**The Concept Space Explorer**  
*A Design Case Knowledge Retrieval System*

LIN Chieh-Yen\(^1\) and CHIU Mao-Lin\(^2\)  
\(^1\) Department of Interior Design, Tainan Woman’s College of Arts & Technology  
\(^2\) Department of Architecture, National Cheng Kung University, Taiwan

**Keywords:** design knowledge, semantic web, ontology, concept space, case library  
**Abstract:** Design is often initiated by some ideas or concepts to form the design process, while design cases provide the impetus for design inspiration and conception. This paper is aimed at establishing a knowledge retrieval mechanism based on design concepts and their semantic relations. This research is based on a case library (CBA) implemented on the web as a repository of house cases and a learning environment for case-based design. The implementation and discussion is presented.

1 **INTRODUCTION**

Design precedents or cases are important keys for design exploration and problem solving (Chiu 2003). The case-based system usually tends to treat a design case as a representation of design knowledge that includes past design problems and their solving experiences. However, most of cased-based systems were established based on the database methods and devoted to precisely store past design solutions rather than past design problems and the design knowledge where those design solutions were inferred from (Maher et al. 1995). The usefulness of computer databases can help users easily to retrieve design cases by comparing their feature attributes to find rapid solutions. Nevertheless, we still need to analyze the contents of relevant design cases to understand and to learn the design knowledge stored in those cases when we hope to learn or apply those design knowledge to infer our own solutions rather than just imitate those solutions.

1.1 **The Case Base for Architecture**

The objective of this paper is to establish a knowledge retrieval mechanism based on design concepts and their semantic relations. “Case Base for Architecture” (CBA), a web-based system prototype, is used as a repository of house cases and a learning environment for case-based design. (Lin and Chiu 2003) It currently contains more than 50 cases and hundreds of keywords as a basis for trial tests.
1.2 Design Information in CBA

Cases contained in the CBA system are composed with three levels of information: general, analysis and recommendation including text or non-text information such as drawings and pictures of cases. (Lin and Chiu 2003) Different cases applying intentions may lead to different searching methods and modalities. It is found that CBA is easy to retrieve cases through the feature-based searching function by comparing those attributes of general information with the new design problem because all of explicit features usually were stored in a will-structured and hierarchical database. However, in the intention of “formal application” and “conceptual application”, cases are difficult to retrieve by searching and matching those attributes in a database. Most of these intentions often expect to retrieve abstract ideas or design concepts and their implementations that usually were kept in the ill-structured information of cases such as drawings, pictures, textual analysis or recommendation. Therefore, we need a mechanism that can extract abstract ideas or design concepts from the ill-structured information and clearly represent them to users in order to help them to search, retrieve and reuse useful cases.

1.3 Keyword Based Methodology of CBA

Most of non-text information such as drawings and pictures are ill-structured and difficult to be converted into indexical and searchable formats because the constraints of the present technology. The ordinary alternative method to overcome this problem is to encode non-text information by attaching textual or symbolic annotations to represent the features of them. Consequently, we found that users still relied on the textual or text-like symbolic clues to present their intentions and retrieve abstract concepts they wanted.

The ordinary method to retrieve textual information in database usually relies on “keywords”. The keyword-based method has two major problems that lead to the inefficiency in the retrieval of abstract concepts. First, the inaccuracy between the queries inputted by users and the data stored in the system. Secondly, the ill-structured information in system may cause too many useless results of searching. Despite the existed problems, the keyword-based methodology is still the easiest way to identify and represent an abstract idea or concept, and it is generally applied in most of knowledge management systems. A “keyword” or a “key-term” actually can be recognized as a key to a clue or an index of abstract concepts and design knowledge in CBA. However, it only has limited power to answer various questions from users. Therefore, we applied the data mining technique to extract a list of keywords of design concepts and knowledge from the textual information of CBA. By clustering and ranking the semantic relations of those keywords on the strength of machine and addressing them reverse to selected cases, we reorganized the semantic relations among cases. (Lin and Chiu 2003) It is important to visualize the semantic relations for establishing a knowledge retrieval mechanism to help users to rapidly retrieve, learn and recognize the useful concepts and knowledge in cases.
2 DESIGN KNOWLEDGE RETRIEVAL

To convert the keyword list of CBA into a knowledge retrieval mechanism, we applied the ontology methodology in AI domain to reorganize and to simplify the semantic relations among keywords. Furthermore, the concept mapping representation provided a blue print for us to visualize these semantic relations.

2.1 Ontology Methodology of Knowledge Representation

The ontology methodology is a technique developed in the artificial intelligent domain to facilitate knowledge sharing and reuse. "An ontology is a formal, explicit specification of a shared conceptualization." (Gruber 1993) The purpose of developing the ontology is to provide a machine-processable semantics of information sources that can be communicated between different agents included machine and human. One of most famous lexical ontology is the WordNet that provide a thesaurus contains more than 1,000,000 English terms explained in natural language. Another advanced version of the lexical ontology is the Suggested Upper Merged Ontology (SUMO), which provide a hierarchy mapping and an ontology portal of all terms in WordNet. However, the concepts and knowledge in WordNet or SUMO are too general and more satisfied to present common sense rather any specific domain knowledge, and we do not attempt to establish a general ontology of every knowledge domain in architectural design for now. Therefore, we applied the basic semantic relations in WordNet and SUMO to establish our own thesaurus and used them as a rapid thesaurus when any new term appeared over the limited contents of our own thesaurus.

2.2 Ontology Based Semantics

For general linguistic purpose, WordNet (http://www.cogsci.princeton.edu/~wn/) contains all function of English words included: nouns, verbs, adjectives and adverbs, and all of words are organized into synonym sets to represent underlying lexical concepts. The semantic relations among words in WordNet are also organized in different synonym sets. These basic semantic relations are:

- Synonymy: The similarities in meanings of words are used to build a concept represented by the set of words;
- Antonymy: The dichotomies in meanings of words usually are used to organize adjectives and verbs;
- Hyponymy: The “A-kind-of” relationships among concepts ensure the inheritance of properties from super-concepts to sub-concepts;
- Meronymy: The “Part-of” and “Has-a” relations among concepts ensure the parts and attribute relations to the whole;
- Morphological: The relations are used to reduce word forms.
The Concept Space Explorer

However, the noun keywords usually are more significant for retrieving and representing an abstract concept or indicating the design knowledge. Therefore, as a prototype system, CBA firstly only organized the noun keywords with four basic relationships: synonymy (with a “≒” mark on link line), hyponymy (“∈”) and meronymy (“⊂” for “Part-of” and “∑” for “Has-a” relations) in order to simplify the representation of semantic relations among design knowledge. We also attached an arrow to every link line to indicate where the super class is (Figure 1).

![Figure 1 The Basic Semantic Relations of Concepts in CBA](image)

2.3 Mapping of the Concept Space

Obviously, any concept is generally difficult to be directly represented by a single keyword, term or symbol especially for anyone who is unfamiliar with those words. It needs other relevant words, terms, symbols, and more descriptions of their correlations to help us to understand, learn and remember the concept represented by those words. Therefore, the representation of correlations among words should be more significant and helpful for presenting concepts underlying those words. One of the ordinary methods to visually represent correlations among concepts in knowledge management systems is the concept mapping. The concept mapping is a technique for representing knowledge in graph, i.e. the network of concepts. The network of the concept mapping is consisted with nodes to present concepts and links to present the relations among them. Then we tried to visualize the semantic relations of noun keywords to represent design concepts kept in cases based on the concept mapping representation.

Based on the idea of simplifying the semantic relations among keywords, there are only four basic semantic correlations in CBA for now. However, we found these simplified hierarchical relations were useful represent the common sense in architectural design domain such as the structure and construction systems of cases which were very similar the explicit features in general information of CBA. It was still very difficult to represent some important design concepts such as “Hi-tech”, “Green”, “Ecological”, “Smart” buildings or other interesting design issues for users, because those issues usually were classified in different levels of hierarchy without any obviously direct semantic relations (Figure 2).
Figure 2 Parts of the Hierarchical Concept Map in CBA

Therefore, we applied the data mining technique to give weights on the associative relations of keywords under special issues based on their co-occurrence frequencies in cases. Based on the Zipf’s distribution theory of the information retrieval, however, the most and lowest frequent words usually are less significant. Therefore, users’ recognition feedbacks were collected to revise those weights (Figure 3 left).

Figure 3 Parts of Associative Weights of Keywords (Left) and the Classifications of Associative Concepts (Right) about “Hi-tech” Style Issue

The original distribution of keyword’s weights revised by users’ feedbacks seemed to be very random and without important significances before we classified them with their super concepts in ontology (Figures 3 right). After the classification of keywords, we found it became a more clear and explainable representation of an abstract concept. Then we reorganized relevant concepts based on the associative weights and the semantic relations in ontology to help users to retrieve, recognize and learn useful design knowledge in cases. Therefore, this complex networks consisted with keywords and their semantic relations becomes a spatial representation of abstract concepts within design cases, i.e. the “Concept Space” in which abstract concepts can be coordinated and measured by nodes of keywords and weights of semantic or associative relations among them.
3 SYSTEM IMPLEMENTATION

In order to build a system to represent those design concepts kept in design cases and help users to find and understand those concepts, we apply the machine learning method to analyze the information in CBA. By using the data mining algorithm, we have accumulated a list of keywords about some design concepts from the textual “analysis” and “recommendation” information of CBA, such as “Hi-tech building” and “Green building”, and established their semantic ontology by clustering their semantic and sentence structural relations from previous studies. Finally, we tried to visualize the semantic ontology of design concepts by a graphical interface to help users to explore and recognize the concept space of those design cases in CBA.

3.1 Visualization of Concept Space

The early prototype of CBA was implemented with Microsoft Access and ASP dynamic web page technology. However, because of the constraints of early ASP technology, the database was converted into MySQL database with PHP dynamic web page technology. The current version of CBA is integrated with Macromedia Flash because of its dynamic and interactive visualization interface.

Our basic idea of visualizing the concept space is applying the technique of the concept mapping graph. A single rectangle node attached with one keyword, which was extracted from users’ queries, or the highest ranking of semantic or associative relations in all synonymous keywords, will be used to represent the concept of all synonymous keywords. Then, the links attached on different position of nodes will be used to present different semantic relations: left edge for “Part-of” parts relations, right edge for “Has-a” attributes or associative relations, up edge for links to upper concepts or keywords and bottom edge for links to “Is-a” sub concepts (Figure 4).

![Figure 4 The Conceptual Diagram for Visualizing the Concept Space](image)

Meanwhile, we divided the display space into two areas by the central keyword that was selected by users or extracted from users’ queries for effectively utilizing finite display screen. The upper area is used to reveal upper concepts and keywords; the lower area is used to reveal sub concepts and keywords. Based on the ideas of semantic links, every area will be divided into three sections in order to reveal the “Part-of”, the “Has-a” and the “Is-a” sub or upper concepts.
3.2 The Primary Ontology of Design Concepts

To establish a general ontology included all domain knowledge of architectural design should be a very important but also an extremely arduous work. As a prototype system, we did not attempt to build up a complete ontology for now, and we first just collected around 300 keywords about the early state of architectural design, house and office building type from the architectural dictionaries and important cases bases such as Great Buildings Online and ArchINFORM. Then we built up their semantic relations as a primary ontology of common sense of architectural design knowledge in order to help us to establish a basic concept map of keywords.

However, as we found from previous studies, this ontology could help to represent the explicit features of cases but no help to present the design issue concepts (Figure 5 left). Therefore we attached the associative weights of keywords extracted from cases with selected issue keywords as the “Has-a” meronymy semantic relations to build up a concept map of selected issues when a specific issue was selected (Figure 5 right).

![Figure 5 The Primary Ontology (Left) and the Ontology of “Hi-Tech” Issue (Right)](image)

3.3 Preliminary System Evaluation

To examine the uses of the concept space, few common approaches of design conception were studied. For instance, “light” is often used to initiate the design concept such as “light as the form giver” to shape the building form by changing the façade, the opening, and the outside-inside spatial relationships. Figure 6 demonstrates the visualization of design concepts retrieved from the architect Steven Holl’s “Light House” and Tadao Ando’s “Koshino House”. CBA provides the visual interface to enhance the associations of keywords by highlighting nodes and links.
These two cases have different relevant concepts about a same design issue: “light”. Light is received through “strip wood” of walls in Holl’s “Light House” while it is received through “slit” opening of “curve walls” in Ando’s “Koshino House”. Through retrieving most relevant keywords of major keyword “light”, such as “strip, wood” in “Light House”, “slit, wall” in “Koshino House”, and their upper classes, the visual interface can reveal different concepts about “light” issue in two cases (Figure 7). The nodes, links and the depth of the network become important clues to understand the semantics of design knowledge, i.e. how “light” is shaping the form.

Figure 7 Different Design Concepts in Two Cases about “Light” Design Issue

4 DISCUSSION

The implementation of a knowledge retrieval mechanism of a case library described above provides the ground for the following discussions.
Concept Nodes and Their Links: The ontology or the shared conceptualization of specific domain knowledge would not only be revealed within the concept nodes but also within the linkage of those nodes. For example, the “architecture” node is a hyponymy of “Artifact” class in SUMO or a synonymy of “Building” class in CBA, but also is an attribute hyponymy of “System” class in computer science domain. Through comparing the different links of similar nodes, such as “light” node and its links in “Light House” and “Koshino House”, we could rapidly and visually find the differences and similarities among relevant cases. When the nodes and links are increased, the concept space and the capability of associations are also expanded. However, too many insignificant nodes and links still became a major obstruction for rapid recognition. How to filter out noise nodes and to detect critical links to reveal significant concepts to users still needs more investigations in the future.

Measure of Concept Space in Ontology Networks: Our approach was devoted to converting design concepts, kept in the case contents, into semantic ontology networks. Therefore, it should be possible to measure conceptual distances among individual cases by computing how long the routes of concept nodes in a case are connecting to another case and further to measure the associative relations among individual cases. Then to cluster cases and concepts kept in them by those conceptual distances should be very helpful for experts to discover explicit knowledge kept in cases and beginners to recognize and learn those concepts. Therefore, an effective algorithm for weighting the links of relevant nodes and for measuring conceptual distances among cases will be one of our future works.

Detection of User’s Retrieving Intentions: By utilizing our visual interface, users not only can interactively explore relevant design knowledge and concepts kept in cases, but also can learn and recognize the meanings of abstract concepts through those visual links of relevant keywords and design cases. Through the navigation of our system, the user’s exploring processes in the concept space can provide more clues to show how a human to retrieve, recognize and learn design concepts in cases and how a human to represent their design intentions and find useful cases. To analyze user’s behaviours in the system will be one of our future works that should be helpful for us to understand human mental behaviour during design thinking and knowledge application.

5 CONCLUSION

To recognize the importance of design knowledge in cases and associations of design semantics, this study implements a web-based interface by the data mining technique to retrieve design knowledge based on the semantic and associative relations of keywords. It represents design knowledge by introducing design concepts or issues as nodes and links to associate relevant keywords. Based on the prior discussions, we found our system was useful for rapidly representing and retrieving design knowledge and concepts kept in textural contents of cases, and also helpful for converting ill-structured and implicit information into a hierarchical and explicit format.
ACKNOWLEDGEMENT

This research is supported by the National Science Council in Taiwan, NSC-92-2211-E-006-081.

REFERENCES


