DEVELOPING AND IMPLEMENTING A CONCEPTUAL COMPOSITE MATERIALS DESIGN DATABASE

by William J. Rasdorf¹

Member, ASME, and

Lisa K. Spainhour

ABSTRACT

Researchers and materials engineers require a greater understanding of the problems and solutions that emerge when integrating composite materials data with computer technology so that utilitarian composite materials databases can be developed to effectively and efficiently support analysis and design software. Composite materials constitute a representational challenge due to their composition and use. However, this paper suggests that a conceptual composite material data model and application software interfaces must be developed to support the dissemination and use of composite materials data.

This paper primarily serves to analyze several of the problems facing developers of composite materials databases, evolving from the complexity of the materials themselves and from the current lack of testing and data representation standards. Without a clear understanding of the scope and nature of these problems, there is no possibility of designing concise yet comprehensive composites data models, yet we feel that such an understanding is presently lacking. In addition, an effort is made to present possible solutions to these difficulties being suggested and/or implemented both by the authors and by other researchers in the field. Such an effort provides a firm foundation upon which future research may be based.

1 INTRODUCTION

The emergence and use of composite materials in engineering applications has increased considerably. A need therefore exists to make information available about composite materials, their derivation, their properties, and their use, to researchers and material engineers. To do so requires a greater understanding of the problems and solutions that emerge when integrating composite materials data with computer database technology [McDonald 90]. Gaining such an understanding will enable the eventual development and operation of utilitarian composite materials databases designed to support a collection of analysis and design software.

In dealing with most types of engineering and scientific data, there is ultimately a need for data, object, and knowledge management. Traditional applications where large numbers of fixed format records must be stored and accessed require the transaction management and query facilities of data management. Object management entails efficiently storing and manipulating nontraditional data types such as bitmaps, diagrams, text, and graphs. Finally, knowledge management includes the ability to

¹ Department of Civil Engineering, North Carolina State University, Raleigh, NC 27695

store and enforce a collection of rules that are part of the semantics of an application, describing integrity constraints as well as allowing the derivation of data [Stonebraker 91].

Managing composite materials data requires the incorporation of data, object, and knowledge management. Individual data items, such as stiffnesses, strengths, and dimensions need to be stored and maintained using traditional data management concepts. Additionally, plies, laminates, and components can be thought of and manipulated as objects. Furthermore, many other nontraditional objects can be used to describe composites, such as pictures, diagrams, and graphs; these require extended object management capabilities. Finally, numerous rules and constraints govern material behaviors, test conditions, and design criteria. A facility for supporting such rules and constraints is desirable.

The focus of this paper is on the data needs of composite materials. The description of material property data is discussed, including material and composite properties, geometries and topologies, test methods and procedures, processing methods and procedures, and loadings and conditions of use. Designed components themselves will not be addressed. However, both object and knowledge management are briefly considered in discussions on image management and the role of knowledge-based systems.

The objective of this paper is to establish the requirements for creating a composite materials design database and data retrieval system for use in structural analysis and/or design applications. The target database is fundamentally an engineering properties database that is designed to support the needs of analysis, design, and data retrieval software. The eventual objective is to develop a prototype database environment to optimally support these needs: to do so requires a combination of the fields of materials, engineering, and computer science. Primarily then, this paper serves to analyze several of the problems facing developers of composite materials databases. Without a clear understanding of the scope and nature of these problems, there is no possibility of designing a concise yet comprehensive composites data model.

2 COMPOSITE MATERIAL PROPERTY DATA

The study and use of composite materials has grown dramatically over the last twenty years. Composites are mechanical mixtures of distinctive constituents which can be combined in order to achieve a synergistic blend of properties that cannot be obtained through the use of any of the constituents alone. Composite materials present a very special challenge to efforts to computerize their properties and performance specifications, because they are uniquely "designed" for specific applications, i.e., they are difficult to classify since they are as various as the engineering circumstances they serve.

In its most basic sense, a composite consists of a matrix, which is contiguous throughout the section, and one or more filler materials, which are embedded in the matrix. Depending upon the required application and the desired characteristics of the composite, the matrix may be metallic, ceramic, or a polymer resin. The filler may also be composed of one of these material types, although carbon, aramid, and glass fillers are the most commonly used. Again, depending on the desired end use of the composite and the characteristics of the filler, the material may be formed into particles of various shapes and sizes, whiskers, short fibers, or continuous fibers. The filler may then be placed randomly, or laminated, woven, or braided in regular patterns throughout the matrix.

2.1 Factors Inhibiting Composite Database Development

A number of factors significantly inhibit the development of a database environment to fully support composite materials. These include but are not limited to the following:

- 1) Limited access to existing data resources.
- 2) Inability to search for data that satisfies constraints on specific material properties.
- 3) Diversity of material property descriptions.
- 4) Lack of suitable data types for engineering data.
- 5) Broad range of data needs over the composite's life cycle.
- 6) Lack of user friendly standard interfaces to engineering applications software.
- 7) Diversity in quality assurance for data.

Several of these issues are discussed in more detail throughout the remainder of this paper. Furthermore, a variety of approaches to overcoming these problems are presented and discussed. While we do not claim to have solutions to all of the difficulties facing developers of composite materials databases, we feel that the approaches presented herein provide a firm foundation on which to base future research efforts.

2.2 Composite Material Data Description

Composite materials constitute a representational challenge both because of the properties of their micro-constituents and because of their arrangement. The global properties of composites, for example, are generally anisotropic such that their response depends on the direction of loading relative to certain special directions defined by the composition. The immediate consequence of this situation from the standpoint of representation is that a single property, such as stiffness or strength, becomes

an array of properties which requires multiple entries to provide a complete description of the material characteristic. Moreover, the micro-constituents of a composite are often combined to make macroscopic structures, which are again combined in various ways to obtain composites with vastly different properties.

An example from the domain of textile composites involves several levels of combination: fibers are joined together to make yarns, yarns are braided or woven into preforms, and preforms are injected with resin to make fiber-reinforced plastic composites. From the standpoint of database representations, this set of multiple levels is in stark contrast to the material properties and characteristics that are commonly associated with "traditional materials." Such properties and characteristics provide a unique and permanent definition of the engineering utility of those materials and are obviously much easier to represent with the tabular arrays in a database. For a discussion comparing the features of various database models, see Section 4.3.

Another feature of composite materials is the need to characterize the properties and response of these materials under unusual conditions. Such conditions frequently involve combinations of severe environments such as cyclic or dynamic mechanical loading, extremely high or low temperature, and severely aggressive chemical exposure. The arrays of properties and performance characteristics mentioned above must be determined as a function of various combinations of these influences that are of interest to specific applications.

Finally, the challenge is further heightened by the necessity to characterize changes in properties and performance as a function of the history of application of the influences commonly applied to these materials. For example, the most important information for materials to be used at very high temperatures is their rate of degradation under a specific set of service and loading conditions, since the utility of the materials is limited by those rates. Moreover, these service-related characteristics depend on the manner in which material properties and performance change during service. The quantification and description of these changes is essential to a complete description of composite materials.

Obviously, the concern for studying material response over time and under a variety of loadings and environmental conditions is not unique to composite materials; however, with composite materials there are several additional factors to consider. First, the response of composites to a variety of simultaneous influences is not as well understood as the behavior of more conventional materials. Secondly, even where the response of composite materials is fairly well understood, both the anisotropy and the macroscopic nature of the material causes unexpected effects. Therefore, the

response of composites under complex loading and/or service conditions must be carefully and thoroughly studied, and appropriate data structures must be developed to store and maintain the resulting data.

2.3 Features of a Composites Data Model

Thus, we see that there are many factors inherent in composite materials property data and current DBMS technology which hinder the development of composite materials databases. As a result, an appropriate database representation for composite materials must account for many difficulties, and provide numerous features. At minimum, the data model should be able to maintain data describing and/or documenting the following:

- 1) The microstructural configuration and shape of the composite and the direction of loading,
- 2) Property definitions of the constituent materials,
- 3) The environment in which they will be used,
- 4) Fabrication and synthesis techniques,
- 5) Standardized and/or nonstandardized test methods and characterization schemes for the determination of properties and performance, and
- 6) The combined response of such materials to multiple excitations and circumstances, such as mechanical, thermal, and chemical loadings.

It is very difficult to find a general representation, be it an empirical formula, mathematical model, or a numerical enumeration, of these highly variable circumstances and combinations of constituent properties and their configurations. Yet consideration of such a general representation will result in a better understanding of representations that will, in turn, lead to improvements in database design for specific applications. However, even if many such general representations are found, an appropriate database model must still be able to represent the very complex nature of composite materials. This statement means that a number of unique circumstances regarding composite materials must be considered:

- 1) In addition to individual data items, arrays of data items will be needed to represent many composite material properties,
- 2) Changes in the structure of the composite material, in its environment of use, and in its loadings may result in different array values, and
- 3) Array values may be time dependent as well.

The key challenge from a database design perspective, and the problem which must be addressed by both the selection of a database management system (DBMS) and the overall database design, is, therefore, the representation of appropriately connected arrays of data item values.

2.4 Determination and Use of Material Properties

Beyond the complexity and sophistication of the descriptive representations suggested above, efforts to establish databases for composite materials are further frustrated by the fact that test methods to determine material properties and characteristics are still under development and thus are not all firmly standardized or accepted as consensus methods for property and performance determinations.

Many end-users of materials data claim that maintaining either actual data on parts in service, or at least laboratory simulations of actual service conditions, provide more useful measures of properties and performance than standard lab procedures currently used. It is hoped that such data will provide insight into previously untested factors including processability, appearance, and manufacturing costs. However, because manufacturing processes and service conditions are influenced by multiple parameters, it is very difficult to compare results and detect trends [Westbrook 86]. Largely because of the complexity of the situation, test methods are commonly developed for circumstances of particular interest to various applications, and variations in those test methods are frequently made as variations in the applications suggest. A significant proliferation of new test methods, along with a proliferation of new materials, thus complicates the data interpretation problem.

Although notable efforts are being made to establish test methods, especially by professional societies such as American Society for Testing and Materials (ASTM), these efforts seriously lag the development of composite materials. Consequently, the data arrays of time-dependent material properties mentioned above frequently suffer from absent entries, and suffer even more greatly from multiple entries which differ because of the differences in methods of determination rather than because of differences associated with the material system of performance. Properties such as stiffness and strength suffer less from this malady than do performance characteristic definitions for phenomena such as creep rupture, static and dynamic fatigue, and damage accumulation.

Another important feature related to the determination of properties and performance is the importance of statistical variations of these characteristics in composite materials. It was mentioned earlier that processing plays a significant role in the final character of such materials. In general, not only is processing difficult to control and difficult to repeat, but it presents a number of difficulties in and of itself. Consequently, variations in properties are surprisingly large from "batch to batch," or

even within a given batch, or in some cases within a specific specimen or component. It is impossible to properly discuss the properties and performance of these materials without taking a rigorous quantitative approach to such statistical variations. Indeed, the design of structures with composite materials, and their subsequent certification for service, is frequently based on allowable levels of applied loading determined from a statistical analysis of an allowable rate of failure determined from these statistical variations in material properties and performance.

2.5 Documenting and Tailoring the Data Set

To address these concerns about test methods and statistical variations, information regarding the manner in which data was obtained and the manner in which materials were prepared is essential. This data is commonly known as metadata, or data about the data. Several authors have noted the importance of storing metadata, including notes relative to the identification, authorization, and reliability of data to be placed in a database [Westbrook 86]. The proper use of metadata helps to ensure that the input data is relevant and is of acceptable form and quality.

To completely characterize a material and a particular test, Harold Mindlin of the Mechanical Properties Data Center (MPDC) [Mindlin 86] suggests that the following items be recorded for each set of test data:

- 1) Data Source (reference, laboratory, traceability),
- 2) Material description (source, lot number, composition, heat-treatment, etc.),
- 3) Test description (standard methods, details of non-standard procedures, equipment, loading procedures, etc.),
- 4) Test environment (temperature, chamber, corrosive chemistry, etc.), and
- 5) Raw data (test values, etc.).

While raw data does not specifically qualify as metadata, it is important to maintain in a materials database, to the extent that space allows. First, raw data might be necessary to validate derived data at a later date. Secondly, raw data can be used to derive new data that was not originally conceived of when the test was performed. Note also that Mindlin suggests that details of non-standard test procedures should be stored. Efforts to achieve standardization of test methods are critical, but nonstandard situations must be expected and properly handled., particularly while standards are under development.

Committee E-49 on Computerization of Material Property Data, whose mission is to develop standard classifications, guides, practices, and terminology for building and accessing material property databases is developing similar guidelines to aid organizations and individuals that are

involved in building databases or intelligent knowledge systems [Kaufman 88]. E-49 guidelines suggest that the following types of information will be useful for test data sources: general test type; specific test method; specific dimensions, location, and orientation; test independent variables; property descriptors; specific test results (including raw data); and all validity requirements and any individual data or assessments needed to establish validity. The recommendations of these and other standards organizations should be supported and enhanced through the development of a precise and accurate data model for composite materials metadata.

The complexity and sophistication of computerized descriptions of composite materials may exact somewhat peculiar requirements on such databases. Since these materials are so strongly driven by applications, it may be that databases will be driven by groups of applications as well. Some of the databases that have currently been established are tailored to groups of users who have peculiar applications or classes of applications which are best served by the structure of the database involved. Since the cost of these databases is substantial, and cost recovery is a persistent and potentially disabling problem associated with these databases, it may be that tailoring databases to a specific market is a viable development route. However, in order to achieve generality, gateway arrangements are necessary to combine databases developed for specific purposes.

Thus, relative to the development of an appropriate representation for composite materials three additional considerations emerge and must be accounted for by the materials data model:

- 1) The lack of standardization and quality control in test methods means that some data item array values are unreliable and that some are missing,
- 2) Metadata and raw test data must be maintained to the extent possible to ensure the quality and reliability of the database and data, and
- 3) The use of data may be application dependent and a centralized general data representation may not be achievable, cost effective, or necessary.

3 A CONCEPTUAL COMPOSITE MATERIAL DATA MODEL

The purpose of this section is to consider the conceptual requirements of a composite materials database. To do so requires an exploration of the structure of materials information. A classifier hierarchy provides a sensible methodology through which the vast amount of materials data can be analyzed.

3.1 A Classifier Hierarchy for Materials Data

In general, materials data can be separated into a classifier hierarchy with three primary levels, as represented in Figure 1. The information at each level of the hierarchy is further subdivided

into relevant classes and areas of interest. At the lowest level is the data describing the material itself, e.g. steel, aluminum, concrete, fiber-reinforced composites, etc. The second level describes the shapes or components which are formed using these materials. Examples here include the AISC steel sections, and the standard composite shapes produced by many pultrusion companies.

The third and final level includes data describing the structures, systems, and assemblies which are manufactured using the components of the second level. Note that complexity increases as one travels up the hierarchy, as the structures at the top level may be composed of numerous shapes which, in turn, may be composed of one or more materials. The physical manifestation of the completed hierarchy will be the schema of the database itself.

Figure 1: Generic material classifier hierarchy

Taking the level of structural shapes, for example, one can subdivide the hierarchy to identify observed objects (components), performances (limit states), and functions (stress states), among others. In this context we define: an object as the physical entity being addressed; a component as a constituent part of an object; a limit state as a state of incipient unsatisfactory behavior in a specific mode (such as limiting loads at which failure occurs); a performance as a measure of the function that a component must provide for its users; a function as the action for which a component is specially fitted or used; and a stress state as a type of induced force per unit area on a component. Using these categories of classification, one can begin a descriptive characterization like the hierarchy in Figure 2.

Figure 2: Detail of classifier hierarchy describing structural shapes.

As an example, consider that a composite I-beam has several members (two flanges and a web), limit states (ultimate capacity in tension, maximum unbraced length, etc.), and, under a specific loading condition, various stress states (compression, shear, torsion, etc.). This portion of the hierarchy, however, is not complete; it does not categorize attributes of objects either in terms of qualities and properties or in terms of processes and behavior; it does not categorize relations and interactions between objects; it does not categorize operations on objects; nor does it categorize location, condition, or time.

The intent of the hierarchy is merely to demonstrate the large number of data items of interest in the area of engineering materials and to show some of the hierarchical relationships between those data items. It is also important to realize that, with traditional materials, most of the data to be maintained falls under the realm of component or shape data. The basic behavior of a material such as aluminum and its response to loading, environmental exposure, and/or high temperature is well documented and well understood. One of the most outstanding features of composite materials data, however, is the large number of classifiers or descriptive items which are necessary to precisely identify characteristics observed at the "material" level of the above hierarchy.

3.2 Modeling Special Types of Materials Data

An important facet of materials data, especially composite materials data, is that it comes in numerous forms, each of which presents particular modeling difficulties to the database designer. Data

in numerous forms such as text, graphs, and images are elements of the overall data model. Ways to store and maintain this data are presented within this section.

3.2.1 Text and Bibliographic References

An important category of descriptive information that is vital to the interpretation of composite materials data is that of textual notes. These are also best modeled by a classifier category. For search purposes, it is best to enter textual notes as specific attributes, if possible. In this way, textual notes might include a keyword, such as "warning" or "application," followed by a brief descriptive phrase. For example:

exposure = suitable for chlorine exposure

Particularly for the newest and most specialized materials, for which experience and understanding has not reached a point of convenience and complete classification, there may be a need for such descriptive attributes.

As a domain unto itself, the collection of textual notes may be modeled as a one-dimensional vector. This vector establishes an open-ended list with the text of the notes and the number of notes both varying. The data describing a particular material then references the note vector, identifying each characteristic attribute as being present, absent, or undetermined. Thus, whenever two different materials share an identical attribute, the same comment is referenced. Distinct vectors may be appropriate for different material categories, which may have vastly different applications and behavior characteristics.

Textual or bibliographic references are also an important class of data to be managed in a composite materials database, because it is through references to external literature sources that the information in a materials database is documented, verified, and supplemented. Textual references might be included as literal strings, containing short quotations from relevant materials literature. They may also be indirect, namely references to published literature. Of course, if the reference is to a large object, for example, a reference to a textbook, only a small part of the referenced information will be relevant. For an actual implementation, direct and indirect information may be combined: the abstract itself may be directly available; for more detail, one may indirectly access further data via the author's name and publication information. In either case key words play a major role in the access process. Thus, the system can not only process attribute/value pairs but it can also locate documents and descriptive text based on key words from which the user can further extract values.

3.2.2 Graphs and Equations

Two additional categories of descriptive information are particularly noteworthy and important: graphs and equations. A graph represents a collection of data points and can, in fact, be represented as such. However, to appropriately computerize the data by generating a graphical representation on-screen, applications such as automatic and manual scaling and axis labeling must be implemented. Curve fitting and generation applications are also necessary. It would also be appropriate to be able to derive data directly from a graph. Items such as slope, maximum value, etc. could be maintained and accessed as distinct attributes, yet obtain their values from a stored graph. Complex graphical representations are very important in understanding the behavior of composite materials.

It is also, at times, necessary to store equations. In one form, equations can be stored in a text field. Such a representation, however, is useful for reference only and is completely non-executable. One alternative is to embed equations in application programs which draw their constants, coefficients, and variable values from the database. A database design is required, therefore, that will minimize the number of accesses to different database tables and that will provide values for all of the constants and variables necessary for the execution of the equation. Another alternative is to store equations as constraints. However, this requires the development of a constraint management system to parallel the database management system to enable the definition, instantiation, updating, and invocation of constraints. Such a system is recommended and described in [Rasdorf 87].

3.2.3 Images

Much of the data gathered in test laboratories today is currently maintained in a paper form, as complex charts, diagrams and sketches, and two- and three-dimensional gray scale images. The use of images is highly important in understanding the nature of composite materials, because both the microscopic nature of certain constituent materials and the nonhomogeneity of the entire composite often cause unpredictable behaviors which are difficult to monitor by traditional means. In one study of braided textile-reinforced plastics, a combination of geometric modeling and optical microscopy was used to study the role of braid reinforcement microgeometry and secondary fabrication effects on the final yarn orientation and content [Masters 92]. Another instance where optical microscopy is commonly used to identify and explain composite material behaviors is in the area of fracture, deformation, and failure.

More research directed at computerizing such physical data representations is necessary so that they can be utilized in conjunction with an engineering database. An extended relational data model, described in more detail in Section 4.3, might be used to store the types of graphical and image data discussed above. With such a model, individual data items could still be stored in tabular fields. Data

items of other classes, such as grey scale images and two- and three-dimensional graphs are used to generate and/or verify specific numeric values, and should be maintained as metadata. These diverse data items would be stored in properly formatted files; the corresponding file names would then be stored in relational fields. When the individual data item is selected from the relational table, the relevant file names are accessed, and the supporting data in those files displayed on the screen. Obviously, as with the graphical data described above, an appropriate set of operators needs to be coded to support such extensions within the chosen database model.

3.3 Data Value Density: The Size of the Data Set

We now describe the elements of a data model. While the concept of a relational data schema is intuitively simple, it still lends itself to a degree of complexity because of the diversity and complexity of the material hierarchies described in Section 3.1. For any realistic set of attributes, the schema is infeasible and large. 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 data set rarely consists of large numbers of instances of fixed format records, as is common with business data processing. Instead, there may exist data sets from many different tests, each in a different format, recording different numbers of attributes and different sets of metadata.

Any materials database schema that attempts to be comprehensive will be both large and diverse. This becomes a problem as users attempt to access and use the data. The speed with which data can be retrieved will quickly deteriorate as the size of the database increases. Further, without some control over the manner in which the highly diverse data is presented to the user, the information contained within will be unintelligible. Two schema characteristics, Interpolation and Reduction, can be utilized to reduce the size of the database schema, but not without introducing some complexity.

3.3.1 Interpolation and Extrapolation

Because of the large and diverse data set associated with composite materials, some type of condensation of data must occur for the effective use of a database. Associating an interpolation function with an attribute can greatly reduce the number of entries in its domain; only critical points may need to be explicitly represented. However, more than one function may be needed for distinct ranges such as phase changes or non linear structural behavior. For example, the raw data of a stress-strain plot may be retained as metadata, given sufficient storage space, but for input data to many analyses it would be more appropriate to represent the behavior as a piecewise linear approximation. Thus, while maintaining only the critical data points, the overall material behavior is quickly and easily accessed.

Using the same function for extrapolation may also cause a reduction in the number of explicit data items, and these may be directly proportional to the number and naturalness of the parameters used to describe the function. Naturalness means a function which mimics the physical behavior of the materials system, rather than a general polynomial function which is adequate by numeric measures but which requires ever increasing terms to span wider ranges of values.

Extrapolation is quite a different matter than interpolation in that the bounds of uncertainty can usually be precisely stated for the latter, but are large and frequently indeterminate for the former. Furthermore, the appropriate interpolation or extrapolation function and its adequacy (both in the mathematical and naturalness senses) varies from property to property. This is not a problem with interpolation, unless an adequate interpolation function is as complex as the original data set. However, extrapolation must always be approached carefully, because, since one is inferring values over an unobserved interval, it is difficult to know whether or not a chosen function is more or less adequate than another.

3.3.2 Reduction

Reduction indicates the decrease in data content when one or more independent attributes are not recorded. In the context of the database it simply means a reduction in the number of data items recorded. Reduction may occur because the underlying information may not exist, or it may be necessary because the available data is very sparse. There may not always be enough data on heat treatment or test-sample-configuration, for example, to justify its presence in the database. Furthermore, when considering the identification of a composite, it may not always be possible to have information about the specific batch number.

The problem with data reduction is, of course, the loss of information for those variables which depend on all of the missing attribute values, i.e., although sparsely available, there may be conditions under which those attributes are highly useful in determining values for other data items. For this reason, reduction must also be used cautiously, weighing the benefit of quicker data access against the cost of a less comprehensive database. Note that reduction is an example of a decrease in data vertically and may be contrasted with the horizontal data reduction provided by interpolation. That is, interpolation causes a sparsity of data at one level in the data hierarchy whereas reduction causes a sparsity of data that propagates through multiple levels of the hierarchy.

3.4 Data Value Modifiers: The Accuracy of the Data Set

The set of data itself is rarely absolute. There will be variation in the material, variation in the measurements, questions of trustworthiness of the data, as well as outright errors. For example, most

data is obtained by measurements. Associated with specific sets of measurements there may be degrees of tolerance and uncertainty arising from the tolerances of the test machine, analog to digital or other conversions, or errors in reading dials and gauges or interpreting test parameters.

Uncertainty and tolerances may be handled by relying on data types such as ranges where exact values are not available. However, we cannot currently deal well with errors using existing database management systems, except by setting error bounds, which will cause erroneous data to be rejected upon input. Several authors note that data validation and quality assurance are important factors that will require more attention in the future. Most feel that this is a labor-intensive process which must be performed manually and is the responsibility of the materials community as a whole [Landgraf 80, Future 85]. Kaufman suggests maintaining the proper metadata necessary to evaluate whether or not a test met the standard conditions specified by ASTM or ISO, for example, then applying validity criteria and establishing how well the test conditions met the specifications [Kaufman 86].

The users' belief in the validity of the data might also be measured and stored. A specific user might associate a belief assessment with the data source, the material source, or a specific test set. ASTM Committee E49 describes numerous criteria on which to evaluate the quality and reliability of a set of data [Kaufman 88]. Data can be evaluated at numerous levels within a database, beginning with the results of individual tests and progressing to data sets (groups of test results describing a measured characteristic), average or typical properties, and design values. Quality can be assessed by considering, along with other factors, the experience and integrity of the testing organization, the completeness of the test method description, the statistical procedures used to derive material properties, and the degree of evaluation by expert groups.

While ASTM suggests numerous criteria on which data quality can be assessed, they do not provide any guidelines for statistically or methodologically analyzing the belief evaluation. There are several well known methodologies which can be used for evaluating and documenting belief, including fuzzy logic, probability theory, various statistical analyses, and simple ad hoc assessments based on the available criteria. A complete materials data model would provide the framework and necessary operators for computationally assessing quality by several commonly accepted methodologies. This would allow the users to express their belief in the validity of the data, at any of the above levels, in whatever manner they find appropriate for the particular data set.

A general model of user belief would associate the belief with a crossproduct of independent attributes such as Source x Material. That is, a number of test sets, from tensile tests on a particular composite, for example, could be gathered; the tests might have been performed on a variety of

specimens by a number of different testing organizations. By evaluating the available assessment criteria using one of the belief methodologies, a matrix could be formed where each coefficient indicates the measure of quality, as determined by the users, of one test performed on a specific test sample by one of the organizations. Auxiliary tables, keyed to the appropriate attributes, unnotched tensile strengths for example, might then represent the users' interpretation of belief in the overall collection of data sets for that property.

3.5 Materials Database Model Components

We now describe the variables required within the database schema and the elements which might be included in the database model. While the information in this section is consistent with, and in some cases summarizes, the ideas presented in previous subsections, it is also drawn from the NSF Materials Database Workshop paper cited in reference [Wiederhold 87]. According to that report, a database can contain at least four types of variable or data:

- 1) Identifiers (graphite/epoxy, T300/5208, ASTM Test D-403),
- 2) Descriptive characteristics (suitable for horizontal tail skin, avoid contact with jet fuel),
- 3) Independent variables (cure temperature, test environment, strain rate), and
- 4) Dependent variables (ultimate tensile strength, percent elongation).

In this example, identifiers determine descriptive characteristics, and both the identifiers and the independent variables determine the dependent variables.

Additionally, there are a number of components which might be included in a materials data model. The relative importance of the components can differ greatly; some may not exist at all in some materials databases. If an existing database system approach is imposed on the materials database, then it is likely that some components will be missing or under-emphasized. For example, the first normal form restrictions of a traditional relational DBMS require that only atomic values (rather than sets) can be stored in individual data cells.

The components which might be included in a materials database model are:

- A base matrix of data values
- A set of values within a data cell of the base data matrix
- The independent variables identifying the data values within the matrix

- Sets of valid domain values for use in identifying data values, e.g. real numbers
- Data values at higher levels of abstraction, that is reduced data, omitting one or more identifiers
- Interpolation/Extrapolation functions to permit computation of domain values
- Modifiers for data values, representing belief functions and/or uncertainty within the data
- Textual identifiers containing descriptive notes of specific attributes
- Literature references cited by textual identifiers

The following relationships connecting the components of the database model can be identified:

- 1) The data values are determined by the set of their identifiers,
- 2) Each data cell is composed of a (small) set of data values, where each set could either be different properties or different measurements of the same property,
- 3) Data may be described using text, graphs, functions, images, or equations,
- 4) There is a set of textual descriptors from which zero or more may be selected to describe a given attribute,
- 5) Some textual descriptors reference documentation,
- 6) Data may be reduced if certain variables, either dependent or independent, are missing or sparse, and
- 7) Domains may have interpolation functions for values.

4 IMPLEMENTATION AND DATA ACCESS ISSUES

The previous section discussed the development of a conceptual materials data model. This section briefly discusses how to implement that model and access the data stored within. There are many factors, both technical and societal, which act as barriers to the development and implementation of composite materials databases. These issues are presented and analyzed in this section. In addition, the applicability and roles of several software models and tools to composite materials data is discussed.

4.1 Significant Aspects of Materials Databases

While the components of a composite materials database (CMDB) can be mapped to components of other types of databases, some differences stand out and make it difficult to implement an effective CMDB on current commercial, off-the-shelf systems or general Database Management Systems (DBMS's) [Wiederhold 87]. These are summarized below where materials database system (MDBS) terms are related to DBMS terms (in parentheses):

- 1) The number of identifying (key) attributes is much greater than those seen in current commercial applications. A MDBS requires a large number of keys where a commercial DBMS normally requires only 1, 2, or 3,
- 2) Much of the actual data in a MDBS does not exist at the base levels, but only at higher levels of abstraction,
- 3) Textual notes, descriptors, graphs, and equations are all important for understanding the applicability of the data,
- 4) Interpolation functions may be attached to the domains, and
- 5) The number of dependent attributes per cell (tuple) is small, but the specific attributes vary from material to material.

Broadly speaking, one consults a computerized materials database or system of databases for one of three purposes [Wiederhold 87]:

- 1) To retrieve the properties data for a particular material,
- 2) To find a short list of candidate materials with a specified set of properties, and
- 3) To sort particular materials into groups serving some functional purpose or application area.

For the purposes of designation, an identification scheme is required which will define and differentiate materials with significantly different properties. Each of the unique aspects of the MDBS complicates this need significantly. There is also, of course, the practical consideration of special features or special properties unique to small groups of materials, and the consideration of new, specialized materials. The classification of these groups may yet be inadequate or incomplete for translation into existing and general databases.

4.2 Non-Technical Aspects of Materials Database Design and Use

It is important to note that many experts consider the barriers to the development and use of materials databases in general and composite materials databases in particular to be more societal rather than technical. One major problem is that developing a materials database is an interdisciplinary project, involving materials scientists, computer scientists, and engineers in all fields who design, analyze, and use various composite structures. Another problem is encouraging the development and use of composite materials testing and data exchange standards. This, in turn, requires the cooperation of standards organizations, such as ASTM and ISO, as well as engineers and scientists [Ostberg 86]. Presently, great advances are being made in this area.

Finally, a great deal of organization and administration is involved in developing and maintaining a computerized materials database. In developing a database, it is necessary to first

evaluate the validity and quality of the available data. Some decision must be made concerning copyrighted and proprietary data, e.g., is data from a company to be provided without identification or only in a compilation of data from similar companies, etc. An important consideration is financing: the initial investment and operational costs of an on-line database are quite high. Once established, a database must be maintained and kept updated; the viability of developing on-line help and/or users manuals within the database system must be determined [Westbrook 86, Scharff 86].

As a whole, the development, maintenance, and use of a computerized composite materials database involves many considerations beyond simply technical competence. There are many complex problems which must be overcome. However, the most important factor in doing so is first establishing a high level of commitment and cooperation among those who expect to benefit from the materials database, that is, the materials scientists, computer scientists, and engineers.

4.3 Choosing a Database Management System

An important facet of developing a computerized composite materials database is choosing an appropriate database management system (DBMS). The most common type of DBMS used today is the relational DBMS, because it allows sufficient flexibility in data representation while permitting different applications to use the same data in a program-independent fashion. The relational DBMS is well suited to representing tabular data, and also well-founded, in that operators are based upon relational algebra and calculus, and upon the structured query language SQL. However, the strict tabular format of relational databases while good for the arrays of composites data described in Section 2.2, is not always conducive to engineering materials data. In keeping with the restrictions of first normal form, each attribute of a relation can only contain a singular item, but there are numerous engineering examples where a set, vector, or matrix should be treated as a single attribute.

Object-oriented database management systems, on the other hand, usually support numerous abstract data types, including types such as sets and ordered lists. These features, along with the concept of storing groups of data hierarchically as objects, subobjects, and superobjects, strongly support scientific and engineering data. The difficulty with object-oriented systems is the lack of a user-friendly interface based on a rigorous mathematical background. Because an object can be anything, object-oriented DBMS's do not provide a standard set of operators. Instead, the database designer must define and implement all procedures necessary to access and manipulate the object [Cattell 91]. While the separation of objects and procedures (methods) sufficiently modularizes the data, it is difficult and time-consuming to implement and is not well formed and universally applicable as is SQL.

A possible compromise between these two database models is emerging with the development of extended relational databases, which seem to combine the best features of both relational and object-oriented DBMS's. One extension of the relational model is the nested relational model (NRM). This model relaxes the first normal form restriction of the traditional relational DBMS, thus allowing attributes of a relation to be a set of values, another relation, an image, etc. The NRM provides an intuitive mapping of complex physical systems into the model, allowing data to remain in hierarchical form and supporting inheritance within the database schema. The model may thus actually be an excellent underlying data model for mapping the complex objects inherent to engineering and scientific databases while retaining the theoretical and mathematical benefits of the relational model [Harvey 91].

4.4 Data Retrieval Applications

Two broad categories of data users exist in the context of a composite materials information system; designers and analysts on one hand, and application programmers on the other. Relative to user interaction with a database, interactive, real-time processing is the mode of choice. It is highly desirable to enable users to directly query the database and extract the information they need. Although general database browsing is not as common in engineering as in other disciplines, it is necessary and must be supported. Data query languages and forms, as well as report generators and interrogation languages are provided by most commercially available database management systems for this purpose and are generally sufficient for composite materials information needs. These facilities increase interactive processing and direct user involvement, thereby increasing the value of stored data.

To provide the required flexibility in a user interface, a layered implementation of the user interface is proposed. Such an interface would be conducive to many levels of access, partitioning the data both horizontally, by application area, and vertically, by amount of detail. The NSF Workshop on scientific databases further suggests that both menu-driven and command language interfaces be provided at minimum [French 90]. In this way, both rare or first time users and more experienced users can access the data quickly and easily. The first time user is lead through the required steps by means of menus and prompts, while the regular user simply issues the desired commands directly. A set of standard query action plans could also be customized to support commonly accessed data.

4.5 The Role of Knowledge-Based Systems

With respect to both material property data storage and composite materials processing, analysis, and design, we feel that it is necessary to integrate expert systems technology with the overall composite materials computational environment. Intelligent Knowledge Systems (IKS) combine knowledge-based systems with database technology by building some analytical and/or expert

knowledge into the front end of a data base. Such systems allow databases to go well beyond acting as simple data sources, instead supporting analytical and diagnostic knowledge bases by which users' inquiries are dealt with by expert logic containing candidate solutions to material science problems.

Uses for these systems may include applications such as preliminary material selection based on critical design parameters for specific applications though quality control and failure analysis. An IKS would include the user interface, the inference engine (containing the logic and analytical capabilities), and the knowledge base (materials database) as key elements. Another important component is the expert interactions necessary to expand the IKS and keep it current. Without reliable scientific and engineering databases to provide support, even the most sophisticated artificial intelligence innovations are of limited practical value: expert systems are only as valuable as the underlying database [Kaufman 88].

Incorporating one or more expert systems as an integral component of the composite materials computational system will facilitate both communication and consultative applications. Among other advantages, knowledge base technology will allow users to assess the validity of data and obtain estimates of missing information. Additionally, an expert system can aid in the selection of existing composite materials to satisfy design criteria and in the design of additional composites from constituent materials to fit unusual structural requirements. Two examples of applications which are specific to composite materials analysis and design are micromechanics and laminate design. In both design situations, a combination of knowledge-based technology and user specifications would allow the design of components to satisfy a variety of specific property requirements. Additional applications might arise from other sources, including functional relations and external codes, standards, and specifications. One or more commercially available expert system shells could be used to develop the knowledge-based applications to link with the composite materials database.

5 CONCLUSIONS

To summarize, this paper has focused on the requirements for creating a composite materials design database and data retrieval system for use in structural analysis and/or design applications. The purpose of the paper is to analyze in detail some of the issues facing developers of composite materials databases, and to present a collection of possible solutions to these conceptual and implementation problems. It is our hope to provide a clear understanding of the scope and nature of these problems, thereby furthering the agenda of materials database research by enabling the design, implementation, and use of a comprehensive yet well-formed composite materials database. In the remainder of this section, we briefly summarize our findings and present a methodology for composite materials database development.

It is clear that existing database technology is not enough in its capabilities to fully accommodate the needs of composite materials data representation and use. Three approaches will emerge as viable to address this issue. The first is the development of new technology. Work is being done in this area in the computer science community where object-oriented data representations are evolving. Additional work is proceeding on database and expert system coupling techniques and on the representation of the heuristic aspects of knowledge and data. As more robust, general systems emerge for storing and processing data, the capabilities of these systems must be evaluated and compared to the needs of composite materials data processing systems. The natural evolution of such systems will provide incremental improvements that warrant consideration for adoption.

The second approach to achieving a technology suited to the needs of composite materials data representation and use is to modify or extend the capabilities of existing database technology. This approach has been suggested in reference [Rasdorf 87] which offers a solution for at least one of the problems enumerated herein. Other researchers have also suggested modifications and extensions to current technology to treat the special needs of engineering databases [Bray 87, French 90]. Such developments must be encouraged.

The adaptation of existing database technology or the development of new database technology both provide data representation and use solutions which will emerge over time. To begin to solve these problems and to satisfy immediate needs, however, requires the direct use of existing technology. This approach is recommended for the design of composite materials databases so that immediate prototyping can begin. Prototyping will facilitate the acquisition of experience and understanding and also facilitate the dissemination and use of composite materials databases. Specific recommendations for creating a successful materials database system include:

- 1) Clearly define compatibility and integration needs between the data and its users and applications,
- 2) Standardize, to the extent possible, testing and data generation procedures,
- 3) Standardize, to the extent possible, a conceptual data model,
- 4) Develop trial database schema implementations based upon the conceptual data model,
- 5) Populate and test databases for trial applications to determine their robustness and flexibility, and
- 6) Monitor the development of database technology and identify alternative technologies as they emerge and mature.

ACKNOWLEDGEMENT

Portions of this paper were based on or drawn from the results of the National Institute for Standards and Technology and National Science Foundation's Materials Database Workshop paper cited in reference [Wiederhold 87]. The author wishes to acknowledge this paper, the National Science Foundation, and the National Institute for Standards and Technology. This research was supported, in part, by the National Science Foundation and by the U.S. Army Ballistic Research Laboratory at Aberdeen Proving Ground, MD. The support of NSF and BRL is gratefully acknowledged.

REFERENCES

- [Bray 87] Bray, O.H., "Engineering Database Management Requirements: The Next Generation," *Proceedings of the Symposium on Engineering Data Management*, 1987 International Computers in Engineering Conference, American Society of Mechanical Engineers, New York, NY (August, 1987); also in *Managing Engineering Data: The Competitive Edge*, R. Fulton, Editor, American Society of Mechanical Engineers, New York, NY, Pages 27-32 (1987).
- [Cattell 91] Cattell, R.G.G., "Next-Generation Database Systems: Introduction," *Communications of the ACM*, Volume 34, Number 10, Pages 31-33 (October, 1991).
- [Dathe 85] Dathe, G., "Peculiarities and Problems of Materials Engineering Data," *The Role of Data in Scientific Progress*, P.S. Glaeser, Editor, CODATA, Pages 233-240, 1985.
- [French 90] French, J.C., Jones, A.K., and Pfaltz, J.L., "Scientific Database Management," Invitational NSF Workshop on Scientific Database Management Report, Technical Report 90-22, Department of Computer Science, University of Virginia, Charlottesville, VA (August, 1990).
- [Future 85] "The Future of Data Banks in Science and Technology," Summary of a Panel Debate, *Nonbibliographic Data Banks in Science and Technology*, Schwarz, S., et al, Editors, International Council of Scientific Unions, Paris, France, Pages 209-210, 1985.
- [Hahn 81] Hahn, H.T., Hwang, D.G., and Cheng, H.C., "Design Data Manual for Composite Materials - Mechanical Properties of Unidirectional Composites," Volume 1, Dept. of Mechanical Engineering and Materials Research Laboratory, Washington University, St. Louis, MO, 1981.
- [Harvey 91] Harvey, T.M., Schnepf, C.W., and Roth, M.A., "The Design of the Triton Nested Relational Database System," *SIGMOD Record*, Volume 20, Number 3, Pages 62-72, September, 1991.
- [Hoskin 86] Hoskin, B.C. and Baker, A.A., Editors, *Composite Materials for Aircraft Structures*, AIAA Education Series, American Institute of Aeronautics and Astronautics, New York, NY (1986).

- [Kaufman 86] Kaufman, J.G., "Data Validation for User Reliability," *Materials Property Data: Applications and Access*, Kaufman, J.G., Editor, American Society of Mechanical Engineers, New York, NY, Pages 17-25, 1986.
- [Kaufman 88] Kaufman, J.G., "Sources and Standards for Computerized Materials Property Data and Intelligent Knowledge Systems," *Engineering with Computers*, Springer-Verlag, New York, NY, Volume 4, Pages 75-85 (Winter, 1988).
- [Landgraf 80] Landgraf, R.W. and Conle, A., "The Development and Use of Mechanical Properties in Industry," *Use of Computers in Managing Material Property Data*, Graham, J.A., Editor, American Society of Mechanical Engineers, New York, NY, Pages 1-12, 1980.
- [Masters 92] Masters, J.E., Foye, R.L., Pastore, C.M., and Gowayed, Y.A., "Mechanical Properties of Triaxially Braided Composites: Experimental and Analytical Results," NASA Contractor Report 189572, National Aeronautics and Space Administration, Langley Research Center, Hampton, VA (1992).
- [McDonald 90] McDonald, R.E., "The Critical Importance of Database Management in Industrial Automation," *Proceedings of the Fourth Symposium on Engineering Databases: The Technology for Integration*, 1990 International Computers in Engineering Conference, American Society of Mechanical Engineers, Boston, MA, Pages 131-144 (August, 1990).
- [MIL 88] U.S. Department of Defense, *Polymer Matrix Composites, Volume 1, Guidelines*, MIL-HDBK-17B, Washington, DC, February, 1988.
- [Mindlin 86] Mindlin, H. and Jaske, C.E., "Development of a Material Properties Database For the Use in High Temperature Design," *Materials Property Data: Applications and Access*, Kaufman, J.G., Editor, American Society of Mechanical Engineers, New York, NY, Pages 17-25, 1986.
- [Ostberg 86] Ostberg, G., "A Paradox in the Development of Computerized Materials Data Systems," *Nonbibliographic Data Banks in Science and Technology*, Schwarz, S., et al, Editors, International Council of Scientific Unions, Paris, France, Pages 99-104, 1985.
- [Pritchard 89] Pritchard, G., *Advanced Composites: Chapter 5*, Partridge, I., Editor, Elsevier Applied Science, London, England (1989).
- [Rasdorf 87] Rasdorf, W.J., "Database Management System Extensions for Engineering Applications," *Computers in Mechanical Engineering (CIME)*, American Society of Mechanical Engineers, Volume 5, Number 5, Pages 62-69, March, 1987.
- [Scharff 86] Scharff, L., "Basic Demands of an On-line Service Center," *Nonbibliographic Data Banks in Science and Technology*, Schwarz, S., et al, Editors, International Council of Scientific Unions, Paris, France, Pages 167-173, 1985.
- [Stonebraker 91] Stonebraker, M. and Kemnitz, G., "The POSTGRES Next-Generation Database Management System," *Communications of the ACM*, Volume 34, Number 10, Pages 78-92, October, 1991.

- [Swanson 88] Swanson, S.R., "Failure of Laminates Under Multiaxial Stress," Presentation at the Symposium on the Failure of Thick Composites, US Army Ballistic Research Laboratory, Aberdeen Proving Grounds, MD, July, 1988.
- [Tsai 87] Tsai, S.W., *Composites Design*, Think Composites, Dayton, OH, 1987.
- [Tsai 88] Tsai, S. and Kim, R., "Macro Failure Theories: Classical Approaches, New Compression/Tension Test Approaches," Presentation at the Symposium on the Failure of Thick Composites, US Army Ballistic Research Laboratory, Aberdeen Proving Grounds, MD, July, 1988.
- [Wiederhold 87] Wiederhold, G., Caindec, L., McCarthy, J., and Reifsnider, K., "A Model for Materials Databases," Draft Report, NSF/NBS Materials Database Workshop, Stanford University, Stanford, CA, June, 1987.

ABSTRACT

1 INTRODUCTION

2 COMPOSITE MATERIAL PROPERTY DATA

- 2.1 Factors Inhibiting Database Development
- 2.2 Composite Material Data Description
- 2.3 Features of a Composites Data Model
- 2.4 Determination and Use of Material Properties
- 2.5 Documenting and Tailoring the Data Set

3 A CONCEPTUAL COMPOSITE MATERIAL DATA MODEL

- 3.1 A Classifier Hierarchy for Materials Data
- 3.2 Modeling Special Types of Materials Data
 - 3.2.1 Text and Bibliographic References
 - 3.2.2 Graphs and Equations
 - 3.2.3 Images
- 3.3 Data Value Density: The Size of the Data Set
 - 3.3.1 Interpolation and Extrapolation
 - 3.3.2 Reduction
- 3.4 Data Value Modifiers: The Accuracy of the Data Set
- 3.5 Materials Database Model Components

4 IMPLEMENTATION AND DATA ACCESS ISSUES

- 4.1 Significant Aspects of Materials Databases
- 4.2 Non-Technical Aspects of Materials Database Design and Use
- 4.3 Choosing a Database Management System
- 4.4 Data Retrieval Applications
- 4.5 The Role of Knowledge-Based Systems

5 CONCLUSIONS

ACKNOWLEDGEMENT

REFERENCES