Housing Types and Classification Systems in Case-Based Design

Ji-Hyun Lee,
Graduate School of Computational Design
National Yunlin University of Science and Technology
jihyun@yuntech.edu.tw

ABSTRACT
Expert designers typically refer to and re-use past solutions for recurring design problems. Case-based design (CBD) attempts to transfer this natural design reasoning process to computer-aided design using artificial intelligence (AI) methods and databases. The housing design domain is particularly suited for applying the CBD approach because the traditional method of home design already makes extensive use of precedents and solutions are highly standardized in that industry, at least in the U.S. This paper introduces classificatory types of housing precedents that provide a basis for a structured knowledge representation that supports case retrieval. The classificatory types gives to a research prototype an efficient classification and indexing mechanism that combines form- and component-based features and remains flexible (i.e. can be modified and customized by users), and a retrieval mechanism that uses the indexing mechanism.

1. MOTIVATION AND GOALS
Case-based reasoning (CBR) is a paradigm for re-using past experience. As a part of the broader field known as artificial intelligence (AI), it is a form of analogical reasoning, a central inference method in human cognition (Carbonell, 1983). People like lawyers, doctors, mechanics, and managers usually remember similar past experiences when they face a new problem and apply this experience to the new problem. CBR is an approach to transfer this natural human reasoning process to the computer using AI methods and database technology. The specific knowledge of previously encountered problem situations is organized into a computerized case-base.

CBR is, at its core, as a problem-solving process, but what a problem and a solution is varies from application to application. In design, the problem is generally a functional specification that includes goals and a set of requirements to be satisfied. The solution in design is a description of an artifact to solve a design problem. Case-based design (CBD), an application of CBR to design, promises an efficient way of finding complex design solutions by minimal search, provided that problems presented to the system have strong similarities to known cases for which solutions exist. A case-based design system supports this process by offering designers a data base of past solutions, called cases, which are appropriately indexed so that promising cases can be retrieved in response to the current situation.

1.1 Housing Types
To date, CBR has been widely used in a great variety of application domains such as mechanical engineering, medicine, and business administration. Nevertheless, its use is not common in architectural design, let alone housing design, despite the fact that CBR appears particularly appropriate for this domain because the traditional method of home design already makes extensive use of precedents, and solutions are highly standardized in the industry. In the US, the housing industry’s design process is largely clientless. Houses are generally developer-built products, sold on the open market just like cars or shoes (Rouda, 1999). In the absence of individual clients
demanding a custom design, homebuilders usually base the plans they use on drawings they find in magazines or journals or rely on stock plans, which can be purchased through magazine advertisements or catalogues (Gutman, 1985, pp 1-2). Designers and developers anticipate the reaction of the housing market based on their past experience and select the designs accordingly. They may also look at other projects that are locally under way; if this work has promise for their own business, they will attempt to learn from it and apply it with just a slight variation. The only notable exceptions can be found in custom-designed houses.

In architectural design, precedents from the past are often used to deal with similar current problems: “Typologies, generic solutions, and prototypes are used to help clarify the nature of problems during the intelligence phase, as a basis for generating solutions during the design phase, and as a yardstick for comparison during the choice phase of praxis” (Lang, 1987, pp 62). An essential aspect of cognition is the ability to categorize: to judge that a particular thing is or is not an instance of a particular category (Jackendoff, 1994, pp 135). Types, in the more generic sense, are categories of thought that can be organized in generation hierarchies (Aygen, 1998). From the eighteenth century on, type is used as a classifying tool, as in Linnaeus’ famous plant classification system. The notion of type entered the architectural discourse based on this meaning (Leupen, 1997, pp 133). “In the fields of environmental design, building, landscape, and urban design, typologies are classifications of built structures according to the similarity of their purposes and/or their formal structure” (Lang, 1987, pp 61-62). The present paper considers only typologies of houses able to support a classification of housing precedents useful for CBD. I use the term classification as the name or label given to a type.

1.2 Case-based Design for Housing

In CBR, a case is defined as “a contextualized piece of knowledge representing an experience that teaches a lesson fundamental to achieving the goals of the reasoner” (Kolodner, 1993, pp 13). A case typically consists of the following pieces of information (Flemming and Lee, 2001):

- a problem description, which is often used as index; that is, the problem descriptions attached to the cases in the database can be compared with the description of the current problem so that the system is able to find and retrieve a case that fits the present situation most closely;
- a solution description;
- an outcome description that can be used in various ways: to select the best cases among those that match the current situation or to provide guidance to the designer after a case has been retrieved.

The CBD approach is particularly promising for the housing industry because of the wide-spread use of standardized plans that reflect commonly accepted plan or house types in that industry, at least in the U.S. An organized case-base as mentioned above becomes more powerful over time as designers solve each new problem and add the solution to the case-base to be re-used when needed. Given such a case-base and a flexible indexing and retrieval mechanism, designers can explore various alternatives with ease, including ones they have not encountered before. If the cases can be imported into a CAD system that facilitates adaptation, customization can also happen in an efficient manner, which in turn, may add new cases to the case base. This can also be helpful to the homebuyer looking for a more customized solution. In this case, tools to visualize a case or its adaptation in 3D would be especially helpful. The case base, in short, can overcome many limitations of traditional paper-based and static media.
Provided such a case base is widely accessible within the industry, it can also help in overcoming some limitations imposed on the industry by vertical and horizontal fragmentation. First, each group can access this case-base and make use of the newest data more easily and quickly. That is, hindrance of communications among each vertically fragmented group can be overcome by this industry-wide accessible case-base. Second, a common case-base would allow each group to share the data that it created with others. Communication among each horizontally fragmented group or segment can be improved by the rich collection of precedents in the case base, which may spread information and new knowledge more rapidly throughout a segment. The case base, per se, does not address vertical fragmentation directly. But if it is set up so that comments or annotations can be added to a case as it passes through the different development phases, information about advantages or disadvantages associated with a specific case may become available up- or down-stream in a timelier manner.

1.3 Goals
The two-fold motivation outlined in the two preceding sections leads to the following general goals:

- investigate promising typologies and classifications that can be used for an efficient and flexible indexing and retrieval mechanism in CBD systems
- based on the investigation, define classificatory types in housing design
- develop an efficient classification and indexing mechanism derived from the classificatory types for a CBD research prototype to retrieve housing precedents

2. Types in Housing Design
Designers have made extensive re-use of precedents through analogical reasoning for a long time. Especially in the housing domain, workbooks (for example Schneider 1997; Sherwood 1994) are organized systematically, comparing and evaluating housing precedents. Typology in housing is used to extract common characteristics and compositional principles from housing precedents and to classify them through the comparative and analysis based on these characteristics and principles.

In order to allow the re-use of precedents for a computational system, a structured knowledge representation is needed. The notion of type in housing may provide a basis for arriving at such a structured knowledge representation. Based on a survey of the literature, I define two main classes of concepts in housing design: form-based classifications and component-based classifications.

2.1 Form-based Classifications
A form-based classification addresses higher levels of spatial organization with focus on the outline of the house plan and its context. This classification reflects site information; access method; the shape, orientation, and size of the floor plan as well as elevations and section; and style.

The shape of the house plan overall leads to some basic types: the horizontal and vertical rectangle type, the square type, the linear type, the L type, the T type, the U type, and the courtyard type.

Exterior design or articulation determines the housing style. It is common for home buyers that they have already decided what their favorite housing style is when they choose to purchase a house. Exterior style may determine the mood that a house conveys and the basic layout and design of the interior of the house (Kicklighter and Kicklighter, 1998, pp 375). Traditional house styles in the United States include Native American, Spanish, Swedish, Dutch, German, French, English,
English/Colonial, Salt Box, Garrison, Cape Cod, Georgian, Federal, Greek Revival, Southern Colonial, Italianate, and Victorian (Lewis, 1994, pp 375-384). Housing styles that have been developed in the recent past are called modern (Lewis, 1994, pp 111). Most modern housing styles are variations of one of two basic designs: the ranch and the split-level (Kicklighter and Kicklighter, 1998, pp 383).

The ranch style, inspired by ranchers’ homes in the southwest, was ideal for that region because of the informal lifestyle, open land areas, and warm climate. Now it has become popular throughout the country. Basic features of the ranch include a one-story design with no stairs and a low-pitched, gable or hipped roof with a wide overhang. The structure underneath may be rectangular or have an irregular shape, such as L, T, U, or H. Ranch houses also tend to have large window areas and sliding-glass patio doors. These houses are easy to maintain for outside tasks such as painting, cleaning gutters, or replacing window screens. They are also easily expanded and pose fewer problems of accessibility because they have no stairs. However, they cover large areas and are less energy-efficient than other housing styles because of their long, rambling configuration (Kicklighter and Kicklighter, 1998, pp 383, Lewis, 1994, pp 112-114). Extensive foundations and roofs cause an increase in construction costs compared to multi-story houses.

Variations of the ranch include the hillside ranch and the raised ranch. The hillside ranch is built on a hill so that part of the basement is exposed. Depending on the layout of the lot, the exposed part may be anything from a living area to a garage. The raised ranch, also called the split-entry ranch, has the top part of the basement and garage above ground. This allows light to enter the basement through windows so that the living area in the basement, like a den, can be pleasant if it is well-insulated and waterproof. The main living quarters occupy the floor above the basement, hence the term ‘raised ranch’. The split-level label refers to the fact that one enters the house a half level above the basement and below the main floor so that a short flight of stairs can take one up or down.

The split-level house is designed for a sloping or hilly site. It has either three or four different levels that are vertically offset from adjacent levels by half a floor. The general arrangement places the social, private, and service areas of the house on different levels, for which many variations exist. The three main variations of the split-level design are the side-to-side, the front-to-back, and the back-to-front arrangement. Advantages of split-level houses are that they provide separation of functions within the house and that they are easily adapted to all kinds of sloping sites. On the other hand, they are often more expensive to build than two-story or ranch homes because of the complicated section. Heating may also be difficult because of the different levels (Kicklighter and Kicklighter, 1998, pp 384, Lewis, 1994, pp. 114).

Since a single-family house is often adapted to the site on which it is built, the shape and orientation of the site affect the size, shape, and orientation of the house plan as well as the number of stories and the means of access to the house. For example, if a developer wants to build a house on a sloping site, the developer can consider a hillside ranch or a split-level house and take advantage of the natural slope of the site to make efficient use of space. Depending on the shape of the lot, some variations may exist within the chosen housing style. The following is an example of how housing style, access for cars and people, and number of stories can reflect the natural constraints of the site. Sites sloping sideways are best suited for the front-to-back style. This type of house appears as a ranch from the front and as a two-story house from the back. A lot that is low in front and high in
back requires a back-to-front design. In this style, the living area is typically at the rear of the house, giving it direct access to the outdoor area (Kicklighter and Kicklighter, 1998, pp 384).

2.2 Component-based Classifications

Some geometries have been discovered that deal not with surfaces of uniform curvature, but with surfaces which are bent, twisted, magnified, shrunk or otherwise distorted. The study of such shapes falls under the general heading of topology, the mathematics of position (geometria situs) and of distortion, which deals not with the bending, twisting, and so on themselves, but with the properties of objects which are so fundamental that no amount of such distortions alters them (Broadbent, 1973, pp 224). Among these, connectivity is especially important in a design context.

Collections of connected objects are often represented by graphs. A graph comprises a set of points, called nodes, that are connected by edges. In a topological plan analysis, each habitable space is typically represented by a node and the possibility of movement between spaces or a direct connection is represented by an edge. The graph can represent all interior spaces of a building and also the surrounding external spaces. A graph-based analysis of house plans can lead to a topological interpretation of the organization of rooms in diverse buildings that may uncover common connectivity features despite widely varying shapes.

This method can show that a number of buildings which appear to have very different configurations share nevertheless an underlying structural pattern. March and Steadman (1971) demonstrated this in their analysis of three houses designed by Frank Lloyd Wright (Fig. 1). This example serves to illustrate that a topological analysis of building configurations is not merely a means of visual representation, but a method of capturing spatial organization for comparative analysis (Lewrence, 1987, pp 51-52).

The component-based classification in the planning process for housing is based on some sort of topological analysis that classifies various ways of organizing internal spaces through the circulation system of the house plan and the connectivity among spaces. People’s lifestyle, needs, and wants affect preferred adjacencies and zoning within the house. There are three primary zones in a house: social, private, and support areas. The social zone in any home encompasses the areas where members of the household gather and where friends are entertained, such as living room and dining room. Private spaces include sleeping and dressing, and hygiene areas like bedroom, bathroom, and dressing areas. The kitchen is usually the center for the support zone (Nissen et al., 1994, pp 206-280).

The spatial arrangements of the house plan can be classified based on the topological relationships between the zoning and the circulation system such as halls, corridors, and stairways. According to Schneider (1997), some significant types of spatial arrangements are the following:
Fig. 1 March and Steadman (1971) show how three Frank Lloyd Wright houses, designed for different sites, share underlying spatial arrangements of rooms

- The corridor type: the floor plan is organized according to a circulation axis and the rooms are lined up on one or both sides.
- The insert box: the floor plan is visually interpreted as a large, open space with an inserted cube (or inserted walls). This is also called the core type.
- The living room as centerpoint: the floor plan develops around the living room. There is a variation where the central living room is combined with the corridor or hall.
- The flowing floor plan: The rooms are rarely separated from the circulation area and only slightly separated from each other.
- The hall type: the floor plan develops around a hall that is directly connected to the entry. There is a variation of combining a central hall with the corridor.

Once the arrangement of blocks of spaces is completed, designers consider the relationships between individual rooms. The relations between living room, dining room and kitchen are among the most important connections in the house. The requirements for each room or area depend on the number of inhabitants and the residential profile. The number of bedrooms, bathrooms and garages is a major consideration and often directly reflected in the classification of a home, like in “3-
bedroom split-level.” Needs for special rooms such as a powder room are also important because they make certain house plans more popular than others for people with this preference.

Note that certain configurational features can be derived automatically from a structured, graph-based representation of cases; i.e. they do not have to be captured explicitly through classification labels.

2.3 Classification of Housing Precedents in CBD

Form-based and component-based classifications can be used to classify housing precedents. Fig. 2 shows parts of a possible type hierarchy for housing precedents. Since a single precedent can combine features represented by several types, it can be grouped under different types. This makes the classification of housing precedents complex and requires multiple classification taxonomies. The classificatory types are able to serve as an indexing scheme for the retrieval of housing precedents with desired characteristics or features in a CBD system. I show in next sections how housing precedents can be multiply indexed based on form-based and component-based classifications.

![Fig. 2 An example of type hierarchy for the classification of housing precedents](image-url)
3. INDEXING AND RETRIEVAL IN CBD

3.1 Indexing and Retrieving Design Cases

An index is “a pointer (or indicator) to which a keyword (or label) is assigned and which leads to information about a specific and related topic in a large collection of data” (Rivard, 1997, pp 112). The indexing process is described by Kolodner (1993) as “assigning labels to cases at the time that they are entered into the case library to ensure that they can be retrieved at appropriate times.”

Retrieving the most relevant design cases in a CBR system depends on an indexing mechanism with an efficient memory organization. Typical indexing problems have two parts: one part is the indexing vocabulary problem, which is how to decide what descriptors should be used for some classes of cases. This problem needs a careful domain analysis to find an appropriate terminology. The second part is the indexing assignment problem, which is the process of choosing identifying descriptors for a particular case (Kolodner, 1993). The classification scheme presented in the previous Chapter shows us how to solve the vocabulary problem in the present context. The assignment problem has to be handled by the prototype system developed in this paper.

Retrieval consists of selecting the most relevant case to a current design situation. Based on appropriately constructed indices, the retrieval process searches case memory to find the relevant design cases (Maher et al., 1995). To find the most relevant ones, the case selection process may use various strategies and algorithms for search and similarity measurement.

3.2 Housing Retrieval in CBD

Classificatory types can provide a promising indexing scheme for the retrieval of housing precedents. The classification of housing precedents can incorporate orthogonal taxonomies so that a memory organization that supports multiple classifications is needed in order to implement such taxonomies. According to the housing types I defined in section 2.3, the same house can be retrieved as a raised ranch, a two-story house, or a 3-bedroom unit depending on which types designers or clients are interested in for a specific design problem. Note also that it is possible to combine classifications for retrieval, for example, search for a 3-bedroom raised ranch.

Supporting flexible classifications is also important. Each party involved in housing design may have its own perspectives and interests when looking at a design case and may not want to use classifications already made by others. Therefore, it will be useful to have an extensible or adjustable indexing scheme to support different indexing vocabularies for different parties. That is, individuals can add new classification instances to the system, and also can modify existing instances. For example, both a designer and a sales person may want to use the retrieval function in CBD. However, the sales person’s interest may differ from the designer’s. The sales person may want to retrieve best-selling house plans and to show them to a specific client quickly. On the other hand, the designer may be interested in retrieving house plans based on cost constraints.

4. APPLYING FOR A CBD APPLICATION

The case-based design system prototype I use is the Software Environment to Support the Early Phases in Building Design (SEED) (Flemming and Woodbury, 1995), which includes a case base (Flemming et al., 1997). Fig. 3 shows the system architecture of the first prototype implementation. SEED’s case base consists of the following components (Flemming and Lee, 2001):

- 390 -
- an **object database** that can be used to save persistently those objects and object configurations that must be available to a SEED module beyond a single session. These data typically comprise problem specifications and solutions.

- a **classification knowledge base** (CKB), which allows individual users to develop multiple classification taxonomies; to combine classificatory concepts from different taxonomies into a classification; and to attach such a classification to an object in the database “by proxy”.

- a **case base** (CB) proper that allows the users of a module to attach further information to classified objects, if this is needed for indexing and retrieval.

I describe below the CKB and CB in greater detail.

**4.1 Classification Knowledge Base**

In order to deal with multiple and flexible classification described above, SEED contains a CKB that is independent of the object database. This “hybrid” approach separates precedent instances stored in the object database from the concepts they represent, which are expressed in the CKB.

The classification labels in the CKB may belong to multiple classification taxonomies that imply subsumption relations. That is, the SEED-CKB engine can infer subsumption relations from the classificatory taxonomies. The separation between CKB and object database also makes the SEED-CKB very flexible. Each user can create his/her own classifications and keep the respective knowledge bases separate from those of others according to his/her needs and interests.

![Fig. 3 System architecture for the first prototype implementation](image)

The terminology used in CKB is based on the CLASSIC knowledge representation (Borgida et al., 1992). A classification label is called a **primitive** in CLASSIC. It corresponds to the name of a classificatory type described in Section 2.3. Primitives can be arranged in taxonomies where higher level primitives subsume lower level ones. For example, the primitive ‘split-level’ can be defined so that it is subsumed by the primitive ‘residential’. A classification engine modeled after CLASSIC will retrieve any object labeled ‘split-level’ when it is asked to find any objects labeled ‘residential’.

Fig. 4 displays an example of a tree view of the primitive hierarchy in the currently active CKB.
Primitives can be combined into *descriptions*, which may comprise primitives from different taxonomies. For example, we may construct a description combining the primitives ‘split-level’ and ‘sloped’, where the first belongs to a section-based taxonomy and the second to a topography-based taxonomy. Furthermore, descriptions can be restricted to selected classes of objects, called *host types* in CLASSIC.

CLASSIC and the CKB engine based on it are able to take a description and *normalize* it in the sense that it is augmented by all primitives that subsume the ones in the original description without redundancies. Such a construct is called a *classification*. Classifications can be modified by means of adding or retracting primitives (Aygen, 1998). Fig. 5 shows a classification editing settings box to add or retract primitives and host types (restrictions). A classification can be attached to an object in the object database. This happens by proxy; that is, the CKB uses the unique object identifiers in the SEED-Database as references for the objects to which a classification is attached. In this way, different CKBs can attach different classifications to the same object. This is the basis for the flexibility with which SEED’s classification component can be used by different designers.

![Knowledge Base (SLCKB) Diagram](image)

**Fig. 4** A tree view of the primitive hierarchy
4.2 Case Base

SEED has a case base with additional capabilities beyond the object database and CKB. The case base schema defines the following concepts: case, target, and match operators.

What constitutes a case is highly domain-dependent, which may contain applications dealing with architectural programming or structural design in addition to SEED-Layout. Application developers are able to define cases as they use the CKB to classify them. But such classifications may not be sufficient to represent all aspects that an application may want to use for retrieval. CB therefore provides a target that can be used by any SEED-based application to attach additional information to a case. The SEED developers also realized that certain retrievals should take configurational aspects into account that may require matching on components independently of any classifications. The CB component therefore allows application developers to add match operators to a particular CB.

A prototype system will consider cases mainly in the form of Layouts that can be retrieved through form- or component-based classifications attached to them. These classifications identify, at an abstract level, the type of problem the Layout solves. That is, these classifications represent the problem specifications part of a case, while the Layout represents the solution part.
Under this scheme, retrieval can be implemented in a computationally very efficient way because the CKB and CB engines are able to retrieve the object proxies associated with specific classifications or targets very efficiently. Another advantage of this scheme is that query indices can be computed automatically from a given Layout Problem if we allow Layout Problems to have the same classification as the Layouts that solve them: in this situation, the prototype can use the classification attached to the current Layout Problem as search index to find in the case base all Layouts with the same classification.

I do use the target to capture component-based classifications. In the current prototype, they are restricted to simple attributes like the number of bedrooms or bathrooms in a Layout, which can be automatically computed if the respective Functional Units in a Layout Problem are appropriately classified (Fig. 6). Note that the case base engine can maintain multiple case-bases in which different designers define and populate their cases, targets, and match operators and register the related objects from the object database as proxies.

5. CONCLUSION
This paper identifies classificatory types of housing precedents that are most useful for a CBD system that retrieves precedents based on form- and component-based classifications. These types serve as an indexing scheme for the retrieval of housing precedents in a CBD system. Component-based queries in the first prototype implemented in this paper are restricted to consider only three basic room types: bedrooms, bathrooms, and workrooms based on given API. Extended and more complex component-based queries that take an open-ended set of room types, location and spatial relations into account are conceivable, but introduce a significant increase in algorithmic complexity that deserves a careful investigation in its own right.

Fig. 6 Retrieved by index settings box
ACKNOWLEDGEMENTS
This research was sponsored by the US Army Corps of Engineers Construction Engineering Research Laboratory (USACERL). Additional funds were provided by NSF through the former Engineering Design Research Center (EDRC), an NSF Engineering Research Center at Carnegie Mellon University.

REFERENCES