Entity-Relationship Modeling of Composite Materials Data

by

Lisa K. Spainhour
FAMU/FSU College of Engineering
Tallahassee, FL 32316-2175

William J. Rasdorf
North Carolina State University
Raleigh, NC 27695-7908

Abstract: Material properties data, particularly that of advanced composite materials such as fiber-reinforced plastics, provides a very complex domain which is difficult to computerize. To aid in this process, a formal conceptual data model was designed. The unique contribution of this model is that it encompasses the data needs of civil and structural engineers throughout the life cycle of a composite material component. The data model was created using Entity-Relationship methodology, and can be used to support automated data exchange and database development. Other issues addressed in this paper include assessing the suitability of the Entity-Relationship model for composite materials data, assessing and maintaining the quality of a data set, and modeling special types of data applicable to composite materials.

1. Introduction

Research into the potential use of composite materials is substantially increasing in the civil engineering community, both in this country and abroad, yet very little of this research is being directed toward data gathering, modeling, management, and integration or toward system integration. The domain of composite materials is very large, complex, and diverse. This complexity inhibits the development of material properties databases to support the use of composites in engineering analysis, design, and construction. Without ready availability of complete, trustworthy properties data, there is an uncertainty on the part of engineering analysts and designers regarding the behavior and performance of composite materials. This uncertainty is unacceptable because civil engineers deal with the public safety, health, and well-being; there is, therefore, some unwillingness to use these materials in engineering design [Stanton88]. A key first need exists, therefore, to make information about composite materials, their derivation, their properties, and their use available to researchers, materials scientists, and engineers.

The number of databases supporting the use of engineering materials is slowly increasing, but most are being developed in an "ad hoc" manner, without the use of formal modeling techniques. This is unfortunate because a good database implementation depends upon how accurately and

---

completely the data involved in the domain can be modeled. One way to ensure the development of a good data model is by using a formal modeling methodology that enables a database schema to be designed in a systematic fashion. Using such an approach improves both the quality and timeliness of the database design procedure.

Conceptual modeling, as defined in [Loucopoulos92], is the activity of formally describing some aspects of the physical and social world around us for purposes of understanding and communication. Modeling the functional application requirements and information system components at a conceptual level is even more important given the growing demand for information systems of ever-increasing size, scope, and complexity. Many empirical studies have shown that a poor initial understanding of data and functional requirements can result in huge expenses later in correcting errors and making changes to a system during operation [Loucopoulos92].

Conceptual modeling is particularly important in the domain of composite materials because of the complexity of the domain and because of the diversity of data contained within it. The spectrum of data within the scope of fiber-reinforced plastics is broad, ranging from raw materials data, to manufacturing data, to systems design information, to function and performance, to specifications and codes. Composites are formed by combining two or more materials, in varying amounts, subject to an almost infinite number of configurations. Simply understanding the data and relationships in the domain is difficult and requires the use of a formal framework in which to cast the system, so that a picture of the domain emerges.

The primary focus of this research was on developing a formal conceptual model of the data in the domain of composite materials. First, we looked at materials data and its use in the engineering domain, focusing on the data needed and supplied by various applications over the entire life cycle of the material. From this information, we were able to develop an information model, which is a high-level model of the domain area; this model is described in reference [??]. We investigated the general constructs provided by the extended ER models, the specific implementations available, and their suitability for composite materials data. Finally, we utilized a collection of Entity-Relationship diagrams to create a modular, extensible conceptual model which can be used as the basis of one or more composite materials database.

This work provides a unique contribution, in that very few people are investigating approaches to modeling materials data, and that no one has developed a rigorous, formal, product model for composite materials data. Several related efforts are discussed in the next section; however, our Entity-Relationship model is the only one which attempts to define and describe the entire set of data and metadata required to document composite materials.

2. Related Work

Standards organizations such as the American Society for Testing and Materials (ASTM), the International Organization for Standardization (ISO), and the Committee on Data for Science and Technology (CODATA) have various efforts focused on modeling and computerizing materials data. The STandard for the Exchange of Product data (STEP) effort is key in the development of
a data exchange language for engineering product data, including material properties. The American Concrete Institute (ACI) and the American Society of Mechanical Engineers (ASME) are attempting, respectively, to gather archival material properties data on concrete that has been in service for up to fifty years, and to build a database to maintain characteristics of the 20,000 types of steel available internationally [Oland90, ASME93]. Key aspects of these efforts include developing data formats for structuring materials data, and gathering properties on materials that have been in service for many years. We hope to avoid the latter scenario by recognizing the data needs of composite materials and developing data structures to support those while the materials technology is still emerging.

The work which is most relevant to that described herein is the work by ASTM Committee E-49 on Computerization of Material Property Data in developing standard data formats [ASTM93]. These ASTM standards, enumerated below, recommend required and optional fields for identifying composite materials and recording test data.

- **E1434**: Standard Guide for Development of Standard Data Records for Computerization of Mechanical Test Data for High-Modulus Fiber-Reinforced Composite Materials
- **E1309**: Standard Guide for Identification of Composite Materials in Computerized Material Property Databases
- **E1471**: Standard Guide for Identification of Fibers, Fillers, and Core Materials in Computerized Material Property Databases

Each standard provides a generic format for computerizing the data within its domain. The standards provide fields that are likely to be valuable to certain engineers and material specialists and therefore should be recorded and included when a set of identification or test data is transferred to another organization or entity. These standards provide the beginning of a data model definition for composite materials.

To contrast these data formats with the conceptual data model described within this report, first note that while the standards are recommendations for what data should be included in computerizing data, they do not provide guidelines for structuring the data within a database. Furthermore, the standards are generic, attempting to be as concise as possible, while accounting for the data needs of many tests and types of composites. Because of their generality, the data fields are often not as precise as possible. For example, ASTM E1309 attempts to address the features of all types of composites, including particulate, short fiber, metal matrix, and ceramic matrix composites, in a useful manner within one format.

Furthermore, the standard leans toward the needs of laminated composites. In doing so, it tends to leave out information pertinent to other types of composites or use generic note fields where specific parameters might be more appropriate. For example, the field *structural detail* is very generic, and might be replaced with the field *winding angle* if a standard format were developed specifically for filament wound composites. The fields *ply count* and *lay-up code* are relevant only to laminates.
3. Entity-Relationship Methodology

An Entity-Relationship (ER) diagram distills the information content of a domain area down to the concepts of 1) entities or objects, 2) relationships between those entities, and 3) attributes or properties of the entities and relationships [Chen76]. An Entity-Relationship diagram is illustrated in Figure 8. In this diagram, the names enclosed by rectangles represent entities (e.g. Batch, Sample, Population), those enclosed by diamonds represent relationships between entities (e.g. a Sample is part of a Batch).

Those names not enclosed but joined with either an entity or relationship by a straight line represent attributes (e.g. the batch mean and standard deviation are properties of a batch). Key attributes, which are indicated by boldface type, serve to uniquely identify each instance of an entity. The data type of each attribute, although not required by the ER methodology, is included after the attribute name. The numbers enclosed in parentheses are cardinality numbers and they represent the minimum and maximum number of each entity that can participate in the relationship. For example, each sample can undergo from one to many MNR tests, but each MNR test applies to one and only one sample.

As a group, extended Entity-Relationship (EER) models are similar enough to the classical model to be easily understandable and close to user perception of the domain; however, they provide more expressive power. In general, they extend the traditional ER model by incorporating one or more of the concepts of 1) data abstraction, 2) complex entities and attributes, and 3) time dependency. Data abstraction and complex entities are described in this section because both are relevant to the domain; time dependency is not as we are not reporting on models of temporal databases at this time. The most common types of data abstraction are listed below [Loucopoulos92].

- **Classification**: The grouping of entities that share common characteristics into a class over which uniform conditions hold.
- **Generalization**: The relationship (is-a) between a general object and a more specialized object, e.g. a tension test "is a" test.
- **Aggregation**: The relationship (is-part-of) between a component and its subcomponents, e.g. a fiber "is part of" a yam.
- **Grouping**: The relationship (is-member-of) between a member object and a set object, e.g. an inspector "is a member of" an inspection team.

Classification enables the separation of objects into groups that have common properties and can be placed into the same category. Only very basic classes are usually defined. The remaining three abstraction forms are specific types of relationships. Depending on the EER implementation, they are usually represented with a symbol other than the generic relationship symbol, which is a diamond connected to the entities by solid lines. The addition of complex entities is an extension of the last two abstraction forms, aggregation and grouping.
A complex entity is composed of one or more subentities related by the aggregation and/or generalization abstractions. These entities can in turn be complex, building up a hierarchy. For example, a pultruded component might be composed of resin plus tapes and yarns, each of which is composed of fibers. Complex attributes are aggregations or groupings of other attributes. An example of this is the attribute Dimensions, which might be the aggregation of the attributes Length, Width, and Thickness.

Certain extended ER models provide for additional semantic constraints on the diagram at the level of entities. In the EARDL model, boolean operators are applied to the relationships between a single primary entity and two or more other entities in the ER diagram [Davis90]. The ND constraint, for instance, indicates that one or none, but no more than one, of the entities may be instantiated at the same time as the primary entity. For example, a person (primary entity) may be a doctor, a lawyer, or neither, but a person may not be both a doctor and a lawyer. We found these features to be useful in further constraining the ER models for composite materials.

4. Suitability of ER Model for Composite Materials

In choosing a modeling methodology for this work, our key requirement was that the methodology include a graphical model that is easily understood. In addition, we wanted to select a conceptual modeling methodology that is well recognized and often used in the computer science community. As part of this research, we evaluated the suitability of the ER model and its extensions for composite materials data, and assessing the applicability, strengths, weaknesses, and limitations of the model and its extensions. Several conclusions from this study are presented below.

4.1. Lack of Standardized Model

One of the difficulties in using the ER model is that there is no one correct version of the ER diagram or methodology. The ER model is in the public domain; therefore, it is not standardized. Since its origin in 1976, various researchers, including the initial developers, have developed new features and added them to the model. This is certainly the case in the Extended ER models, but is also true to a lesser extent with the classical ER model. For example, certain researchers utilize pairs of cardinality numbers, which indicate the minimum and maximum numbers of an entity which can participate in a relationship [Theodoulidis92]. Others use only a single cardinality number to state the maximum number of participants (usually either one or many) and use an alternative to the solid outline of the entity rectangle to indicate that its existence is not mandatory. The lack of standardization within the model leads to confusion concerning the meaning of certain symbols.

4.2. Usefulness of Extended Entity-Relationship Models

It is important for developers of conceptual models and composites databases to understand the abstraction concepts even if the symbols of the extended ER model are not used. That is, even if the classical model is selected instead of an EER model, the concepts of aggregation and
generalization and their relevance to composite materials data are important to understand. If used, the symbolic representations provided by the EER models clarify the abstraction concepts and can simplify the model diagrams because fewer symbols are often required to present an idea; therefore there is benefit in using an extended model.

However, the limited number of symbols in the classical model is much less confusing to interpret, particularly if one is not as familiar with the modeling methodology. Because the abstraction types are specific examples of commonly occurring relationships, they can also be modeled using the original relationship symbols. As a result, they do not actually add any new capabilities or functionality to the ER model. They do, however, enable additional expressiveness by providing unique symbols for these events, simplifying the diagrams and clarifying the concepts.

### 4.3. Need for Additional Data Types

The complex attributes provided by the EER models are an aggregation of enumerated attributes, either simple or complex. For example, the Date attribute might be composed of the Day, Month, and Year attributes. With complex types, the number of subattributes is predetermined and cannot be varied. We feel that additional nonatomic attributes, such as attributes with a set, ordered list, or matrix data type, are required to fully model composite materials and other engineering data. However, these additional field types are actually a function of the database model and its implementation capabilities rather than the conceptual model. In other words, stating that an attribute has a specific field type within an ER model does not require the ability to implement that type. Although the field types enumerated below are commonly provided by programming languages, most database management systems have very limited type capabilities.

- User-defined abstract field types
- Specialized atomic types
- Enumerated type
- Boolean type
- Computed type
- Set, list, and bag types
- Complex type
- Matrix type

### 4.4. Difficulty of Modeling Complex Domain Areas

Developing an ER model becomes less intuitive as the domain becomes increasingly complex. This is considered to be a drawback of formal models in general, rather than specifically the ER model. A typical example of an ER diagram models the academic environment using entities representing Teachers, Students, Books, Classes, and Classrooms. The realm of materials data is much more abstract, resulting in some uncertainty regarding what defines an entity or object. For example, a Test is fairly distinguishable as an entity, but a Property could be regarded as either an entity or an attribute. Processes, dimensions, coordinates, environmental conditions,
and failure must be considered to successfully model composite materials data. Drawing Entity-Relationship diagrams is also increasingly difficult for larger and more complex domains. For systems that exhibit multiple possible interactions, the diagrams can be difficult to create and to understand.

4.5. Conclusions Regarding Suitability

In general, we feel that the ER methodology is very suitable for composite materials data. The abstraction concepts aid the data modeler in obtaining a fundamental understanding of the concepts underlying the data. With the addition of nonatomic and complex attributes, the resulting model is much simpler and more reflective of the actual domain being modeled. We feel that, because of the high level of abstraction of the ER model, it is very useful in understanding and visualizing the underlying concepts in the domain area, although more formal models such as the NIAM model might provide a better formulation for translating into a computational model.

Of the abstraction concepts, we found that the "is-a" relationship is required most often. In addition, "is-part-of" expresses the relationship between the component and structure. We did not use a distinct symbol to represent the "is-a" relationship, but found that we were able to remove the cardinality numbers from the relationship. The understood cardinality numbers are (0,1) on the subclass and (1,1) on the superclass. This occurs because the subclass always corresponds to one and only one instance of the superclass, yet the superclass instance relates to only one of the subclass instances, meaning that the rest of the subclasses must have no relationship with that instance. The "is-a" relationship was used most often with the XR constraint, to state that for each superclass instance, there is a corresponding instance in one and only one of the subclasses.

Of the nonatomic field types considered in this research, the type required most often is the "set of complex" type. Because the type is complex, numerous subattributes can be defined; because the set type is incorporated, multiple instances of the subattributes can be stored for each corresponding instance of the primary attributes. In effect, this field type creates a nested entity within the original entity, in a similar manner to a nested relational DBMS. This field type, and the similar "list of complex," therefore, enables complex entities to be modeled, which eliminates the need for separate entities related to the original with many-to-one relationships.

5. Overall Conceptual Model for Composite Materials Data

It was immediately apparent that no single database design could (or should) encompass the data requirements of all types of fiber-reinforced composites, or the data needs of all engineering application programs and users over the entire life cycle of the composite. The resulting database would be very large, inefficient, and most users would require only a minimal portion of the data included within. What we have provided, therefore, is a modular, extensible ER model, which can be modified to suit any number of purposes. Various entities representing, for example, different types of composites or alternate test methods can be selected or discarded from the model based upon the needs of the end users. In a similar manner, the model is
extensible: additional entities and attributes can be added as new materials and manufacturing methods are developed and as different organizational priorities are established. In this way, the underlying data model can be customized for a specific organization, then used to develop a specialized custom database.

In addition, we have provided model components which define phenomena and data types pertinent to composite materials, e.g. dimensions, graph data, and bitmap data. These components can be altered to suit the needs of particular applications and attached to the overall model at various points. For example, bitmap representation of photographs and diagrams might be used to explain test setups, fracture surfaces, and/or component geometries; therefore, the bitmap model component can be attached to the global model at numerous locations. The remainder of the paper illustrates several components of a conceptual model for composites. The entire collection of Entity-Relationship diagrams can be found in reference [Spainhour94].

5.1. Overall ER Diagram

Figure 1 presents the overall structure of the Entity-Relationship model for the composite materials database. The model is loosely based on a hierarchical view of materials data, allowing the user to maintain data on each of the material types in the hierarchy. At the top level is a generic material, also known as a material class. Generic materials are broad categories of materials; their properties are characteristic of all materials in that class, e.g. structural steel, or graphite/epoxy prepreg. Generic materials may be instances of other generic materials; for example, high-strength concrete is an instance of concrete. Following the left branch of the hierarchy, available materials are those which are obtainable from a specific manufacturer, which have specific properties that supersede those of the related generic material.

However, each material is generally available in several forms, e.g. various resin contents, filament counts, or fabric widths, and the properties of available materials must account for this variability. On-site materials are those which are available on-site for manufacturing; one of the available forms has been selected from the manufacturer, yet the material is usually supplied in bulk form. Thus, while it is immediately available, it is not a component until some shape, form, or dimensions have been applied to the material through machining, cutting, or other processing. Finally, the components can be combined into structures, which can further be combined into more complex structures.

On the other branch of the hierarchy are designed components and designed structures. They are identical to those on the left branch, except that their existence is conceptual. That is, components may be dimensioned, designed, and analyzed in a theoretical framework, using either hand calculations or computer analysis. Because designed structures do not necessarily exist, they may be based on properties of either generic materials or available materials. Tying together both branches of the hierarchy, actual structures may or may not be based on the designs in the theoretical hierarchy.

The ER diagram models each material type or category in a separate entity. A generalization hierarchy is established for generic materials; aggregation hierarchies are established for actual
structures and designed structures. Cardinality numbers are used to constrain the model to valid situations. For example, each component, because it is a unique material product, may be used in only one structure, but a structure is constructed of at least one and up to many different components. A more complex structure may optionally be constructed from many other structures, but again each substructure may only be used in one structure.

In the right branch of the model, each designed component may be created from either a generic or available material, but not both. This restriction is indicated by the XR (exclusive or) constraint across the relationship lines, a feature adopted from the EARL model [Davis90]. However, because the components are theoretical, they may form the basis of many designed structures. Likewise, each designed structure may be used many times as the substructure of a more complex structure. Finally, each designed structure may be the basis of many actual structures, and each actual structure may be based on only one design.

5.2. ER Diagrams for Test Data

To illustrate the conceptual model, we will include a portion of the ER diagram representing properties data obtained from either standard or non-standard tests. Figure 2 illustrates some of the entities required to model tests. The figure depicts the entities as being related to the On-site materials entity, shown in Figure 1; the same entities are used to describe tests on components and structures, although not depicted in this figure because of space limitations. A complete set of diagrams and attributes can be found in [Spaihno94].

A test, described by the Test entity, is either performed by a worker on-site or by a worker at a remote laboratory. The laboratory has a quality assessment measure applied to it as a means of indicating its accuracy and reliability. The Test entity, as well as the Raw data, Specimen property, and Test property entities described below, may also have quality measures applied, although the entities are not shown in the diagrams.

In addition, each test may be described by one or more images, including photographs of the equipment and/or diagrams of the test setup or required specimen. A description of the image and information on how to locate it, or the image file itself, are stored in the Image entity, which has a many-to-many relationship with the Test entity. The Environment entity is used to describe the variation in temperature, humidity, and chemical exposure during the test. The entity has a many-to-many relationship with the Test entity because the environments may be altered many different times during a test, yet each environment record may apply to many different tests.

At this point, the model branches in two primary directions, one for standard tests and the other for non-standard test methods. Because of the complexity involved in both branches, each is considered separately. Note that the processing which the material itself undergoes is not documented in the test entity. It is expected that this information will be stored in entities describing the component or structure itself.
5.2.1. Non-Standard Test Model Structure

Figure 3 depicts the entities required to describe a non-standard material properties test. Non-standard tests are very important in the domain of composite materials because many standard test methods are evolving rapidly or just emerging. The Test description and Procedure steps entities document the test so that its methodology can be used repeatedly to test many different samples. In this way, even though the test method may not be approved by a specification agency, some degree of repeatability can be established for the test results.

To further document the test, the Equipment entity is used to identify the specific pieces of equipment used and to describe the equipment settings. A set-typed attribute is used to describe the settings, because each parameter may be varied many times during a test. The Parameter entity is used in a similar manner to describe test parameters which do not depend on a specific piece of equipment, such as the gage length. The remainder of the entities pertain to the Specimen entity, which describes the specimens on which the test is performed.

Attributes of the Specimen entity include the start and end time of the specimen test and any observations about that specimen. Note that the start and end times of the specimen test can be correlated with the start and end times of the Environment entity to determine the exact test environment for each specimen. The Specimen parameter entity is used like the Parameter entity to document parameters that vary for each specimen. One or more images can be associated with a specimen through the relationship with the Image entity.

The coordinate system for each specimen is described with the Coordinate system entity. With this entity, the origin of the system is defined, as well as the physical meaning of the three orthogonal directions. The Extraction entity describes the position and orientation of the specimen within the original material batch, component or structure, the method of removing the specimen, and any processing procedures applied to the specimen after removal, e.g. polishing, sanding, etc.

Each specimen has a set of dimensions documented in the Dimensions entity, which is not described in this paper. The Dimensions entity represents a link to an ER model component that incorporates dimensions data through a generalization hierarchy. This model allows dimension data to be stored in a minimal amount of space, while providing the most intuitive set of attributes possible for an object shape. In addition, many tests require that the specimen have tabs and/or a hole, which are described with the Tabs and Hole entities and dimensioned with the Dimensions entity. The Gage entity describes the location and orientation of any gages placed physically on the specimen. Finally, if a specimen fails during the test, it can be documented with the Failure entity, which includes attributes to describe the failure mode, and whether or not that mode is considered to be a normal or acceptable failure mode for the test. A relationship with the Image entity is included so that photographs of the failure surfaces can be stored.
Test data and results are documented in several locations. First, arrays of raw data may be defined for one or more individual specimens. Attributes which might be described using the Raw data entity include the variation in load, time, deflection, and/or weight of the sample(s). In addition, atomic properties, such as ultimate stress, breaking strain, or electrical resistivity may be measured for each specimen. The attribute name and measured value are stored in the Specimen property entity.

The overall properties for each test may be reduced from the properties of the individual specimens using a basic statistical procedure. In this case, the reduced properties, including the maximum, minimum, average, and median values and the standard deviation, are stored in the Test property entity. (The number of samples is an attribute of the Non-standard test entity.)

5.2.2. Standard Test Model Structure

The tests represented in this model are many of the standard test methods approved by ASTM Committee D30 and documented in reference [ASTM93]. The test methods included in the model are those dealing with fiber-reinforced polymers; tests describing metal-matrix composites and test methods for sandwich composites and core materials are not considered. The ER model for standard test data is based on the generic model for non-standard tests; however, it is specialized to document the specific parameters required by each test. In addition, the test methods were examined for commonalities in the required data items.

We discovered that certain attributes and entities are common to all tests, others are common to all tests of a specific type, e.g. flexural tests, impact tests, or composite properties tests, and still others are unique to a single test. For example, all of the tension test methods were found to have common data fields, such as speed of testing and tensile strength. However, ASTM D2289 required the work done on the specimen, while the other tension tests studied did not. This finding enabled us to reduce the size of the ER model by using the concept of the generalization hierarchy. For purposes of user interface, the levels of the model can be compressed, so that the user perceives all information pertaining to a single test as being on the same level. The following lists show the tests covered by this research and the category they were placed in when classified according to type of test:

- Tests using chemical reagents or solvents (D543, D4102, D3530, D3529, D3171, C613)
- Flexural tests (D648, D671, D790, D2344)
- Compression tests (D2586, D3410, D695, D3846, D2990)
- Tension tests (D2290, D3039, D3379, D3479, D4018, D638, D2289, D2343, D3163)
- Impact tests (D256, D1822, D3029)
- Composite properties tests (D1423, D3800, D792, D1505, D2584, D2587, D2734, D3355, D3418, D3531, D3532, D696)

ASTM D2990, a creep-rupture test, is classified with compressive tests, although it can be performed on specimens loaded in compression, tension, or flexure. Figure 4 illustrates the top level of the model, which includes data pertaining to all tests, regardless of type, and a portion of
the second level, indicating properties common to each test type. The entities on the second level have additional relationships that are not indicated on this diagram. Because the "is-a" relationship is always a one-to-one relationship, cardinality numbers are omitted from these relationships on this and other diagrams. The Test method entity in Figure 4 contains attributes describing the standard test methods, including such attributes as test title and number, approval date, and a summary of the method.

To document the exact procedure required by the test specification, a relationship could be created between the Test method entity and the Procedure steps entity, described in the previous subsection. The Standard test entity contains information documenting a standard test and, moreover, information which is common to all tests. Deviations from the standard procedure specified by the test in the Methods entity are described in the Standard test entity. The Standard test entity and Test method entity have a one-to-many relationship, because each test is described by one method, yet there can be many instances of each test method in the database. Finally, the Test specimen entity contains information that is common to all specimens regardless of test type; attributes include the start and end time of the specimen test and any observations about that specimen.

Figure 5 shows the second level of the model for the set of tension tests provided by ASTM. Note that the entity names identify the level and branch of the hierarchy. This naming convention is used throughout the ER model. The existence of the third and final level of detail, that of the individual test specifications, is indicated on this diagram by the entities named after those test specifications dealing with tension tests. The other entities at the second level describe properties and parameters which are common to all tension tests.

The Tension specimen entity includes attributes describing the method of preparation and gage length of the specimen. Specimen level parameters, properties, and tab information are modeled in individual entities. Information that is common to all tension tests includes test speed (sometimes referred to a cross-head speed), tensile strength, maximum load applied, and tab material. The tab dimensions are indicated by a relationship with the Dimension entity, so that a variety of tab shapes can be accommodated. In addition, overall test properties, including the mean value and standard deviation for tensile strength and elastic modulus, are modeled in the Tension test property entity.

Figure 6 shows the third and final level of model for ASTM D2289, one of the tension tests on the bottom level of the hierarchy. Information which is specific to ASTM D2289 includes test properties, specimen information, specimen properties, and raw data, each modeled in an entity by the same name. The dimensions of the ASTM D2289 specimen are sufficiently complex to require a link with the Dimensions entity, described previously.

6. ER Models for Common Data Forms and Concepts

An important facet of materials data, especially composite materials data, is that it is recorded in numerous forms, each of which presents particular modeling difficulties to the database designer.
Establishing a relationship between conceptual modeling concepts and technology and common data formats for eventual implementation purposes is necessary. Advanced data types, such as those commonly provided with programming languages, simplify the task of modeling composite materials data, and should be provided within database management systems. These types were listed and discussed in Section 5.4.

In addition, when investigating the types of data required to computerize composite materials information, we discovered several broad concepts that are relevant to various aspects of composite materials data and thus are worth considering individually. The shape and dimensions of objects are important when describing test specimens and manufactured components. Equations and constraints are worthy of separate consideration because they are used in describing parameters, processes, tolerances, tests, design scenarios, and all types of engineering data. Finally, as with most types of engineering data, units are required in almost every aspect of composite materials data. As an example of a concept which is used to describe various aspects of composite materials data, techniques for modeling graph data with Entity-Relationship diagrams are presented in this section.

6.1 Two- and Three-Dimensional Graphs

This subsection presents techniques for modeling graph data with Entity-Relationship diagrams. These components can be altered to suit the needs of particular purposes and attached to the overall model at various points. For example, graphs might be used to describe test results, loading patterns, or surface geometries; therefore, the graph model component can be modified and attached to the global model at numerous locations. Note that the term “graph” is defined more specifically when dealing with materials data than is common in the computer science literature, where a graph is a group of nodes with some connectivity between them. We define a graph as a figure with axes demarcating either two- or three-dimensional space with points plotted in the space.

Graphs have traditionally been very important in understanding various material behaviors. It is much easier to evaluate material properties such as the elastic modulus, yield stress, and ultimate stress using a graphical representation of the stress-strain behavior of the material, as opposed to a tabular representation. A graph, as defined here, is a physical representation of a collection of data points in two or three dimensions and can, in fact, be maintained in a database as such. However, we are interested in storing other information in order to display the graphical data in a manner desired by the user, e.g. data describing the format of the axes, the points, and the title, for example. Types of graphs which are important in modeling materials data, including various types of splines, regression curves, and curves with error bars, are considered in the model.

For modeling purposes, we have distinguished between line graphs and surface graphs. Line graphs consist of one or more individual curves which are not connected in any way. The points may have either two or three orthogonal coordinates. All points on a surface graph have two curves passing through them in orthogonal directions. The curves connecting the points form a representation of a three-dimensional surface. For this reason, surface graphs have only three-dimensional points. However, model components are included for surface graphs which consist
of arbitrary (x, y, z) points and those in which the z-coordinate of a point is always a function of the x- and y-coordinates.

Figure 7 provides ER model components for the data required to describe a graph. Entities are included in the figure; attributes are not enumerated because of space limitations. The Concept entity represents the numerous points at which an image can be attached to the overall ER model. In the full diagram, the concept might be a Component, a Property, or anything which might be supported by a graph. The many-to-many relationship indicates that each concept may be described by many graphs, and that each graph may be used many times to support different concepts in the database.

Each graph has from two to many axes, and from zero to many text blocks, which can hold any type of descriptive or explanatory information. In addition, a graph type, title, and textual description may also be defined. For surface graphs in which the z-coordinate is a function of the x- and y-coordinates, the x- and y-coordinates are stored in one-dimensional arrays, represented by the list of real type. The z-coordinates are then stored in a two-dimensional matrix; the row and column indices of each z-coordinate value are equal to the indices for the corresponding values in the x- and y-vectors, respectively. In other words, the x- and y-coordinates define a two-dimensional grid of uniform structure, and each z-coordinate corresponds to a single location on that grid.

The underlying structure of the remaining three types of graphs states that graphs are made up of lines, which are made up of points. Attributes are included to document the type of surface, lines, markers, and points used in the graph. The line type describes the manner in which the points are connected to form the line; point-to-point lines are permitted, as well as splines and regression curves. For line graphs, the point coordinates may either be stored in the graph entity, or extracted from other numeric attributes in the database.

Sufficient information is stored throughout the ER model to enable an application program to construct various types of splines, regression curves, and curves with error bars. The m, n, and k parameters, used in constructing B-splines and Bezier curves are modeled in the Surface graph and Line graph entities. The applicable equations for regression functions and blending functions, also used in creating B-splines and Bezier curves, are stored in equation form in the Blending function and Regression function entities.

7. ER Models for Reducing Size of Data Sets

As indicated previously, the data in the domain of composite materials is both large and diverse. Moreover, because of the broad range of data needs over the composite's life cycle and the variety of experimental and service data which can be obtained, the set of domain data rarely consists of large numbers of instances of fixed format records, as is common with business data processing. To address the issue of data reduction, we investigated various techniques applicable to or currently being used in the field of composite materials. Entity-Relationship model components were developed to support these data reduction activities.
An important component of data reduction with respect to composite materials is the reduction of experimental properties into design properties which describe the behavior of entire populations. Experimental property values are often measured from multiple samples, which are extracted from one or more batches drawn from the population at large. Complicated statistical data reduction techniques are applied to these large numbers of experimental property values to compute reduced property values (average, mean, B-basis) which are applicable to the entire population. One well-established procedure for evaluating property values from individual tests is provided by MIL-HDBK-17, the military handbook describing polymer matrix composites [Army92].

The data reduction procedure outlined in MIL-HDBK-17 is that which must be followed to include composite material properties in the handbook. It is applied primarily to strength and strain-to-failure properties of composite laminates. However, the presented techniques are applicable to a broad variety of materials and properties. From the set of experimental values, the following reduced properties are computed:

- Typical/mean/average value,
- Minimum value,
- Maximum value,
- Standard deviation, and
- B-basis value.

The B-basis value is the property value above which at least 90 percent of the population is expected to fall, with a confidence of 95 percent. An A-basis value can also be evaluated, which is the property value above which at least 99 percent of the population is expected to fall, with a confidence of 95 percent. Therefore, if the B-basis value for the tensile strength of a sufficient collection of [0/45/-45/90]₄s graphite-epoxy laminates is evaluated to be 40 ksi, then 90 percent of all similar graphite-epoxy laminates are expected to have a tensile strength of at least 40 ksi, with 95 percent confidence.

We feel that the data reduction procedure presented in the military handbook is well-founded and can be used as the basis for any statistical data reduction which accompanies the data in a composite materials database. For this reason, we used this procedure as the basis for the ER model for statistical reduction of experimental properties, depicted in Figure 8. The entire MIL-HDBK-17 data reduction procedure was examined and the resulting entities, attributes, and relationships were evaluated. This information was used to generate the ER model for the statistical analysis and reduction of multiple samples. A summary of the procedure, which can be correlated with the ER diagram in Figure 8, may be found in reference [Army92].

8. Conclusions

The first conclusion arrived at from developing a conceptual data model for composite materials is that there is no one "correct" model. Again, it was very clear that no single data model could encompass the data requirements of all types of fiber-reinforced composites, or the data needs of all engineering application programs and users. Instead of providing one comprehensive model,
therefore, we developed a modular, extensible conceptual model from which components can be selected and modified as needed. In addition, new components can be developed and added to the model in a relatively straightforward manner.

Regarding the suitability of the ER methodology, in general, we feel that it is very suitable for composite materials data, although certainly not the only suitable model. The high level of abstraction in viewing the data aids the modeler in obtaining a fundamental understanding of the concepts underlying the domain. With the addition of nonatomic and complex attributes, the resulting model is much simpler and more reflective of the actual domain being modeled. However, with a complex domain area, such as required for composite materials, it is difficult to conceptualize the domain and reduce it to the level of entities, attributes, and relationships. We feel that this complaint is not specific to the Entity-Relationship model, but rather the usefulness of any formal methodology is inversely proportional to the size and complexity of the domain to be modeled.

To implement the composite materials conceptual data model, we recommend that the extended relational database model could be used. The extended relational models provide an intuitive mapping of complex physical systems into the model, allowing data to remain in hierarchical form, if necessary, and supporting inheritance within the database schema. Abstract data types, sets, and lists are often permitted for defining field types. However, other relationships including data abstractions are supported through the inherent "relation" association. In addition, an extended SQL interface is provided, so that the database designer does not need to implement methods to define and manipulate the data, as required with an object-oriented system.

In modeling composite materials data, one must conclude that there is an inherent difficulty in modeling composite materials data because of the extreme variability in shape and form of the composites. In the Standard Guide for Identification of Composite Materials in Computerized Material Property Databases (E1309), ASTM describes the difficulty in modeling a composite material as follows: "Classification of composite materials is complicated by the fact that composites are formed by combining different materials in varying amounts and configurations, resulting in an infinite number of possibilities. An effective identification scheme must be capable of describing the majority of possible combinations without overburdening the system with details relevant to only a limited number of material systems [ASTM93]." We elected to model the structure of composite materials with the "Component made-into Structure" and "Structure made-into Structure" relationships, shown in Figure 1, which allow the modeling of infinitely complex structures. The details of the resulting shape and form are modeled in the attached entities which describe the dimensions of the components and structures and the processing required, including pultrusion, braiding, weaving, and other processes which apply shape and form to the component or structure.

A tradeoff always exists between the size and the representative ability of a data model. It is possible to develop a small model with few entities which are very generically defined and therefore applicable to many situations. Another difficulty in modeling composite materials arises from the vastly different amount of information available on different materials, and from
the vastly different forms in which it is provided. For example, while one supplier might only provide a base property, e.g. tension strength, for a given material, another might indicate the test method, sample size, and test conditions used to generate the data, as well as statistical information, such as the average and median strength of a specific number of samples. As a information system modeler, the challenge is to provide a conceptual model with fields for as many attributes as possible, without wasting a large amount of space on data which is not commonly provided or reasonably available.

Secondly, data is often available from suppliers, tests, and external databases in entirely different formats. One supplier might indicate the maximum available sheet width for a fabric, another might supply a set of available widths, and a third might simply state that custom widths are available. To model each of these forms of data requires, respectively, a single real attribute (e.g. Maximum width), a set of real attributes (e.g. Available widths), and a text attribute in which the custom widths can be discussed (e.g. Width). Here, the challenge is to select a model which reflects the needs of the majority of the data in content and in format, while providing enough variability to include atypical information which might prove very important.

We have addressed this issue by including a highly normalized data model in which each concept is modeled in a unique entity or set of entities. This isolates individual concepts, so that each material and property can be described to the highest degree possible according to the amount of information available, yet space is not wasted in the resulting database by attributes which are frequently uninstantiated. In addition, note fields are provided throughout the data model to allow the user to describe special circumstances and requirements.

In conclusion, this work represents a first attempt to provide a comprehensive look at fiber-reinforced composite materials and the data needed and supplied during their use in civil and structural applications. One of the contributions of this work is to formally enumerate the domain. This effort is summarized in reference [Spainhour95]. Other contributions include subsequently developing the information and conceptual models for composite materials, as described in this paper. Finally, the overall contribution of this research is to support an automated design process by providing a foundation for the development of historical composite material databases and computational systems.

References


[ASME93] "ASME Compiling Data on 20,000 Alloys," Mechanical Engineering, American Society of Mechanical Engineers, May 1993, pp. 70.


Figure 1: Overall structure of Entity-Relationship model for materials database

Figure 2: Entity-Relationship model for test data
Figure 3: Entity-Relationship model for non-standard test data

Figure 4: Top level of Entity-Relationship model for standard test data
Figure 5: Second level of Entity-Relationship model for tension test data

Figure 6: Third level of Entity-Relationship model for ASTM D2289
Figure 7: Entity-Relationship model components for graph data
Figure 8: Entity-Relationship model for statistical reduction of sample data