At the core of any computational system that can support design development, analysis, and evaluation is a building representation which should be able to represent all the different components that make up a building, along with the manner in which they come together. In other words, the representation must be informationally complete and semantically rich. The paper discusses these two criteria in detail, and briefly reviews other research efforts aimed at developing building representations for CADD that attempt to meet them. Our solution to this problem is then presented. It is aimed primarily at the schematic design phase, the rationale for which is also stated. Taking the view that buildings are unique assemblages of discrete, mostly standardized components, our representation is clearly divided into two components: the Object Database (ODR) which stores detailed information about various building elements, and the Project Database (PDB) which holds information about how these elements are assembled to make up a particular building. An ODR may be shared by many building projects, while the PDB must necessarily be unique to each. The data schemes of both the PDB and the ODR are described in detail and their computational implementation, to the extent that it has been completed, is illustrated.
Introduction

The use of computers to enhance the quality of the processes used for designing buildings, as well as the quality of the built environment itself, has been the subject matter of considerable research for more than three decades. Early efforts concentrated primarily on evaluating and optimizing various building functions, such as their layout and energy consumption (see, for example, Armour & Buffa 1968, Shaviv & Shamir 1975). These efforts, while effective in some narrow domains, failed to bring about the sought improvements for the building industry in general. Since the late 1970s, many researchers have recognized that the fundamental obstacle to the desired improvements is the lack of an adequate computational representation of buildings.

Traditional representation methods used by architects and engineers for hundreds of years, such as scale drawings, renderings, and three-dimensional scale models, contain only a small part of the information needed to interpret and assess the quality of the design. Drawings, for example, represent only the geometry of the building in highly symbolic form. Even when this information is accompanied by textual labels and specifications (materials, products, construction techniques, etc.), it still relies heavily on the intelligence and professional training of the (human) observer, who must augment the information with his/her own knowledge of the field.

The almost direct importation of these traditional representation methods to computer-aided design have propagated the same informational deficiencies. Augmentation of the data with computer-readable “knowledge” is difficult, given that inference of meaning is one of the most difficult abilities to impart to computers (witness the difficulties encountered in the field of Artificial Intelligence). Ad-hoc efforts to add intelligence to specific applications were generally unsuccessful in enhancing the utility of computers in building design, except in some very well defined, narrow domains. What is needed is an overall more “intelligent” representation, which will embody some of the knowledge added to the interpretation of drawings by the human observers. The history of research in architectural CAD, from the late 1970s on, has been marked by the quest for this “holy grail” of intelligent building representation.

What is an “intelligent” building representation?

While different researchers proffer different definitions of what an “intelligent” building representation should be (see, for example, Carrara & Kalay 1995), they generally agree that a computational representation that can support the qualitative aspects of architectural design needs to have the following two qualifications:

1. It must include information that describes all aspects of the building.
2. This information must be semantically meaningful.

informational completeness

The first qualification, whose importance is self-evident, must further meet the following requirements:

1. Well formedness: The data must be self-consistent. For example, a window representation in a plan drawing must have the same location and dimensions as it has in an elevation or a section.
2. Generality: The data must include the necessary information to assess the building from many different points of view. For example, it must be able to support evaluation that pertains to spaces as well as the enclosure, and to structural qualities as well as architectural ones.

semantics

The second qualification is more difficult to define and to accomplish. How, and how much of our own “understanding” can we impart to computers? A similar set of questions...
have frustrated AI researchers in architecture as well as in other fields for the past 40 years (see, for example, Coyne et al 1989). Yet, when considered from a representational point of view, these questions take on a somewhat less formidable character: it is not necessary to make the computer be able to "understand" the meaning of the data, only to be able to identify it, for the benefit of other programs or human experts that will use the data intelligently. For example, it is not necessary to make the computer understand the essence of "door-ness," as a portal that enables social processes like "entry" and "exit." From a representational point of view, the computer only needs to be able to identify a door when one exists in a given location. This ability will support qualitative assessments such as access, egress, and privacy, which will be performed by expert systems or other types of programs, to be developed separately from the building representation.

Other Efforts

Considerable research effort throughout the CAS community has been devoted to developing a representation that can provide adequate support for computer-aided building design. One of the foremost and most widely known of these efforts is STEP: a product model concerned with defining a standard for the representation and exchange of data (Bjork and Wix, 1991). STEP, however, does not focus only on building information but product information in general. Therefore, it is too general-purpose and too limited, as far as semantic content is concerned, to provide an efficient and workable solution to the problem of building representation.

The COMBINE project, working with the STEP standard, deals with a neutral file exchange of data among various application programs—primarily energy and HVAC (Amor et al 1991). It does have, at its core, an integrated data model for the building description, which gets translated to the applications. As this data model was developed by synthesizing the individual data schemas of the various design tools it wanted to integrate, each time a new tool has to be added to the system, the data model has to be re-worked. It therefore, does not qualify as an informationally complete building representation. Similarly, systems like ONSYS (Richens, 1977) and IBIDE (Fennes et al. 1996) that deal with hospital buildings and high-rise office buildings respectively are also too specialized to provide the solution to a generic building representation.

One effort directed specifically towards enhancing the semantic content of building representations was KAAD (Carrara et al 1995). It employed an object-oriented approach, where entities called space units, building units and functional elements possessed parameters that defined specific attributes of the represented objects. (Figure 1) In addition, the objects included methods, which are functions that calculate values assigned to an object, based on values already assigned to other objects. Objects also included several kinds of semantic links, such as...
as inheritance links (AKO, ISA) and association links (IMS). Using these attributes, methods, and links, KAAD was able to help in assessing the quality of nursing units in Italian hospitals in terms of circulation, space allocation, fire egress, energy consumption, code compliance, and other performances. It was, however, a closed system, useful only as a concept demonstration rather than an actual tool.

Probably the effort that comes closest to our definition of a generic building representation is the EDM (Eastman and Siahibis 1995). It was developed to address issues in the design of buildings with wide ranges of design abstractions, construction technologies, and variations in building use, and has representations for each of these in the form of the constructs BOUNDED SPACE, CONSTRUCTED FORM, and ACTIVITY, respectively. It is meant to be open-ended so that additional descriptions can be added. EDM thus aims to be general-purpose as well as complete. EDM, however, separates the geometric representation of the building from the integrity constraints that endow it with semantic validity. The semantics are computed as and when needed by each application program individually, based on its own pre-conditions and its domain of interest. In contrast, we take the view that most, if not all, of the semantic integrity should be built into the topological representation of the building itself, so that no computation is needed to establish the semantic validity of the model once it has been built.

**our approach**

Having witnessed the successes and failures of these efforts, we set out to develop a semantically rich building representation that will be capable of supporting the interpretation and assessment of buildings. This representation is based on the following three principles:

- It addresses only one phase in the overall design process, specifically the schematic design phase (not just for architecture, but for all the building-related disciplines).
- It recognizes that building design is, in many ways, the process of assembling universal components into a unique composition.
- It makes a conscious effort to integrate the representation of spaces and enclosures in one unified representation.

**schematic design**

Schematic design is the phase where the gross features of the building are defined, including its overall shape, size, structure, adjacencies, circulation, and the materials it is made of. While each feature will, no doubt, undergo refinement and elaboration in subsequent design phases, schematic design is the phase when most of the cardinal properties of the building are determined, or at least are first being considered.

It is, of course, difficult to define precisely what constitutes schematic design. Different architects will likely consider different parts of their work as "schematic" vs. "design development" or even "detailed design." It is even possible that such designations change from one project to another. While we do not attempt to define "schematic design," we are cognizant of other efforts that tried to do so. For example, a systematic effort to characterize the elements of different phases in the design of energy-efficient buildings was made by Shavis & Kalay (1994). In that study, various energy-related design parameters were identified by the design phase in which they are considered first.

We have chosen to limit our efforts to support this particular phase of the design process because we consider it the phase when designers are most in need of the assistance that could be offered by computers in terms of assessing whether the areas that were allocated are appropriate, the adjacencies among them proper, whether the proposed structure will stand up, whether the budget has not been exceeded, whether the form that is being con-
considered can support all the functional needs, and so on.

Our choice of design phase was also influenced by the simplification our decision brought to some of the representational aspects we need to deal with. The major simplification is due to the geometry: at the schematic phase, only the gross geometric features of the building need to be specified. As such, our data structure is limited to representing only prismatic forms (ones where there is no change in geometry along the Z axis) which are bounded by planes. Furthermore, we ignore small geometric changes, such as the ones along walls that would be caused by columns and window sills. This allows us to also simplify the tools that will be used to create and edit the representation at the schematic phase. We assume that the design developed at this phase will eventually be moved on to some commercial CAD system where its form will be further elaborated, even modified within some semblance of the schematic form.

Another simplification that accrues from limiting the representation to the schematic design phase pertains to the semantics that need to be embedded in the database to facilitate meaningful analysis. These semantics need not account for all the details of the building. In fact, most CAD research efforts in the area of evaluation and prediction emphasize the importance of analyzing the building in the preliminary design phase when all the major design decisions are made. Therefore, several evaluation tools are specifically targeted at the schematic phase for providing useful feedback to the respective building professionals on various design aspects such as spatial allocation, cost, structures, energy consumption, fire safety, and so on without requiring detailed design input. We strongly believe in this approach to evaluation and our representation is, therefore, aimed at addressing the needs of such evaluation tools.

Two kinds of information

Buildings, like many other artifacts, are unique assemblages of discrete, mostly standardized components. In developing a computational building representation, therefore, it would be prudent to separate the information into two categories:

- Information that pertains to the components
- Information that pertains to the assembly

We call the information pertaining to the components Object Data, because it typically represents discrete objects (doors, walls, windows, etc.). We call the information pertaining to the assembly Project Data, because it is typically project-specific. A certain project may use several object databases, and an object database can serve many projects. (Figure 2).

The first kind of information is typically the province of the manufacturers or the specialists who produce building components. The second kind of information is the province of the building design professionals. In many ways, such a separation already exists today in the paper-based practice of building design, where general product catalogs like Sweets and Thomae are routinely used along with standardized office details to complete the information pertaining to a specific design.
Other research efforts in the field do not make the separation between assembly and object and have both kinds of information integrated within a single representation (see, for example, the component-based building representation effort by Harfmann & Chen 1993). In contrast, we have found that by separating the components (objects) from the assemblies in which they are used, we can simplify both representations (of the assemblies and the components), and make the databases easier to implement and to maintain. Yet, the ability of the separate representations to encode all the information necessary for representing a building is not affected, because the two come together through the applications that evaluate the proposed design (Figure 2), similar to the manner it is currently done in the construction industry.

**object database (ODBs)**

ODBs are domain-specific knowledge bases of project-independent information that is needed to generate and to evaluate objects of the kind they represent. These can be spatial objects (rooms, external spaces) or physical building components (e.g. walls, slabs, doors, windows, furniture and equipment, etc.). For example, rather than define a hung door every place in every project that uses one, it is described only once in the appropriate ODB, then instantiated whenever needed. Only attributes and requirements that are specific to the instantiating project need to be added to the generic definition, such as the location of the door and its association with other objects in the project (e.g., wall, spaces, etc.).

The practice of augmenting drawings with paper-based product catalogs, and more recently with ca-rams-based electronic catalogs, is not new. Currently, this useful practice is being transformed again, using the World Wide Web in lieu of ca-rams (see, for example, www.constructor.com, www.sweets.com, and www.productbuilding.com/product/product.html). The Web, which enables sharing media-rich information across platforms without regard to the location of the information, provides much more up-to-date product information while saving product suppliers the cost of producing many copies (and versions) of the ca-rams. Furthermore, when the information provided by the product manufacturers is made compatible with the ca-ram systems and the applications that use them, product data carried by the Web could be integrated directly with the project data. This integration would allow for dynamic updates (e.g., for costing and scheduling), early detection of incompatibilities among products, substitution of similar but less expensive products where possible, and generally linking the project database directly to the manufacturers of the components that are needed to realize it.

However, before such benefits can accrue, two functional and several organizational problems need to be solved. The functional problems are:

- **Usability**: To benefit from the full potential of the electronic implementation of an object database, the content of the objects in that database must first be made semantically meaningful. Then they need to be linked to the project database, so that the information they contain is accessible to the applications that combine the two separate databases into one whole representation. These applications will have to be able to parse not only the building data itself (the PDB), but also the object data (ODB), so they could take off quantities from the PDB and correlate them with the relevant objects in the ODBs.

- **Searching**: There is a vast number of building-related components, and an even larger number of manufacturers who produce them. Each manufacturer uses different terms to describe the attributes of their products, which makes it difficult to identify the sought components and to compare them to each other. Some kind of a standardized organization of all this data is needed.
The organizational problems associated with this idea pertain to the legal responsibilities associated with distributed information systems, their maintenance and update, and of course, the reliability of the information in the ODBs. Although we have some thoughts about how to solve these organizational issues, our research addresses only the functional issues.

**Content and structure of the ODBs**

Each object in an ODB must contain all the information needed to define and to evaluate objects of the kind it represents. This information includes:

- Attributes, such as the object's name, its form (shape), the materials it is made of, and other properties (e.g., manufacturer, cost, etc.).

- Classification relationships, which define an inheritance path for properties this object shares with other, more generalized objects of the same type (e.g., that a hung-door is a kind of door, which is a kind of opening).

- Information that describes the function of the object (e.g., that a door facilitates entrance and egress to spaces).

- Integrity constraints, which are logical propositions associated with attribute values and define generic expectations from the object (e.g., that the door must lie within a wall).

- Cases, in multimedia form, which provide anecdotal information about the object.

Unlike other product databases (e.g., EDM), our ODB includes only two types of relations:

- Extra-object relationships of two kinds:
  a. The relationship between sub-classes and their super-class in the abstraction hierarchy. We call this relationship \text{AKO} (A Kind Of).
  b. The relationship between an instance and its defining class. We call this relationship IS-A.

- Inner-object relationships: The relationship between an object and its attributes.

Notice the absence of assembly (Part Of) relationships, which would be useful to define composite objects (a gypsum partition, a brick wall, a window), but would also significantly complicate the representation (witness similar efforts such as COMBINE and EDM). Instead, we embed assembly information in the representation of the object itself, and make extensive use of the classification hierarchy to compensate for the absence of Part-Of links. Thus, a Window object, for example, can be described as a single entity with certain properties and performance characteristics, or as a composition of the materials and products of which it is made. Detailed versions of the window, which include information about specific glass and finishes might be stored as sub-classes of the window. In this manner, the components of the window will be treated as properties of the more specific window object, and the assembly information that is considered integral to the object definition will be embedded in the representation of the object itself.

**Classification system for the ODBs**

Objects are stored in hierarchical classification structures, which facilitate compact, non-redundant, and consistent representation a method that has now been popularized by object-oriented programming languages. Objects with similar properties and functions are grouped into classes. A class is defined at large and then refined into subclasses. The properties of objects located higher in the abstraction hierarchy (superclass) are inherited by the objects located lower in the hierarchy (subclass). Subclasses add their unique properties to the definition of the superclass. Common attributes are stored at the highest level in the tree that is still shareable by all the objects that can benefit from this information. An inheritance mechanism makes higher-level information available to objects that are stored in lower levels of the hierarchy.

An example of a Hung Door object and its
Figure 1: A Hung door inheriting properties in an "Openings ODB." 

inheritance hierarchy is depicted in Figure 3. 
This object shares many attributes with other doors, such as revolving, sliding, and composite doors, in that it consists of a screen, usually made of wood, metal, glass, or a combination of materials, installed to swing, fold, slide, or roll in order to close an opening to a room or 
building. A Hung Door adds to these properties its own unique attributes, such as that it 
has some kind of a hinged assembly. 

The structure of the classification tree 

introduces uniformity into the database (e.g., it 
unifies the number of steps from general to spec-
ific objects), and provides good ground for 

extends the database. Moreover, the consid-
erable effort that is needed to develop a class-
ification system can be spared. 

MasterFormat and SFB organize information 

Hierarchically, from more general (broad in scope) to the specific (narrow in scope), as layer 
guidelines use tags, rather than a hierarchy and, 
therefore, cannot be used for the ODB. 

MasterFormat is the construction coordi-
nating system used in the USA and in Canada. 
It was developed by construction specifiers to 

provide a standard format for specifications 

and a cataloging system for product literature. 
MasterFormat allows coordination between the 
specifications and the manufacturers' litera-
ture and makes it easier to reuse the specifi-
cations. It supports classification of the physical 
elements of the building (construction systems, 
building elements, materials), but does not have 
provisions to classify spatial properties and 
requirements. 

SFB provides for representing information 

about spaces in addition to physical building 

components. It was developed in Sweden, and 
is broadly used in most European countries, 
especially in Great Britain. SFB was developed as a "common language" which can be used by all 
the professionals involved in the design and 

construction of buildings to assist commu-

nication and to make it easier to find informa-
tion. Contrary to MasterFormat, SFB has been used 

for coding drawings. Furthermore, it has been 

used at the briefing stages of the design process, 
to represent data on user requirements for 
spaces (e.g., kitchens, bathrooms).
We have been implementing the ODIs using two kinds of objects:
- Generic objects
- Catalog objects

Generic objects form the upper levels of the inheritance hierarchy. They define all the properties of the objects. The Catalog objects indicate specific instances of the objects that are actually manufactured by specific producers. (Figure 5)

generic objects

Each generic object represents a building component such as a wall, a door, or a window. It describes the nature and basic qualities of that component. Generic objects do not specify default values for any of the attributes. Shahiv and Kalay (1990) showed that default values, or assumptions about most frequently used materials, construction techniques and dimensions, may lead to incorrect results in energy evaluation and other applications. Figure 6 shows an example of what a Generic Object in our implementation looks like.

Our earlier decision not to include assembly information in the ODIs may lead to exponential proliferation of unique combinations of specific objects. For example, there can be a wide variety of swinging "hung" doors, because construction types, door types, operation types and core materials of swinging doors appear in nearly all combinations. Door types may be superimposed on any door construction. Operation types, such as hinged, pivoting, balanced, pivoting double acting, may be combined with almost all door types. To avoid having to represent explicitly each combination, we have provided a mechanism for creating "ad-hoc" objects that represent a unique combination of attributes. This mechanism is in the form of a template that describes the variable attributes of the object (e.g., construction), along with the allowable values that may be assigned to the attribute (e.g., open, flush, etc.). Users can "create" a "new" object by choosing a value for each attribute. (Figure 7). The database ensures that objects created with the template are well-formed by limiting the value choices to only those that are consistent with the ones already selected.

catalog objects

Catalog objects link the generic object hierarchy to the manufacturers’ own product cata-

Figure 5. Object Database: Generic objects and Catalog objects.

Figure 6. Door as a Generic object.
This includes the responsibility for ensuring the three conditions of disjointedness, connection, and subordination. The first is when objects must not occupy the same space (at the same time), the second is when they are supposed to do so (e.g., pipes in a wall), and the third is when one object must be included in another (e.g., window). The ODB, in contrast, carries other kinds of semantics, such as what is a door, a window, a wall, etc. In this way, the semantics of both the PDB and the ODBs complement each other.

In developing a suitable data structure for the PDB, one of the key issues that needs to be considered is the relationship between the spaces and the structure of the building.

**The space-structure dilemma**

Space and structure in a building are not just two different categories of components, they are of a different “substance” altogether. Spaces are abstract entities; they do not have a physical reality. In contrast, the structure of the building is definitely physical slabs, beams, columns, walls, windows, and so on; actually exist and have a shape, size, material characteristics, etc.

The difficulties in building representation arise from the fact that the physical structure and the abstract space are highly interdependent. Depending upon how the building is viewed, the structure defines the space, but the space, in turn, also determines the structure that bounds it. Building professionals such as the engineer and construction manager are primarily concerned with the physical fabric of the building. In contrast, for professionals such as the architect, the spatial qualities of the building are paramount. There are still others such as the energy consultant for whom both are equally important.

We have identified four main challenges for a building representation as far as space and structure are concerned:

- Both space and structure must be explicitly represented, rather than leaving one to be inferred or derived from the other. Most com-
merial CAO systems represent only the structure, leaving the spaces it encloses to be inferred implicitly by the designers.

—Both space and structure must be represented on an equal footing rather than giving one precedence over the other; since the representation is aimed at serving the needs of all the building design professionals rather than just one group over another.

—The representation should be non-redundant. Thus, if a wall is represented in the structural assembly of the building, it should not be re-represented in the spatial assembly.

—The representation must reflect the strong interdependency between space and structure. Spaces are bounded by walls and slabs which are supported by beams which in turn are supported by columns. If the size and shape of a space is changed, it must bring about a corresponding change in all of these elements. Similarly, any change in the location of a wall or slab, for instance, must affect the shape of the spaces they enclose. In short, the representation should not allow space and structure to be created or modified independently of each other.

No research effort has emerged so far in the area of building representation that has yielded a satisfactory solution to the space-structure problem just discussed. We will now describe our solution to this problem that has been developed by addressing all of the four issues stated above.

**version 1 of the PDB**

The first version of the PDB (Kalay et al. 1975) represented in a non-redundant, independent, complementary way, the spaces and the partial structure of the building. Slabs and beams were not represented at all, while columns were represented indirectly rather than explicitly. Nevertheless, it served as a good starting point and enabled us to test it with architecture-specific evaluation tools such as spatial layout, adjacencies, path-finding, etc. (Khemlani & Kalay 1997, Lee 1997).

The dominant feature of the first version of the PDB was that it had one key building element around which the representation of all the other entities was built. This was the **segment**, defined as that abstract surface that encloses a space, but which has some semblance of physical reality by being linked to a wall and having the attribute of surface finish associated with it. The “inspiration” for this idea came from the well known Split-Edge Data Structure, which in turn is a variant of the Winged-Edge Data Structure, both of which have been very successfully implemented in geometric modeling systems (Baumgart 1979, Bae et al. 1979, Kalay 1985, Kalay 1989). These data structures use a simple and elegant relational scheme to solve the edge-face problem in planar polyhedra (analogous to the space-structure problem in buildings) in a compact, non-redundant, consistent, and efficient manner.

**version 2 of the PDB**

The current version (Version 2) of the PDB extends the earlier representation to explicitly include the structural entities of the building such as columns, slabs, and beams. While the Split-Edge data structure concept is still embodied in the representation, now as a surface rather than a segment, it is no longer the key element of the representation. The key is now a revamped vertex entity which is purely geometric (as opposed to using it to also represent columns as we did in Version 1), but which is used to physically locate and bring together all the various elements of the building. The simplicity and elegance of the Split-Edge data structure, however, continues to “inspire” the development of the representation, and is very much in evidence in the current version.

Now that the structure is more comprehensively represented, it necessitated a redefinition of what constitutes space, as opposed to what consti
tures structure, and how they are inter-related. We found it convenient to adopt the approach illustrated in Figure 8 for our representation. A space is regarded as a volume bounded by surfaces (not walls). Surfaces can have openings associated with them for doors and windows. Columns, slabs, and beams form another group of interconnected elements. Walls may or may not be integrated with the structure. The space and the structure are related to each other in a number of ways: vertical surfaces are linked to walls, horizontal surfaces are linked to slabs, slabs are linked to beams, walls are linked to beams, and surfaces are also linked to columns.

We will now explain the data scheme of the various elements that constitute the PDB, with reference to a simple building as an example, illustrated in Figure 9. Since the database is relational, the data will be represented by means of tables for each element. Where the element is also an object that can be referenced in an ODB, one field in the table is allocated for storing the tag which will link the element to the corresponding object in the ODB.

**Table 1: The BUILDING table**

<table>
<thead>
<tr>
<th>building ID</th>
<th>first level</th>
<th>building ODB link</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLD1</td>
<td>LF1</td>
<td>top name</td>
</tr>
</tbody>
</table>

Buildings rather than just a single building. We therefore have a **building** table that can have multiple entries for buildings as shown in Table 1.

Within a building, the highest category is a **level**. A building can be composed of many levels. (Figure 9) For each building in the table however, only the first level is identified. Thereafter, information about subsequent levels of the building can be obtained from the levels entity, which will include a pointer to the next level of that building. (see Table 2). If some information about that building type is available in a “Building ODB”, there is a place in the building table for storing that link.

A level is more of a functional than a physical grouping of spaces, i.e. spaces at different physical heights (from the ground plane) can still be considered at one level. In this respect, the concept of a level is similar to that of a floor plan as defined by the architect – it can accommodate slight changes in floor heights of the spaces that are considered to “belong” to that floor.

A level has two main components: spaces and slabs. This is because there is no predefined correspondence between the two; one cannot be derived from the other. Consider the example

**Figure 9. A sample building for demonstrating the PDB representation.**
of the building BL1 shown in Figure 9. The level LV1A consists of three offices on two different floor slabs but with the same ceiling slab. LV1A consists of two offices with the same floor and ceiling slabs. (We consider stairs and stairwells to be a special element, and will therefore ignore them for now.)

For a given level entity, we would be interested in knowing what spaces it is comprised of, as well as which are the physical slabs that make up that level. Therefore, we store the first space and the first floor slab in the level table as shown in Table 2. (It does not matter which space or slab is indicated as the first.) The responsibility of pointing to the next space or slab in the list is delegated, following the same principle mentioned earlier, to the space and slab entities respectively.

The level entity also includes a back pointer to indicate the building it belongs to. This information can be derived from the building table by doing a search, but it is more convenient and efficient to store it with each level.

**spaces, surfaces, walls, and openings**

One of the two basic components of a level is the space. What defines a space are its surfaces, both vertical and horizontal. A vertical surface is essentially a wall surface there will necessarily be many of these. For a given space, if we list its first surface (again, it does not matter which is designated as the first), we can associate a next surface to each one of the surfaces and thereby collate, in a sequence, all the vertical boundaries of the space. (Table 3.) For horizontal surfaces, we assume (in this version) that there will be only one floor surface and one ceiling surface associated with every space. The area and volume of the space can be derived once all its surface information has been obtained. (Table 4)

As shown in Table 3, the outside of the building is also included as a space for reasons that will become clear once we look at the surface table. (Table 5)

Floor and ceiling surfaces, on their part, have a link to the slabs they rest on or hang off respectively. The only other information they include, as shown in Table 4, is their distance from the associated slab. Their geometric loca-
tion and configuration can be derived (as the intersection of the vertical wall surfaces with the horizontal slabs) so it does not need to be explicitly stored.

We have already stated that our representation is geared towards the schematic design phase. Therefore, we have simplified the representation by assuming that wall surfaces are always vertical and never inclined. Then, given that the height of the space can be derived from the slab information linked to the floor and ceiling, what is needed to locate the wall surface geometrically in space is simply a vertex for which the X, Y, and Z coordinates are specified. Since any surface always has a link to the surface immediately succeeding it, the vertex associated with that surface as well as its own vertex will define the horizontal segment that has to be extruded by the space height to define the surface. Figure 11 shows the vertices linked to the surfaces that appeared in Figure 10.
The surface table showing the entries for the surfaces shown in Figure 11 is given in Table 6. The scheme is based on the split-edge data structure as stated earlier. Thus, every surface necessarily has a next surface, as well as an opposite surface. Since every surface must have one and only one opposite surface, a double-height space, for instance, might have an additional surface located directly above a surface. If such a situation occurs (as it will for the "Outside" of the building in the current example), it can be represented by the "Surface Above" field in the table.

Surfaces do not exist in isolation. They have openings for doors and windows, and are also associated with walls. In version 1 of the PDB (Khemlani & Kalay 1997), there was an exact correspondence between a pair of surfaces and the wall they were on opposite sides of. However, this lead to situations where a continuous wall was broken down into several parts when it was spanning across several spaces. In the current version, the one-to-one correspondence between surfaces and walls has been given up for the sake of preserving the continuity of the wall as a physical element. Thus, the wall W1 shown in Figure 12 will be referred to by four surfaces S07, S07, S07, S07; rather than a single pair.

The wall is geometrically located by specifying its two vertices and its height. A single external wall can extend up to several levels, therefore the height of a space cannot be used to determine the height of a wall. The wall table carrying all this information is given in Table 6.

The walls are also linked to the structure of the building by indicating if there is a beam running across the top of the wall. It is important to note that there is no one-to-one relationship between walls and beams not all walls have beams above them, and not all beams have walls below them. If that was the case, we could have used one entity to derive the other. In the absence of such a relationship, we simply list the beam associated with a wall, if any, by having a "Beam" field in the wall table.

The opening entity is relatively straightforward. Each surface carries the information of the first opening located within it; thereafter, each opening indicates which is the next opening on that surface. In this manner, all the openings located within a single surface can be collected. Also, a single opening will necessarily be associated with both surfaces that make up the pair. Thus, the opening WING shown in Figure 11 is referenced by both surfaces S07 and S07 in their table entries. The opening table is shown (Table 7) with the entries for this particular window.

As we have seen, there will always be a pair of surfaces referencing any given opening. But the single surface listed in the opening table determines which direction the door or window opens in. That is why S07 is the surface referred to by WING rather than S07. The "Distance" field

---

**Table 6: The WALL table.**

<table>
<thead>
<tr>
<th>wall ID</th>
<th>vertex 1</th>
<th>vertex 2</th>
<th>height</th>
<th>beam</th>
<th>wall ODB link</th>
</tr>
</thead>
<tbody>
<tr>
<td>W01</td>
<td>V02</td>
<td>V03</td>
<td>15</td>
<td>B12</td>
<td>tag name</td>
</tr>
<tr>
<td>W01</td>
<td>V02</td>
<td>V03</td>
<td>15</td>
<td>B13</td>
<td>tag name</td>
</tr>
<tr>
<td>W12</td>
<td>V06</td>
<td>V03</td>
<td>15</td>
<td>B12</td>
<td>tag name</td>
</tr>
<tr>
<td>W13</td>
<td>V31</td>
<td>V05</td>
<td>15</td>
<td>B13</td>
<td>tag name</td>
</tr>
<tr>
<td>W14</td>
<td>V41</td>
<td>V05</td>
<td>15</td>
<td>B13</td>
<td>tag name</td>
</tr>
</tbody>
</table>
locates the opening geometrically on the surface by specifying the distance from the center of the opening to the vertex associated with that surface. The values in the “Width” and “Height” fields would be useful if the opening was not yet linked to a specific door or window object in the ODB, a situation very common in actual practice. Once the linking was done, however, the “Width” and “Height” values would be replaced by those specified in the ODB for that particular object.

### Slabs, Beams and Columns

The second main component of the **LEVEL** entity, as introduced earlier, is the slab. A slab can be represented in terms of the beams that support the slab. Figure 13 shows the slab SLB3 and the beams that support it. The boundary of this slab can be obtained by collecting all its peripheral beams. Thus, we can simply define the first beam for every slab in the **SLAB** table as shown in Table 8, and associate a next beam with every beam in the **BEAM** table. For each slab, we also have a back pointer to indicate which level it belongs to.

For the **BEAM** entity, we can locate a beam geometrically by defining its two vertices as shown in Figure 14.

We have seen that we also need to associate a next beam with every beam to define the slab. However, there is a lot more information about the structure in Figure 14 that we need to capture. We would like to have the full picture - the primary beams that rest on columns, secondary beams that rest on the primary beams,
tertiary beams that rest on the secondary beams, and so on. Trying to derive this information from the vertex information of the beams alone would be highly inefficient and tedious. However, we can make the BEAM entity carry some more information in addition to which is the next beam along the slab—such as the next parallel beam, and the next continuous beam (see the entries for beams shown in Table 9). It should be noted that all these lists are circular, so that "previous" beams for any category need not be stored.

The final building component of our basic representation is the column. A column can simply be represented as a sequence of vertices as shown in Figure 5. Thus, we can associate a starting vertex to every column, and let each vertex carry the information about which is the next vertex in the vertical sequence. (There will be only one such vertex in the vertical direction, where there can be many vertices that can qualify as the next vertex of a particular vertex.)

The only complication arises in the case where one vertex is shared by two columns. For instance, the vertex V5 in Figure 5 is shared by the columns C1 (a foundation column) and C3 (a superstructure column). Therefore, it also becomes necessary to indicate for every column, the last vertex in addition to the first vertex, as shown in the Table 9 for the column entity.

vertices

All that remains now is the specification of the vertices. Each vertex carries information about the next vertex in the vertical sequence. Also, if there is a column passing through that

<table>
<thead>
<tr>
<th>beam ID</th>
<th>vertex 1</th>
<th>vertex 2</th>
<th>next beam along slab</th>
<th>next parallel beam</th>
<th>next continuous beam</th>
<th>BEAM OBD link</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>B10</td>
<td>V26</td>
<td>V30</td>
<td>B19</td>
<td>B22</td>
<td>—</td>
<td>tag name</td>
</tr>
<tr>
<td>B19</td>
<td>V30</td>
<td>V45</td>
<td>B20</td>
<td>B24</td>
<td>—</td>
<td>tag name</td>
</tr>
<tr>
<td>B20</td>
<td>V45</td>
<td>V43</td>
<td>B21</td>
<td>B18</td>
<td>—</td>
<td>tag name</td>
</tr>
<tr>
<td>B21</td>
<td>V43</td>
<td>V26</td>
<td>B18</td>
<td>B19</td>
<td>—</td>
<td>tag name</td>
</tr>
<tr>
<td>B22</td>
<td>V31</td>
<td>V35</td>
<td>—</td>
<td>B23</td>
<td>—</td>
<td>tag name</td>
</tr>
<tr>
<td>B23</td>
<td>V36</td>
<td>V40</td>
<td>—</td>
<td>B20</td>
<td>—</td>
<td>tag name</td>
</tr>
<tr>
<td>B24</td>
<td>V39</td>
<td>V34</td>
<td>—</td>
<td>B25</td>
<td>B27</td>
<td>tag name</td>
</tr>
<tr>
<td>B25</td>
<td>V38</td>
<td>V33</td>
<td>—</td>
<td>B26</td>
<td>B28</td>
<td>tag name</td>
</tr>
<tr>
<td>B26</td>
<td>V37</td>
<td>V32</td>
<td>—</td>
<td>B21</td>
<td>B29</td>
<td>tag name</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

Table 9: The BEAM table.
Table 10. The COLHAY table.

Table 11. The VERTEX table.

The overall data schema of the PDB, showing how all the entities are related to each other, is illustrated in Figure 6. Back pointers between entities are not shown. The connections to the various ODBs are also indicated.

Implementation of the PDB

We had developed a CAD editor which could be used to create a graphical description of the building in Version 1 of the PDB (Figure 7). We are currently updating it to work with Version 2. The attempt was not to re-create a full-blown, sophisticated drafting or modeling software such as those available commercially. Rather, we needed a reasonably efficient and user-friendly representation tool that would enable PDBs of buildings to be quickly generated for testing by various evaluation programs. It was particularly important for the editor to provide the facility of carrying out a variety of semantically-meaningful operations that currently available general-purpose CAD programs do not provide. For example, deletion of the peripheral room of a building should automatically update the walls of the adjoining rooms from being "internal" to "external" walls.

In the current implementation of the editor, the links from the PDB to the various ODBs are made by associating "types" to each of the PDB entities levels, spaces, walls, surfaces, columns, and openings. Figure 8 shows this link being made through a dialogue box that opens up by clicking on one of the spaces.

In future implementations, we envision the link between the PDB and the ODBs to be
made differently. On the graphic editor screen where the PDB representation is being generated, there will be tabbed palettes of instantiateable objects (walls, doors, windows, rooms, etc.). Objects will be grouped logically, e.g. classrooms in one palette, offices in another, sanitary equipment on a third, and so on. Each symbol in the palettes will represent and will be linked to one object in an ODB. Users can link a graphic entity in the PDB directly to a generic object located at any level of abstraction hierarchy. In this way all properties of the generic object become attributed to the graphic entity, and can be easily referenced by evaluation tools.

Conclusions and further work

We have shown that our proposed representation, based on the separation of the building assembly from its components, can adequately support the needs of design generation, analysis, and evaluation at its most crucial stage—the preliminary design phase. We have already tested the representation with some evaluation tools for architectural design. These tools are just a precursor to more design and evaluation tools for other building-related disciplines that we intend to develop. We see the tool-building process as one of the best ways to ratify and further develop the underlying data structure of the representation. To this end, we plan to invite various disciplinary experts to use our proposed representation as the basis for the development of their design aids.

At the same time, we will continue to develop further both the PDB and ODB components of the database. It is important to emphasize that...
what is presented in this paper is work in progress. There is room for evolution and improvement, even substantial change, if inconsistencies or errors in our data schemes are revealed in the course of future implementation.

We are also working on an integrated building environment, which will include not just the PDB and the ODBs, but also a content database (CDB) and an issues database (IDB). We will develop utilities that overlay all these databases and provide application programs with the information they need, so that the creators of these programs do not need to be concerned about the internal data structures of the building representation. We also intend to focus on the communication and control issues, knowledge representation issues, and user interface issues as the research progresses.

References


locates the opening geometrically on the surface by specifying the distance from the center of the opening to the vertex associated with that surface. The values in the “Width” and “Height” fields would be useful if the opening was not yet linked to a specific door or window object in the ODB, a situation very common in actual practice. Once the linking was done, however, the “Width” and “Height” values would be replaced by those specified in the ODB for that particular object.

slabs, beams and columns