A CASE-BASED REASONING APPROACH FOR CONSTRUCTION PLANNING

Nie-Jia Yau* and Jyh-Bin Yang**

* Associate Professor, Department of Civil Engineering, National Central University, Chungli, Taiwan, R.O.C.
** Ph.D Candidate, Department of Civil Engineering, National Central University, Chungli, Taiwan, R.O.C.

ABSTRACT

This research effort aims to use a Case-Based Reasoning (CBR) approach for construction planning and attempts to investigate its advantages over traditional expert systems approach. Many expert systems for construction planning have been developed, but the existing planning systems are incapable of dealing with the situations when the required input information is incomplete or when the input information is slightly different from the applicable rules in their knowledge bases. A CBR approach is thus explored to overcome such problems. In this paper, a commercially available CBR software - ESTEEM™ is employed as the platform and a set of hypothetical planning cases are provided as the case base. The hypothetical cases are generated by a previously developed planning expert system that estimates duration and costs for building construction at the preliminary design stage. Test results have shown that CBR is a promising tool to compensate the drawbacks of traditional expert systems.

1. INTRODUCTION

Construction industry is an experience-oriented industry. In construction engineering and management domain, experienced experts' knowledge is essential for solving problems and providing suggestions. Recently, many researchers are trying to use rule-based expert systems approach to model the process of construction planning and scheduling. However, the rule-based expert systems need comprehensive input in order to obtain valid results. Besides, traditional expert systems cannot store previous results for future uses. These drawbacks are the limitations of traditional planning expert systems that need to be improved.

Over the past few years, case-based reasoning has grown from a rather specific and isolated research area to a field of wild spread interest in the artificial intelligence domain. CBR is able to compensate some drawbacks of traditional expert systems by remembering and learning from previous solutions (or experience) to solve or provide suggestions for current problems. In this research we use a CBR approach to fast estimate construction duration and costs from previous construction cases with minimum input of characteristics of the construction project.

The rest part of this paper is organized as follows. In the next section we briefly introduce the basic concept of CBR. Section 3 describes how the case base in the developed CBR system is created while section 4 describes the CBR system architecture. The input/output of several test projects are discussed in section 5, and some conclusions are outlined in the final section.

2. CHARACTERISTICS OF CASE-BASE REASONING

CBR is a rapidly growth research area in cognitive science and artificial intelligence. It is a new approach for representing knowledge and using that knowledge to help user solve problems. It's basic idea is that "a case-based reasoner solves new problems by adapting solutions that were used to solve old problems" [7]. Through the techniques of CBR, the experience can be captured and organized as a set of historical cases, cumulated in a case base, used to help problem solving or suggestions providing by recalling similar cases.

The stored case base is similar to a database system with some particular features (fields), but it is more than a database system because it does its retrieval based on the specifics of a situation and finds partially matching cases that can be used to answer the specific question of the user. Moreover, it does not require the full matching of features, nor does it require a database administrator to formulate queries [5]. Thus, a CBR approach is quite suitable for the domains that are experience-rich and probably are hard to define features.

In traditional rule-based expert systems, several problems are found as follows [2,3,4,8]:

1. they do not store nor reason by previously similar cases;
2. human don't usually reason in terms of rules;
3. rules are hard to simulate exceptional cases; and
4. rule-based system are not robust.

Most of these problems can be solved by CBR. For the domains that are very difficult to extract information by rules, CBR provides an approach which makes knowledge representation more flexible, and it also allows for storing cases and solutions for later on retrievals.

However, CBR has some native weakness: it must require cases for reasoning; it might be tempted to use old cases blindly, relying on previous experience without validating it in the new situation [6]. It can be foreseen that the later frailty could be avoided, if proper expert rules are added on the CBR system.

3. GENERATING A CONSTRUCTION CASE BASE FOR CBR APPROACH

In this research, we try to incorporate CBR techniques to estimate construction duration and costs for building construction at the preliminary design stage, since the duration and cost play a key role in deciding which design is feasible or beneficial to the owner at the early stage of project planning.

A collection of real construction project data, which is the most difficult part, to form a case base is needed for this CBR approach. Unfortunately, cases or data of actual construction projects are hard to obtain. To proceed the test of CBR capability, we manage to generate a hypothetical construction case base through a construction planning expert system, Time/Cost Integrated System (TCIS) [9].

Running on Kappa PC, a Windows-based expert system shell, TCIS is an object-oriented expert system for automatic scheduling and cost estimating for building construction at the preliminary design stage. TCIS incorporates rules from experienced construction experts and the Means Cost Data to produce duration and cost estimations. The input to TCIS is a set of 24 design parameters of typical building construction, such as site area, number of floors above ground, number of floors below ground, average floor height, etc. The outputs of TCIS are a construction schedule and construction costs that are aggregated into various groups or levels.

In this research, the construction case base consists of 60 hypothetical projects generated by TCIS from randomly inputting of major features. Before generating the hypothetical projects, the following assumptions are made to simplify complexity of the construction projects.

1. The projects are limited to mid-rise concrete office buildings.
2. The type of foundation is pile foundation with fixed pile size; there are no old buildings or trees on the site; and the quality of lighting and plumbing fixture are medium.
3. The type of interior partition is drywall; the type of exterior partition is cmu wall; the type of floor covering is carpet; the type of heating system is heating pump; and the type of fire protection system is sprinklers.
4. The output activities in the schedule are neglected except the total project duration.
5. The output costs are labor, material, and equipment total costs.

An example case with input and output information is shown in Table 1.

Table 1 Features of An Example Project in the Case Base
(bold-italic represents TCIS output information)

feature	example value	feature	example value
project name	pj01	percentage of window	20
project start	1/1/1997	no. of elevator	4
project finish	12/31/1997	percentage of pavement	30
site area (in acre)	1	percentage of landscaping	20
no. of floors above ground	5	***duration (in days)***	***255***
no. of floors below ground	1	***equipment cost (in dollar)***	***191900***
average floor height (in feet)	11	***material cost (in dollar)***	***1992613***
slope of roof	0	***labor cost (in dollar)***	***1677331***

4. THE CBR SYSTEM ARCHITECTURE

In the following, a Case-Based Reasoning system for building Construction dURation and cost Estimation (CBR-CURE) is presented. The system architecture is shown in Figure 1.

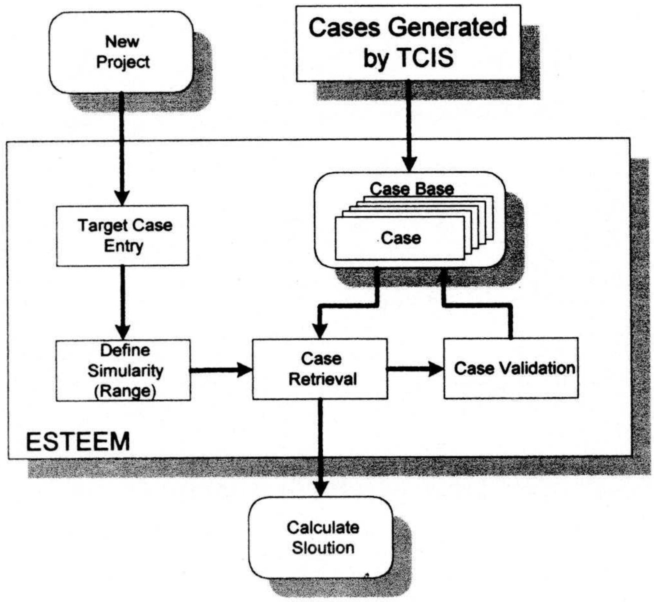

Figure 1 System Architecture of CBR-CURE

CBR-CURE is built on a commercially available CBR software - ESTEEM™, which is a windows-based tool for developing decision enabling applications built through the use of previous problem-solving experiences (cases) [1]. As described in the previous section, CBR-CURE uses 16 features shown in Table 1 to represent a case in the case base. For a new project, 9 main features are entered through a user interface, as shown in Figure 2, into the system to retrieve similar cases. The 9 features are shown in Table 2. Note that each of the features has a weight that is assigned subjectively by the authors, so as the range of matching. These weight and range of matching can be changed if necessary. As soon as a new project is entered as the target case, the system will retrieve similar cases from the case base according to the range of matching in each feature.

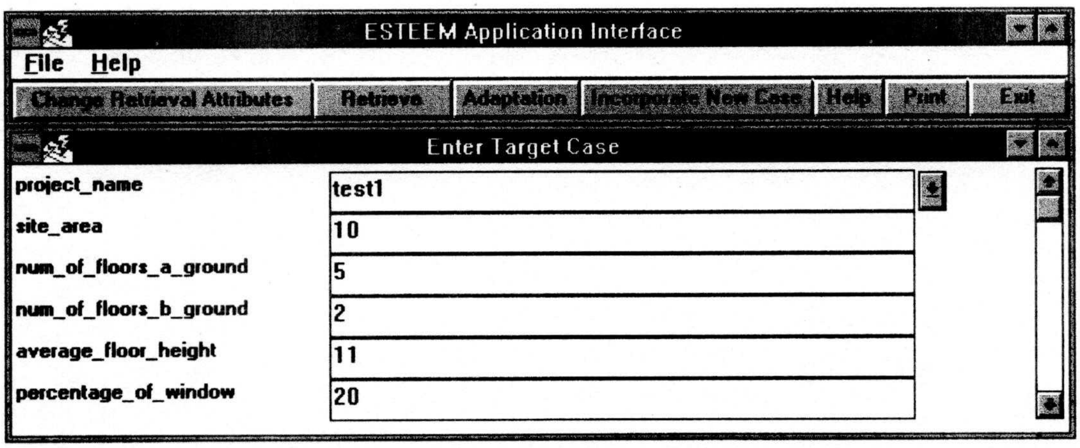

Figure 2 User Interface in CBR-CURE

Table 2 Characteristics of Reasoning Features for Matching Cases

feature	entered range	range for matching (similarity)	weight for similarity evaluation
site area (in acre)	1-50	+/- 5	9
no. of floors above ground	2-10	equal	7
no. of floors below ground	1-5	equal	7
no. of elevator	2-10	equal	5
average floor height	7-15	+/- 10	3
percentage of window	0-100	+/- 10	3
percentage of pavement	0-100	+/- 10	3
percentage of landscaping	0-100	+/- 10	3
slope of roof	0-89	+/- 10	1

In the matching process, a case is deemed similar and retrievable if its score of matching is higher than the minimum score (criterion) input by the user. The score of a case is calculated in such a way that if a feature value of the case falls into the matching range shown in Table 2, the feature weight is accounted for into the score, which will be finally normalized in a scale of 100. I.e., if the features of a case in the case base match every input feature, the case will have a score of 100. There are circumstances that the minimum score is too high to retrieve a similar case from the case base. In such situation, the minimum score needs to be lowered. In other words, the matching criterion is quite subjective, depending on the number of cases retrieved.

After the similar cases are selected from the case base, the intended construction duration and costs (target values) for the new project are computed by the equation shown below. The equation is defined in such a way that the score of each selected case decides how much the case would contribute to the target value.

$$\text{Target Value} = \frac{\sum(\text{score of selected case} \times \text{value(duration or cost) of selected case})}{\sum \text{score of selected cases}}$$

The input features of the new project as well as the calculated construction duration and costs, can be added into the case base if these features are validated and accepted.

5. TEST OF PROJECTS AND RESULTS

Three test projects were used to illustrate and evaluate the performance of CBR-CURE. To examine the capability of CBR that it still can obtain results with various input information, the number of input features of each test project was reduced from 9, 7, 5, to 3; shown as scenario 1, 2, 3, and 4; respectively, in Table 3. In the meantime, construction duration and costs for these three test projects were also generated by TCIS for comparison.

Table 3 Test Results of the CBR-CURE System

Test Project	Project 1				Project 2				Project 3			
Scenario	1	2	3	4	1	2	3	4	1	2	3	4
site area (in acre)	5	5	5	5	10	10	10	10	2	2	2	2
no. of floors above ground	6	6	6	6	5	5	5	5	4	4	4	4
no. of floors below ground	1	1	1	1	2	2	2	2	1	1	1	1
average floor height	11	11	11	-	11	11	11	-	11	11	11	-
percentage of window	15	15	15	-	15	15	15	-	15	15	15	-
percentage of landscaping	20	20	-	-	15	15	-	-	20	20	-	-
percentage of pavement	30	30	-	-	30	30	-	-	30	30	-	-
no. of elevator	5	-	-	-	6	-	-	-	3	-	-	-
slope of roof	20	-	-	-	0	-	-	-	0	-	-	-
selected case id (similarity score)	5(61) 26(56) 53(51)	5(61) 26(56)	5(46) 26(41)	5(39) 26(34)	10(56) 32(51)	10(54) 32(49)	10(46) 32(41)	10(39) 32(34)	2(63) 25(59) 52(54)	2(61) 25(56)	2(46) 25(41)	2(39) 25(34)
duration difference (%)	7.8	6.3	6.3	6.3	4.6	4.6	4.6	4.6	-7.4	-5.7	-5.7	-5.7
equipment cost diff.(%)	11.3	9.6	9.7	9.7	11.7	11.7	11.7	11.7	-12.0	-9.9	-10.0	-10.1
machine cost diff.(%).	9.6	8.2	8.3	8.4	8.2	8.2	8.3	8.3	-18.1	-16.1	-16.2	-16.3
labor cost diff.(%)	11.1	9.6	9.7	9.8	10.0	10.0	10.1	10.1	-13.8	-11.7	-11.8	-11.9
TCIS(dur., e.c., m.c., l.c.)	285	233117	2350794	2006048	280	248914	2334079	2021369	230	165865	1592441	1395638

$$\text{Note: difference (\%)} = \frac{\text{TCIS(duration, cost)} - \text{generated(duration, cost)}}{\text{TCIS(duration, cost)}} \times 100\%$$

It can be seen from Table 3 that, for the same project, the same cases were selected even if the number of input features was reduced to only 3, except that the less features input, the less scores obtained. Since the case base was randomly generated, selection of the same cases might prove that CBR is able to retrieve some similar cases and to produce a solution with incomplete input information. However, the premise of such approach is that major features should be properly defined first. If this can be done, for the problems which require lengthy input, the CBR approach should be able to reduce the process time to a minimum.

Compared to the duration and costs generated by TCIS, the results obtained by CRB-CURE are quite acceptable. The duration differences of the three test projects are less than 10% while cost differences are significantly less than 20%. The higher cost differences could be a result of low matching scores. Note that little variances exist in the costs even though the same cases were selected. This is because that target values are computed based on similarity scores of the selected cases, not on the cases only.

In this study, only quantitative figures are used for matching. But this does not meant that qualitative or descriptive information cannot be used as features for matching. In addition to that, there is a function provided by ESTEEM™ but was not used in CBR-CURE. The function allows for incorporating special rules for retrieval and similarity evaluation during the matching process. For example, suppose that "short of iron workers" is the key factor of a new project, the user may define a rule which directly retrieves similar cases having such feature without considering all the other features, then the matching process would be further shortened.

Currently, the output of CBR-CURE contains only duration and cost information which is also numerical. If the output is in text format, such as construction activities and their logical sequence, the equation for numerical computation is not applicable; more rules need to be incorporated in the system in order to decide the final results based on the selected cases. This part of research will be further explored in the future.

6. CONCLUSION

This study shows that CBR is a promising tool for solving problems if previous cases are available. In the developed CBR-CURE system, the most impressive capability demonstrated is that it can retrieve previous cases by various numbers of input features. It is then concluded that CBR is helpful when knowledge is incomplete, or evidence is sparse; a CBR approach is obviously superior to traditional expert systems in this aspect.

The basic function of CBR approach is retrieving of previous cases based on which the current problem is solved. How the previous cases are retrieved from the case base is the most important thing and needs most consideration for a successful CBR application. When developing a CBR application, features of a case base as well as the features' weights for matching need to be defined carefully. If they are not defined properly, the retrieved cases may not be the ones leading to correct solutions.

From the tests shown in this study, we can conclude that CBR approach is one of the acceptable alternatives for fast construction planning. This is mainly due to the experience-driven nature of construction industry and the ability of CBR approach to mimic the decision making of human planner. The CBR-CURE is a pilot system for integrating CBR approach into the construction planning domain. To validate the system, test of a real construction project is needed in the future.

Results of this study also provide new research directions to incorporate the CBR approach for other construction management problems; such as schedule generation, site layout, construction method selection, resource management, and time/cost predictions. Other computer technologies such as neural networks, expert systems, and multiple criteria decision making software, may also be incorporated into the CBR system to enhance its problem solving capability.

REFERENCES

[1] ESTEEM Documentation, Version 1.4, Esteem Software Inc., 1995.

[2] Harmon, P., "Case-Based Reasoning", *Intelligent Software Strategies*, Vol. 12, No. 11, pp.1-4, 1991.

[3] Harmon, P., "Case-Based Reasoning", *Intelligent Software Strategies*, Vol. 12, No. 12, pp.1-9, 1991.

[4] Gupta, U. G., "How Case-Based Reasoning Solves New Problems", *Interfaces*, Vol. 24, No. 6, pp.110-119, 1994.

[5] Kolodner, J. L., "Improving Human Decision Making through Case-Based Decision Aiding", *AI Magazine* Vol. 12, No. 2, pp.52-68, 1991.
[6] Kolodner, J. L., *Case-Based Reasoning*, Morgan Kaufman Publishers, Inc., 1993.
[7] Riesbeck, C. K., and Schank, R. S., *Inside Case-Based Reasoning*, Northvale, NJ:Erlbaum, 1989.
[8] Slade, S., "Case-Based Reasoning : A Research Paradigm", *AI Magzine*, Vol. 12, No. 1, pp.42-55, 1991.
[9] Yau, N.-J., "An Object-Oriented Project Model for Integrating Building Design, Construction Scheduling, and Cost Estimating for Mid-rise Construction", Ph.D. thesis, Department of Civil Engineering, University of Illinois at Urbana-Champaign, 1992.