

# **Chapter 2: Entity-Relationship Model**

- Entity Sets
- Relationship Sets
- Design Issues
- Mapping Constraints
- Keys
- E-R Diagram
- Extended E-R Features
- Design of an E-R Database Schema
- Reduction of an E-R Schema to Tables



**Database System Concepts** 

2.1



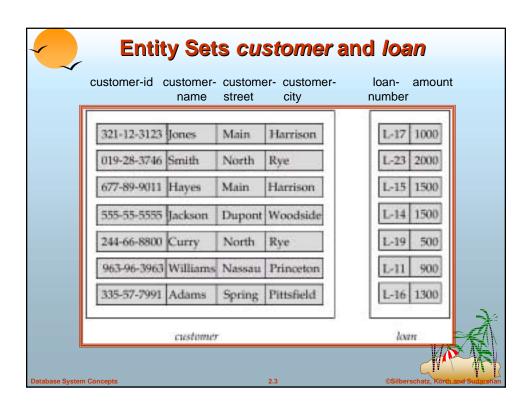
## **Entity Sets**

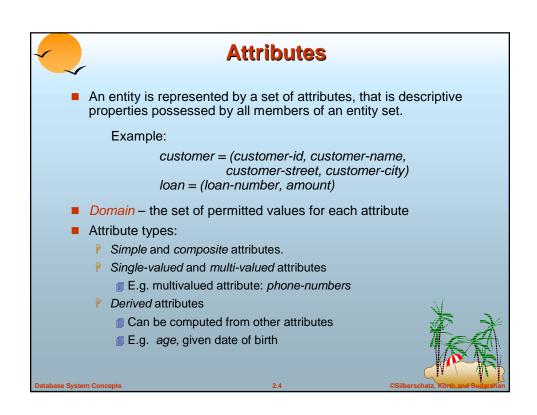
- A database can be modeled as:
  - a collection of entities,
  - relationship among entities.
- An entity is an object that exists and is distinguishable from other objects.
  - Example: specific person, company, event, plant
- Entities have attributes
  - Example: people have names and addresses
- An entity set is a set of entities of the same type that share the same properties.
  - Example: set of all persons, companies, trees, holidays

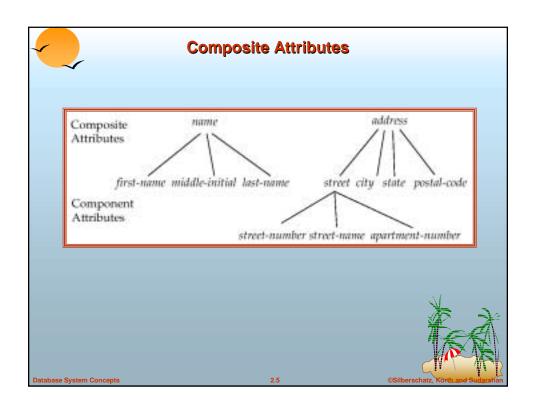


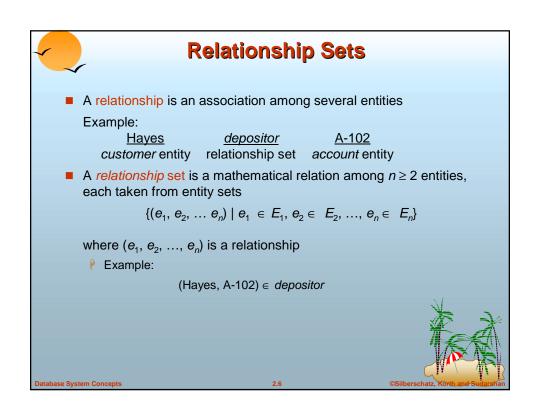
atabase System Concepts

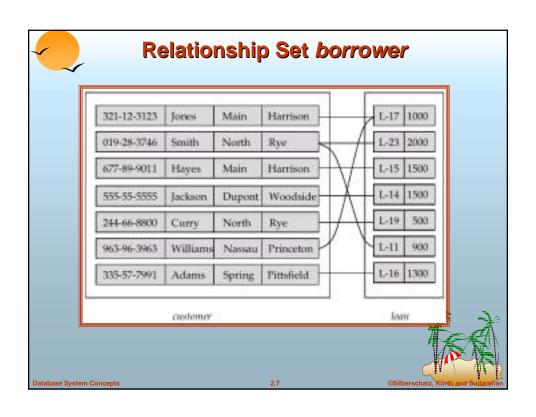
2

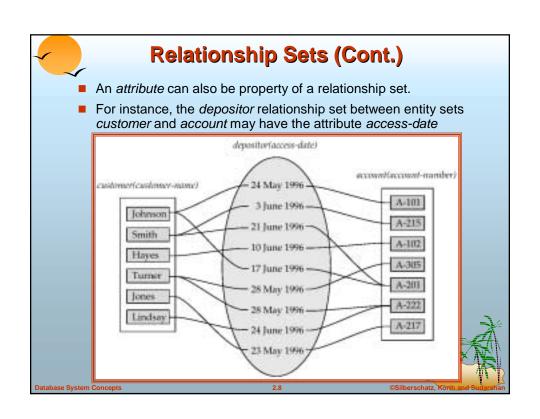














### **Degree of a Relationship Set**

- Refers to number of entity sets that participate in a relationship set.
- Relationship sets that involve two entity sets are binary (or degree two). Generally, most relationship sets in a database system are binary.
- Relationship sets may involve more than two entity sets.
  - Fe.g. Suppose employees of a bank may have jobs (responsibilities) at multiple branches, with different jobs at different branches. Then there is a ternary relationship set between entity sets *employee*, *job and branch*
- Relationships between more than two entity sets are rare. Most relationships are binary. (More on this later.)

**Database System Concepts** 

2.9



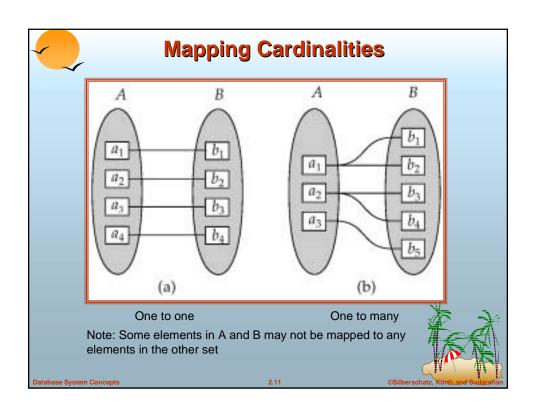


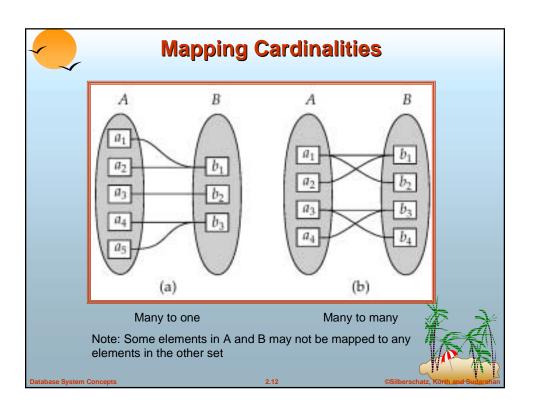
# **Mapping Cardinalities**

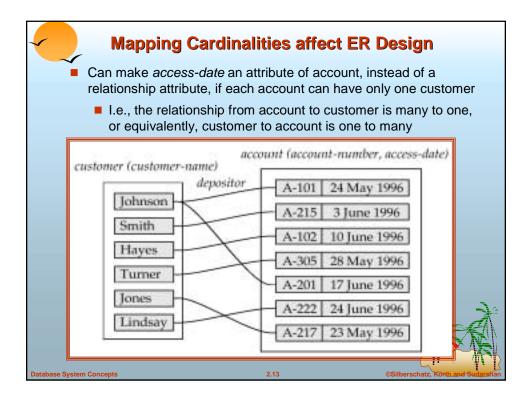
- Express the number of entities to which another entity can be associated via a relationship set.
- Most useful in describing binary relationship sets.
- For a binary relationship set the mapping cardinality must be one of the following types:
  - One to one
  - One to many
  - Many to one
  - Many to many

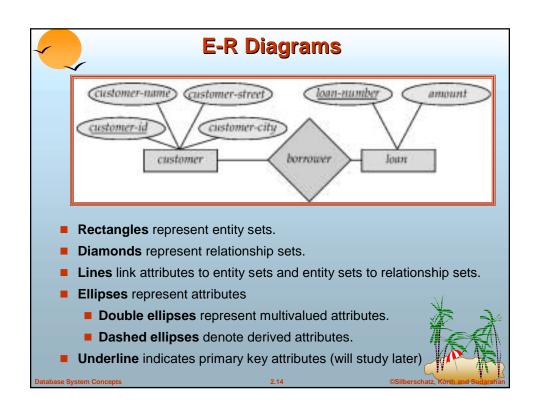


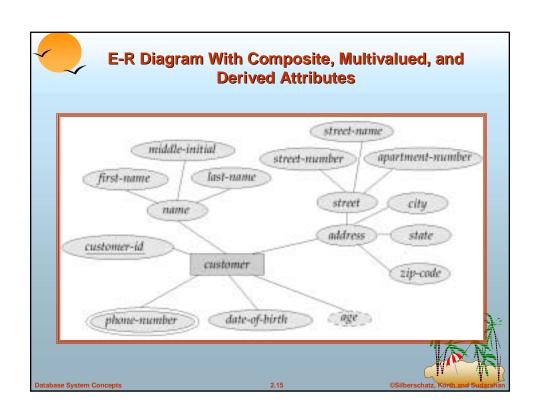
atabase System Concepts

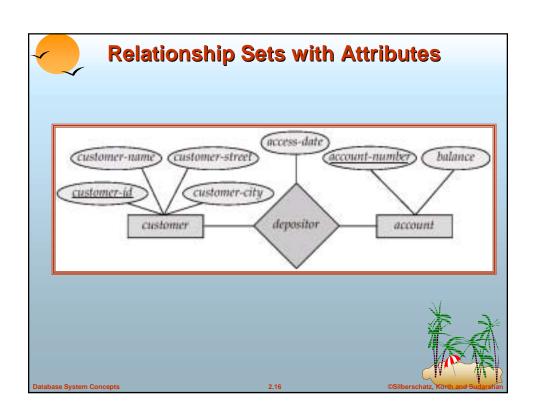








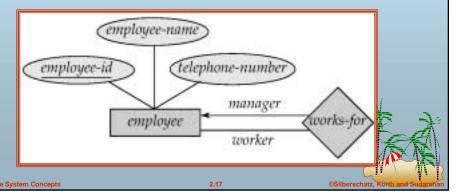






#### **Roles**

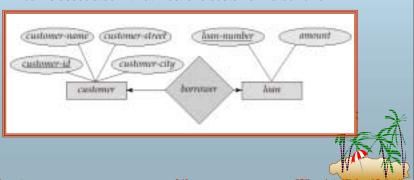
- Entity sets of a relationship need not be distinct
- The labels "manager" and "worker" are called roles; they specify how employee entities interact via the works-for relationship set.
- Roles are indicated in E-R diagrams by labeling the lines that connect diamonds to rectangles.
- Role labels are optional, and are used to clarify semantics of the relationship

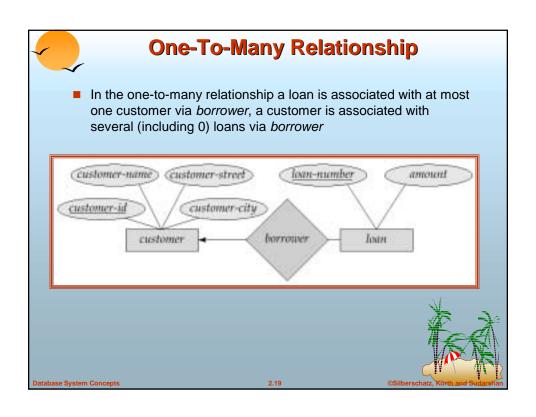


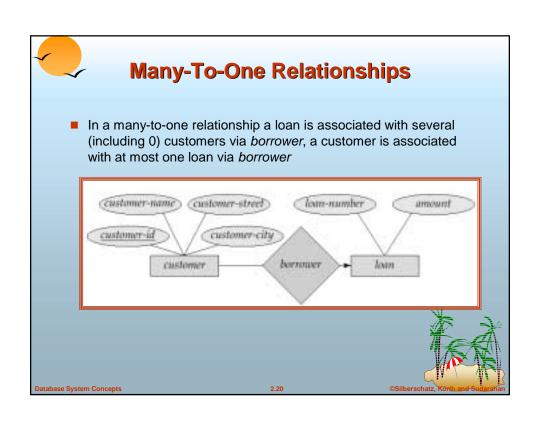


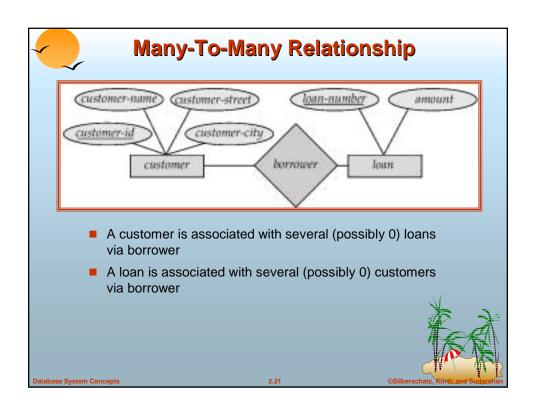
# **Cardinality Constraints**

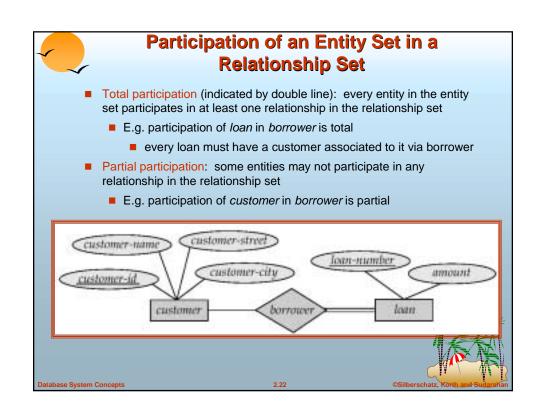
- We express cardinality constraints by drawing either a directed line (→), signifying "one," or an undirected line (—), signifying "many," between the relationship set and the entity set.
- E.g.: One-to-one relationship:
  - A customer is associated with at most one loan via the relationship borrower
  - A loan is associated with at most one customer via borrower

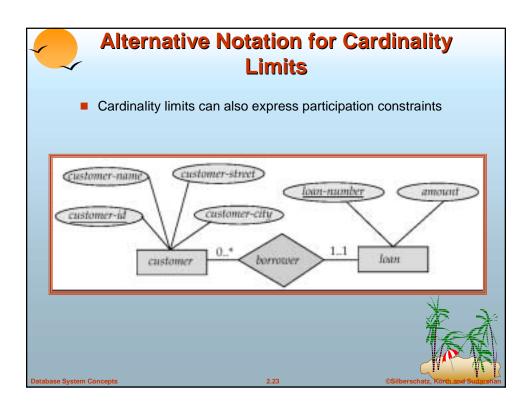


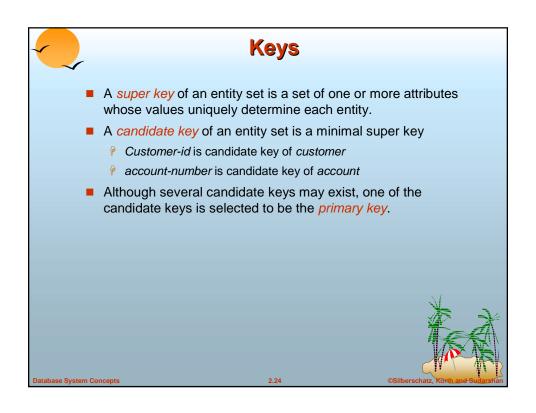














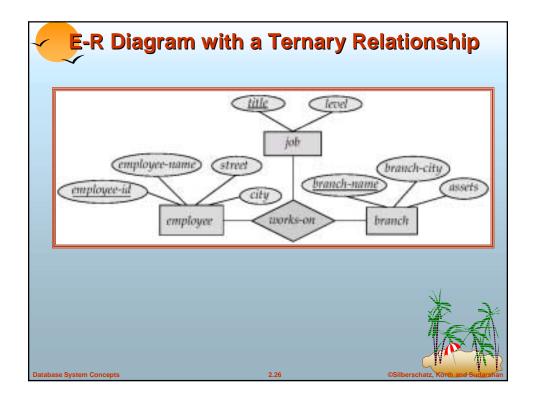
### **Keys for Relationship Sets**

- The combination of primary keys of the participating entity sets forms a super key of a relationship set.
  - (customer-id, account-number) is the super key of depositor
  - NOTE: this means a pair of entity sets can have at most one relationship in a particular relationship set.
    - E.g. if we wish to track all access-dates to each account by each customer, we cannot assume a relationship for each access. We can use a multivalued attribute though
- Must consider the mapping cardinality of the relationship set when deciding the what are the candidate keys
- Need to consider semantics of relationship set in selecting the primary key in case of more than one candidate key

**Database System Concepts** 

2.25

Silherschatz Korth and Sudarsk





# Cardinality Constraints on Ternary Relationship

- We allow at most one arrow out of a ternary (or greater degree) relationship to indicate a cardinality constraint
- E.g. an arrow from *works-on* to *job* indicates each employee works on at most one job at any branch.
- If there is more than one arrow, there are two ways of defining the meaning.
  - E.g a ternary relationship R between A, B and C with arrows to B and C could mean
  - 1. each A entity is associated with a unique entity from B and C or
  - 2. each pair of entities from (A, B) is associated with a unique C entity, and each pair (A, C) is associated with a unique B
  - Each alternative has been used in different formalisms
  - To avoid confusion we outlaw more than one arrow



**Database System Concepts** 

2.27

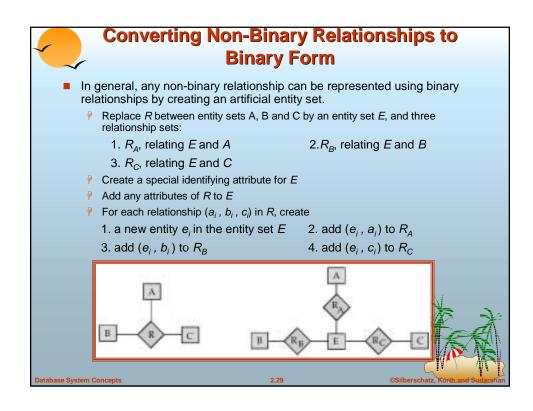


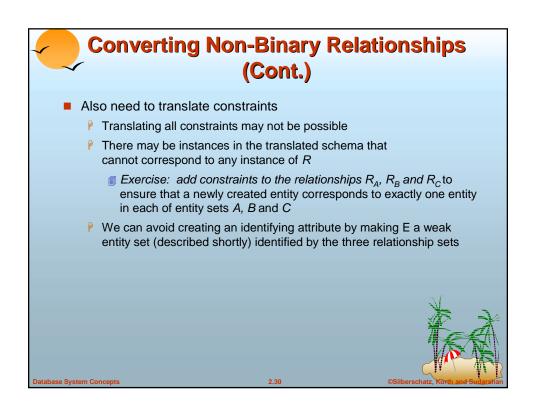
# **Binary Vs. Non-Binary Relationships**

- Some relationships that appear to be non-binary may be better represented using binary relationships
  - E.g. A ternary relationship *parents*, relating a child to his/her father and mother, is best replaced by two binary relationships, *father* and *mother* 
    - Using two binary relationships allows partial information (e.g. only mother being know)
  - But there are some relationships that are naturally non-binary
    - E.g. works-on



atabase System Concepts







#### **Design Issues**

- Use of entity sets vs. attributes Choice mainly depends on the structure of the enterprise being modeled, and on the semantics associated with the attribute in question.
- Use of entity sets vs. relationship sets
   Possible guideline is to designate a relationship set to describe an action that occurs between entities
- Binary versus n-ary relationship sets Although it is possible to replace any nonbinary (n-ary, for n > 2) relationship set by a number of distinct binary relationship sets, a nary relationship set shows more clearly that several entities participate in a single relationship.
- Placement of relationship attributes



**Database System Concepts** 

2.31

How about doing an ER design interactively on the board?
Suggest an application to be modeled.



### **Weak Entity Sets**

- An entity set that does not have a primary key is referred to as a weak entity set.
- The existence of a weak entity set depends on the existence of a identifying entity set
  - it must relate to the identifying entity set via a total, one-to-many relationship set from the identifying to the weak entity set
  - ldentifying relationship depicted using a double diamond
- The discriminator (or partial key) of a weak entity set is the set of attributes that distinguishes among all the entities of a weak entity set.
- The primary key of a weak entity set is formed by the primary key of the strong entity set on which the weak entity set is existence dependent, plus the weak entity set's discriminator.

**Database System Concepts** 

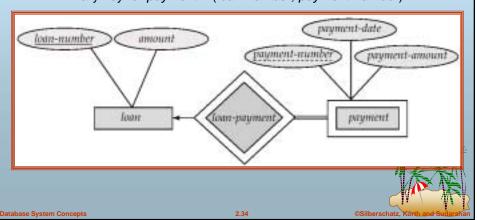
2.33

Silberschatz, Korth and Sudars



# **Weak Entity Sets (Cont.)**

- We depict a weak entity set by double rectangles.
- We underline the discriminator of a weak entity set with a dashed line.
- payment-number discriminator of the payment entity set
- Primary key for payment (loan-number, payment-number)





### **Weak Entity Sets (Cont.)**

- Note: the primary key of the strong entity set is not explicitly stored with the weak entity set, since it is implicit in the identifying relationship.
- If *loan-number* were explicitly stored, *payment* could be made a strong entity, but then the relationship between *payment* and *loan* would be duplicated by an implicit relationship defined by the attribute *loan-number* common to *payment* and *loan*



**Database System Concepts** 

2.35



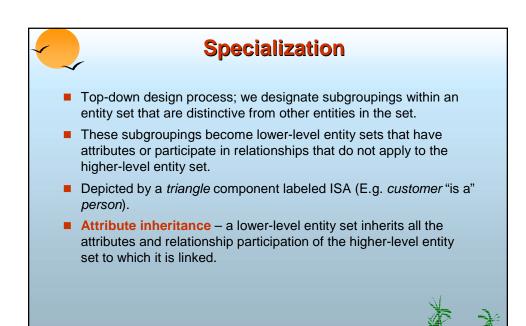
# **More Weak Entity Set Examples**

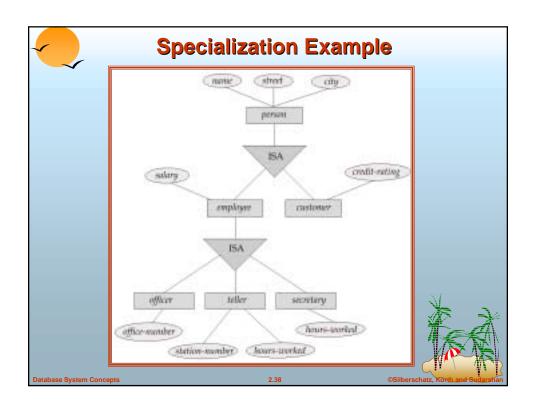
- In a university, a course is a strong entity and a course-offering can be modeled as a weak entity
- The discriminator of *course-offering* would be *semester* (including year) and *section-number* (if there is more than one section)
- If we model *course-offering* as a strong entity we would model *course-number* as an attribute.

Then the relationship with *course* would be implicit in the *course-number* attribute



Oatabase System Concepts







#### Generalization

- A bottom-up design process combine a number of entity sets that share the same features into a higher-level entity set.
- Specialization and generalization are simple inversions of each other; they are represented in an E-R diagram in the same way.
- The terms specialization and generalization are used interchangeably.



**Database System Concepts** 

2 30



# Specialization and Generalization (Contd.)

- Can have multiple specializations of an entity set based on different features.
- E.g. permanent-employee vs. temporary-employee, in addition to officer vs. secretary vs. teller
- Each particular employee would be
  - a member of one of permanent-employee or temporary-employee,
  - and also a member of one of officer, secretary, or teller
- The ISA relationship also referred to as superclass subclass relationship



Oatabase System Concepts



### Design Constraints on a Specialization/Generalization

- Constraint on which entities can be members of a given lower-level entity set.
  - condition-defined
    - **I** E.g. all customers over 65 years are members of senior-citizen entity set; senior-citizen ISA person.
  - user-defined
- Constraint on whether or not entities may belong to more than one lower-level entity set within a single generalization.
  - Disioint
    - an entity can belong to only one lower-level entity set
    - Noted in E-R diagram by writing disjoint next to the ISA triangle
  - Overlapping
    - an entity can belong to more than one lower-level entity



**Database System Concepts** 

2.41

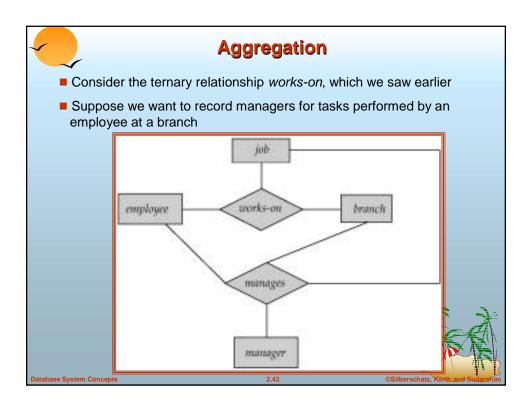


# Design Constraints on a Specialization/Generalization (Contd.)

- Completeness constraint -- specifies whether or not an entity in the higher-level entity set must belong to at least one of the lower-level entity sets within a generalization.
  - for total: an entity must belong to one of the lower-level entity sets
  - partial: an entity need not belong to one of the lower-level entity sets



Oatabase System Concepts



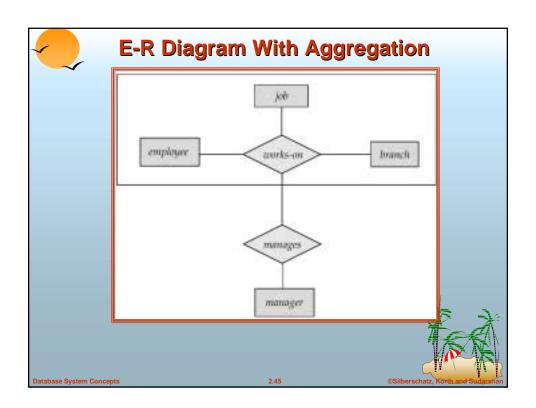


# **Aggregation (Cont.)**

- Relationship sets works-on and manages represent overlapping information
  - Every manages relationship corresponds to a works-on relationship
  - However, some *works-on* relationships may not correspond to any *manages* relationships
    - So we can't discard the works-on relationship
- Eliminate this redundancy via aggregation
  - Treat relationship as an abstract entity
  - Allows relationships between relationships
  - Abstraction of relationship into new entity
- Without introducing redundancy, the following diagram represents:
  - An employee works on a particular job at a particular branch
  - An employee, branch, job combination may have an associated manager

Database System Concepts





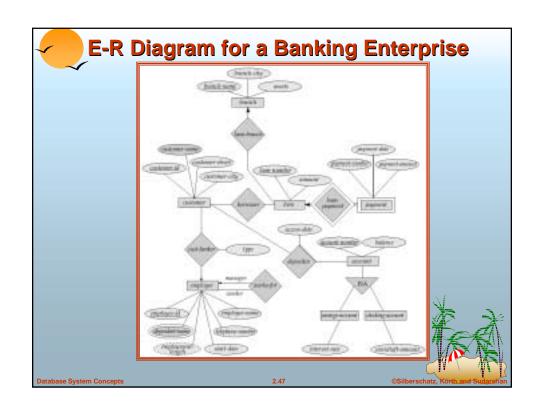


# **E-R Design Decisions**

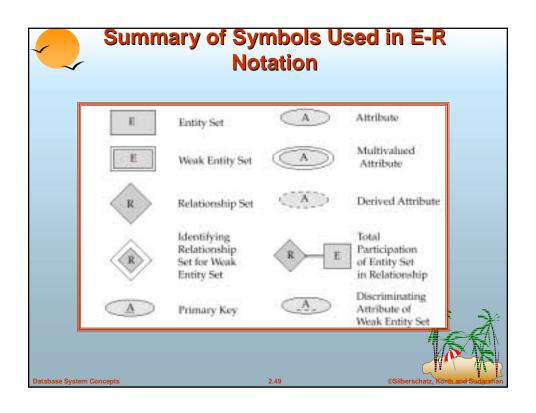
- The use of an attribute or entity set to represent an object.
- Whether a real-world concept is best expressed by an entity set or a relationship set.
- The use of a ternary relationship versus a pair of binary relationships.
- The use of a strong or weak entity set.
- The use of specialization/generalization contributes to modularity in the design.
- The use of aggregation can treat the aggregate entity set as a single unit without concern for the details of its internal structure.

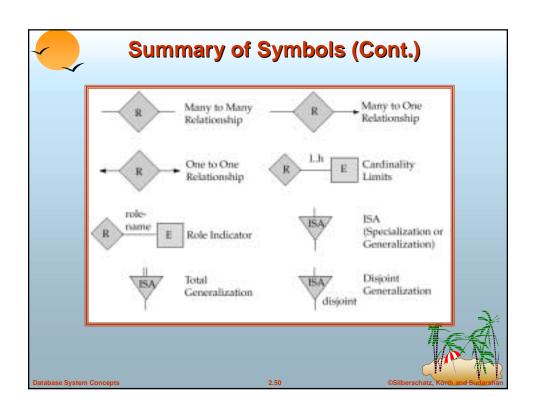


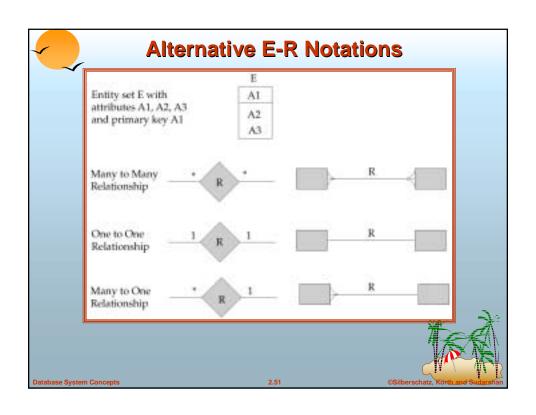
Database System Concepts

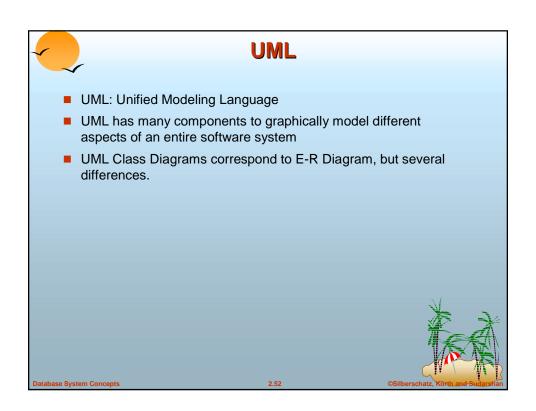


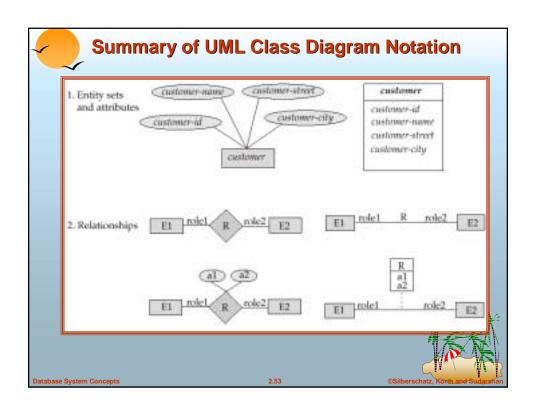
How about doing another ER design interactively on the board?









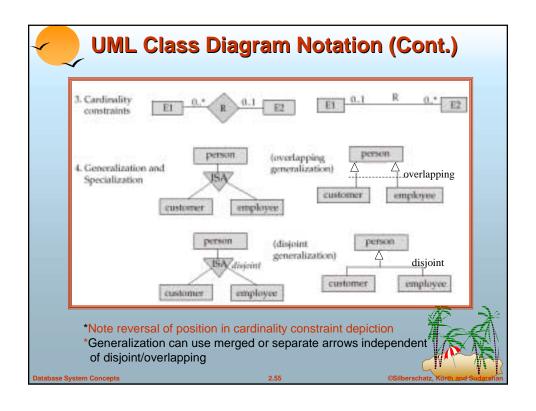




# **UML Class Diagrams (Contd.)**

- Entity sets are shown as boxes, and attributes are shown within the box, rather than as separate ellipses in E-R diagrams.
- Binary relationship sets are represented in UML by just drawing a line connecting the entity sets. The relationship set name is written adjacent to the line.
- The role played by an entity set in a relationship set may also be specified by writing the role name on the line, adjacent to the entity set.
- The relationship set name may alternatively be written in a box, along with attributes of the relationship set, and the box is connected, using a dotted line, to the line depicting the relationship set.
- Non-binary relationships drawn using diamonds, just as in E diagrams

Database System Concepts





### **UML Class Diagrams (Contd.)**

- Cardinality constraints are specified in the form *l..h*, where *l* denotes the minimum and *h* the maximum number of relationships an entity can participate in.
- Beware: the positioning of the constraints is exactly the reverse of the positioning of constraints in E-R diagrams.
- The constraint 0..\* on the *E*2 side and 0..1 on the *E*1 side means that each *E*2 entity can participate in at most one relationship, whereas each *E*1 entity can participate in many relationships; in other words, the relationship is many to one from *E*2 to *E*1.
- Single values, such as 1 or \* may be written on edges; The single value 1 on an edge is treated as equivalent to 1..1, while \* is equivalent to 0..\*.

**Database System Concepts** 



### Reduction of an E-R Schema to Tables

- Primary keys allow entity sets and relationship sets to be expressed uniformly as tables which represent the contents of the database.
- A database which conforms to an E-R diagram can be represented by a collection of tables.
- For each entity set and relationship set there is a unique table which is assigned the name of the corresponding entity set or relationship set.
- Each table has a number of columns (generally corresponding to attributes), which have unique names.
- Converting an E-R diagram to a table format is the basis for deriving a relational database design from an E-R diagram.

**Database System Concepts** 

2.57





# **Representing Entity Sets as Tables**

A strong entity set reduces to a table with the same attributes.

customer-id	customer-name	customer-street	customer-city
019-28-3746	Smith	North	Rye
182-73-6091	Turner	Putnam	Stamford
192-83-7465	Johnson	Alma	Palo Alto
244-66-8800	Curry	North	Rye
321-12-3123	Jones	Main	Harrison
335-57-7991	Adams	Spring	Pittsfield
336-66-9999	Lindsay	Park	Pittsfield
677-89-9011	Hayes	Main	Harrison
963-96-3963	Williams	Nassau	Princeton

Database System Concepts



## **Composite and Multivalued Attributes**

- Composite attributes are flattened out by creating a separate attribute for each component attribute
  - E.g. given entity set *custome*r with composite attribute *name* with component attributes *first-name* and *last-name* the table corresponding to the entity set has two attributes name.first-name and name.last-name
- A multivalued attribute M of an entity E is represented by a separate table EM
  - Table EM has attributes corresponding to the primary key of E and an attribute corresponding to multivalued attribute M
  - E.g. Multivalued attribute dependent-names of employee is represented by a table employee-dependent-names( employee-id, dname)
  - Each value of the multivalued attribute maps to a separate row of the table EM
    - E.g., an employee entity with primary key John and dependents Johnson and Johndotir maps to two rows: (John, Johnson) and (John, Johndotir)

**Database System Concepts** 

2.59





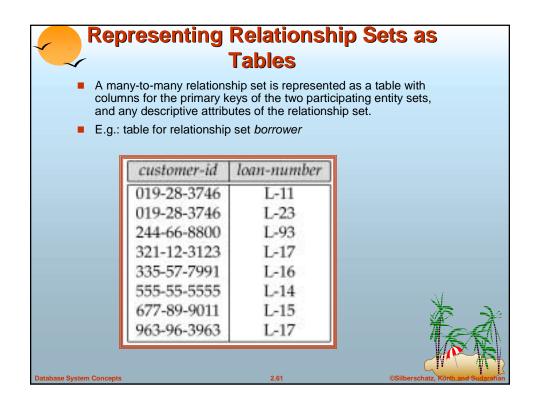
# **Representing Weak Entity Sets**

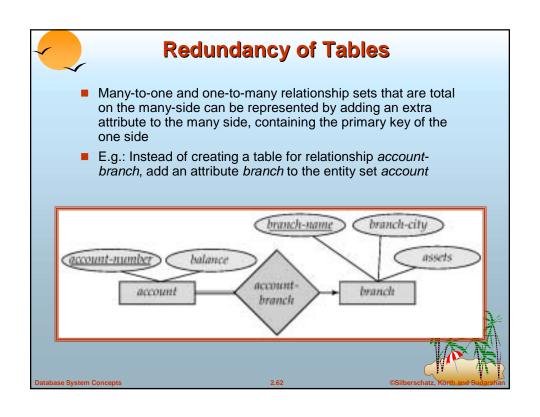
A weak entity set becomes a table that includes a column for the primary key of the identifying strong entity set

loan-number	payment-number	payment-date	payment-amount
L-11	53	7 June 2001	125
L-14	69	28 May 2001	500
L-15	22	23 May 2001	300
L-16	58	18 June 2001	135
L-17	5	10 May 2001	50
L-17	6	7 June 2001	50
L-17	7	17 June 2001	100
L-23	11	17 May 2001	75
L-93	103	3 June 2001	900
L-93	104	13 June 2001	200

Database System Concepts









### **Redundancy of Tables (Cont.)**

- For one-to-one relationship sets, either side can be chosen to act as the "many" side
  - That is, extra attribute can be added to either of the tables corresponding to the two entity sets
- If participation is partial on the many side, replacing a table by an extra attribute in the relation corresponding to the "many" side could result in null values
- The table corresponding to a relationship set linking a weak entity set to its identifying strong entity set is redundant.
  - E.g. The *payment* table already contains the information that would appear in the *loan-payment* table (i.e., the columns loan-number and *payment-number*).



**Database System Concepts** 

2.63



### **Representing Specialization as Tables**

- Method 1:
  - Form a table for the higher level entity
  - Form a table for each lower level entity set, include primary key of higher level entity set and local attributes

table	table attributes	
person	name, street, city	
customer	name, credit-rating	
emplovee	name, salary	

P Drawback: getting information about, e.g., employee requires accessing two tables



Oatabase System Concepts



# Representing Specialization as Tables (Cont.)

#### Method 2:

Form a table for each entity set with all local and inherited attributes

table	table attributes
person	name, street, city
customer	name, street, city, credit-rating
employee	name, street, city, salary

- If specialization is total, table for generalized entity (*person*) not required to store information
  - Can be defined as a "view" relation containing union of specialization tables
  - But explicit table may still be needed for foreign key constraints
- P Drawback: street and city may be stored redundantly for persons who are both customers and employees

**Database System Concepts** 

2.65

ilherschatz Korth and Sudars



### Relations Corresponding to Aggregation

- To represent aggregation, create a table containing
  - primary key of the aggregated relationship,
  - the primary key of the associated entity set
  - Any descriptive attributes



Database System Concepts

