An example of expert system on numerical modelling system in coastal processes

K.W. Chau and W. Chen
Department of Civil & Structural Engineering, Hong Kong Polytechnic University, Hung hom, Kowloon, Hong Kong

Abstract
With the advent of artificial intelligence technology as well as the widespread popularity of desktop microcomputers in recent years, integration of this new technology with the traditional numerical modelling system becomes a current trend in order to solve various engineering problems. It renders a more intelligent and user-friendly system on the problem domain. In this paper, a knowledge-based expert system on numerical modelling system for coastal water processes is delineated. Expert system application, as a key branch of artificial intelligence technology, is integrated with traditional numerical modelling for simulating flow and water quality phenomenon in coastal waters. The knowledge bases are classified into five major types, namely, a variety of models, relations between various model parameters and real physical conditions, feasible options of model parameters, question base as an user-interface directing the user to depict the actual physical conditions, and the rules of inference deducing the feasible choice of model and its parameters. A hybrid expert system shell, Visual Rule Studio, is employed as an ActiveX Designer under Microsoft Visual Basic environment because it combines the advantages of both production rules and object-oriented programming technology. Both forward chaining and backward chaining are used collectively during the inference process, which is mainly driven by premises and conditions with the highest factors of confidence. The inference engine will drive the decision tree to explore the most probable option of numerical model and parameters matching the real problem specifications. It is shown that the application and integration of the knowledge-based expert system technology into numerical modelling for coastal processes can provide substantial assistance to novice users for selection of numerical model as well as parameters.

Keywords
Coastal water process, expert system, knowledge base, numerical modelling, object-oriented programming technique, production rule

Introduction
In order to mimic and predict accurately the changes of flow and water quality conditions in coastal water processes, numerical modelling techniques have often been used as the simulation tool. Over the past few decades, with the advancement of numerical modelling technique, enormous amount of mathematical models have been developed and applied to various hydraulic engineering and environmental engineering problems. However, extensive and detailed expertise knowledge is required before one can distinguish the special features and limitations of these individual numerical models and then select the appropriate model to apply in a particular circumstance. It is found that these existing models always lack in some user-friendly aid for their selection, which will be very useful to novice engineers or engineering students as the design aid or training tool. Nowadays, the development of numerical modelling system has been developed to such a state that it tends to encapsulate more and more recent and advanced computer technology.
Abbott (1989), Abbott (1991) and Cunge (1989) have classified different models into various stages of computer generation. The initial models, which are developed for very specialized scope of problem, can only be used by the developer and other special users after having undergone training of long duration. Gradually the models incorporate more user-friendly tools such as menu of parameters specification, automatic grid formation, pre-processing, post-processing, management of collected field data, etc. It is intended for use by a much wider range of general audience or users to solve a much wider scope of engineering problems in different nature. The current trend is that new models often incorporate recent artificial intelligence technology into the mathematical modelling to form a single system.

Since the last decade, some researchers have started to integrate expert system technology with numerical hydraulic and environmental modelling (Blanpain and Chocat, 1994; Chau and Yang, 1992; Chau and Yang, 1993; Chau and Yang, 1994; Chau and Zhang, 1995; Knight and Petridis, 1992; Uzel et. al., 1988). These integrated systems for coastal water processes only dealt with the simplest one-dimensional modelling situation. However, even for that simplest case, the symbolic programming for the knowledge representation and selection procedure required enormous effort. The integration of expert system and two-dimensional or three-dimensional numerical modelling will be much more complicated. The basic requirement is that the system should be able to provide expert advice on selection of the most appropriate model as well as the entailed model parameters under that particular scenario. Since the numerical modelling programs have often been developed in some traditional programming languages such as Fortran, Pascal, C, etc., it is considered not cost-effective to re-write and replace these well-proven and validated programs whose development involved long hours of concerted effort. Instead, the program is often written in some embedded source codes and the integrated system has to be accompanied by a usage wizard, which provides assistance and guidance for use and direction of non-expert users. As such, the use of the recent expert system technology, through establishing the requisite knowledge bases for the problem domain as well as utilizing a proper reasoning inference engine, is found to have high potential in providing assistance for both selection and manipulation of numerical models. Up to date, research studies on the incorporation of expert system technology into manipulation of numerical modelling are, as a matter of fact, very scarce.

The key objective of this study is thus to develop an integrated system for coastal water processes, which incorporates the recent expert system technology into traditional numerical models together with some pre-processing and post-processing tools. One of the major requirements of the integrated system for numerical simulation of coastal hydrodynamic and water quality processes is to serve as design aid and training tool for novice users. It will furnish expert advice on the selection of simulation methods as well as their pertinent model parameters. In additional to the usual components in a typical expert system, namely, knowledge bases encapsulating the domain knowledge, inference engine, context, user-interface, knowledge acquisition and explanation modules, it also incorporates executable numerical models and databases. The architecture of the integrated system is shown in Figure 1.

Direct user-friendly medium of interaction should be set up between the user and the system so that the user can modify certain domain knowledge inside the knowledge base at his/her will. Output results obtained by executing various numerical models should be interpreted through the aid of post-processing graphic expedients. Based on the knowledge base, which acts as the repository of all the domain knowledge, comparison and evaluation of the
performance of different models and thus their relative advantages, applicability as well as limitations can be thoroughly explored. Aids on manipulation of numerical modelling should also be included. The system is verified and validated by application to some real prototype problems in Hong Kong coastal waters, which can, of course, be applicable to any other coastal zones in the world.

The application of the expert system technology for selection of mathematical model is presented and delineated in this paper. Many modules with different functions will be integrated into the whole system and some of the tools are linked by employing add-in tools. It is found that it is difficult, if not impossible, to incorporate all these modules into traditional standalone expert system shells. Hence the expert system shell, Visual Rule Studio (VRS), which runs as an ActiveX Designer under the windows based programming language environment Microsoft Visual Basic 6.0 (VB), is employed here. VRS is a hybrid expert system shell, which combines the advantages of both production rules and object-oriented programming paradigm. (Rule Machines Corporation, 1998) VB, being an object-oriented programming language, includes all the usual control objects of the common interface under Windows 98 or 2000 environments such as command button, textbox, option button, checkbox, etc. The user-friendliness of the Windows-based VB interface is one of the main reasons in choosing this shell. Besides, VB has the capability to link or execute external programs written in other conventional programming languages such as Fortran 90, Pascal or C. More Internet techniques are also incorporated into the system so as to pave the way for its future Internet version.

Knowledge acquisition

The physics of flow and water quality problems in coastal water processes are often represented by mathematical models, which solve numerically the simulated and discretized partial differential equations. The procedure in undertaking numerical modelling is usually divided into a few steps. The physical law or equations representing the problems is first written out and then discretized. The proper boundary conditions, coefficients of turbulence and other requisite model parameters are then chosen. The appropriate problem-solving algorithm is then adopted to solve the discretized equations subjected to the above constraints. In the above steps, the number of parameters involved in the selection procedure is 40 or so. Hence, based on these steps, a full decision tree of model selection procedure can be composed.

The domain knowledge entailed in the development of this expert system has been encoded mainly from literature review and interviews with expert numerical modellers. Experience from the experts are gleaned on a variety of aspects including the logical reasoning in making decision, relative importance of various parameters, individual comments on various mathematical models, their preference choice of models, etc. More discrepancies in the selection manners were found between well-experienced modellers and novice users than what is expected. This also reinforces the potentials of an expert system in assisting the selection of numerical model for coastal water processes.

The combination of theoretical knowledge from literature and their own practical experience in numerical modelling render the model selection made by well-experienced modellers more reasonable. They can strike a better balance between accuracy and effectiveness in model simulation. The novice users lack in the crucial practical experience and hence attempt only to follow strictly what the literatures describe. In so doing, a good balance between accuracy
and effectiveness cannot be easily achieved. As such, it is very imperative to incorporate more heuristic experience of numerical modelling experts into the expert system. As a matter of fact, it is very difficult to glean experience of various expert modellers in the world. However, the expert system can be further developed and updated through frequent usage and feedback from the users as well as through validation of personal conclusions and experience in the previous literatures on numerical modelling for coastal water processes.

In this prototype system, the involved parameters are classified into a total of seven types, namely, dimension and method, numerical scheme, grid formulation, boundary conditions, initial conditions, eddy turbulence, and miscellaneous, as detailed in Table 1. The classification method is designed so as to facilitate the query and reasoning processes in searching the knowledge bases. Besides, it also effects the knowledge compilation and representation with regard to numerical modelling for coastal water processes. The initiative to select a numerical model in solving a physical problem for coastal water processes usually comes from the user. In choosing the numerical model and its pertinent set of parameters, the key factors to be considered are the main purpose of the user in applying the numerical modelling, the real conditions of the physical problem, the previous experience of the user in numerical modelling and the capability of the user in the interpretation of model results. The first two factors, i.e. problem specification and physical conditions, have been considered in most of the previous expert system in numerical modelling. However, in this system, the experience of the user is also taken into account through soliciting his/her comments and deriving some conclusions from his/her responses. The fourth factor, i.e. the interpretation of the model results by the user, depends on the experience of the user and will become important in manipulation of numerical models, which is treated separately. Table 2 shows the classification of user specifications into these three aspects.

The selection of the most feasible numerical model becomes a matching process to relate the model parameters listed in Table 1 with the physical and other conditions specified by the user in Table 2. Hence the use of expert system technology is mainly on the establishment of the connection between Table 1 and Table 2, through knowledge bases, rule base and an appropriate reasoning inference engine. Since any numerical model incorporates a specific set of model parameters, the selection of numerical model will inevitably also involves the selection of model parameters. The whole procedure of model selection is represented by a decision tree, which is gradually filled as the selection procedure proceeds forward until the solution to the problem is accomplished.

At present, this prototype system incorporates only a few popular and commonly used numerical models. Nevertheless, the selection process based on Table 1 and Table 2 already includes all the succinct features and heuristic experience in model and parameter selection. Anyway, it demonstrates the skeleton of how the expert system technology can be integrated into numerical modelling for coastal water processes. As a future extension, more numerical models can be incorporated into the system and, in that situation, the selection process will be developed with mathematical basis. The rules and the knowledge bases are designed to be transparent such that it can be easily updated with new additional knowledge. That is in fact one of the purposes in adopting expert system approach.

One of the key issues in determining the selection of numerical models is to strike a balance between two intrinsically conflicting components, namely, accuracy and effectiveness. Some general concepts on this aspect are also considered and included into the knowledge bases. For example, it is believed that the effectiveness of finite element methods is usually lower.
than that of finite difference method. Another example is that larger spatial step or larger time step both contribute to higher effectiveness and lower accuracy. Through asking for the requirement of the user on the relative significance of accuracy and effectiveness, the system can search the rule base and recommend the most feasible selection. The knowledge-based expert system acts only as an assistant to guide the decision making process and does not automatically balance the accuracy and effectiveness. The relevant knowledge is encoded into the knowledge base with varying factor of confidence. The selection decision is designed to be driven by the pertinent rule with the highest factor of confidence. Nevertheless, the recent rates of advancement for both the speed and memory size of microcomputers are so fast that the accuracy consideration is expected to be attained relatively easily than the effectiveness aspect in the near future. The balance between accuracy and effectiveness may then be shifted. Flexibility should be allowed for modifications of the knowledge base to reflect that new balance.

Knowledge bases

The knowledge base of the expert system is divided into a few sub-bases, namely, relation base, selection base, question base, rule base and model base. The relation base contains the knowledge concerning all the factors related to each parameter, which can be represented in the form of relation tree. Figure 2 shows part of the relation tree for model parameters. Although, in general, the tree nodes are usually represented in one layer, they can also be represented in two layers if the related factor is related to some other factors as well.

The selection base stores the information regarding the possible options for each parameter and can be represented under the structure of a selection tree. Figure 3 displays part of the selection tree for model parameters. Each node of the selection tree represents an individual parameter and the tree branches represent available options of that parameter. Hence the relation tree and the selection tree are designed to assist in the knowledge representation of the parameter relations and options respectively. During the query process, all the related factors are determined by searching the relation tree. In the decision making process, all the feasible options of each parameter are determined from the selection tree. In fact, most expertise knowledge can be represented in the form of a knowledge tree, which is a static tree with its structure defined before the model selection.

Questions used to direct the user to input the problem specifications are grouped together into the question base. The questions are not static and the ensuing questions will depend on the answers made by the user on the previous questions. The relation base is involved to provide the required relevant parameters in order of priority. Figure 4 gives an example of the inference process of an individual parameter to be specified by the user. The queries together with the possible set of answers are derived. Once the user makes a response on the query, it is feedback to the system and the inference engine will drive the selection decision process forward until the final solution is attained. Figure 5 gives a typical example of the inference direction from a specification by the user to the inference by the system.

The knowledge representation of the rule base is intrinsically different from those of the relation base or selection base. It encapsulates the whole set of inference production rules linking the specifications made by the user and the selection of parameters. The following example gives a typical production rule, which incorporates the fuzzy description:

RULE to determine dimensions of model: 2 of 8
IF the water depth is deep AND the density stratification in the vertical direction is not significant
THEN the 2-dimensional horizontal numerical model is selected with a confidence factor of 70

The IF statement of the above rule statement describes the premises or conditions that the water depth and the density stratification should be deep and not significantly respectively for the conclusion to be fired. The THEN statement of the rule gives the conclusion that the 2-dimensional horizontal numerical model is selected with a confidence factor 70. There are 300 or so rule statements with fuzzy relations in the rule base of the system. The confidence factor is employed as the determining factor to control the inference process and the selection of each parameter. The range of the factor is basically from 0 to 100 representing the degree of confidence with which the statement is known. The rule base is designed to provide the link between the specifications made by the user and the recommended parameter selection by matching the highest confidence factor. In the above example rule statement, the water depth and the vertical density difference are expressed as a fuzzy description. As an alternative method, the user can also enter their exact numerical values during the query process. The system can then transform the numerical values into the corresponding fuzzy description by a fuzzy member curve, which computes the pertinent confidence of membership prior to searching the rule base for conclusions. Figure 6 shows the fuzzy description of water depth, in which curves A, B and C depict the fuzzy definitions of “very deep”, “deep” and “shallow” respectively.

The model base contains a number of popular and well-proven numerical models for simulating coastal water processes. Typical examples of case studies around the world employing numerical models are stored here and can be retrieved easily for study. As such, it effectively gives some references for numerical modelling with prototype examples. When the user completes the selection process of all modelling parameters in a particular case, the system will then attempt to match the models in the model base. If an existing model in the model base is found to have similar characteristics to the specified selection process by the user, the user will be prompted whether or not to adopt the selection based on the existing model.

Inference engine

The confidence factor is the key determinant used by the inference engine to drive the selection of various parameters. A default value of 50 is set initially for the confidence factors of all parameter selection, which represents a half-and-half chance for it to proceed successfully. When the whole picture of specifications is gradually completed by the user through the query process, the inference process continues and the new confidence factors of various parameter selection are determined through the matching process. In general, a few selection options will have confidence factors greater than 50. In this case, the system will recommend the user to opt for the one with the largest value of confidence factor. However, since the questions asked by the system are dynamic and the descriptions specified by the user are varying, it is difficult to ensure that a solution exists in every case. If no rule can be matched with the user specifications or no higher confidence factor can be obtained, the selection procedure will not be triggered.

The confidence factor of each attribute’s value is recorded and maintained as part of the state of the attribute in the session context. The complete state of an attribute consists of its actual
value as well as its confidence factor. Confidence factors determined as a result of the firing of a rule are based on a product space calculation. The confidence factor of an antecedent preceded by NOT is 100 minus the confidence of the non-negated antecedent. When an antecedent of a rule is evaluated, the confidence resulting from the evaluation of the antecedent is first determined by comparing the minimum confidence of all antecedents joined by AND with the maximum confidence of all antecedents joined by OR. A single confidence value of the antecedents is determined to be the lower of the above two confidence factors. The conclusion of the rule is given a final confidence that is calculated from the single confidence factor of the antecedents and the confidence specified in the knowledge base for the conclusion. The product of this multiplication is normalized by dividing with 100, and the result is then assigned as the final confidence in the session context.

Thus, the inference engine is designed as a set of codes to search and match for a way, which can lead to increasing confidence factors of parameter selection. The user is also allowed to execute some existing models employing the default parameter values. Of course, the existing model may not be the most appropriate in this particular instance. If the actual physical data differs too much from the bathymetry of the existing model, a low confidence factor of the model selection will be acquired. If the user can provide the specifications in greater details, there is a higher possibility for the inference engine to match the parameter selection with higher confidence factor and the resulting suggestions become much trustworthy. The main purpose of the query process is thus to glean the purpose of the user and the physical conditions of the problem so as to increase the confidence of the model selection. Hence, if it is anticipated that any questions or answers will have very small contribution to the confidence of parameter selection, it may not enter into the interrogative dialogue interface. As mentioned earlier, each parameter is related to some other factors through the relation base. However, several physical conditions are usually involved in order to determine each of these related parameters. As a result, the system will need to ask an enormous amount of questions for the users to fill in prior to completely specify the particular case. The mechanism is designed such that the system will simulate the possible conditions and determine the possible largest change in the confidence of selection through matching with the rule base before the query process. Questions with higher potential to contribute to the confidence of selection will be asked with a higher priority.

The concept of decision tree is employed as a pre-set route to drive forward the process of the inference engine. It is in stark contrast with the static knowledge tree mentioned earlier for the relation base and the selection base although it is also represented in tree form. The key differences are that decision tree is dynamic and the structure of the tree branches will change according to the specifications provided by the user. Here the set of parameters is expressed under the structure of a decision tree. The procedure of parameter selection then involves searching through the decision tree and filling in all the tree branches after querying the user to specify myriad conditions. Hence the size of the decision tree depends on the complexity of the problem. As the query process proceeds and specifications by the user to some conditions have been added, the inference engine searches the rule base and computes the new confidence factors for all branches of the decision tree.

It is designed that the classification method will cover all the involving parameters as far as possible in numerical modelling of coastal water processes. According to the present classification method, it is inevitable that conflicts may exist between parameters when some parameters preclude the existence of another parameter. For example, if a 2-dimensional
model is selected, the vertical grid spacing will no longer be used. In order to eliminate these conflicts in the expert system, the rule base is designed carefully to filter the required parameters in the decision tree. The most improbable branches of selection, defined as those with confidence factor less than 20, are cut away from the decision tree. Thus the decision tree can also be considered as the guide in a tree skeleton format, directing the inference engine to query the user as well as to make decision based on the responses. This function is provided by the control object “TreeView” under the programming environment Visual Basic 6.0. It has many useful features of tree adding, tree cutting, tree searching, tree shape changing, and tree display on the user-interface.

Two types of inferencing strategy, namely, forward chaining and backward chaining, are often used in an expert system to control the matching between the facts and the rules. In this system, forward chaining is employed to search from responses made to the query by the user and to modify the decision tree whilst backward chaining is used to find the requisite parameters and then to determine the query to the user. Figure 7 shows the cyclic relationship amongst the inference engine and the various knowledge bases. The inference engine first searches the decision tree to determine the parameter, which will have highest potential confidence to model selection. The relation base is then involved to find the problem specifications or physical conditions related to that particular parameter. The question base is then searched to prompt the query together with the possible answers to the user to glean the problem specifications. The rule base then enter the scene to match the parameter and the specification, to cut away the impossible tree branches of the decision tree, to compute the confidence factors for a variety of selection options, and to recommend the one with the highest confidence factor. The new decision tree as well as the newly selected parameter is recorded. Another cycle of inferencing process is then ready to run again, until the whole decision tree is filled up.

**Conclusions**

Through the successful development of this prototype system, it has been proved that the expert system technology can be integrated into the numerical modelling for simulating flow and water quality phenomenon in coastal waters in order to provide assistance on the selection of model and its pertinent parameters. The integration renders a more intelligent and user-friendly system in the problem domain. The use of the hybrid expert system shell, Visual Rule Studio, which runs together with Microsoft Visual Basic 6.0 is found to be very effective in producing the system under the popular Windows environment. It combines the advantages of both production rules and object-oriented programming paradigm. Of course, expertise knowledge on more extensive types of numerical models can be added to enrich further this prototype system. Nevertheless, the existing architecture of the expert system has shown that it is useful to achieve the above objective. Further research effort will be spent on the representation of the more difficult encoded dynamic knowledge, knowledge elicitation of expertise questionnaire through World Wide Web, verification and validation of effectiveness of artificial intelligence technology in aiding numerical modelling.

**Acknowledgements**

The work described in this paper was substantially supported by a grant from the Research Grant Council of the Hong Kong Special Administrative Region (Project No. PolyU5084/97E).

**References:**
### Table 1. The classification of parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Dimension and Method</th>
<th>Numerical Scheme</th>
<th>Grid Formulation</th>
<th>Boundary Conditions</th>
<th>Initial Conditions</th>
<th>Eddy Turbulence</th>
<th>Miscellaneous</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
<td>Dimension Numerical method Co-ordinate system Numerical method in vertical direction Vertical coordinate system</td>
<td>Error of scheme Stability Explicit or implicit Alternating direction algorithm Advection term</td>
<td>Node number per element Shape of grid Horizontal grid spacing Vertical grid spacing Uniform or non-uniform grid Type of point setting</td>
<td>Tidal constituents at open boundary Discharge at open boundary Water quality variables at open boundary Variation in open boundary Variation in close boundary High order variation in boundary</td>
<td>Initial conditions of elevation Initial conditions of current Initial conditions of water quality</td>
<td>Turbulence model Vertical eddy diffusion Horizontal eddy diffusion Vertical eddy viscosity Horizontal eddy viscosity</td>
<td>Coliolis force Geometry Bathymetry Wind at surface Bottom drag coefficient</td>
</tr>
</tbody>
</table>

### Table 2. Classification of the user specifications

<table>
<thead>
<tr>
<th>Purpose of study</th>
<th>Physical characteristics</th>
<th>Previous experience of user</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tidal propagation</td>
<td>Spatial scale</td>
<td>Grid formulation</td>
</tr>
<tr>
<td>Water quality study</td>
<td>Time scale</td>
<td>Accuracy of scheme</td>
</tr>
<tr>
<td>Salt water intrusion</td>
<td>Water depth</td>
<td>Effectiveness of scheme</td>
</tr>
<tr>
<td>Storm surge</td>
<td>Complexity of bathymetry</td>
<td>Stability of scheme</td>
</tr>
<tr>
<td>Hot water discharge</td>
<td>Ratio of tidal elevation to depth</td>
<td>Initial condition</td>
</tr>
<tr>
<td>Sediment transport</td>
<td>Longitude and latitude of location</td>
<td>Interaction of boundary conditions if adjacent to each other</td>
</tr>
<tr>
<td>Seasonal structure of current in coastal zone</td>
<td>Existence of river</td>
<td>Representation of the driving forces</td>
</tr>
<tr>
<td>Water exchange between water bodies</td>
<td>Vertical variation of density</td>
<td>Boundary at surface and bottom</td>
</tr>
<tr>
<td>Research on eddy turbulence</td>
<td>Vertical variation of current</td>
<td>Open boundary at river end and ocean end</td>
</tr>
<tr>
<td>Waste water discharge</td>
<td>Vertical mixing</td>
<td>Turbulence model and vertical variation</td>
</tr>
</tbody>
</table>
Figure 1 Architecture of the expert system
Figure 2. Part of the relation tree of model parameters
Figure 3. Part of the selection tree of model parameters
Figure 4. An example of the inference direction from a parameter to questions and answers
Figure 5. An example of the inference direction from the user’s specifications through the inference engine
Figure 6. Fuzzy description of the water depth (A---very deep) (B---deep) (C---shallow)

Figure 7. The inter-relationships amongst the inference engine and various knowledge bases