Kuwait J. Sci. & Eng. **24**(2), 1997

An expert systems approach for analyzing and forecasting construction project performance

HASHEM AL-TABTABAI

Civil Engineering Department, Kuwait University, P.O. Box 5969, Safat 13060, Kuwait
*e-mail: tabtabai@*kuc01.*kuniv.edu.kw.*

ABSTRACT

Analyzing construction project performance is a diagnostic process, and forecasting project performance is a cognitive process. Both processes require the exercise of expert judgment and intelligence. Artificial Intelligence programming techniques such as expert systems and neural networks provide assistance to construction project managing personnel in capturing the construction knowledge needed for analyzing and forecasting construction projects. This paper describes the development of a framework titled PAFEX (Performance Analysis and Forecasting Expert System). The framework has two main modules. The first module is a diagnostic expert system that has the objective of analyzing construction project performance and adopts causal diagnostic reasoning as a strategy, and object-oriented programming technology as a knowledge representation scheme. The second module has the objective of forecasting the construction performance and applies artificial neural networks as a modeling strategy. The advantages and limitations of these two modeling processes are discussed. Issues regarding integrating the proposed approach with existing project management systems are also discussed.

INTRODUCTION

The project control phase is vital to the successful execution of construction projects. During this phase, the project manager evaluates status reports to analyze project performance and predict any variations from planned parameters over the duration of the project. Traditional performance measurement techniques such as the Earned Value method are satisfactory tools in producing progress and variance analysis reports and in making predictions of the performance of construction project. However, large and complex projects can involve hundreds of tasks in which each has its own character, and therefore requires the expertise of the project manager in making the necessary analyses and predictions. Effective control and accurate predictions not only require the project manager to analyze past performance trend analysis, and compare it against the original baseline estimate, but also require him/her to identify problem areas within construction tasks and incorporate his/her expertise to make an intuitive estimate of their effect of these areas on the expected performance of these tasks, so that possible solutions can be identified.

Many of the existing automated project management systems perform most of the project management functions (i.e. cost estimating, scheduling, cash-flow analysis, cost control and forecasting). However, they are mainly isolated tools that are useful at one stage of the project, and project managers utilize them when it would not be practical to use manual techniques. On the other hand, forecasting techniques, including multiple regression analysis, time series and econometric methods, involve development of an empirical model from historical data to produce predictions of variables of interest. They are based on the assumption that the future is indistinguishable from the past, except for the specific variables identified as affecting the likelihood of future outcome. According to Willis (1987), a forecast should rely on past performance predictions as long as the pattern of changes in the environment is steady. However, in the construction industry, which operates in an environment fraught with complex and ill-defined problems, past history is not always an accurate predictor of the future (Al-Tabtabai & Diekmann 1992). Under these circumstances, forecasts must be based on expert judgment. Moreover, this expertise needs to be modeled and incorporated in current forecasting techniques.

The ability to use construction knowledge and expertise and apply mitigating strategies to analyze project performance deviations and then forecast project cost and schedule completion represents possibly the most important feature not accommodated to date in project management systems. This is because performance analysis and forecasting of construction projects are handled quantitatively by these systems. Artificial Intelligence (AI) modeling tools, such as expert systems and artificial neural networks (ANNs), which have been an active area of research in recent years, can be used to capture human expertise and decision processes in different construction domains. Expert systems are characterized by their use of large bodies of domain knowledge, facts and procedures gleaned from human experts that have proved useful for solving typical problems in their domain (Dym & Levitt 1991). ANNs are dynamic systems that can learn from experience like humans, and can recognize a complex relationship without definition. However, ANNs represent intelligence from a different angle and therefore, can be used to complement expert systems. The application of expert systems to represent the subjective performance analysis and forecasting process of a construction project manager expert is introduced in this paper. A prototype expert system titled PAFEX (Performance Analysis and Forecasting Expert System) is developed that contains two modules. The first module is a diagnostic module that adopts causal diagnosis reasoning as a strategy, and utilizes object-oriented modeling technology and production rules to represent, control and process the acquired knowledge in identifying deviations in performance parameters for each construction task. The second module uses the ANN modeling approach to forecast the performance of each of these tasks in terms of cost and schedule parameters of construction tasks.

The main objectives of this paper can be outlined as follows:

To investigate how to incorporate, model and represent, using expert systems, the intuitive efforts of the project manager in the performance evaluation and forecasting processes.

To illustrate the advantages associated with the object-oriented modeling and the ANN techniques that support integration and facilitate the incorporation of expert systems.

PERFORMANCE VARIANCE ANALYSIS IN CONSTRUCTION

Performance variance analysis of construction projects is a dynamic process involving constant monitoring and frequent updating of costs and schedules. This dynamic nature leads project managers to rely on their knowledge and past experience in project control to evaluate the status of the current project in terms of its work packages (WPs), and then to perform efficient performance variance analysis on these tasks. In typical large construction projects, hundreds of WPs may be in progress simultaneously at different stages, making the monitoring and control functions of each WP's performance tedious and cumbersome.

Project managers usually use numerical techniques that are based on the Earned Value method to monitor, evaluate and then estimate the performance of incomplete and future WPs (Slemaker 1985). This method is a management technique that relates cost planning to schedules and to technical performance requirements. All construction work is planned, budgeted, and scheduled in time-phased "planned value" increments constituting a performance measurement baseline. As work performed, it is "earned" on the basis on which it was planned, in dinars or dollars. Therefore, The Earned Value method introduces three important parameters, namely budgeted cost of work scheduled (BCWS), actual cost of work performance (ACWP), and budgeted cost of work performed (i.e. earned) (BCWP).

The planned value (BCWS) compared with the earned value (BCWP) measures the dinar or dollar volume of work accomplished. Any difference is called an accomplishment or schedule variance (SV). The earned value (BCWP) compared with actual cost incurred (ACWS) for the work performed provides an objective measure of cost performance, and any difference is called a cost variance (CV). Current monthly performance analysis uses these three parameters to diagnose whether a certain WP is performing according to plan in terms of CV and SV measures, and then to forecast the expected cost and schedule performances at project completion by using the cost performance index (CPI) and schedule performance index (SPI), respectively.

Forecasting techniques employed in current Project Management Systems (PMS), such as Primavera™ and Artimis™, use historical performance cost and schedule data as a major source for their predictions. Most of these systems, which are used in managing and controlling construction projects, are based on the assumption that past performance is an accurate representation of future performance. They apply simple regression analysis in forecasting, in which they use monthly performance indices, such as the CPI, to forecast the expected index using a straight line equation (i.e. $Y = a + bX$). The predicted value of the CPI is then used to estimate the required costs to complete the project by using the following equation:

$$EAC = ACWP + (BAC - BCWP)/CPI \tag{1}$$

where EAC is the estimated cost of completion, ACWP is the cumulative actual cost of work performed, BAC is the budgeted cost at completion. BCWP is the cumulative earned cost of work performed, and CPI is the predicted cost performance index at project completion.

RESEARCH METHODOLOGY

Typically, construction performance monitoring and control is an expert-based judgment process. Project managers process a wide spectrum of complex and interrelated information from the project environment, and link it with their own intuition, judgment, commonsense and technical knowledge to analyze the project's performance and produce a forecast. Whenever these experts are faced with a decision-making problem, they try to recall similar problems solved in the past, and map these solutions onto the current problem. Thus, expert decision-making is essentially a cognitive process, in which experts draw conclusions or make decisions about an uncertain event from a set of data or factors that they perceive or which represents the information available to them for making that decision (Al-Tabtabai & Diekmann 1992).

The decision-making process of project manager experts during the construction progress evaluation of a typical WP is assumed to go through three stages. First, they receive various kinds of information from the project environment (e.g. current project status, feedback obtained from the project site, and reports on project progress, etc.). Secondly, these experts usually learn to detect patterns of behavior in work progress, match them with planned or similar behavior they have observed and experienced in past projects, draw conclusions and make a decision about the observed behavior of the project. Thirdly, the experts make predictions about the performance of a construction project on the basis of available information that may influence the performance of the project.

The limitations of existing PMS in capturing expertise can be overcome, to a degree, by the use of a different type of computer-based tool called expert systems. What makes expert systems different from other large computer programs is their ability to formalize and represent knowledge within a specific domain, so that computers can solve problems that include tasks such as planning, diagnosis, structured selection and similar 'cognitive' work. The organization of expert systems is different from that of algorithmic programs. A conventional algorithmic application is organized into data and programs. Expert systems separate the program into an explicit knowledge base that describes the problem, and a control strategy or inference engine that manipulates the knowledge base. Expert system technology is suitable for construction project control and can be applied to the cognitive process used by project managers (Levitt 1985; Al-Tabtabai & Diekmann 1992). The procedures adopted to model the decision-making process of construction project experts in project performance analysis and control consist of the following steps:

the diagnosis of the cause of performance deviation in a given work package through identification of the major factors that are considered by construction project experts,

the prescription of an appropriate action to mitigate the effects of the cause, once it has been diagnosed,

the forecasting of the impact of the deviation on the remaining work in the specific work package,

the prediction of the impact of current, completed and in-progress deviations on those work packages that have not been started, the integration of these results with existing PMS to assist in the generation of revised project plans and schedules.

MODELING EXPERT JUDGMENT

An expert system framework that incorporates and models the decision-making process of construction project experts in analyzing and forecasting the performance of a construction project has been developed. The basic structure is depicted as shown in Fig. 1. A proof-of-concept prototype has been developed and is entitled PAFEX (Performance Analysis and Forecasting Expert System). Two main modules are contained in PAFEX. The first module is the Performance Analysis Module which performs the diagnostic reasoning, and the second module is the Forecast Module which predicts the expected performance of a construction work package. Strategies and knowledge representation schemes for both modules are explained in the following sections.

The performance analysis module

The structure of the Performance Analysis Module, as illustrated on the left hand side of Fig. 1, is similar to the general structuring of expert systems. It has an inference engine, a working memory, and a knowledge base. The module is implemented using LEVEL5 OBJECT™, which is an expert system representation scheme that makes use of objects and rules. The reasoning and the knowledge control structure of this module start with an interactive user interface. The user interface is linked to various elements of the module to provide a flexible and

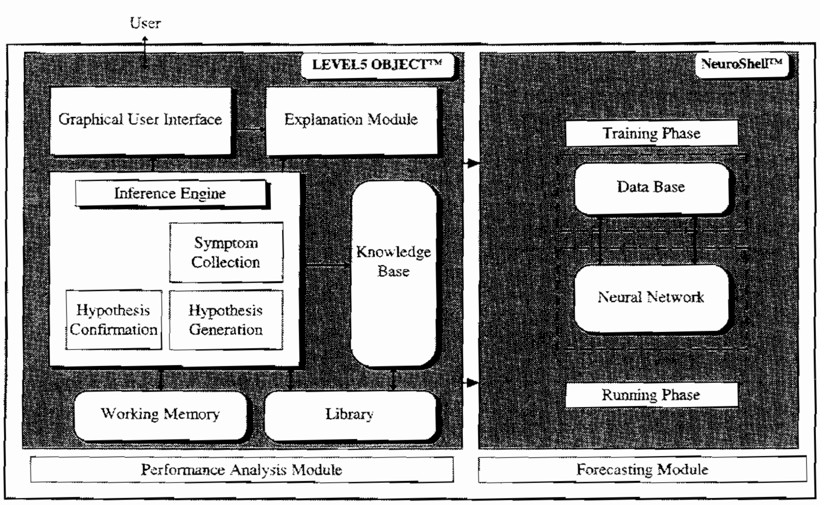


Fig. 1. PAFEX expert system architecture.

friendly working environment. The interface establishes communications channels with a database library, which transfers and uses particular information concerning construction WPs that are in progress for a particular project. Such information includes cost and schedule data (symptoms) that describe a current WP status, WP characteristics and other relevant and project-specific information that is needed for the performance evaluation. Once these data and information have been retrieved, the module then exploits the diagnostic strategy adopted by the inference engine and the knowledge base, to analyze the overall performance in terms of cost and schedule components for in-progress WPs by using the performance variance analysis of the Earned Value method. The eventual judgment classifies a WP's performance into three performance categories: favorable, non-favorable and highly non-favorable performances. If the overall performance for the WP is detected to be favorable, no further analysis is performed and the next in progress WP is proceeded with; otherwise the diagnostic analysis continues, using WP information and characteristics and other project information, to form a hypotheses set of causes. From this set, the expert system analysis continues to confirm or not confirm each hypotheses, and to establish the set of causes for the discrepancy in WP performance.

Diagnostic reasoning strategy:

The performance deviation analysis module is a diagnostic process and therefore can follow the design of diagnostic expert systems. In the design of any diagnostic expert system, there are two distinct approaches: evidential diagnosis and causal diagnosis. In the first approach, the knowledge involved is heuristic and often empirical and is derived from experience, which enables an expert to relate various situation fragments to partial or complete solutions. An example of this knowledge is medical diagnosis, where a physician uses his/her experiential knowledge to relate a symptom to a specific disease. Expert systems using this type of knowledge reduce searches by omitting intermediate relations, because they are unobservable or poorly understood and may not hold up in a specific case (Clancey 1985). For example, in most medical diagnostic expert systems, heuristics typically skip over the causal relations between symptoms and diseases because it may not always be possible to identify the problem structurally (Milne 1987). In these systems, which are mainly rule-based, the diagnosis starts with a list of findings about a patient, and establishes an association between findings and possible diseases ranked by a calculated estimate of likelihood or degree of belief.

In the second approach, the diagnostic expert system in which interest has grown recently, involved reasons in terms of a causal model, or how hypothesized causes bring about symptoms. The diagnostic problem starts with the observation of some behavior. The diagnostician then generates some hypotheses about the cause of the malfunction based on the structure of the device. Although the details of recent approaches of causal diagnosis differ, they all emphasize and justify the need to reason from first principles. To determine why something is performing poorly, using this approach, it is useful to know how it was supposed to perform in the first place (Davis 1988). Therefore, two types of knowledge must be considered in causal diagnostic applications: structure and behavior (Rich 1983). Structure is the physical layout of the component parts, and its knowledge is concerned with the configuration of the system and the way its components work together to perform a specific task. Behavioral knowledge refers to what any of these components is supposed to do. It describes the expected or intended behavior of each component in the causal system. This behavior can be observed or measured and can be defined in terms of the relationship between the information entering and leaving a component, and can be described as a set of rules. The difference between an intended behavior and an observed behavior is termed a discrepancy (Davis 1988). Thus, a fundamental presumption behind causal model diagnosis is the notion that if the model is correct, then all discrepancies arise from defects in the device.

The causal approach was selected as the diagnostic strategy in performance analysis reasoning because it enables the system to reason about how the device, in this case the construction WP, works, and how it fails, by detecting the behavior of the WP's components. Based on the causal approach and its structure and bahavior concepts, a construction WP can be defined by a certain number of parameters or components. Figure 2 illustrates a construction WP and its components as nodes, and inside these components, there are more components or nodes. In this Figure, a WP has a schedule part and a cost part. The schedule part contains construction material procurement, man-hour and equipment schedules as major schedule components. The cost part is divided into labor cost, materials cost, and equipment cost categories. Each cost category can be divided into components. For example, labor cost of any construction WP is a function of quantity of work, productivity, and labor wage rate, and therefore these items can be considered as components of the labor cost category as shown in Fig. 2.

These components can have expected constructional behavioral modes within the WP, and therefore become a basis for WP diagnostic process baselines. In other words, a WP needs to be diagnosed so that its performance can be corrected if it shows "unfavorable behavior" or misbehavior, such as poor cost performance. Misbehavior can be recognized because it exhibits observable symptoms such as unfavorable labor cost performance and unfavorable quantity of work performance. Misbehavior has underlying causes associated with the WP's components and can be corrected to solve or eliminate the observable symptoms. A causal model that

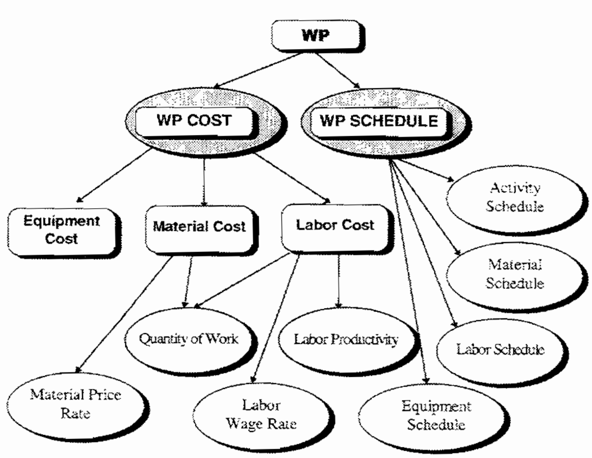


Fig. 2. The causal model of a Construction Work Package (CWP).

links a structure's components with subcomponents can therefore be identified. These subcomponents have expected behavior that can be tested against actual behavior and therefore can be potential causes of the overall discrepancy of the WP. Figure 3 illustrates causal models of the labor cost components (i.e. quantity of work, labor productivity and wage rate) that describe causes for the unfavourable performance of these components. For example, if actual craft mix (actual behavior) has changed from planned craft mix (expected behavior), then it is a possible cause for misbehavior of labor wage rate, which is a possible cause for misbehavior of labor cost.

The causal diagnostic reasoning process that each WP undergoes, and which is adopted by the Performance Analysis Module of PAFEX framework, consists of the following three steps:

Step 1. Symptom collection:

This step represents the performance indicator of each of the WP components introduced in the causal model. The initial symptom collected is the overall cost and schedule performance of the WP using the Earned Value procedures. If, for example, WP cost performance is not favorable, then labor cost, materials cost and equipment cost performance indicators are analyzed and collected.

Step 2. Hypotheses generation:

The task in this step is to study information such as a component's performance indicators, relative degree of severity, WP characteristics, and project specific details, to identify a set of possible causes that could explain the detected discrepancies.

Step 3. Hypotheses confirmation:

The purpose of this step is to establish current causes of the WP's poor performance by testing each hypothised cause in the active hypotheses-set by comparing the

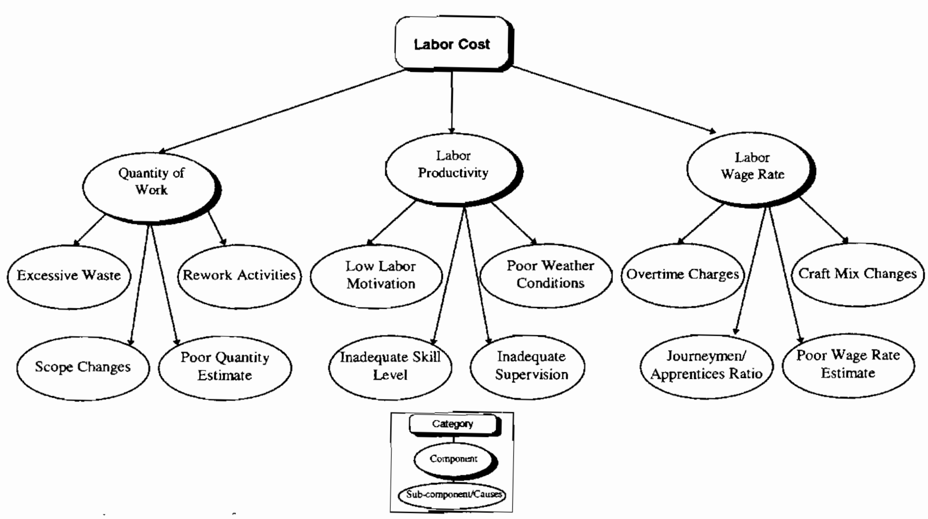


Fig. 3. The causal model for labor cost category.

components' expected behavior with their actual behavior. In this process, hypotheses are eliminated whose behavior is in conformity with expected behavior.

Knowledge representation:
A vital task of the expert system's knowledge base is to construct knowledge models which represent performance deviation propagation within WP components. Application of expert systems in the construction industry is characterized as a slow process. This is because of the inability to integrate expert systems with databases. This inability is due to the incompatibility between data modeling used by expert systems and database management systems, in which expert systems use semantic networks, frames and production rules, while most existing databases are developed using the relational model (Elzarka & Bell 1995). Object-oriented programming (OOP) technology can solve this integration problem by providing a standard model in which each of these systems can use a unified paradigm.

OOP techniques view the world in terms of objects and their behaviors. An object is a self-counted entity composed of data (attributes) and procedures (methods). The attributes describe the object's state while the methods describe its behavior. Methods and attributes are particular to the object. Communication among objects is achieved by sending messages to one another. These messages invoke one or more methods inside the receiving object. Methods can respond to messages by updating local attributes, displaying attributes' values or sending messages to other objects.

In OOP methodology, objects that have the same definition of attributes and methods, while the value of each attribute may differ, are classified under the same class. For example, a construction project may include many work packages that have the same set of attributes (e.g. activity-code, type, start date, quantity of work, budget, float), but the attributes may be different. A particular object then becomes an instance of the class. An object-oriented system therefore consists of various object kinds (including class objects and instance objects) and provides communications with messages among objects that lead to some unique features that makes OOP methodology a powerful programming and software development tool. These features include abstraction, encapsulation, and inheritance (Rumbaugh *et al.* 1991).

Abstraction is the process of deliberately omitting certain details about an object during its definition in order to focusing attention on the essential characteristics of the object relative to the problem domain. Data abstraction is used in data modeling to reduce data complexity and to aid in understanding the essential properties of an object. Encapsulation is the concept of hiding an object's attributes and methods behind the message interface. It allows an object-oriented system's class to control its behavior through methods encapsulated within its structure. For an object, one part of its data attributes and methods within its structure. For an object, one part of its data attributes and methods is hidden from other objects, while the other part is linked with the outside world, allowing changes in the object's state. An object design can be totally changed without affecting the remainder of the program. Inheritance is what distinguishes the OOP approach from other programming systems. Inheritance allows object class to be organized in a class hierarchy, in which each class may have an immediate base class and several immediate derived classes. It allows new classes to be defined from existing classes, and thereby can inherit from its base class the data structures and the methods. New classes can only be specified by their differences from existing classes, rather than having to be totally redefined.

Implementation:

The Performance Analysis Module was developed based on the aforementioned OOP technology. This technology provides the facility to integrate a wide variety of related knowledge elements into one single object, and also provides the inheritance mechanism that can propagate properties from one object to another. Objects are ideally suited for representing structural components of a device, in this case a WP. Classes can be identified by looking into the problem space. A construction WP is part of a system in a construction project that leads to a final physical product. An example is a foundation work package that is part of the substructure system in a building construction project which leads to a footing. The work package represents the effort or process using a group of tasks that use a set of resources (i.e. time, labor, materials, and equipment) and apply a specific planned method to produce the final product. The final product is part of a building system, which is part of a facility. Therefore, a product model, representing the facility and its building systems and final products, is identified.

Each construction WP is measured in terms of cost and schedule performance indicators. Therefore, a process model can be identified to represent the planned effort in terms of projects, systems, work packages and their resources. The process model views any project as a device that has a mission to accomplish something (i.e. a facility). If the device is performing poorly, then a search for a cause is accomplished by diagnosing the performance of its contents, i.e. the work packages. The performance of each WP is measured in terms of the performance of its resources or components (i.e. schedule, labor cost, materials cost and equipment cost). Therefore, a fault model can be defined that describes the behavior of these components and identifies causes of the total device.

To implement the knowledge structure of the Performance Analysis Module using the above description, a computer program or software that supports OOP technology is sought. There are different software development tools based on object-oriented concepts that exist and which implement different methods with different syntactical notations and associated meaning. LEVEL5 OBJECT was selected as the software development tool for the performance Analysis Module because it offers an ability to resolve complex diagnostic problems across a broad range of industries. In this software, knowledge of PAFEX's Performance Analysis Module is represented by two types: passive knowledge and active knowledge. The passive knowledge refers to data that needs to be captured and it is represented using objects and their attributes, while the active knowledge refers to the rules and methods processing the available data and it is represented using rules, event-driven triggers called demons, or procedures written using a high-level grammar in the form of methods, attributes, and objects.

Passive knowledge:

Objects used in the Performance Analysis Module followed the class scheme which was described in the previous section. Major class categories that illustrate the general nature of the model's classes and the relationships between these classes are illustrated in Fig. 4, and can described as follows with examples:

The product model. It consists of those classes which represent the actual product, i.e. the constructed facility and its physical components. Facility class is the base or root class and contains attributes that describe the facility. The facility class has the building-system class as a derived class, and the final-product class as a derived class of the system class. The final-product class has instances or objects such as footing, concrete-slab, concrete-wall, etc. The product model is important in designing and structuring construction performance analysis to extract information particular to end items.

The process model. This model, which consists of classes representing the construction process used to create the facility, starts with a project class as the base (root) class and contains attributes that describe project information that all derived classes can inherit. The system class contains attributes that describe aspects relevant to each system. Only three instances or objects under this class were introduced in the current PAFEX prototype. These systems are the "substructure", the "concrete-work", and the "finishing-work systems. Each system object has a class of work-package that defines these work packages that are associated with a particular system. Figure 5 illustrates an example of a project class, a system class and a work-package class and instances under each class. In this figure, the instance Project 1 describes is general information of a construction project that is being investigated by the expert system. Three instances are illustrated under the system class, which represent substructure, concrete work and finishing work. The object substructure-system has the class WP-substructure to define all work packages that belong to substructure, such as excavation, foundation and *de*-watering work packages. The work-package class has a set of attributes such as WP-code, type, start date, quantity of work, budget,

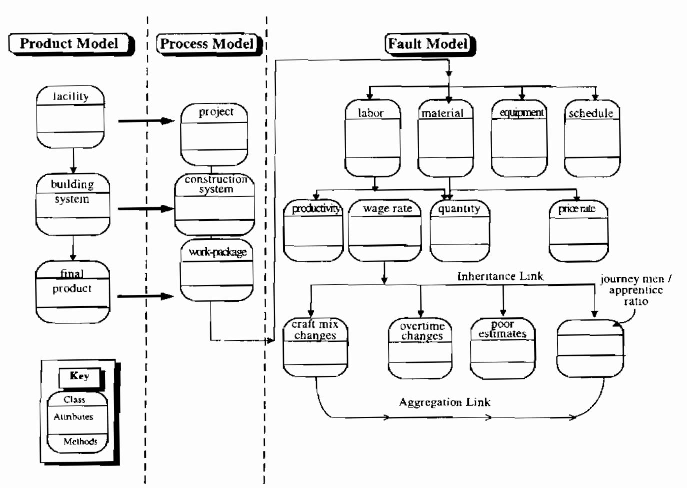
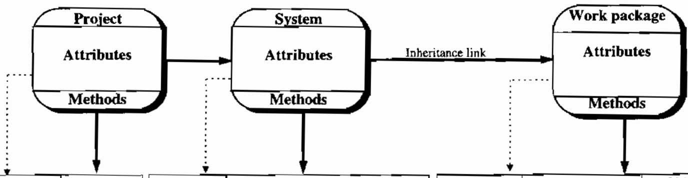


Fig. 4. Object model for performance analysis module.



Attributes	Instances
	Project 1
Contract No	*MPW-95-783*
Budget	12,000.00
Start date	12/20/96
Finish date	01/30/98
Client	KU
Consultant	KEO
Location	Khaldiya
Construction type	Building
-	-
-	-

Attributes	Instances		
	Sub-structure	*Concrete work*	*Finishing work*
Code	3-2000	3-3300	3-5000
Location	F1	C7	R3
Weather sensitivity	Low	High	High
Labor sensitivity	Very Low	Medium	High
-	-	-	-
-	-	-	-

Attributes	Object 1	Object 2	Object 3
	WP 2000	*WP 2300*	*WP 4200*
Description	Excavation at F1-02	Concrete at C7-1	Masonary at B3
Status	Finish	In-progress	Not-started
Start date	01/01/97	01/15/97	03/01/97
End Date	01/14/97	02/28/97	03/30/97
Float (days)	0	4	7
BCWS	1000	750	0
ACWP	1200	950	0
BCWP	1000	700	0
Quantity	100M^3	500M^3	0
C.V	200	250	0
S.V	0	50	0
%completion	100%	32%	0

Fig. 5. Object definition for the process model.

float, etc. Each WP object under this class will carry these attributes after its addition to the system, but with a different mix of values. The inheritance link allows a class to assume the structure and behavior of one or more parent classes by inheriting all of the attributes, methods, and rules.

The fault model:
This model consists of classes representing hypothesized causes for each WP's cost and scheduled resources. Each WP consumes a set of resources; therefore, the classes, schedule, labor, materials, and equipment are identified that describe their relevant information derived from the resource class. Each resource in a WP can misbehave because of causes relevant to each resource. Therefore a class of causes for each resource is identified. Figure 6 shows how a typical hypothesized cause may be represented in the Performance Analysis Module. This figure illustrates the object modeling approach of the causal model of the labor wage rate component, which encompasses all faults attributed to the wage rate as was presented in Fig. 3.

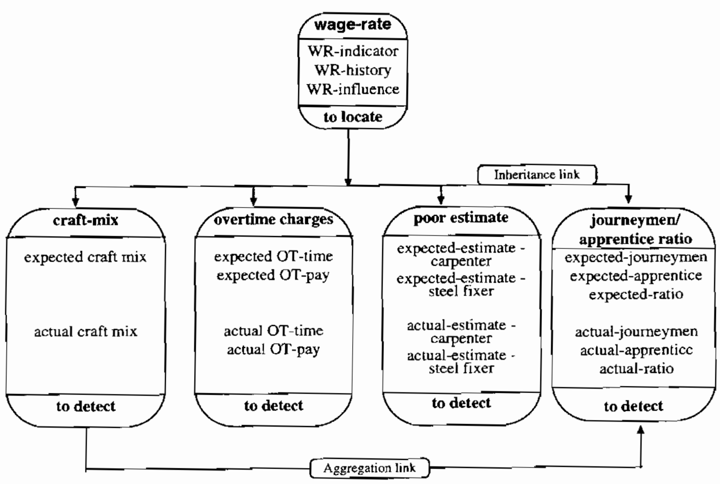


Fig. 6. The object models of WP-labor wage rate faults.

The classes of potential causes for the labor wage are identified: journeymen/apprentice-ratio, craft-mix, overtime, and poor-estimate.

Data attributes and their associated methods are identified for each sub-component or fault class that describe the communications among these classes. For example, the journeymen/apprentice-ratio class contains attributes that describe the expected behavior or planned performance of this sub-component or hypothesized cause. These attributes are expected-journeymen, expected-apprentices, and expected-ratio. The other type of attributes are actual-journeymen, actual-apprentice and actual-ratio, in which they explain the actual behavior of this sub-component or fault in contrast to the planned behavior. In addition, the semantic of the wage-rate causal model is described, as shown in Fig. 6, by the inheritance link and the aggregation link. The inheritance link allows the four classes to inherit the data attributes: WR-indicator, WR-history, and WR-influence and the method to-locate from their base class wage-rate. The aggregation link expresses a hierarchy of composition between the journeymen/apprentice-ratio and craft-mix classes.

Hypothesized causes for the other WP components are depicted using similar object representation in the framework. The number of causes were categorized into general causal groups in order to limit the size of the prototype. For example, the number of causes attributed to unfavourable labor productivity performance were found to be more than 110 causes, however only four generalized causes are employed, as shown in Fig. 3.

Active knowledge:

Data manipulation, i.e. the process of performing calculations, extracting information, or finding specific data within PAFEX Performance Analysis Module, is accomplished by using a series of methods, message sending and rules. Methods and message sending define the attribute's behavior and they are used to determine the attribute's value or execute a series of actions when the attribute's value changes. Methods and message sending are used in this module to retrieve and load values of the attributes concerning WP's components, and activate the necessary diagnostic rules. For example, in diagnosing the quantity of the work component, a message is sent by the object to retrieve and load data values of the attributes concerning quantity of work. Another message is activated, which is to diagnose the present performance using quantity-performance rules. These rules match the retrieved attributes' data with performance rules and then display the performance conclusion for this component, which classifies the performance into favorable and non-favorable states. If performance is favorable, a message is sent to initiate and activate the labor productivity performance object to investigate this component using the same procedures. If quantity performance is unfavorable, the analysis then initiates the quantity sub-component class by sending a message. The next step is to confirm each sub-component or hypothesized cause by activating another message, which then make use of the hypotheses-confirmation rule group to confirm or disconfirm each hypotheses in the list.

Rules incorporated into PAFEX's knowledge base are similar to production rules used in expert systems (i.e. IF-THEN-ELSE). These rules are used by the prototype to represent the diagnostic reasoning process derived from experts and site personnel in the construction project control. For example, if the attributes CV or SV of a particular WP instance have negative performance values, then a method is activated in this instance to determine the potential causes using the rules incorporated in the knowledge base. There are three main rule groups attached to the expert system that follow the causal approach adopted for this prototype. These causal rule groups are: symptom collection rule group, hypotheses generation category, and hypotheses confirmation group.

Symptom collection rule group:
This group represents rules associated with each structural component object of the causal model to describe the performance of each component in the causal model. This analysis uses the Earned Value Method's variances and a set of thresholds to classify the performance of each component in terms of three qualitative states: favorable, nonfavorable and highly non-favorable. An example of this rule group is as follows:

If wp-cost variance is greater than (threshold value: CC1) and
present quantity variance is greater than (threshold value: CQP1) and
cumulative quantity variance is greater than (threshold value: CQC1)
then quantity performance indicator of this work package is: Highly non-favorable.

Hypotheses generation rule group:
If a performance of a WP is unfavorable then a message is sent to initiate the hypotheses generation rule group of this particular WP. These rules make use of the performance qualitative states in terms of the WP's components (i.e. quantity of work, productivity, wage rate, materials price rate, and equipment) and project information and WP characteristics, to establish a hypotheses set or a list of possible causes of the performance deviation. An example of this rule group is as follows:

If the overall labor productivity performance indicator is unfavorable or highly-unfavorable, and
the overall quantity performance of this WP is unfavorable or highly unfavorable, and
this WP is labor intensive, and
one of the causes of quantity poor performance is: scope-changes
then hypotheses-cause = slow-progress-due-scope-changes, and
hypotheses-cause = low-labor-motivation.

Hypotheses confirmation rule group:
These rules attempt to confirm each hypothesized cause in the hypotheses set. If a cause is confirmed, then the prototype assigns this cause as a value to the conclusion slot for each component object. The behavioral comparison strategy explained earlier, in which the observed or actual behavior is compared with the expected behavior of a sub-component, is practiced by this rule group.

Using the labor-wage-rate sub-component as an example after the initial data is retrieved for this object, the performance state of labor-wage-rate is determined. The goal of hypotheses confirmation is to test each fault that appears in the causal network, as shown in Fig. 3, to determine if it accounts for the observed unfavorable labor wage rate behavior. The behavior testing starts with the Journeymen/Apprentices ratio sub-component, in which the journeymen-apprentices-ratio object is initiated. The attributes that describe the expected behavior or planned performance of this sub-component (i.e. expected-journeymen, expected-apprentices, and expected-ratio) are used for behavior testing with the other types of attributes in that they explain the actual behavior (i.e. actual-journeymen, actual-apprentice and actual-ratio). The hypotheses-confirmation rule group associated with this object performs two operations. The first operation sends a message to retrieve values of both expected and actual behavior attributes. The second operation determines the overall behavior of this hypothesized fault or sub-component. The conclusion of this process will either result in the confirmation of one or more hypotheses, or the elimination of all hypotheses due to a match between actual behavior and expected behavior.

Knowledge acquisition:
Knowledge acquisition is one of the most important yet difficult aspects in the development of an expert system (Scott *et al.* 1991). Factual data and rules employed by experts in construction project control were sought. Knowledge required for the Performance Analysis Module was acquired by interviewing an expert in construction project control and using collected literature on the subject of factors affecting project achievements.

The forecast module

Once the Performance Analysis Module has determined causes for any performance deviations in a WP, the next step in this framework is activating the Forecast Module. In this module, the expected final performance in terms of the WP's components, is predicted. This analysis accesses the database library to obtain the information necessary to forecast the expected performance of a project in terms of the schedule and cost components. The components that were chosen for this module are 1) WP schedule, 2) quantity of work, 3) materials price rate, 4) labor productivity, and 5) equipment cost. For each of these components, quantitative and qualitative factors that influence their performance are elicited and identified. The output or solution presented to the user by the system is a description of the individual WP performance and a prediction of future performance in terms of each component.

Following the identification of the quantitative and qualitative factors, an approach that models an intelligent link between these factors and the forecasted output is needed. The a artificial Neural Network based (ANN) forecasting approach is adopted for this module rather than using production rules. The serial architecture and knowledge structuring of rule-based expert systems limit learning by themselves and processing incomplete, partial cues and unseen data while ANNs allow self-learning, self-organization and parallel processing, and are well-suited for problems involving matching of input patterns to a set of output patterns where deep reasoning is not required. ANNs are used extensively for modeling judgements and the thought processes of experts and estimating project control parameters from current project conditions, as illustrated in Moselhi *et al.* (1991), Al-Tabtabai & Alex (1997), and Al-Tabtabai *et al.* (1997).

ANNs are a class of modeling tools inspired by the work of biological neural systems. They are composed of neurons and connections which are organized in three layers: an input layer, middle or hidden layer(s), and an output layer. The signals detailing the problem enter the input layer and flow to the output layer, which represents the networks' solution to that problem through the middle layer(s). The connections between the neurons are associated with numerical values called connection weights which determine the influence of one neuron on the other. The input layer's neurons receive their activation from the environment, while the activation levels of neurons in the middle and output layers are computed as a function of the activation levels of the neurons feeding into them. Typically, this involves the summation of all incoming signals followed by the application of a non-linear function termed as the transfer function. The connection weights are determined or learned by the network by repeatedly supplying training examples. In this process, termed the training process, the connection weights are modified continuously until the error between the desired output and the actual output is minimized. Knowledge is effectively learned and stored by the weights on the connections between the neurons. Once training has been completed, the ANN can generate the required output for other problems not considered during training. Training time is mainly influenced by the number of input neurons and the extent of noise in the training data. It is observed that the neural network could adequately learn the decision process and reproduce the judgements made by the experts for more details; see Lippmann 1987; Caudill 1989.

The ANN technique, therefore, has the ability to learn and adapt a solution to a problem from experience, which can be utilized to capture the functional relationship between the performance of a construction project or WP and the environmental factors that cause this performance. For the current problem representation, the input layer consists of the construction project environment variables, while the amount of variance expected in the schedule and cost components represents the output layer of the neural network. Changes in the key environmental variables affecting the schedule and cost components can be related to the effect they create on the project's schedule and total cost. Once a set of neural networks for the cost components that could be trained with the above relationship is obtained, then each of these networks can act together to predict the expected variance in the total project direct cost. The trained networks together compute the expected variance in the project cost by analyzing the domain information from the project environment.

Knowledge elicitation:

Knowledge required for the Forecast Module was acquired from a series of interviews with twenty-one scheduling experts from the Kuwaiti construction industry. The experts were asked to identify the factors that would cause a performance variance at the completion of this project for each of the schedule and cost components. For example, the following eight factors were identified to influence the forecast of the schedule variance:

1. Performance of the contractor's management (to cope with planning, controlling and technical problems).
2. Cash flow situation.

3. Materials and equipment procurement situation (such as materials availability, equipment procurement difficulties, import restrictions, transit delays and difficulties in the approval of alternate materials).
4. Skilled labor availability and productivity.
5. Weather and environmental influences (weather influences include temperature, humidity, wind speed and rainfall, while environmental factors include remote sites and poor access).
6. Amount of rework, extra work, work difficulty (due to unexpected job difficulties and scope changes).
7. Percentage of work completed.
8. Trend in schedule variance (the pattern or movement of the schedule variance in the last three reporting periods).

The next stage in knowledge elicitation is generating hypothetical case profiles by arranging the previously obtained factors for each component with random values attached to them. Each case profile constitutes a hypothetical construction project situation. The purpose is to create the necessary training cases by eliciting decisions from project experts on these case profiles. Five project managers (experts) exercised their judgements by analyzing a set of case profiles (input) that have mixed values for the factors. The judgement (output) represents the expected change of variance for the described project, and is scaled so that it represents different percentage changes in the variance. The judgements obtained from each expert on these case profiles were distributed among the experts as feedback for modification and revision of judgements. The revised judgements were averaged to reach a smoothed judgment value for each of the case profiles. The procedures for knowledge elicitation and conflict reduction among these experts to reach a unified or smoothed judgement are described in full details in Al-Tabtabai & Alex (1997).

Implementation of the ANN:
The case profiles with the factor values (input) and extracted judgement values (output) were coded to generate a set of training data for each neural network. The generated input-output data pairs were divided into a training set and a test set, one set of profiles to be used for training and the second set of profiles for network validation. A continuous-function mapping network has been adopted since both the input and output parameters are continuously variable. Neuroshell^{TM}2, a commercially available neural network modeling tool, was used for the training of all six neural networks. A three-layer back-propagation neural network was adopted for training. The back-propagation algorithm is a widely-used training technique for continuous-function mapping. It is simple to code and is a theoretically sound method that performs well at modeling non-linear functions (Rumelhart *et al.* 1986). Back-propagation algorithm training develops the input-to-output mapping by minimizing a sum-squared-error cost function measured over a set of training examples. For example, the eight factors that are considered for the prediction of schedule variance became the input variables for the schedule variance representation as shown on the top level of Fig. 7. Therefore, eight input neurons and one output neuron connected through a hidden layer of nine neurons constitutes the neural network arrangement. The training data set was continuously looped through the

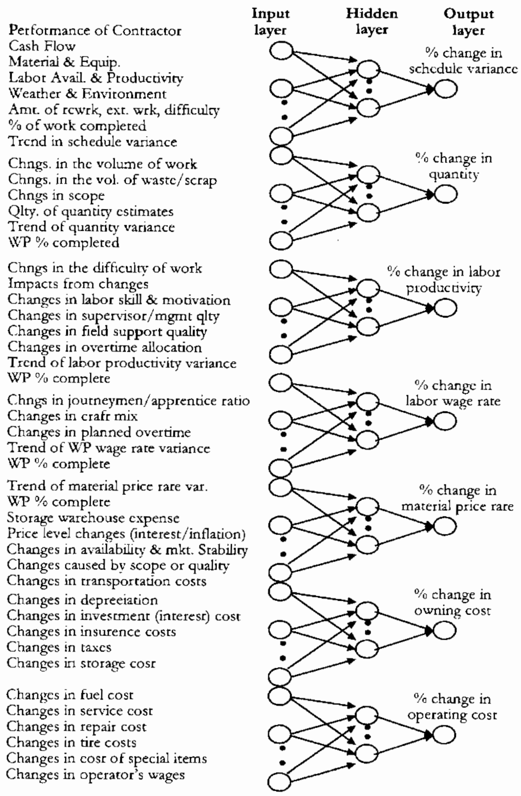


Fig. 7. A typical neural network performance forecasting system.

network and after every predefined number of iterations, the test-set data was passed through the evolved network to generate an output. Mean squared error between the predicted values and the actual values for the test-set were calculated and compared. In the case of schedule variance forecasting, by the end of 20,000 cycles it was found that the error could not be significantly reduced and thus the training was concluded. The complete methodology adopted in the schedule variance forecasting process using ANNs is detailed in Al-Tabtabai *et al.* (1997), while the implementation of the cost forecasting system using ANNs is presented in Al-Tabtabai & Alex (1997).

EVALUATION OF PAFEX

The evaluation process of expert systems and ANNs constitute two processes: validation and verification. Validation refers to the process of determining that the prototype actually solves the problem of interest, while verification focuses on the mechanics of the reasoning process. The validity of the Performance Analysis Module was judged based on the effectiveness of this module while diagnosing a hypothetical case and the ability to perform consistent diagnosis. One expert indicated that the strategies developed using the causal approach are valid and similar to the diagnostic process of a construction project manager. The ability to make consistent diagnosis was focused on quantity of work and labor cost components of a WP using a number of test cases in which the developed prototype was able to make the required diagnostic analysis, and eventually identify hypothesized causes for these components.

Validating the performance of the Forecast Module was conducted by applying this module on a multistory office building project in Kuwait. The estimated budget of this building was KD 10 million (KD 1.00 ≅ US$3.00) with an estimated project duration of 24 months. A progress report was designed to extract the necessary knowledge needed for the performance analysis. The ANN models were applied after 23% of work was actually completed for this project. A project manager evaluated the performance of the project every month and provided the values of independent variables. These values were then applied to the ANN models to predict the expected variances of cost and schedule on the planned completion date. Both the models consistently predicted a large variation from the beginning and provided early warnings to the management regarding the high variance expected on the planned completion date.

LIMITATIONS AND FUTURE WORK

At present, PAFEX remains in the prototype stage and has been operated as a training and research tool only. It must be pointed out that although the presented framework focuses on many issues in the analysis and forecast of construction performance, additional work must be done to complement the limitations. The current version of the developed prototype established the causal diagnosis strategy on quantity of work and labor cost components. To apply this strategy to the other components, it is important to identify those parameters or attributes that constitute an expected and acceptable behavior for each WP component, and then try to establish what values could lead that component to misbehave or to be a cause of unfavorable performance.

Validation results depend on the quality of the cognitive process and the judgments adopted for the development of the expert system and ANN modeling techniques, and the ability to represent project-specific data. The models should be regularly revised according to the change in working environment and future policy changes. It would be appropriate to develop individual models for different classes of construction projects undertaken by a firm (e.g., highway construction, bridge construction, building structures-low, medium and high rise, etc.) and integrate these with existing PMS to act as decision aids for schedule control.

Limitations in the Forecast Module must also be noted. This module needs to address two main issues during the ANN modeling process. The first issue, the accuracy of the models, mainly depends on the soundness of the underlying expert decisions. In other words, the quality of the generated predictions by the models are directly affected by the legitimacy of judgments used for training. The second issue is that the training cases generated from the judgment of experts are not free of bias because of the intuitive and subjective nature of the judgment process of individuals. Therefore, any variables that have a bearing on the WP performance, if omitted during the knowledge representation stage, will affect the validity of predictions.

Further work is needed to make the presented framework PAFEX a better decision support expert system for the analysis and forecast of construction project performance. Future work will include the following activities:

Certainty Factors:
The incorporation of certainty factors handling uncertainties is a task of future work for the system. These factors can be used to calculate and compare each hypothesized cause in the hypotheses generation step. This could narrow down the list so that better performance causal analysis could be achieved.

Real Training Cases:
The ability of ANNs to operate with noisy and incomplete data suggest that the above mentioned limitations can be overcome to an extent. This ability can be enhanced by choosing high quality, well-experienced professionals in the local construction industry as domain experts for knowledge representation. Increasing the number of training cases with a wide representation of various possible situations will enable the network to generalize and learn the problem more accurately. Training cases can also be obtained from actual projects. In this case, the environmental variables and the effects of these variables on the WP are known and can be related to the schedule control domain. Careful observation at regular intervals is necessary to generate such training cases. Thus, a more realistic representation can be obtained for training. The models can be further refined to represent any unique and peculiar characteristics of the current project. New data can be appended to the training set for subsequent updates of the models.

Integration with existing PMS:
Another item of future work is the integration with existing PMS following the development of the expert system. Construction project sites typically have computing facilities that consists of a PC running under the WINDOWS™ environment which uses powerful and inexpensive PMS (e.g. Primavera™ and MS Project™), cost estimating systems (e.g. Timberline Estimator™), construction design (e.g. Stadd™), along with spreadsheets and database management systems. Most of these software systems allow the user to customize various operations by providing at least one programming language support to their software systems. This allows the user to create powerful macros which automate various sequences of actions within the software systems. Most window-based systems also allow the transfer of data between compatible systems using dynamic data exchange (DDE), object linking and embedding (OLE), clipboard transfer, open database connectivity (ODBC) etc., along with the import and export of data files in pre-defined formats. Given these capabilities, it is possible that the developed expert system and the ANN models can be embedded into MS Project™ using the built-in Visual Basic™ programming language. For example, queries regarding the environment factors required by the ANN model can be prompted using dialog boxes created in Visual Basic™ and the user entries about the status of these factors can then be used by the model to generate the forecast. The forecast can be applied automatically to the WPs remaining or in progress, and the resulting changes will be reflected in the project system by rescheduling the tasks based on these predictions.

CONCLUSION

The paper presented a framework to model the decision-making process of the project expert in analyzing and forecasting the performance of construction projects in terms of its work packages. The performance analysis part of the framework adopted the causal approach in diagnostic expert systems. In this approach, hypotheses confirmation has the goal of testing each suspect in the hypotheses set to confirm if it counted for some or all of the unfavorable performance in the component. This testing is done through comparison between a set of expected or planned behavior with actual readings of these parameters, and then programming this testing so that the comparison between actual and expected behaviors are done with user interaction. The current prototype version established this strategy in the quantity of work and labor wage rate components, while the other components depend on user interaction. To complete this strategy on the other components, it is important to identify those parameters or attributes that constitute expected behavior (i.e. acceptable performance indicators or parameters) for each sub-component, and then try to establish what values could lead that sub-component to misbehave or to be a cause of the unfavorable performance of that component. The next version must address this issue and evaluate the alternatives based on causal reasoning. Although this subject is relatively new in the field AI application to construction project control, extensive research in the area of causal reasoning is being conducted on other applications and that can benefit construction domains.

The forecasting part used non-linear modeling based on ANN which provide the following advantages: (a) ability to model the complex non-linear mapping required in the decision-making process, (b) fault tolerance capability in smoothing out the noise in the data collected for training, and (c) ability to generalize on incomplete information.

Modeling construction experience to use in future projects can significantly help achieve project objectives. ANN models that utilize current data and the knowledge of experts has been developed to improve schedule plans at regular intervals. The ANN models allow project personnel to use the most current data to dynamically evaluate the impact of environmental variables on work package schedules and to create fresh schedules instantly. The expert system and the ANN models can be incorporated into current project management procedures so that more realistic and meaningful revised schedules can be generated systematically at any stage of the project. The complexity involved in judging the status of each activity or work package (whether remaining or in progress) can be reduced by the proposed framework. Moreover, the framework can provocatively improve the analysis of construction performance and the accuracy of cost and schedule projections by incorporating actual field data observed from past projects. Project personnel or owners no longer need to wait for weeks to acquire updated project status reports.

The proposed framework deals with the need for designing intelligent PMS and demonstrates that there are potential benefits to applying AI technology, especially expert systems and neural networks to construction project management.

ACKNOWLEDGEMENTS

The author wishes to express his appreciation and gratitude to Kuwait Foundation for the Advancement of Science (KEAS) for the financial support provided to conduct this research under grant 89-09-01. An acknowledgment should also be directed to Kuwait University Research Center for organizing and directing this grant.

REFERENCES

Al-Tabtabai, H.M. & Alex, P.A. (1997). Prediction of cost performance in construction projects using neural networks. Kuwait Journal of Science and Engineering **24(2)**.

Al-Tabtabai, H.M., Kartam, N., Floods, I., & Alex, P.A. (1997). Construction Project Control Using Artificial Neural Networks. Artificial Intelligence for Engineering Design, Analysis and Manufacturing **11(1)**.

Al-Tabtabai, H.M. & Diekmann, J.E. (1992). Judgemental Forecasting in Construction Projects. Construction Management and Economics **10(1)**: 19–30.

Caudill, M. (1989). Neural networks primer. AI Expert, Feb., 61–67.

Clancey, W.J. (1985). Heuristic Classification. Artificial Intelligence **27(3)**: 289–350.

Davis, R. (1988). Model Based Reasoning: Troubleshooting. In **Shorbe, H.E. (Ed.).** Exploring Artificial Intelligence. Morgan Kaufmann Pub., San Mateo, 297–346.

Diekmann, J.E. & Al-Tabtabai, H. (1992). Knowledge Based Approach to Construction Project Control. International Journal of Project management **10(1)**: 23–30.

Dym, C.L. & Levitt, R.E. (1991). Knowledge Based Systems in Engineering. McGraw Hill, New York.

Elzarka, H.M. & Bell, L.C. (1995). Object-Oriented Methodology for Materials-Management Systems. Journal of Construction Engineering and management. **12(4)**: 438–445.

Levitt, R.E. (1985). Using Knowledge of Construction and Project Management for Automated Schedule Updating. Project Management Journal **16(5)**: 57–76.

Lippmann, R.P. (1987). An introduction to computing with neural nets, IEEE ASSP Magazine, April, pp. 4–22.

Milne, R. (1987). Strategies for diagnosis. IEEE Transactions on Systems, Management and Cybernetics **SMC-17(3)**: 333–339.

Moselhi, O., Hegazy, T. & Fazio, P. (1991). Neural networks as tools in construction. Journal of Construction Engineering and Management **117(4)**: 606–625.

Rich, E. (1983). Artificial Intelligence. McGraw Hill, New York.

Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F. & Lorensen, W. (1991). Object-Oriented Modeling and Design. Prentice-Hall, Inc.

Scott, A.C., Clayton, J.E. & Gibson, E.L. (1991). A Practical Guide to Knowledge Acquisition. Addison-Wesley, Reading.

Slemaker, C.M. (1985). The Principles and Practice of Cost/Schedule Control Systems. Petrocelli Books, Inc. Princeton.

Willis, R.E. (1987). A Guide to Forecasting for Planners and Managers. Prentice Hall, Englewood Cliffs.

(Accepted 2 April 1997)

نظام خبير لقياس الأداء الحالي والمستقبلي لمشاريع التشييد

هاشم مساعد الطبطبائي
قسم الهندسة المدنية بجامعة الكويت
ص . ب: 5969، الصفاة 13060 الكويت

خلاصة

يتناول البحث تطبيق الأنظمة الخبيرة Expert Systems في مجال ادارة مشاريع التشييد. والأنظمة الخبيرة تعتبر ادارة لاستعمال الحاسوب في تحصيل وتركيب المعرفة البشرية والخبرة في مجالات لا يمكن تطبيق البرامج الرقمية مثل " Fortran " والمجال التطبيقي للبحث يشرح عمل تطبيق الأنظمة الخبيرة لادارة مشاريع التشييد حيث يدير ويحلل أداء الأنشطة التشييدية مقابل الخطة ومعرفة مسببات أي انحراف في الأداء الحالي لهذه الوحدات العملية. والاستراتيجية المستخدمة في تحليل الأداء هو استخدام " Diagnosis " Causal Analysis وبرمجته باستخدام تقنية " Object Oriented Programming ".
الخطوة الثانية التي يعمل برنامج النظام الخبير هو التنبوء بالأداء المستقبلي لهذه الأنشطة التشييدية من حيث التكلفة والوقت باستخدام استراتيجية " شبكة العصبيات الصناعية " Artificial Neural Network " ويهدف البحث في تجميع وتنظيم المعرفة المتعلقة بتحليل وتنبوء ادارة المشاريع ومناقشة المواضيع المتعلقة باستخدام هذا البحث على أرض الواقع وربطها بأنظمة ادارة المشاريع التقليدية.