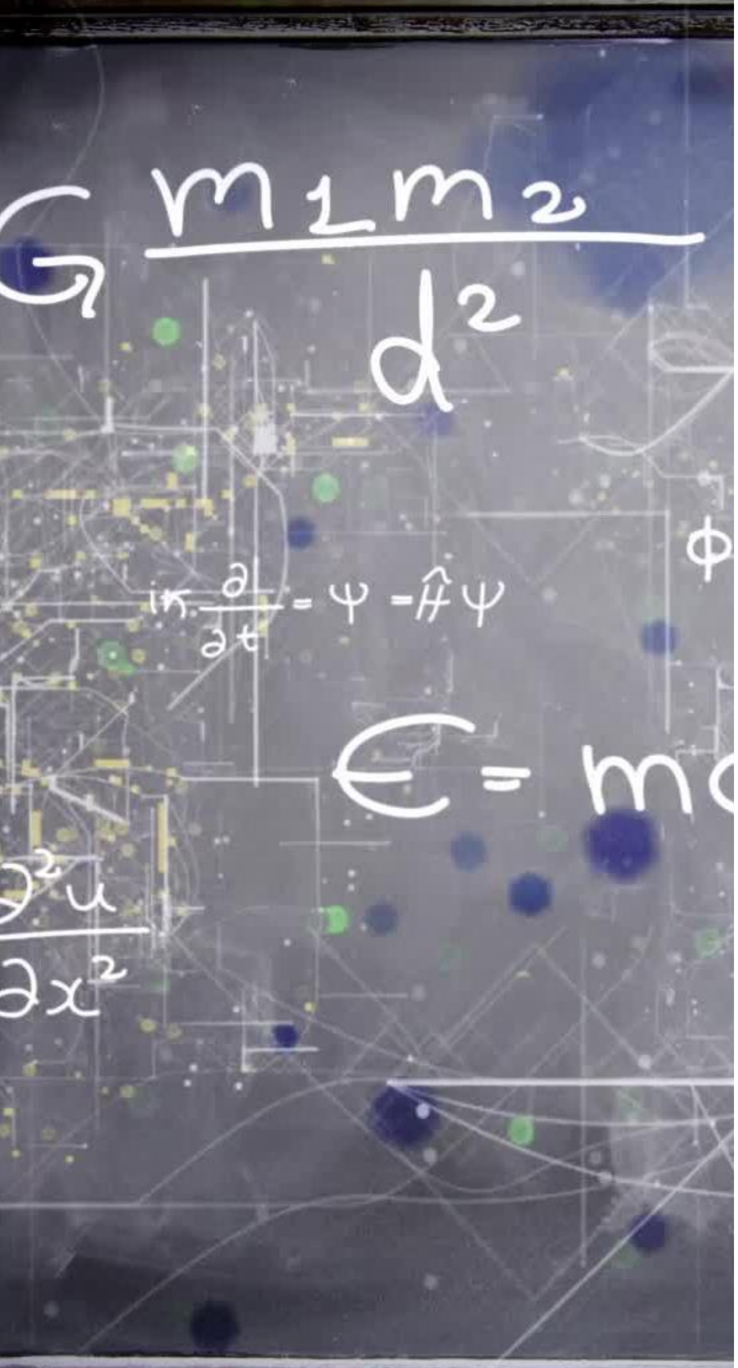
5. Advanced Data Modeling

DATABASE DESIGN & BUSINESS APPLICATION DEVELOPMENT

Learning Objectives

After completing this chapter, you will be able to:

- Describe the main extended entity relationship (EER) model constructs and how they are represented in ERDs and EERDs
- Use entity clusters to represent multiple entities and relationships in an entity relationship diagram (ERD)
- Describe the characteristics of good primary keys and how to select them
- Apply flexible solutions for special data-modeling cases



The Extended Entity Relationship Model (EERM)

Enhanced entity relationship model

- Result of adding more semantic constructs to the original entity relationship (ER) model
- EER diagrams (EERDs) use the EER model

Entity Supertypes and Subtypes

Entity supertype

- Generic entity type related to one or more entity subtypes
- Contains common characteristics

Entity subtype

- Contains unique characteristics of each entity subtype

Criteria to determine usage

- There must be different, identifiable kinds of the entity in the user's environment
- The different kinds of instances should each have one or more attributes that are unique to that kind of instance

Specialization Hierarchy (1 of 2)

Entity supertypes and subtypes are organized in a specialization hierarchy

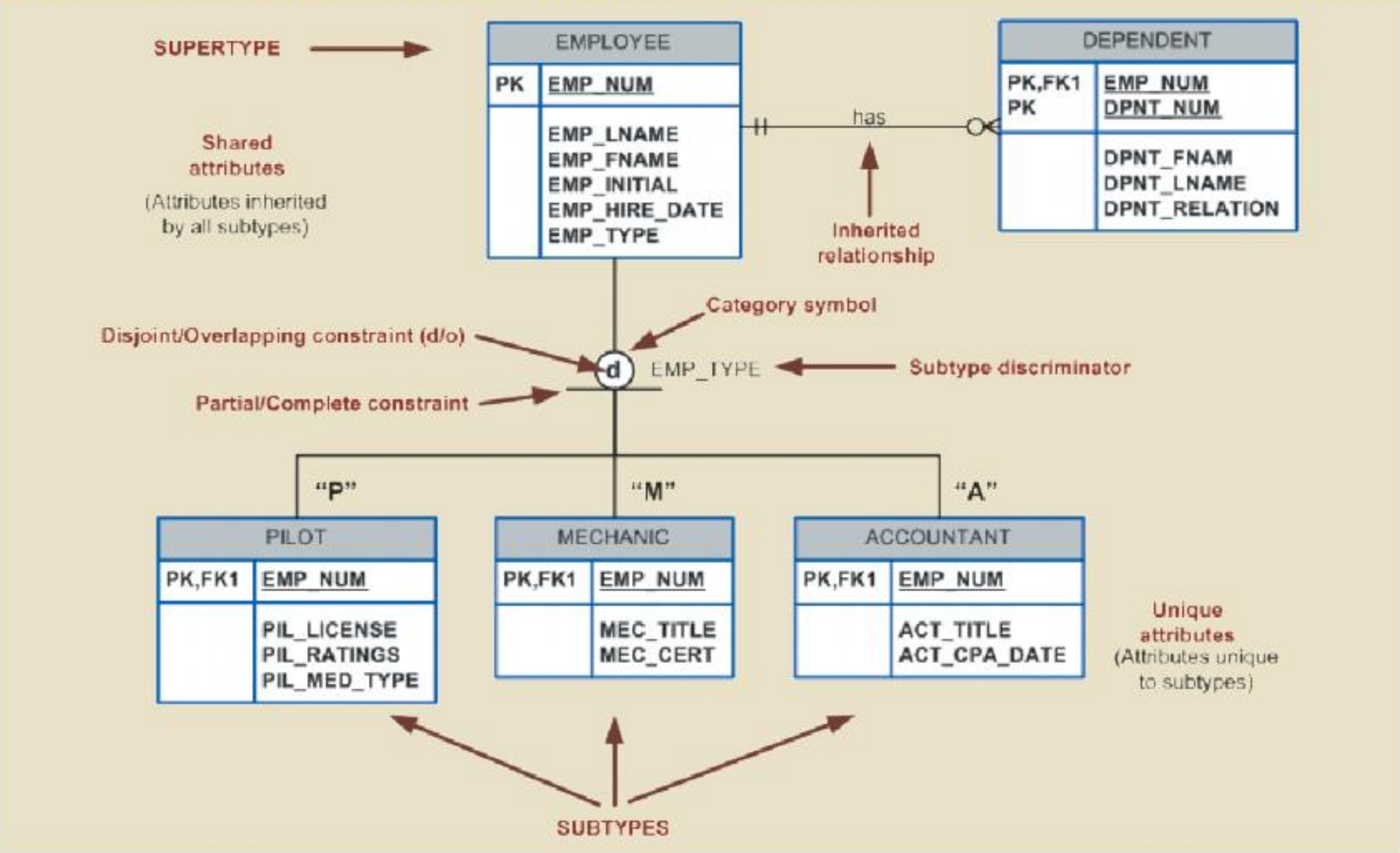
- Depicts arrangement of higher-level entity supertypes and lower-level entity subtypes
- Relationships are described in terms of "is-a" relationships
- Subtype exists within the context of a supertype
- Every subtype has one supertype to which it is directly related
- Supertype can have many subtypes

A specialization hierarchy provides the means to:

- Support attribute inheritance
- Define a special supertype attribute known as the subtype discriminator
- Define disjoint or overlapping constraints and complete or partial constraints

Specialization Hierarchy (2 of 2)

FIGURE 5.2 A SPECIALIZATION HIERARCHY



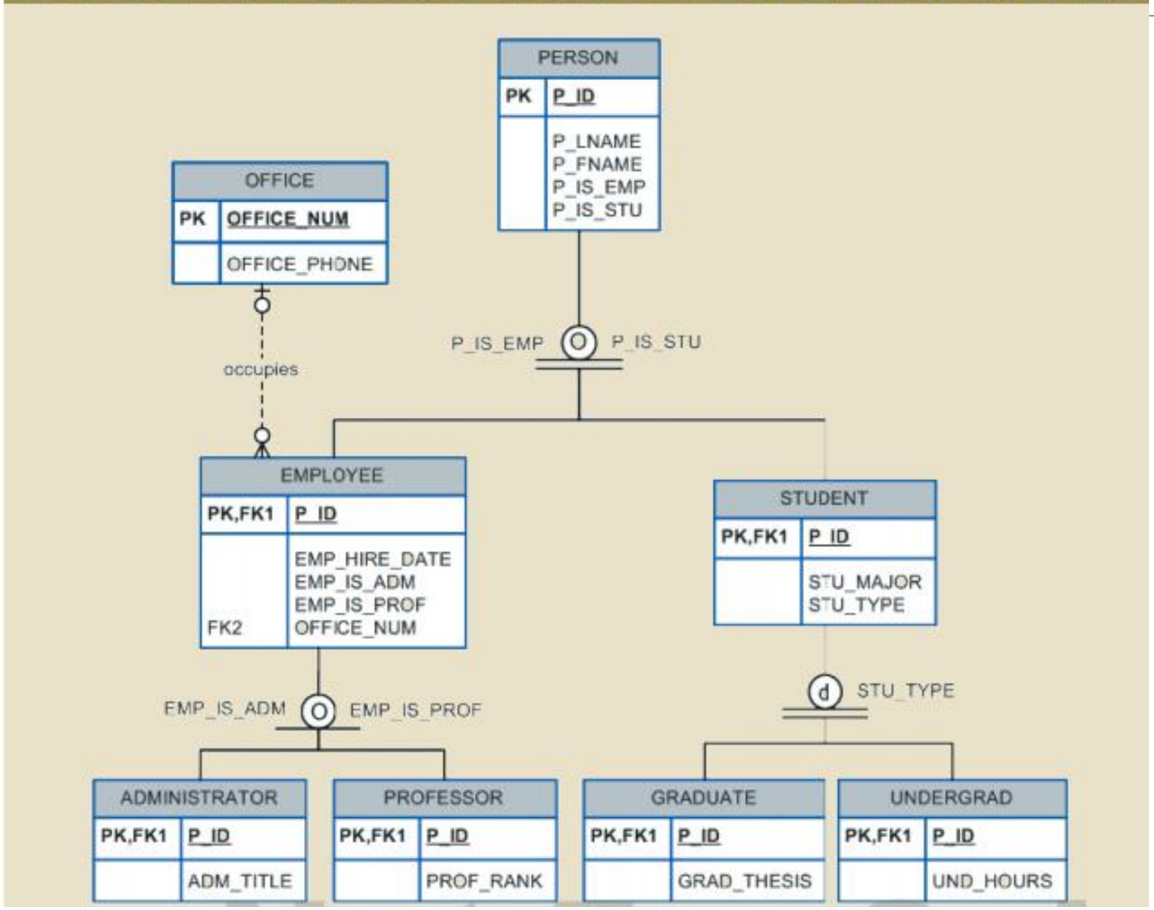
Inheritance (1 of 2)

Enables an entity subtype to inherit attributes and relationships of the supertype

- All entity subtypes inherit their primary key attribute from their supertype
- At the implementation level, supertype and its subtype(s) maintain a 1:1 relationship
- Entity subtypes inherit all relationships in which supertype entity participates
- Lower-level subtypes inherit all attributes and relationships from its upper-level supertypes

Inheritance (2 of 2)

FIGURE 5.4 SPECIALIZATION HIERARCHY WITH OVERLAPPING SUBTYPES



Subtype Discriminator

Attribute in the supertype entity that determines to which entity subtype the supertype occurrence is related

- Default comparison condition is the equality comparison
- In some situations the subtype discriminator is not necessarily based on an equality comparison

Disjoint and Overlapping Constraints (1 of 2)

Disjoint subtypes: contain a unique subset of the supertype entity set

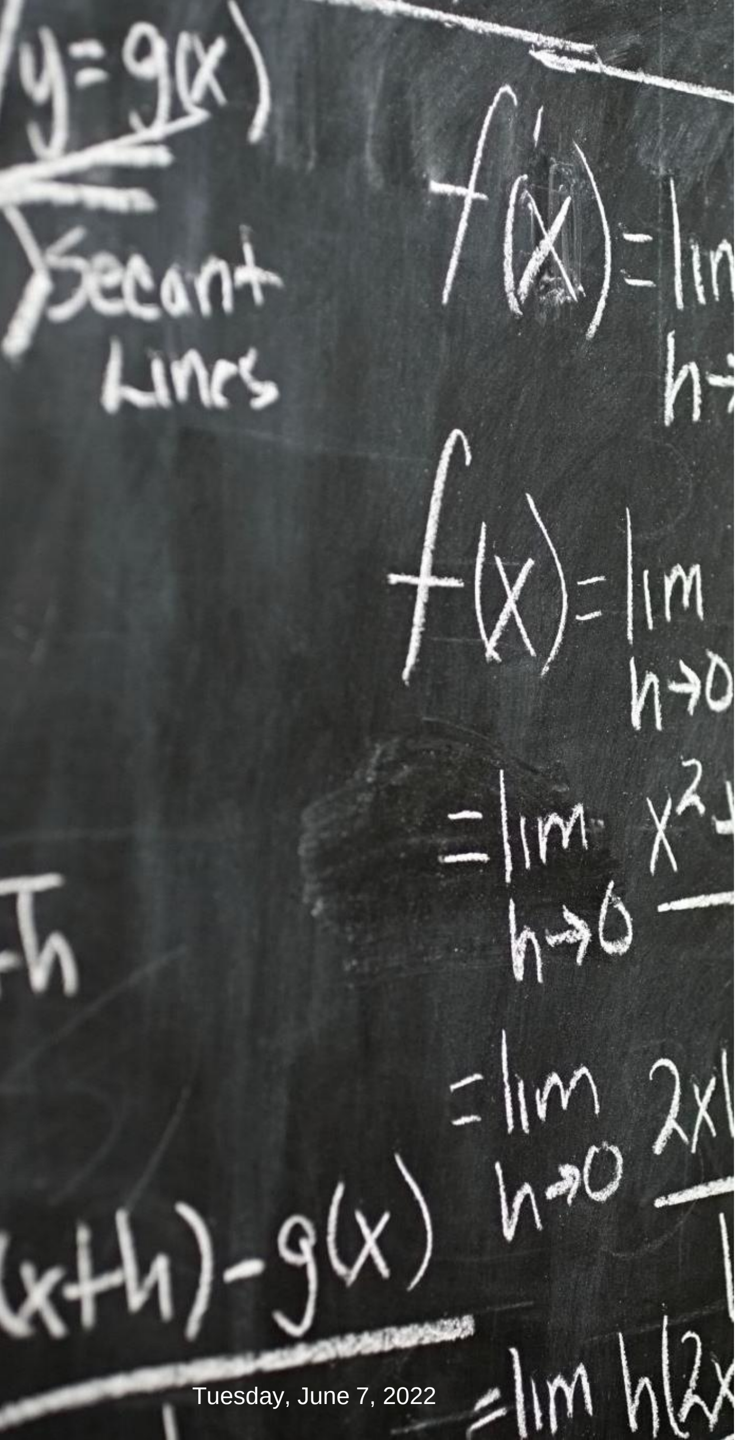
- Known as nonoverlapping subtypes
- Implementation is based on the value of the subtype discriminator attribute in the supertype

Overlapping subtypes: contain nonunique subsets of the supertype entity set

- Implementation requires the use of one discriminator attribute for each subtype

Disjoint and Overlapping Constraints (2 of 2)

TABLE 5.1: Discriminator Attributes with Overlapping Subtypes		
Discriminator Attributes: Professor	**Discriminator Attributes: Administrator**	**Comment**
Y	N	The Employee is a member of the Professor subtype.
N	Y	The Employee is a member of the Administrator subtype.
Y	Y	The Employee is both a Professor and an Administrator.



Completeness Constraint (1 of 2)

Specifies whether each supertype occurrence must also be a member of at least one subtype

- Partial completeness: not every supertype occurrence is a member of a subtype
- Total completeness: every supertype occurrence must be a member of at least one subtypes

Completeness Constraint (2 of 2)

Table 5.2	Specialization Hierarchy Constraint Scenarios	
Type	**Disjoint Constraint**	**Overlapping Constraint**
Partial	Supertype has optional subtypes. Subtype discriminator can be null. Subtype sets are unique.	Supertype has optional subtypes. Subtype discriminators can be null. Subtype sets are not unique.
Total	Every supertype occurrence is a member of only one subtype. Subtype discriminator cannot be null. Subtype sets are unique.	Every supertype occurrence is a member of at least one subtype. Subtype discriminators cannot be null. Subtype sets are not unique.

Specialization and Generalization

Specialization

- Top-down process
- Identifies lower-level, more specific entity subtypes from a higher-level entity supertype
- Based on grouping unique characteristics and relationships of the subtypes

Generalization

- Bottom-up process
- Identifies a higher-level, more generic entity supertype from lower-level entity subtypes
- Based on grouping common characteristics and relationships of the subtypes

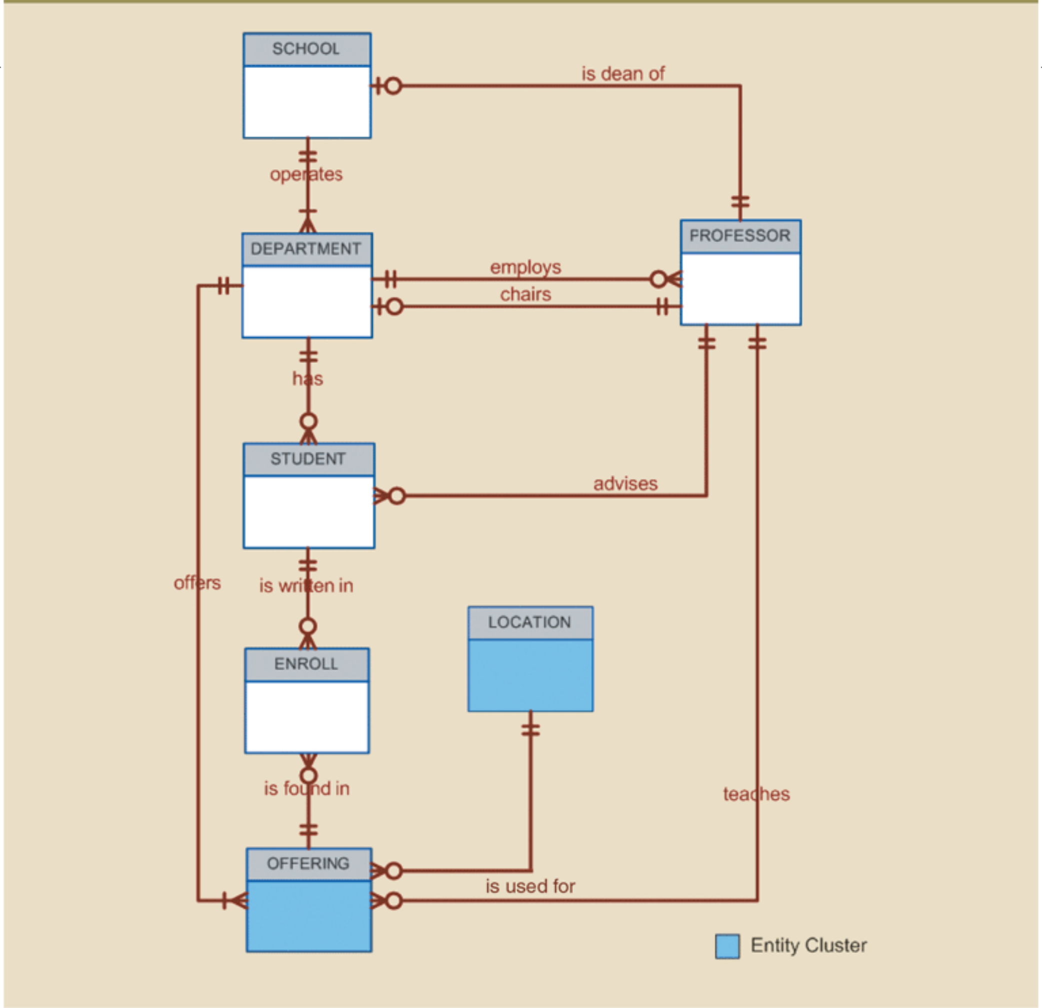
Entity Clustering (1 of 2)

"Virtual" entity type used to represent multiple entities and relationships in ERD

- Formed by combining multiple interrelated entities into a single, abstract entity object
- General rule: avoid the display of attributes to eliminate complications that result when the inheritance rules change

Entity Clustering (2 of 2)

FIGURE 5.6 TINY COLLEGE ERD USING ENTITY CLUSTERS



Entity Integrity: Selecting Primary Keys

Primary keys: single attribute or a combination of attributes

- Uniquely identifies each entity instance
- Guarantees entity integrity
- Works with foreign keys to implement relationships

Natural Keys and Primary Keys

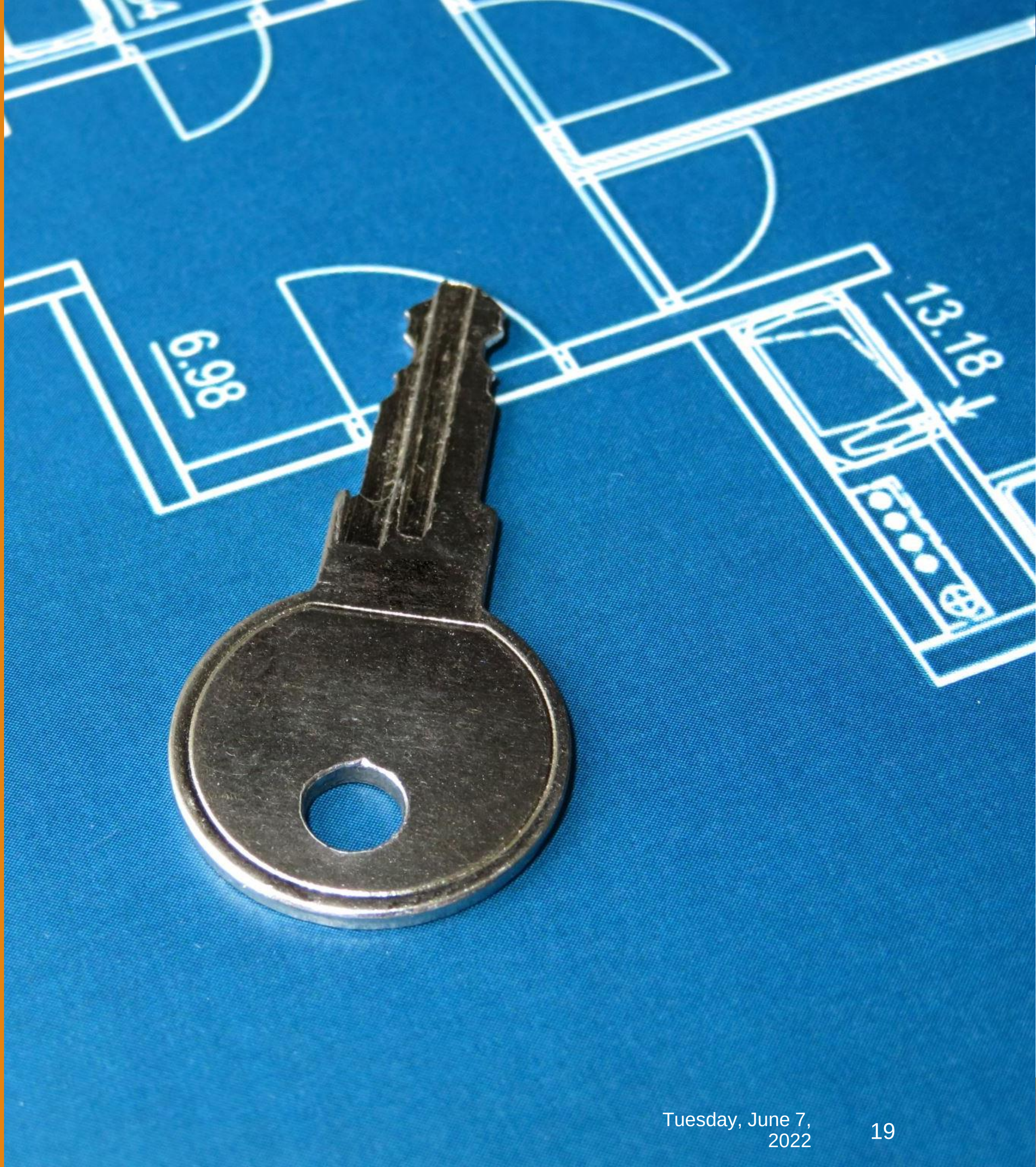
Natural key or natural identifier: real-world identifier used to uniquely identify real-world objects

- Familiar to end users and forms part of their day-to-day business vocabulary
- Used as the primary key of the entity being modeled

Primary Key Guidelines

Desirable primary key characteristics

- Non intelligent
- No change over time
- Preferably single-attribute
- Preferably numeric
- Security-compliant



When to Use Composite Primary Keys (1 of 2)

Identifiers of composite entities

Each primary key combination is allowed once in M:N relationship

Identifiers of weak entities

Strong identifying relationship with the parent entity

Represents a real-world object that is existence-dependent on another real-world object

Represented in the data model as two separate entities in a strong identifying relationship

When to Use Composite Primary Keys (2 of 2)

FIGURE 5.7 THE M:N RELATIONSHIP BETWEEN STUDENT AND CLASS

Database name: Ch05_Tinycollege

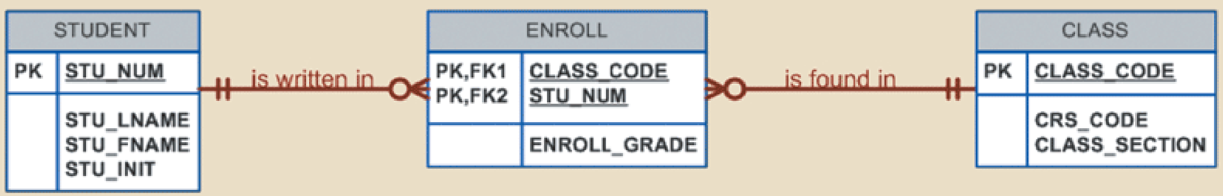


Table name: STUDENT
(first four fields)

STU_NUM	STU_LNAME	STU_FNAME	STU_INIT
321452	Bowser	William	C
324257	Smithson	Anne	K
324258	Brewer	Juliette	
324269	Oblonski	Walter	H
324273	Smith	John	D
324274	Katinga	Raphael	P
324291	Robertson	Gerald	T
324299	Smith	John	B

Table name: ENROLL

CLASS_CODE	STU_NUM	ENROLL_GRADE
10014	321452	C
10014	324257	B
10018	321452	A
10018	324257	B
10021	321452	C
10021	324257	C

Table name: CLASS
(first three fields)

CLASS_CODE	CRS_CODE	CLASS_SECTION
10012	ACCT-211	1
10013	ACCT-211	2
10014	ACCT-211	3
10015	ACCT-212	1
10016	ACCT-212	2
10017	CIS-220	1
10018	CIS-220	2
10019	CIS-220	3
10020	CIS-420	1
10021	QM-261	1
10022	QM-261	2
10023	QM-362	1
10024	QM-362	2
10025	MATH-243	1

When to Use Surrogate Primary Keys (1 of 2)

Primary key used to simplify the identification of entity instances	• Useful when there is no natural key • Helpful if selected candidate key has embedded semantic contents or is too long
Require ensuring that the candidate key of entity in question performs properly	• Use “unique index” and “not null” constraints

When to Use Surrogate Primary Keys (2 of 2)

Table 5.4: Data Used to Keep Track of Events					
DATE	TIME_START	TIME_END	ROOM	EVENT_NAME	PARTY_OF
6/17/2018	11:00 a.m.	2:00 p.m.	Allure	Burton Wedding	60
6/17/2018	11:00 a.m.	2:00 p.m.	Bonanza	Adams Office	12
6/17/2018	3:00 p.m.	5:30 p.m.	Allure	Smith Family	15
6/17/2018	3:30 p.m.	5:30 p.m.	Bonanza	Adams Office	12
6/18/2018	1:00 p.m.	3:00 p.m.	Bonanza	Boy Scouts	33
6/18/2018	11:00 a.m.	2:00 p.m.	Allure	March of Dimes	25
6/18/2018	11:00 a.m.	12:30 p.m.	Bonanza	Smith Family	12

Design Case 1: Implementing 1:1 Relationships (1 of 3)

Foreign keys work with primary keys to properly implement relationships in relational model

- Place primary key of the parent entity on the dependent entity as foreign key

Options for selecting and placing the foreign key

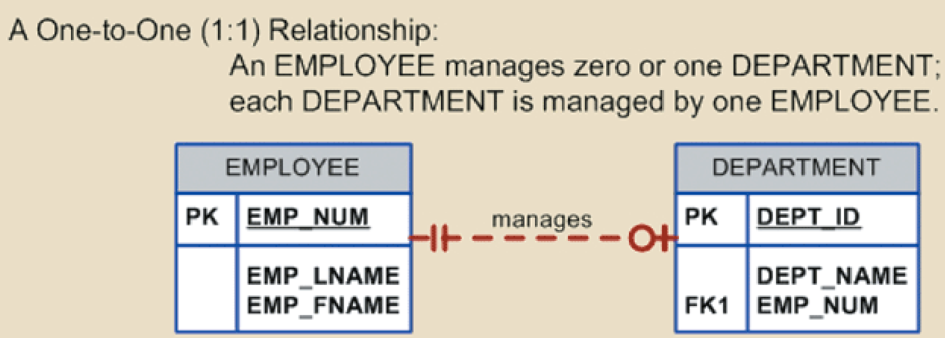
- Place a foreign key in both entities
- Place a foreign key in one of the entities

Design Case 1: Implementing 1:1 Relationships (2 of 3)

Case	ER Relationship Constraints	Action
I	One side is mandatory and the other side	Place the PK of the entity on the mandatory side in the entity on the optional side as a FK, and make the FK mandatory
II	Both sides are optional	Select the FK that causes the fewest nulls, or place the FK in the entity in which the (relationship) role is played
III	Both sides are mandatory	See Case II, or consider revising your model to ensure that the two entities do not belong together in a single entity

Design Case 1: Implementing 1:1 Relationships (3 of 3)

FIGURE 5.8 THE 1:1 RELATIONSHIP BETWEEN DEPARTMENT AND EMPLOYEE



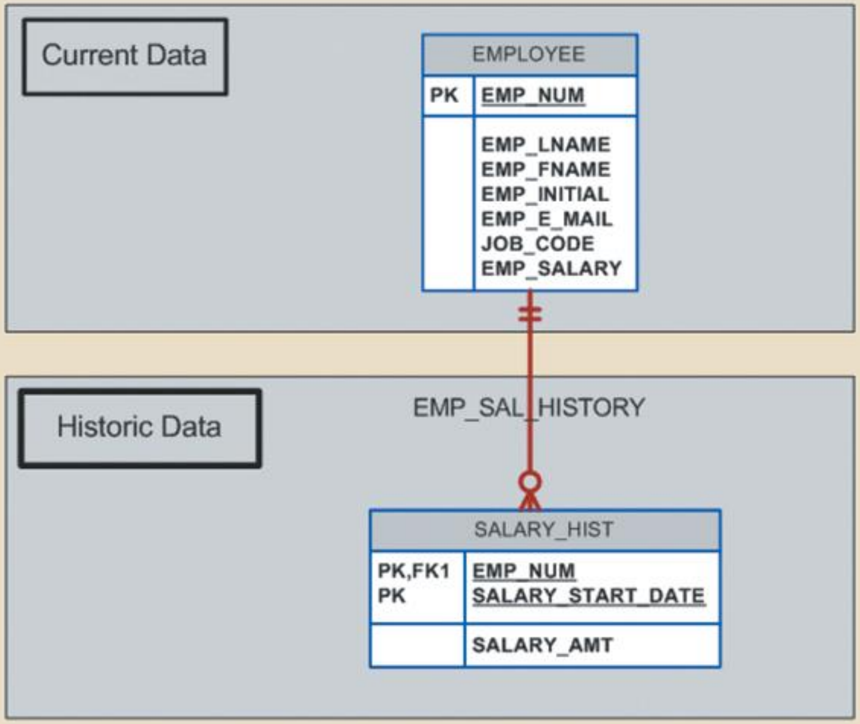
Design Case 2: Maintaining History of Time-Variant Data (1 of 4)

Time-variant data: data whose values change over time and for which a history of the data changes must be retained

- Requires creating a new entity in a 1:M relationship with the original entity
- New entity contains the new value, date of the change, and any other pertinent attribute

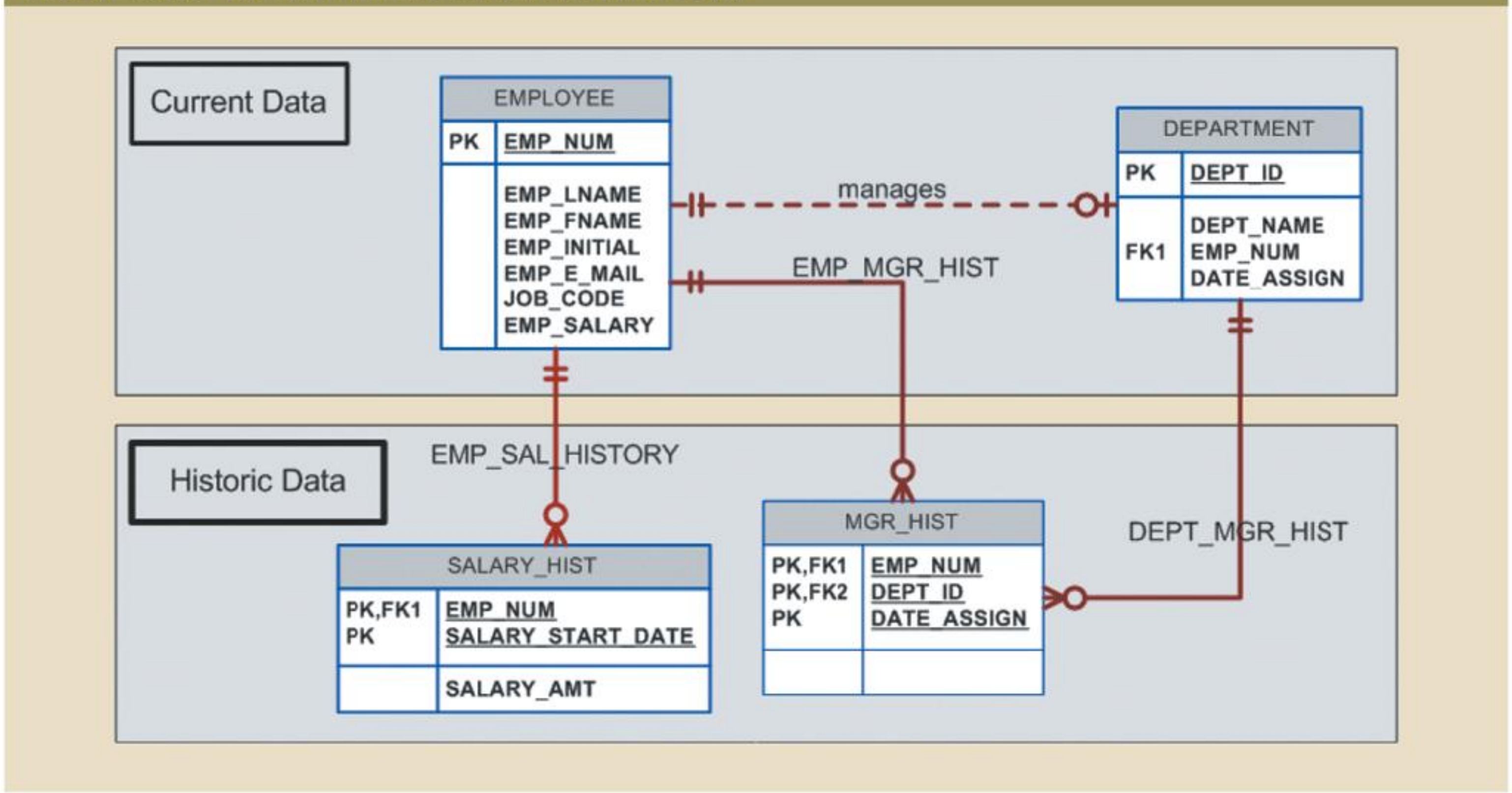
Design Case 2: Maintaining History of Time-Variant Data (2 of 4)

FIGURE 5.9 MAINTAINING SALARY HISTORY



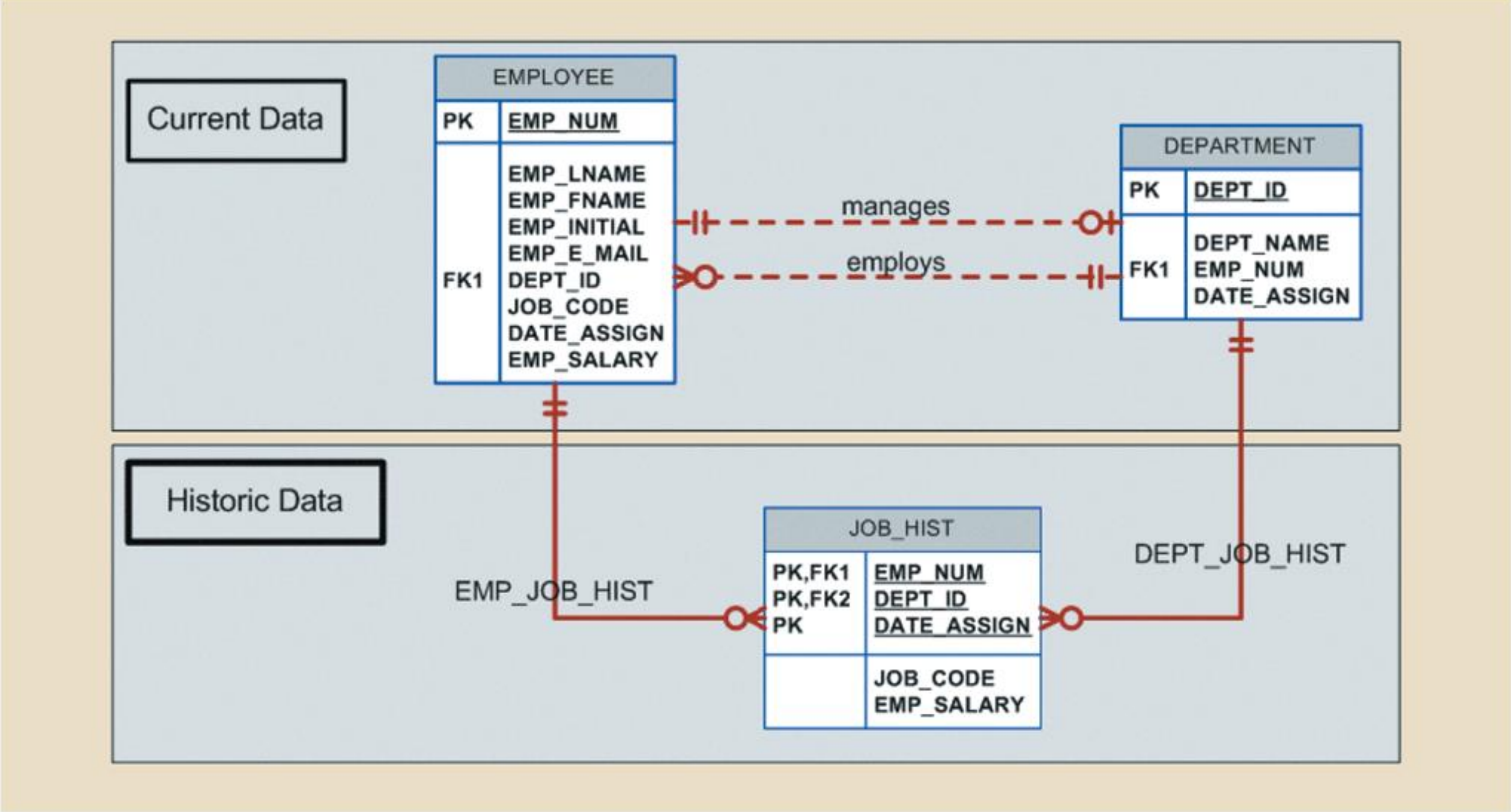
Design Case 2: Maintaining History of Time-Variant Data (3 of 4)

FIGURE 5.10 MAINTAINING MANAGER HISTORY



Design Case 2: Maintaining History of Time-Variant Data (4 of 4)

FIGURE 5.11 MAINTAINING JOB HISTORY



Design Case 3: Fan Traps (1 of 3)

Design trap: occurs when a relationship is improperly or incompletely identified

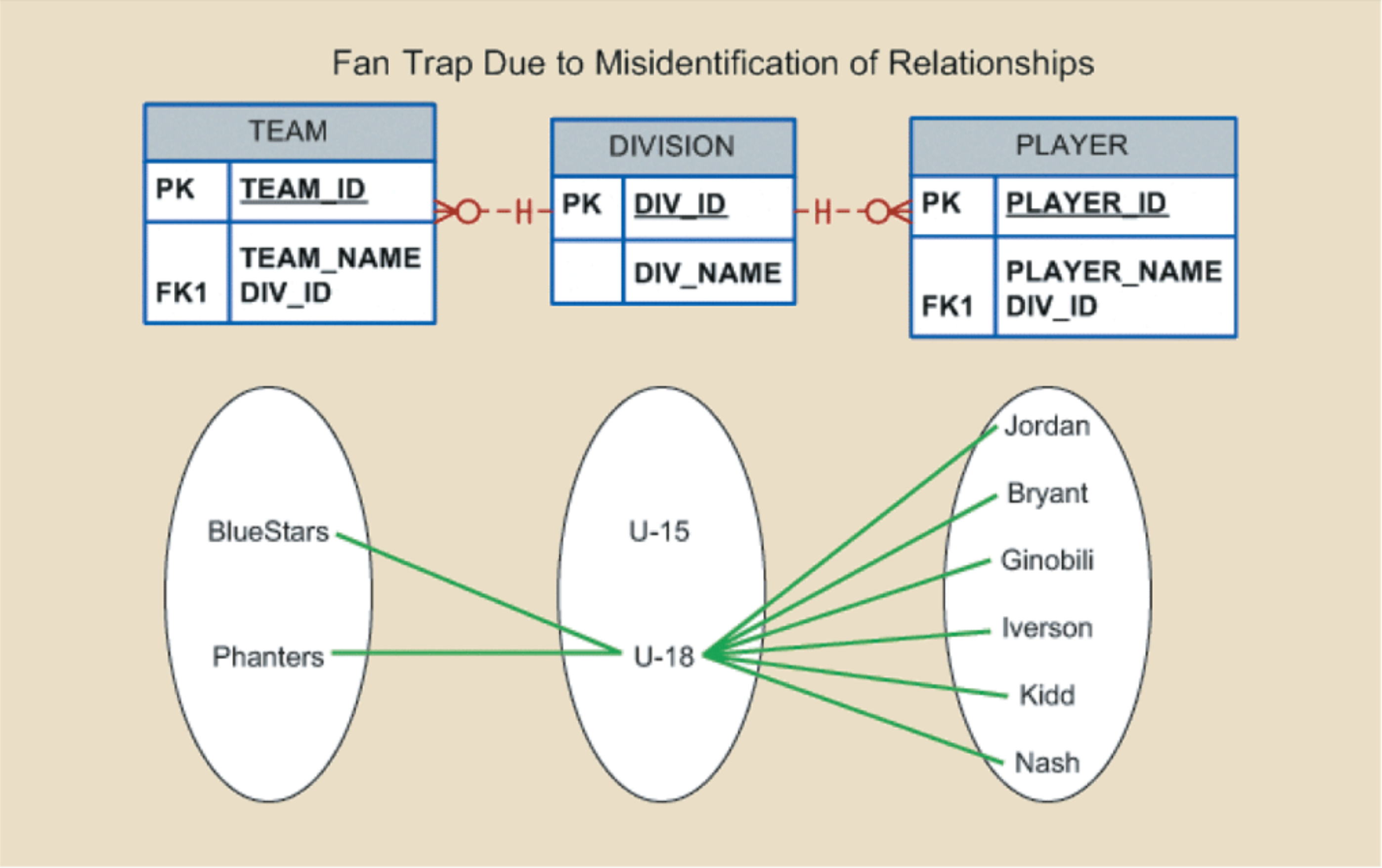
- Represented in a way not consistent with the real world

Fan trap: occurs when one entity is in two 1:M relationships to other entities

- Produces an association among other entities not expressed in the model

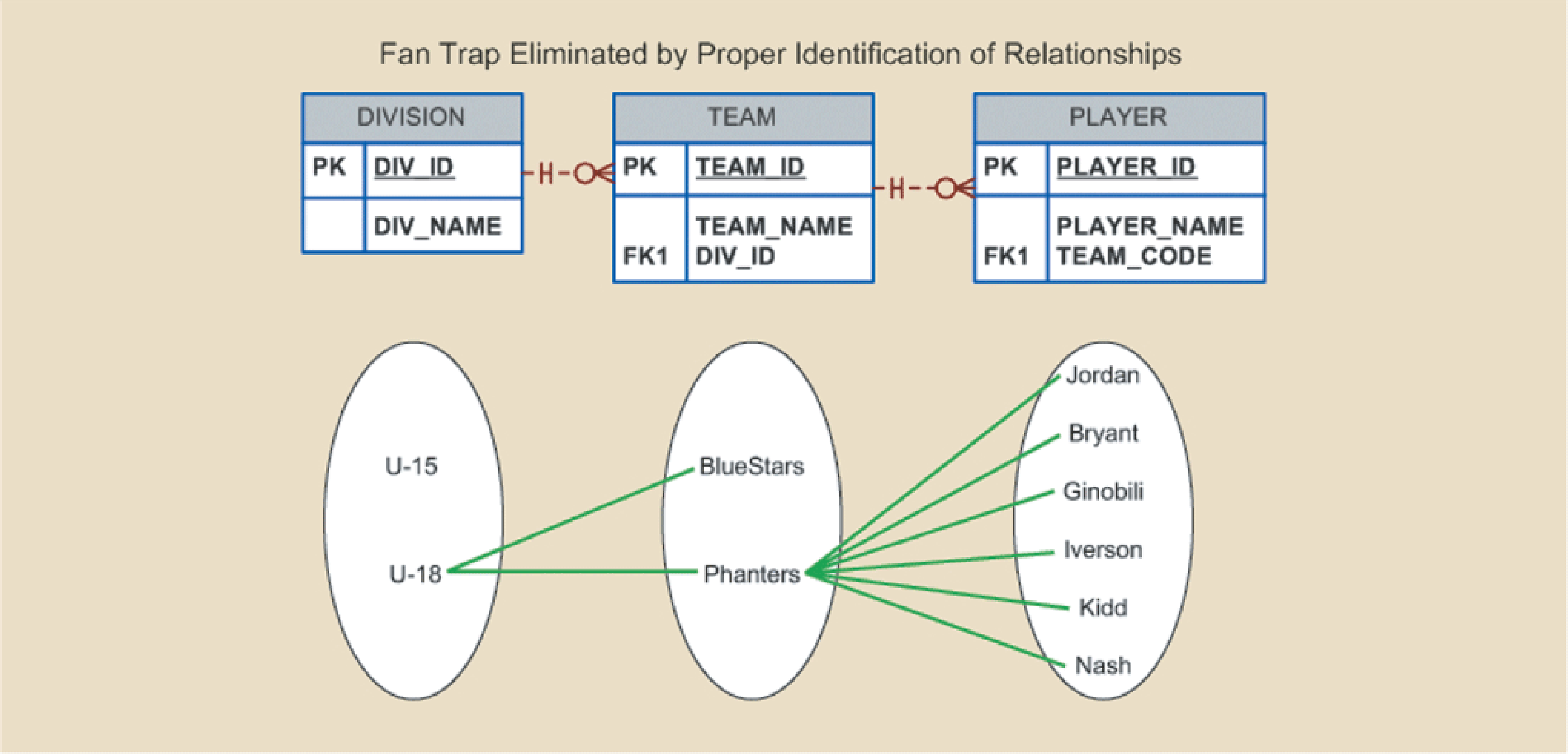
Design Case 3: Fan Traps (2 of 3)

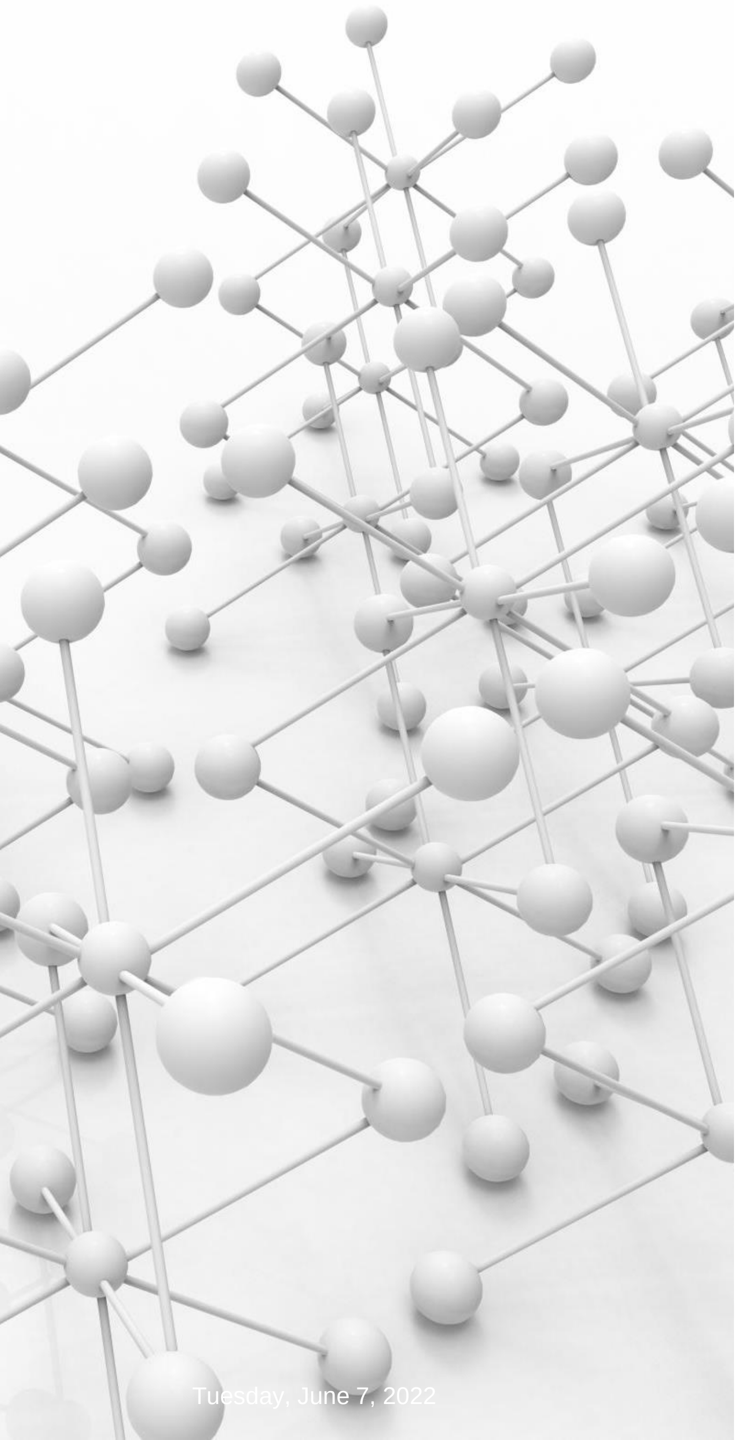
FIGURE 5.12 INCORRECT ERD WITH FAN TRAP PROBLEM



Design Case 3: Fan Traps (3 of 3)

FIGURE 5.13 CORRECTED ERD AFTER REMOVAL OF THE FAN TRAP





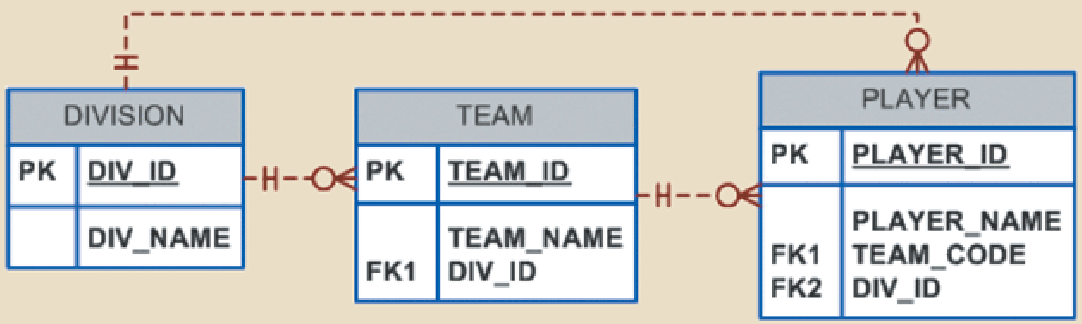
Design Case 4: Redundant Relationships (1 of 2)

Occur when there are multiple relationship paths between related entities

- Must remain consistent across the model
- Help simplify the design

Design Case 4: Redundant Relationships (2 of 2)

FIGURE 5.14 A REDUNDANT RELATIONSHIP



Summary

The extended entity relationship (EER) model adds semantics to the ER model via entity supertypes, subtypes, and clusters

- A specialization hierarchy depicts the arrangement and relationships between entity supertypes and entity subtypes
- An entity cluster is a “virtual” entity type used to represent multiple entities and relationships in the ERD
- Natural keys are identifiers that exist in the real world
- Composite keys are useful to represent M:N relationships and weak (strong identifying) entities
- Surrogate primary keys are useful when there is no natural key that makes a suitable primary key, when the primary key is a composite primary key with multiple data types, or when the primary key is too long to be usable
- Time-variant data refers to data whose values change over time and require that you keep a history of data changes
- A fan trap occurs when you have one entity in two 1:M relationships to other entities, and there is an association among the other entities not expressed in the model

References

Concepts of Database Management, 10th Edition | Lisa Friedrichsen, Lisa Ruffolo, Ellen Monk, Joy Starks, Philip Pratt, Mary Last | ISBN: 978-0357422083 © 2020 | Publisher Course Technology – Cengage Learning

Database Concepts, 9th Edition | David M. Kroenke, David Auer, David J. Auer, Scott L. Vandenberg, Robert C. Yoder | ISBN: 978- 0135188392 © 2020 | Publisher Pearson Education

Database System Concepts 7E, Abraham Silberschatz, Henry F. Korth, S. Sudarshan ©2020 McGraw Hill

DATABASE SYSTEMS Design, Implementation, and Management 13E, Carlos Coronel | Steven Morris, SBN-13: 978-1337627900 © 2018 Cengage Learning, Inc.

Microsoft SQL documentation - https://docs.microsoft.com/en-us/sql/?view=sql-server-ver15

Educational SQL resources - https://docs.microsoft.com/en-us/sql/sql-server/educational-sql-resources?view=sql-server-ver15

SQL Server Technical Documentation - https://docs.microsoft.com/en-us/sql/sql-server/?view=sql-server-ver15

SQL Developer Documentation - https://docs.oracle.com/cd/E12151_01/index.htm