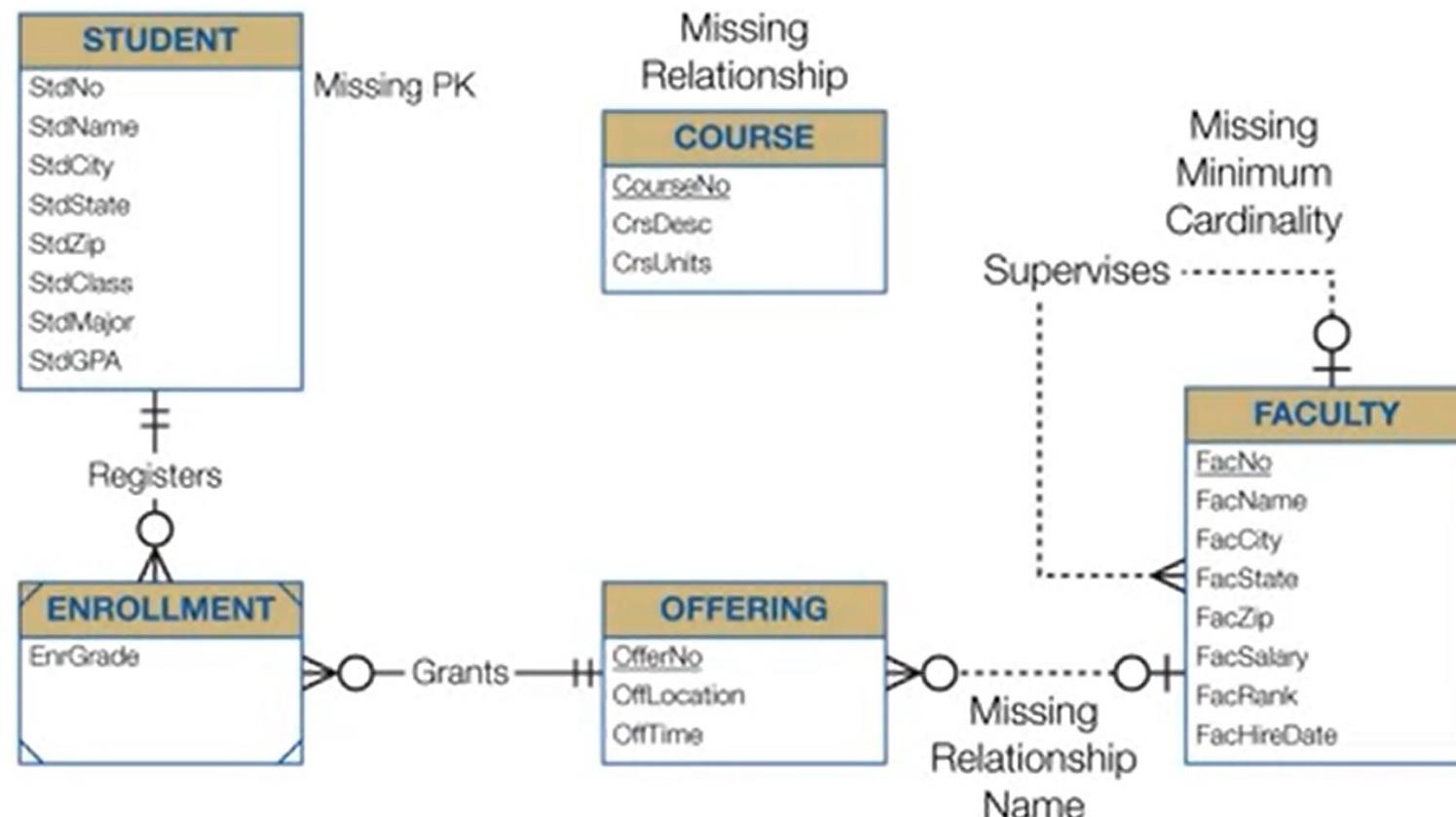


Entity Participation Rule

All entity types participate in at least one relationship.



Completeness Rule Violations



Primary Key Rule Issue

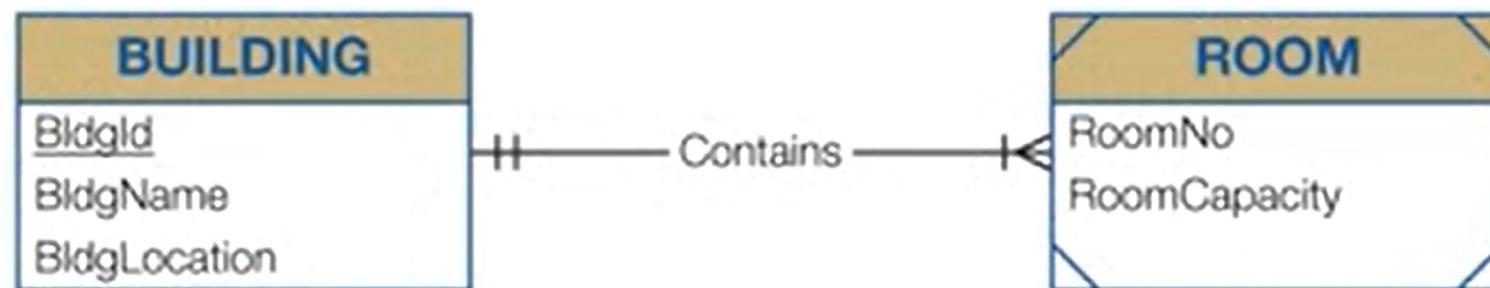
- Primary Key rule is simple in most cases.
- For some weak entity types, the PK rule is subtle
 - Weak entity type with only one 1-M identifying relationship
 - Weak entity type must have a local key to augment the borrowed PK from the parent entity type.
 - Violation of PK rule if local key is missing



PK Rule Violation Example

PK Rule Violation

- A single 1-M identifying relationship
- Room does not have a local key



Attribute Name Rule

Attribute names must be unique within each entity type and relationship.



Unique Relationships Names

Unique Relationship
Names



Easy to Distinguish
Relationships



Broad Goals of Database Development

- Develop a Common Vocabulary
- Define Business Rules
- Ensure Data Quality
- Provide Efficient Implementation



Develop A Common Vocabulary

- Diverse groups of users
- Difficult to obtain acceptance of a common vocabulary
- Compromise to find least objectionable solution
- Unify organisation by establishing a common vocabulary

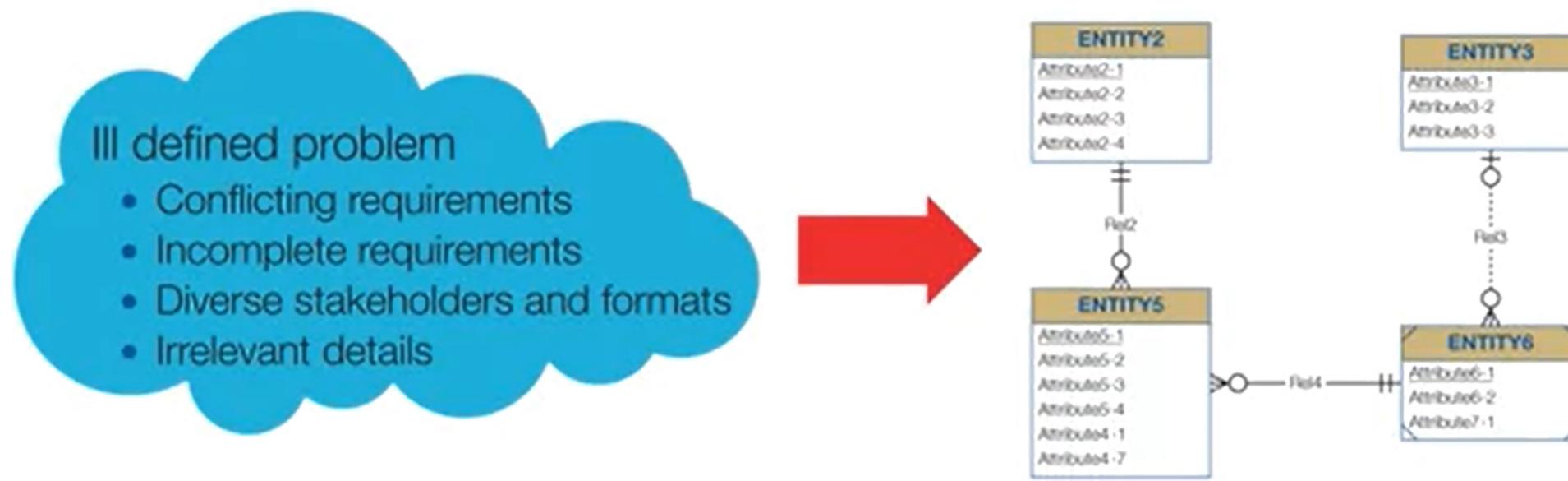


Define Business Rules

- Support Organisational Policies
- Determine restrictiveness
 - *'Too Restrictive'* will reject valid business interactions.
 - *'Too Loose'* will allow erroneous business interactions.
- Provide exceptions for flexibility



Conceptual Data Modelling

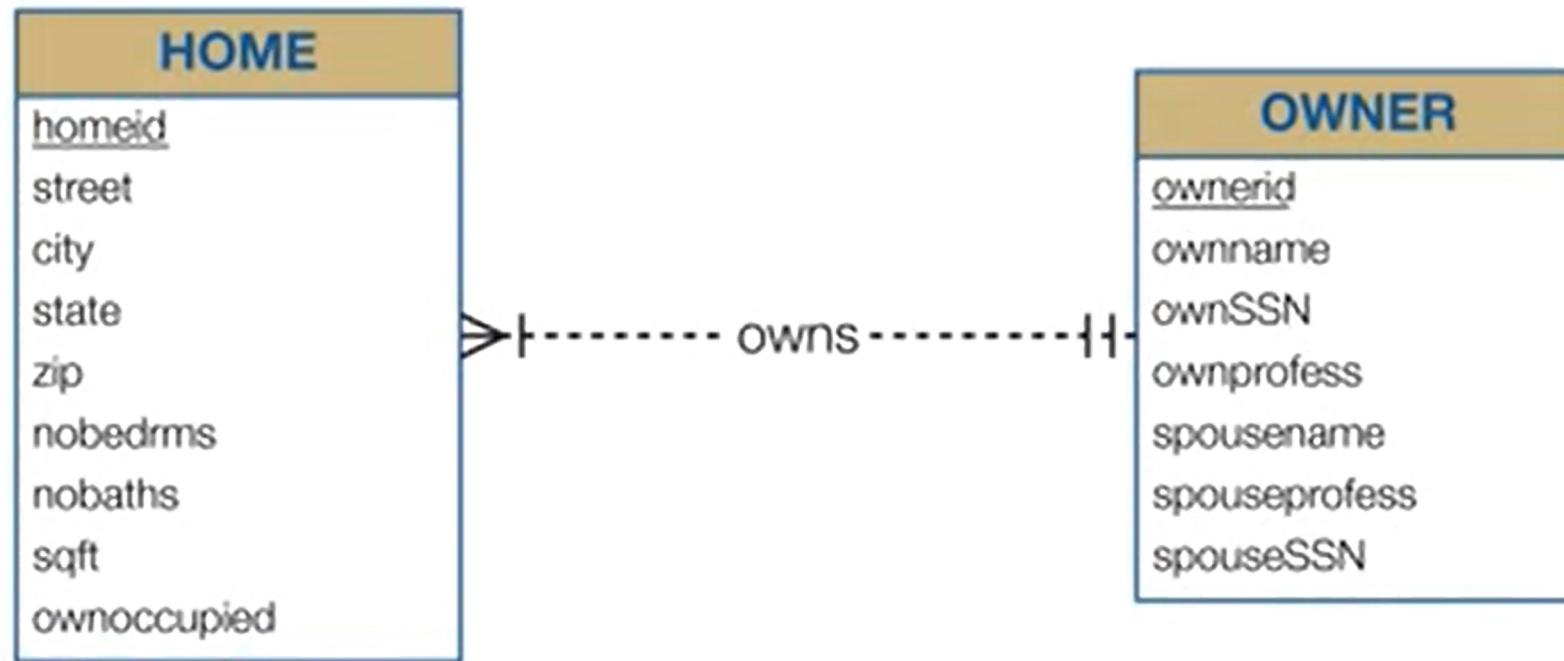


Data Modelling: Problem 1

- Track homes and owners
- Home has unique home identifier, address, characteristics (number of bedrooms, bathrooms, ...) and occupied by owner, tenant or vacant.
- Owner has unique owner number, government identifier, name and spouse details.
- Home has only 1 owner. Owner can possess 1 or more homes.



Problem 1: Solution

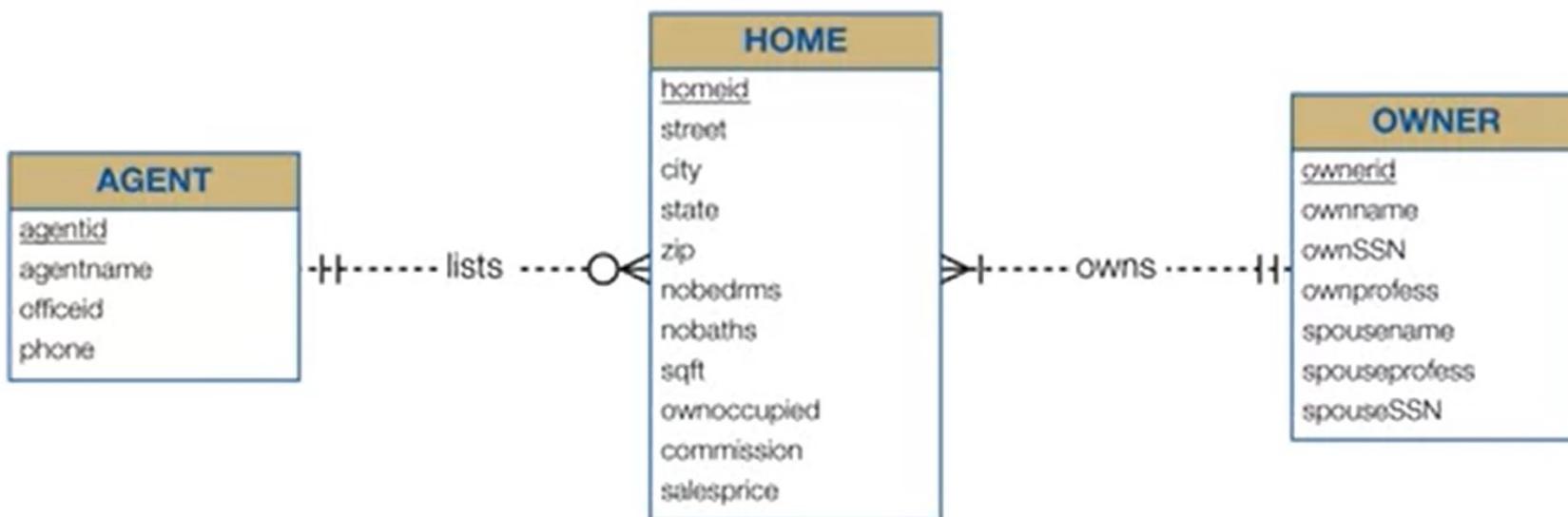


Data Modelling: Problem 2

- Track agents
- Agents represent owners in home sales. An agent can list many homes, but a home is listed by a maximum of one agent.
- Agent has a unique agent identifier, name, office identifier and phone number. details.
- Commission and listing price are determined when a home is listed.



Problem 2: Solution

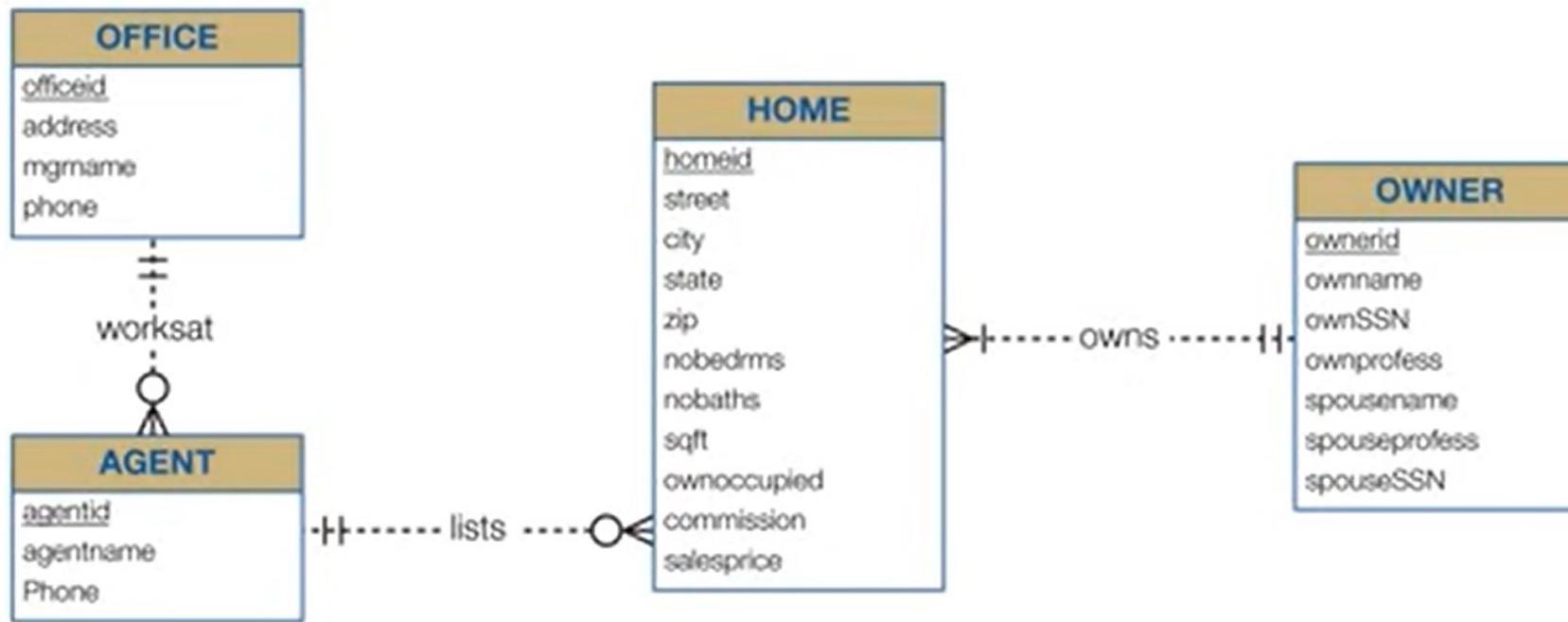


Data Modelling: Problem 3

- Transform office identifier into an entity type
- Office has a phone number, manager name and address.



Problem 3: Solution

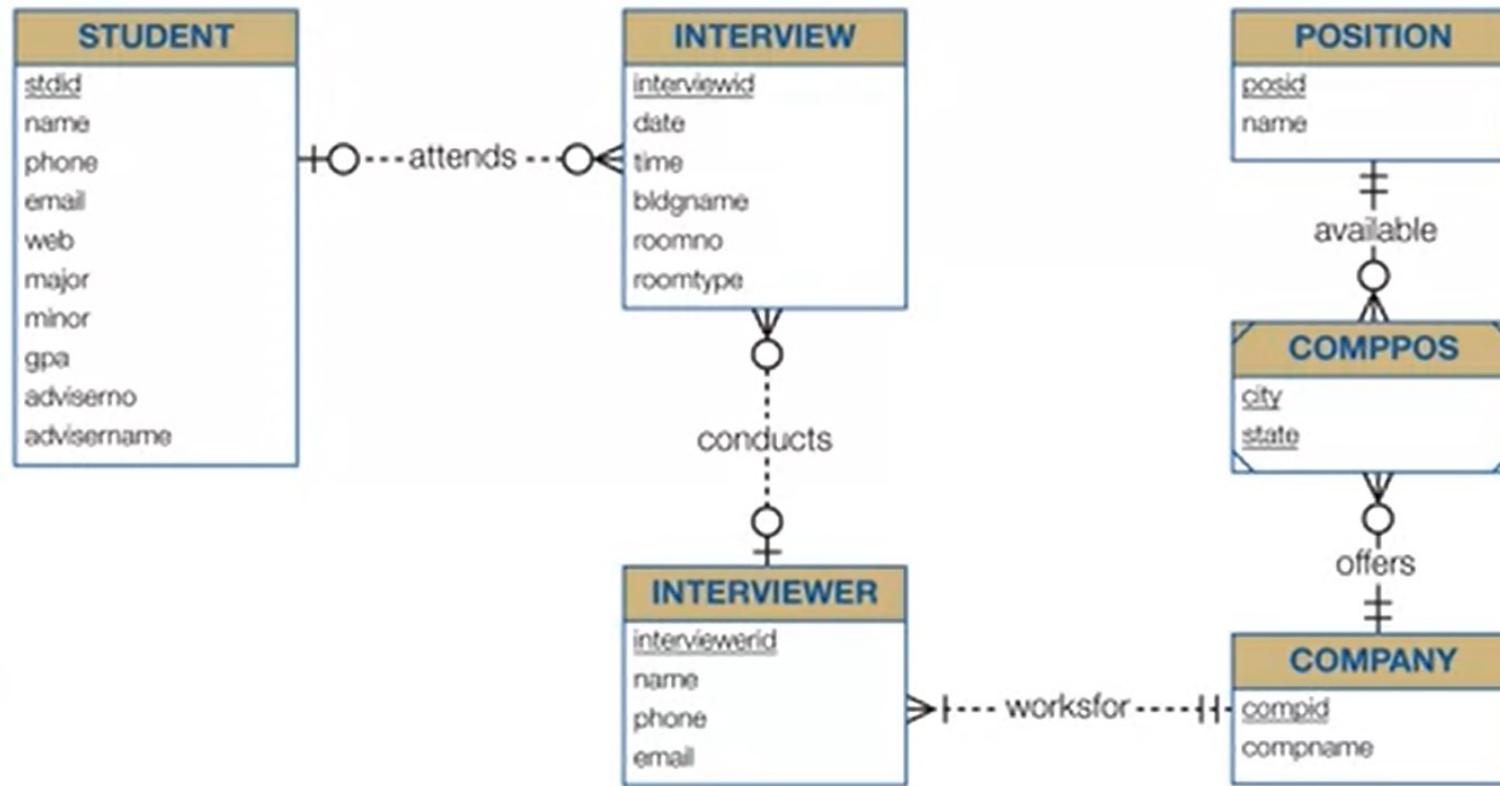


Data Modelling: Problem 4

- Placement office database for interview scheduling and searches
- Major nouns: student, position, company, interviewer, interview



Problem 4: Initial Solution

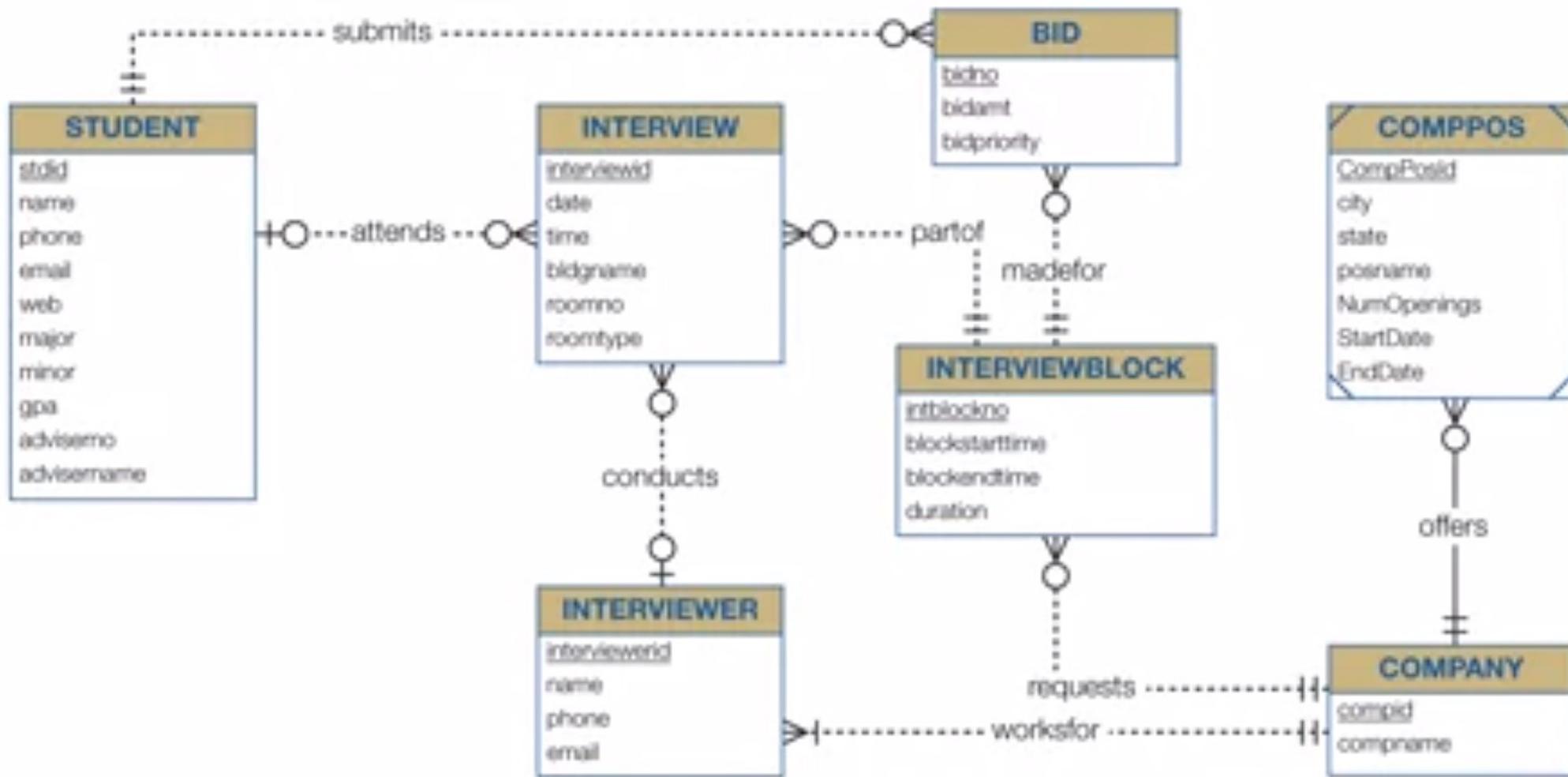


Data Modelling: Problem 4 Revision

- Interview blocks
 - Blocks of time for interviews
 - Companies reserve interview blocks before placement office schedules interviews
- Bids
 - Students can submit many bids for an interview blocks
 - Interview block can receive many bids
- Positions
 - Local to each company
 - No standardised list of positions



Problem 4: Revised Solution



Summary: What We Have Learned

- Crow's Foot ERD notation is widely used. There are subtle different between ERD notation and Relational Data model. ERD is for conceptual modelling. The latter are towards implementation.
- Identification dependency is a specialized existence dependency. Existential dependent entity types has mandatory participation in a relationship. Weak entity types are existential dependent and borrow part or entire primary key.
- The diagram rules are to ensure structural consistency and completeness. Identification dependency is the most common source of errors.
- Data modelling is an essential part of information systems development. The requirements are often ill-defined and ambiguous. The focus is to be on the development goals.
- To solve data modelling problem, we focus on goals of narrative problem analysis especially on unstructured problem statements. The notations are to be used precisely in business data modeling problems.



Q&A

