Chapter 3  The Enhanced E-R Model and Business Rules

Chapter Overview

The purpose of this chapter is to present some important extensions to the E-R model (described in Chapter 2) that are useful in capturing additional business meaning. In particular, we describe two types of extensions to the E-R model. First, the enhanced entity-relationship (EER) model includes constructs for supertype/subtype relationships. Second, the inclusion of new notation for business rules allows the designer to capture a broader range of constraints on the data model than were previously available.

Chapter Objectives

Specific student objectives are included in the beginning of the chapter. From an instructor’s point of view, the objectives of this chapter are to:
1. Introduce the concept of supertype/subtype relationships, and prepare the student to recognize when to use these relationships in data modeling.
2. Describe the use of specialization (top-down perspective) and generalization (bottom-up perspective) as complementary techniques for defining supertype/subtype relationships.
3. Introduce notation for specifying both completeness constraints and disjointness constraints when modeling supertype/subtype relationships.
4. Help students gain sufficient perspective so that they recognize when to use (and when not to use) supertype/subtype relationships in realistic business situations.
5. Describe the basic premises of a business rules paradigm.
6. Discuss the universal data model and its use in packaged data models.

Key Terms

<table>
<thead>
<tr>
<th>Attribute inheritance</th>
<th>Generalization</th>
<th>Subtype discriminator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness constraint</td>
<td>Overlap rule</td>
<td>Supertype</td>
</tr>
<tr>
<td>Disjoint rule</td>
<td>Partial Specialization rule</td>
<td>Supertype/subtype hierarchy</td>
</tr>
<tr>
<td>Disjointness constraint</td>
<td>Specialization</td>
<td>Total specialization rule</td>
</tr>
<tr>
<td>Enhanced entity-relationship (EER) model</td>
<td>Subtype</td>
<td>Universal data model</td>
</tr>
<tr>
<td>Entity cluster</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Instant download and all chapters Solutions Manual Modern Database Management
11th Edition Jeffrey A. Hoffer, V. Ramesh, Heikki Topi
Classroom Ideas

1. Introduce the concept of supertypes and subtypes with a familiar example, such as VEHICLE (subtypes are CAR, TRUCK, SUV, etc.).
2. Introduce the basic notation for supertype/subtype relationships (Figure 1). Use this notation to represent the example you introduced in (1). Introduce your students to all three types of notation.
3. Discuss the EMPLOYEE example with subtypes (Figure 2). Use this figure to introduce the example of attribute inheritance.
4. Use Figure 3 to discuss the two major reasons for introducing supertype/subtype relationships: unique attributes among subtypes, and unique subtype relationships.
5. Contrast generalization and specialization using Figures 4 and 5. Have your students suggest other examples that use each of these techniques.
6. Introduce the completeness constraint using Figure 6. Give other examples where either the total specialization rule or the partial specialization rule is more appropriate.
7. Discuss the disjointness constraint and related notation using Figure 7. For reinforcement, have the students work Problem 8 or 9 (Problems and Exercises) in class.
8. Introduce notation for a subtype discriminator (Figures 8 and 9). Discuss why a different notation is required for the two cases shown in these figures.
9. Discuss entity clustering and illustrate with Figures 13 and 14.
10. Review the extended example of a supertype/subtype hierarchy shown in Figure 10. For reinforcement, ask the students to work Problem 2 (Problems and Exercises) in class.
11. Review universal data models and discuss how these are being used more widely today. Consider inviting an industry guest speaker to discuss how these universal data models are utilized in his/her company. If your students have access to computers during your class session, break the students into small groups and have them complete Review Question 16 then report back to the large class with their findings.
12. Ask your students for examples of other business rules they have encountered recently in their work, school, or home experience that could be modeled with supertype/subtype hierarchies. See if they can diagram these rules using the notation provided in this chapter.
Answers to Review Questions

1. Define each of the following terms:
   a. **Supertype.** A generic entity type that has a relationship with one or more subtypes
   b. **Subtype.** A subgrouping of the entities in an entity type that is meaningful to the organization
   c. **Specialization.** The process of defining one or more subtypes of the supertype, and forming supertype/subtype relationships
   d. **Entity cluster.** A set of one or more entity types and associated relationships grouped into a single abstract entity type
   e. **Completeness constraint.** A type of constraint that addresses the question whether an instance of a supertype must also be a member of at least one subtype. The completeness constraint has two possible rules: total specialization and partial specialization
   f. **Enhanced entity-relationship (EER) model.** The model that has resulted from extending the original E-R model with new modeling constructs such as supertypes and subtypes
   g. **Subtype discriminator.** An attribute of the supertype whose values determine the target supertype or subtypes
   h. **Total specialization rule.** Specifies that each entity instance of the supertype must be a member of some subtype in the relationship
   i. **Generalization.** The process of defining a generalized entity type from a set of more specialized entity types
   j. **Disjoint rule.** Specifies that if an entity instance (of the supertype) is a member of one subtype, it cannot simultaneously be a member of two (or more) subtypes
   k. **Overlap rule.** Specifies that an entity instance can simultaneously be a member of two (or more) subtypes
   l. **Partial specialization rule.** Specifies that an entity instance of the supertype is allowed not to belong to any subtype
   m. **Universal data model.** A generic or template data model that can be reused as a starting point for a data modeling project

2. Match the following terms and definitions:
   d) supertype
   f) entity cluster
   a) subtype
   e) specialization
   g) subtype discriminator
   c) attribute inheritance
   b) overlap rule
3. **Contrast the following terms:**
   a. **Supertype; subtype.** A supertype is a generalized entity type that has one or more subtypes, while a subtype is a subgrouping of the entities in a supertype.
   b. **Generalization; specialization.** Generalization is the process of defining a generalized entity type from a set of more specialized entity types, while specialization is the process of defining one or more subtypes of the supertype.
   c. **Disjoint rule; overlap rule.** With the disjoint rule an instance of a supertype must be a member of only one subtype at a given time. With the overlap rule an instance of a supertype may simultaneously be a member of two or more subtypes.
   d. **Total specialization rule; partial specialization rule.** With the total specialization rule, each instance of the supertype must be a member of some subtype in the relationship. With the partial specialization rule, an instance of the supertype is allowed not to belong to any subtype.
   e. **PARTY; PARTY ROLE.** In a universal data model, PARTY represents persons and organizations independent of the roles they play whereas PARTY ROLE contains information about a party for an associated role.
   f. **Entity; entity cluster.** An entity is a person, place, object, event, or concept in the user environment about which the organization wishes to maintain data. An entity cluster is a set of one or more entity types and associated relationships grouped into a single abstract entity type.

4. **Two conditions for using supertype/subtype relationships:**
   a. There are attributes that apply to some (but not all) of the instances of an entity type.
   b. There are relationships that apply to some (but not all) of the instances of an entity type.

5. **Reasons for using an entity clustering approach:**
   a. Simplifying the presentation of a complex enterprise-wide E-R diagram.
   b. Enabling a hierarchical decomposition of a macro-level data model into finer and finer views of the data.
   c. Desiring to focus part of the model on an area of interest to a community of users.
   d. Creating several different entity cluster segments each with a different focus, such as departments, information system applications, business processes, or corporate divisions.

6. **An example of a supertype/subtype relationship:**
   The supertype PERSON has many possible subtypes: MALE, FEMALE, INFANT, TEENAGER, etc, assuming these different types of persons have somewhat different attributes or participate in different relationships. In an organizational context, PERSON may have subtypes of EMPLOYEE, CONTRACTOR, CUSTOMER, VENDOR, MANAGER, etc.

7. **Attribute inheritance explanation:** Attribute inheritance is a property of the enhanced ER diagram that ensures subtype entities inherit the values of all attributes of their supertype(s). This property is important because it makes it unnecessary to include supertype attributes redundantly with subtypes.
8. *Examples of Supertype/subtype relationship where:*
   a. the disjoint rule applies: PERSON has subtypes MALE and FEMALE.
   b. the overlap rule applies: PERSON has subtypes INSTRUCTOR and STUDENT.

9. *Types of business rules in EER:* The types of business rules that are normally captured in an EER diagram include terms, relationship constraints, and supertype/subtype relationships (see Figure 11).

10. *Subtype discriminator purpose:* The purpose of a subtype discriminator is to determine the target subtype (or subtypes) for each instance of a supertype.

11. *Utility of packaged data model:* A packaged data model is most useful when one can easily customize it to the specific business (that is, the organization is very similar to other organizations for the same industry or purpose or the functional area is roughly the same as that functional area in other organizations). As long as the packaged data model is for the type of business or functional area, then it can generally be customized. The amount of customization depends upon the types of specialized business rules in place for the organization.

12. *Starting project with packaged data model vs. from scratch:* A packaged data model provides the metadata of a standardized, industry-vetted data model usually built with a structured data modeling tool (i.e., ERWin from Computer Associates or Oracle Designer from Oracle Corporation). A data modeling project that starts with a packaged data model is different from one using a model developed from scratch along the following dimensions:
   
a. The project would begin by identifying the parts of the packaged data model that apply to your specific project’s data modeling situation, rather than beginning to draw model elements.
   
b. The identified data elements from the packaged data model would be renamed to terms local to the organization.
   
c. Data in the packaged data model would be mapped to data in current organization databases, with the intent of developing migration plans for converting organizational data.
      
i. Some of the data will not be able to be mapped (e.g., data elements in the package won’t be in the current systems, and likewise). Determine that each non-mapped item is essential and unique, as well as whether these requirements are necessary now or in the future.
      
ii. A purchased data model will have business rules to cover all possible circumstances where your specific local situation may need less flexibility and complexity.
      
iii. The purchased data model can be used to “seed” questions for coverage with the end users of the new system and database, allowing for earlier and more in-depth participation of system users and managers in the data modeling project.
      
iv. The comprehensive nature of the purchased data model will likely force the project to prioritize the staging of systems requirements related to
customization of the overall data model.

13. **Data profiling usage:** Data profiling is a way to statistically analyze data to uncover hidden patterns and flaws. Profiling can find outliers, see shifts in data distribution over time, and identify other phenomenon. Each perturbation of the distribution of data may tell a story, such as showing when major application system changes occurred, or when business rules changed. Often these patterns suggest poorly designed databases (e.g., data for separate entities combined to improve processing speed for a special set of queries but the better structure was never restored). Data profiling can also be used to assess how accurate current data are and anticipate the clean-up effort that will be needed to populate the purchased data model with high-quality data.

14. **Skill needed for packaged data model vs. without:** A data modeling project using a packaged data model requires at least the same amount of skill as a project not using a packaged data model, and in some cases, may require more skill. The primary reason is that when a data modeling project uses a packaged data model, the data modeler must customize the packaged data model to meet local organizational needs and constraints. Thus, a successful data modeler using a packaged data model needs advanced skills and knowledge about the organization’s business rules, complex data modeling formalisms, and the structured data modeling tool used to specify the packaged data model.

15. **Benefit of packaged data model:** A packaged data model provides the metadata of a standardized, industry-vetted data model usually built with a structured data modeling tool (i.e., ERWin from Computer Associates or Oracle Designer from Oracle Corporation). The packaged data model contains a fully populated description of the data model and the structured data modeling tool that permits customization of the data model and printing of several reports from the model. The structured data modeling tool often includes the ability to produce SQL commands for database definition in a variety of database management systems.

16. **Utility of supertype/subtype hierarchy:** A supertype/subtype hierarchy is useful when you have several subtypes that are also supertypes. An example would be for bank accounts. At the first level (supertype), you can have savings, checking and loans. Underneath loans, there are several subtypes, including personal, auto, home, etc.

17. **Supertype/subtype membership:** A member of a supertype is always a member of at least one subtype when the rule of total specialization applies to an EERD.
Solutions to Problems and Exercises
1. A supertype/subtype example listing follows for a GRADUATE STUDENT:

<table>
<thead>
<tr>
<th>Attribute Name</th>
<th>Data Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSN</td>
<td>736-94-1802</td>
</tr>
<tr>
<td>Name</td>
<td>Jessica James</td>
</tr>
<tr>
<td>Address</td>
<td>25 Lake Dr. Medford OR 95106</td>
</tr>
<tr>
<td>Gender</td>
<td>female</td>
</tr>
<tr>
<td>Date_of_Birth</td>
<td>Oct. 23, 1967</td>
</tr>
<tr>
<td>Major_Dept</td>
<td>Computer Science</td>
</tr>
<tr>
<td>Test_Score</td>
<td>986</td>
</tr>
</tbody>
</table>
2. Figure 10 with subtype discriminators

[Diagram showing categorization of PERSON into Employee, Alumnus, and Student, with further subcategories for each type.]
3. 
   a. *Figure 2, revised*

   ![Diagram of Employee types]

   - **HOURLY EMPLOYEE**
     - Hourly Rate
   - **SALARIED EMPLOYEE**
     - Annual Salary
     - Stock Option
   - **CONSULTANT**
     - Contract Number
     - Billing Rate

   Employee Type =

   - "H"
   - "S"
   - "C"

   **b. Figure 3, revised**

   ![Diagram of Patient and Physician relationships]

   - **PATIENT**
     - Patient ID
     - Admit Date
     - Patient Type
     - Patient Type =
       - "O"
       - "I"
   - **RESPONSIBLE PHYSICIAN**
     - Physician ID
   - **OUTPATIENT**
     - Checkout Date
   - **RESIDENT PATIENT**
     - Date Discharged
   - **BED**
     - Bed ID

   Is Cared For
   Is Assigned