United States
Environmental Protection
Agency

Research and Development

Risk Reduction
Engineering Laboratory
Cincinnati, OH 45268

EPA/600/SR-92/218 January 1993

EPA

Project Summary

Use of Expert Systems in a Water Utility

Richard M. Males, Walter M. Grayman, Beth Hertz, Harry Borchers, David Milan, and Judith S. Coyle

A 3-yr cooperative agreement between the North Penn Water Authority (NPWA), Lansdale PA, and the US EPA Drinking Water Research Division to study use of expert systems technology in a water utility, has resulted in the development of two expert systems that demonstrate basic principals that should be broadly applicable to other water utilities. A "customer query expert system" assists non-technical users in handling customer questions relating to water quality. A "pump efficiency watcher" expert system is designed to flag long-term changes in well pump performance, and to calculate a number of pump performance parameters of interest.

The utility of expert system capabilities was found to be of interest primarily as a part of larger systems with broader capabilities desired by the user. For example, the customer query system was desired as a method of automatically logging customer calls, preparing work orders, looking up customers in a database, and sending appropriate form letters. Similarly, the expertise embodied in the pump efficiency watcher was part of a larger effort, providing reports, charts, and statistical analyses of pump efficiency.

The results clearly demonstrated that expert systems that are potentially useful to water utilities can be developed at moderate cost. Significant effort, however, should be expected to move from prototype "proof of concept" expert system demonstrations, to fully-featured, field-usable products.

This Project Summary was developed by EPA's Risk Reduction Engineering Laboratory, Cincinnati, OH, to announce key findings of the research project that is fully documented in a separate report of the same title (see Project Report ordering information at back).

Introduction

In the late 1980s, much attention was given to the promise of "expert system technology." Expert systems is a branch of artificial intelligence in which the knowledge and experience of a human expert in a particular field is first obtained (through interviewing and observation), then structured into a set of rules, and finally placed within the framework of a computer program. This program can then be "consulted" to provide diagnoses, recommendations, or advice, in theory much as the human expert would be used. Expert systems technology had shown some promise, more so than other applications of artificial intelligence. A key tenet of the expert systems approach is that the focus of the effort is on a narrow, restricted, "knowledge domain."

EPA's Drinking Water Research Division (DWRD) has had a longstanding association with the North Penn Water Authority (NPWA), the water service utility in Lansdale, PA, through a series of technical cooperative agreements relating to air stripping of volatile organics from well water and studies of modeling of water quality in water distribution systems. In view of the joint interest on the part of NPWA and DWRD in exploring the technology of expert systems as applied to a drinking water utility, a cooperative agreement was initiated in May of 1988. The project was originally intended to be a 2-yr effort and was funded at a total cost of $73,867. A 1-yr, no-cost extension was later added.

The North Penn Water Authority serves approximately 18,000 residential, commercial, and industrial customers in the Montgomery County area of Pennsylvania, north of Philadelphia. Water is supplied from some 50 wells extending into a fractured-fissured rock aquifer, and from treated river water, purchased from the American Water Works Service Company. Another surface water source, the Forest Park Treatment Plant, using Neshaminy Creek water, has recently been developed and is expected to be increasingly important over time. Depending upon the time of year and demand, water supplied to a particular location may be river water, well water, Forest Park water, or a mixture of these. The well water has a high degree of hardness, and the river water has seasonal odor problems resulting from algal blooms. Quality problems vary throughout the area due to water mixing and varying quality of water available from wells. The nature of the aquifer makes precise knowledge of the groundwater flow situation difficult.

The initial project objective was to explore the utility of expert systems techniques coupled with simulation models of water quality in distribution systems. The original hypothesis was that utility-resident expertise on particular contaminants, coupled with the modeling capability to determine flow directions, would allow for a simple method of answering such questions as "where is the likely source of this contaminant?" This would be an example of "spatial reasoning," i.e., using the spatial knowledge resident in the system network model to provide locational information. This objective was later modified to a broader examination of expert systems usage, as further study proved that the modeling approach would not be fruitful in the NPWA context.

Procedure

The general procedure for development of an expert system involves appropriate selection of a problem, identification of an expert, knowledge acquisition relating to the particular expertise, codification of that expertise, implementation of the codified expertise in a computer program, and numerous steps of testing and revision of the program.

Following a review of the literature concerning general concepts related to expert systems and specific applications involving spatial reasoning, a small problem, of interest to NPWA personnel, and having important spatial and temporal characteristics, was selected to serve as a prototype and proof-of-concept. After experience was gained through the prototype, additional problem areas were investigated.

Six general problem areas were examined in the course of the study: handling of customer queries (the "prototype" problem); examination of the "sources of contaminants" problem; examination of short-term and long-term operational knowledge issues; examination of long-term changes in pump efficiency; efficient use of the groundwater resource; long-term trends in water quality.

Expert systems were implemented in association with the customer query and pump efficiency problems. A conceptual exploration of the use of expert systems for examining the groundwater reservoir was carried out, but project resources did not allow for implementation of an expert system itself. The "sources of contaminants" problem did not prove worthwhile as the basis for an expert system, nor did the short- and long-term operational knowledge issues, but the latter, coupled with the long-term trends in water quality issue, led to the development of the "watcher" concept, in which an analysis and expertise component "watches" a data stream from a laboratory information management system (LIMS) or SCADA system, looking for long-term changes in the data.

During the course of the study, a number of problems were identified and associated knowledge acquisition carried out. During the course of the project, knowledge acquisition sessions were routinely tape-recorded and transcribed into a word processor, providing a computer-readable record. This computer-readable record was then used in two fashions: 1) to aid in organizing and structuring the information through use of an "information manager" computer program that allowed for rapid recall from the transcript of all text that referenced particular items of interest; 2) as input to a computer program that analyzed participant contributions to the discussion over time (through word counts for each individual), in order to examine the process of knowledge acquisition itself. These two methods both proved extremely interesting, and although exploration of them was limited within the project, they are seen as valuable approaches for further exploration.

The initial approach to expert system development was through the use of an "expert system shell." An expert system shell is a "generic" expert system, much as a database management system package is a generic data management system. For the expert system shell, the user provides a database embodying the rules that specify the expertise being represented. The shell then handles all of the necessary interaction with the user and provides inference engine capabilities necessary to use the specific knowledge presented to it. Such shells are extremely popular and can be acquired at relatively low cost. As the project developed, however, the limited capabilities of the selected shell to handle some of the necessary functions (user interface, database lookup), dictated a move to an alternative method, specifically, the use of the C programming language coupled with a set of library routines to handle the desired inferencing from the knowledge base. Abandonment of the shell allowed for the development of much more powerful and "user-friendly" programs. The shell itself was useful for initial development and prototyping of the rule base but was not adequate for the other demands placed on it. It should be noted that many other researchers using expert system shells have recently made similar moves to programming languages.

Results and Discussion

Customer Query Problem

The handling of customer queries relating to water quality was selected by NPWA as the prototype problem, as a real project of significant interest to them. The customer query problem was selected because: it is a real problem of significance to the utility; specific expertise is required to handle the problem, and an expert was available; the problem is interesting, reasonable in size, specific, and is not open-ended; the result, while specific to NPWA, would have generic aspects common to all water utilities; there are spatial and temporal aspects to the problem.

Initially, a series of customer calls were taped and transcribed, simply to provide a "feel" for the nature of the questions and answers supplied. This was followed by a full-day, formal knowledge acquisition session in January of 1989, in which the problem was categorized, and specific responses were identified. The tapes of this meeting were also transcribed in computer-readable format. This information was entered into an "information manager" computer program that allows for selective retrievals and structuring of text information. For example, using this software, it was possible to retrieve all references in the transcript to "brown water" discussed simultaneously with "odor."

The problem was seen to be amenable to a "rule-based" expert system framework, and a set of "if-then" rules that describe the response to particular inquiries was generated. A low-cost expert system shell, VP-Expert* (a product of Paperback Software), was selected for the initial implementation. The prototype system involved some 44 rules and was delivered to NPWA for testing in August of 1989.

As the system development evolved, a number of additional features were requested in the system design, such as the ability to print work orders and form letters and to log calls to a database of previous callers. The initial prototype lacked some of these desired features, in particular the capability to look up a customer in an external database. Other features were implemented in only a rudimentary fashion, e.g., the printing of work orders and form letters, in order to demonstrate the feasibility and method of implementation.

The prototype system was installed and demonstrated at NPWA with the intent that it be verified in use. After a short test period in which the logic apparently performed adequately, the system was set aside as other priorities and personnel changes took place at NPWA. Following these changes, testing resumed with results available in February of 1990. The system performed adequately on most of the problems with only minor logic flaws relating to the handling of odor complaints. An additional category of problem, sickness, or skin irritation was requested, and a number of additional "usability" features were requested (the ability to clear a previous entry, the ability to recall data previously entered, linkage with customer account numbers, etc.). Also desired was the ability to store information about NPWA's response to the problem. Thus, the expert system was clearly being viewed not solely as a consultant/adviser technology, but also as a database and full-functioned system.

While the customer query system was originally envisioned as a prototype and proof of concept and was not expected to be an installed system, interest on the part of NPWA in the system remained high, and a portion of project effort was devoted towards attempting to turn the prototype into an actual working system. Significant efforts were required to move to a working system. As noted previously, a key desired feature of the system was the ability to handle databases of customers and previous callers. After an extensive period of time attempting to implement this capability within the VP-Expert shell, the shell technology was abandoned in favor of a custom-written C program combined with function libraries for user interface, database manipulation, and expert system inference. This approach had been avoided previously, due to the desire to keep a system that, ostensibly, could be revised and maintained without the need for programmers. However, the impossibility of developing an adequate system within the shell eventually dictated the change to the programming language. In discussions with other developers of expert systems in the water field, a similar evolution has taken place, with abandonment of the shells in favor of computer language-based approaches to system development.

Once the move to a programming language was made, it became possible to accomplish the desired modularization of functionality for the system. In particular, a simple menu structure and data entry mechanism could be created, allowing the user to acquire information from the caller, search a customer database and a database of previous callers, select form letter information to be printed, and log the call and response into the previous caller database. The expert system functionality became a module, activated by simply picking a menu item. This results in display of the expert recommendations which can then be changed by the user and printed in a work order.

The expert system rules developed under the VP-Expert shell were readily translated into the syntax required by the C library inference engine, as both are rule-based systems. Thus, the effort that went into the rule development using the shell was not at all wasted, and the use of such shells is still seen as an extremely valuable method for rapid prototyping and development of the rule base.

Sources of Contaminants Problem

In August of 1989, an additional knowledge acquisition session on quality issues was conducted. This session focussed on determination of sources of contaminants identified in system sampling. For the most part, when contaminants are detected they are found in source water sampling, thus the source is known. Thus, it was revealed that modeling would be of little interest in answering these kinds of problems. NPWA personnel did, however, express a desire for some method of examining flow direction graphically, in order to have some idea of the possible sources of brown water problems. As a result, an interactive Turbo Pascal program was developed to display results of a previously-developed hydraulic model for NPWA. The program allows the user to graphically designate a source, and display flow paths to and from the source, time of travel, and concentrations at other nodes of water from the source, as well as to display information results from the hydraulic model.

Operational Knowledge Problem

In June of 1990, a detailed knowledge acquisition session was conducted, examining the issues of operations with a view towards developing the expert system-modelling link originally envisioned for the project. The results of the knowledge acquisition indicated that there are two levels of "knowledge" about system operation of the NPWA distribution system. One level is the "day-to-day" level, involving turning specific pumps on and off to maintain tank levels and adequate supply. The second level is a longer-term orientation, that looks at the day-to-day operation to assess impact on total system cost, efficiency, and availability of water over longer periods. The day-to-day knowledge is embodied in the individual who actually operates the distribution system, while the longer-term knowledge is that of the Executive Director of NPWA. This knowledge acquisition did not go into sufficient detail on either of these two levels to allow definition of specific expert system rules for operations at NPWA; rather, it focused on the general character of the system operational knowledge.

NPWA is in the process of acquiring a Supervisory Control and Data Acquisition (SCADA) system for remote, quasi-automated operation of the utility. The introduction of the SCADA system, in essence, must embody the day-to-day knowledge, as the control programs for the SCADA system must have similar capabilities of starting and stopping pumps based on time of day, tank level, and other factors. The SCADA system also generates a significant amount of information on system operation and conditions that can be logged to a central computer site.

The original intent of the project, prior to this knowledge acquisition session, was to explore the integration of analytical hydraulic network models in conjunction with an expert system relating to day-to-day operations. Based on the knowledge ac-

* Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

quisition, however, it was clear that the day-to-day operational decisions are not (and would not be) based on modeling results, but rather are based on measured readings, in particular the well water levels and the levels in the tanks, and general policies and practices relating to which wells to use based on quality, and which source to use, based on cost and availability. The SCADA system will provide additional telemetry and perform much of the basic day-to-day operations (turning pumps on and off). The building of "expertise" into such a system beyond that normally done by defining set-points for the controllers is a possibility, i.e., the SCADA system could be designed to consult an expert system. This is apparently on the edge of current-day technology for SCADA systems.

While the SCADA system (without an attached expert system) will be designed to handle the day-to-day operational issues, it will not be optimizing at the "long-term" level, i.e., making decisions based on availability of water, total system cost, etc. The knowledge acquisition clearly indicated a desire for some examination, using expert systems, of the longer-term issues, in particular in determining large-scale water management issues [How much water is available? Where should water be extracted (the river, purchased, or groundwater)? Will there be expected water shortages in the next few months? etc.].

The Watcher Concept

The desire for longer-term examination of data resident in the proposed SCADA system led to the development of the "watcher" concept. Watchers are combined expert and analysis systems that examine the data stream and historical data available in automated systems such as SCADA and Laboratory Information Management Systems (LIMS). Such systems may be inherently capable of flagging situations that are outside of pre-determined limits, but generally do not have features for examination of long-term trends, particularly those requiring some longitudinal analysis of the data.

These "watchers" would consist of a set of expert rules, coupled with data analysis and display mechanisms, that would periodically examine the historical databases for operations and quality, looking for long-term trends and consequences. The watchers at present do not propose what to do, but rather indicate that a problem or long-term implication of the current situation exists. Where an automated data stream does not yet exist, as for the yet-to-be implemented SCADA system, the watcher can still be created, but will rather operate on historical, rather than real-time data.

Three "watchers" were conceived in the course of the project: 1) a water quality watcher; 2) a pump efficiency watcher; and 3) a groundwater watcher.

The water quality watcher operates upon the NPWA LIMS database, to track locations where repeated coliform variations exist. The quality watcher was implemented simply as a command file in the R:Base database management system in which the NPWA LIMS is constructed, which is run periodically to scan for these locations. No expert system rules were involved.

The pump efficiency watcher examines wire-to-water efficiency in pumps through analysis of pumpage and electricity usage data. The pump efficiency watcher was implemented as a combination C program and set of expert system rules. The C program provides an excellent user interface, analysis capabilities, data entry, and graphics display. The expert rules examine statistics generated by the program for the wells, and assess, based on these statistics, the likelihood of a problem in well efficiency in the near, intermediate, and long run. The program is general purpose in nature, and detailed documentation is provided.

The groundwater watcher deals with the long-term issues of groundwater management. Project resources did not permit full implementation of the groundwater watcher. Initial steps involved statistical examination of the historical record and experimentation with a groundwater model. A "concept paper" describing the manner in which such a groundwater watcher could be developed and applied was prepared.

Conclusions

The results of the overall research indicate that expert systems technology can be used effectively in various areas of a water utility, although perhaps some of the early claims for the technology are overrated. While the conclusions are based on a limited sample (a single water utility, and the customer query system and pump efficiency watcher as the most developed of the systems), one thing that seems clear is that expert systems technology serves best when embedded in a larger system that provides other user-desired functions (i.e., consulting expertise is not the exclusive goal, but is rather one of the functions of the total system).

The use of an expert system shell to perform initial development of the rule base was worthwhile, although there is clearly a learning curve associated with using expert system technology even with a simple shell. The desire for additional functions beyond the consultation function eventually dictated the move to a programming language. The programming language approach allows for a good deal of modularity to be introduced into the eventual product, something not so easily handled with the shell structure. In particular, programming languages allow for "procedural solutions," i.e., first do this, then do this, then do this. Using the techniques resident in the two shell systems explored as part of this project, ordering the flow of control is much more difficult. Further, the programming language approach, while requiring more specialized skills to use, is more transportable, more flexible, and more efficient.

In terms of the overall process of developing an expert system application, a number of points are clear:

- Reasonable expert systems applications can be developed at moderate funding levels, given certain preconditions. There is clearly a learning curve associated with expert systems technology and its application.
- Prototype systems, demonstrating the concept and validity of the rules, are significantly easier to develop than full-scale, working systems, where issues relating to suitability of the functionality, and ease of user interface become significant.
- Confinement of the problem is essential.
- There is a high need for trust and a high need for familiarity with the problem in the knowledge acquisition process.
- Programming and systems analysis skills are clearly needed in the development of an expert system.
- The development of such systems is clearly iterative.
- Computerized management of the transcript is valuable, and computerized analysis of knowledge acquisition transcripts shows promise for providing insight into the knowledge acquisition process itself.

Recommendations

While the work at NPWA is clearly site-specific, it does indicate that expert systems technology clearly has a role in various applications. That is, the technology is not merely "hype." Nonetheless, small projects are more likely to yield success (or at least smaller failures) than ambitious projects. Projects that embed the expert system technology as a consultative function within a larger framework are more likely to be of value and interest.

Sharing of experience relative to expert systems technology within the water utility industry will lead to fewer individuals having to repeat the same learning curve and mistakes in this arena. The distinction between prototype systems and true fielded systems must be made clear in this regard.

The process of knowledge acquisition itself, using computerized analysis and information structuring tools on computer-readable transcripts of knowledge acquisition sessions, appears to be a fruitful area for further study.

Expert systems shells are valuable as rapid prototyping mechanisms, but excessive effort should not be devoted to attempting to implement full systems within shells.

The full report was submitted in partial fulfillment of Contract No. CR-814918-01-1 by the North Penn Water Authority, under the sponsorship of the U.S. Environmental Protection Agency.

*U.S. Government Printing Office: 1993 — 750-071/60160

Richard M. Males is with RMM Technical Services, Inc., Cincinnati, OH, 45208. Walter M. Grayman is with Walter M. Grayman, Consulting Engineer, Cincinnati, OH 45229. Beth Hertz, Harry Borchers, and David Milar are with the North Penn Water Authority, Lansdale, PA, 19446. Judith A. Coyle is an independent consultant in Ballwin, MO 63021.
***Jeffrey Adams** is the EPA Project Officer (see below).*
The complete report, entitled "Use of Expert Systems in a Water Utility," (Order No. PB93-123081/AS; Cost: $19.50, subject to change) will be available only from:
National Technical Information Service
5285 Port Royal Road
Springfield, VA 22161
Telephone: 703-487-4650
The EPA Project Officer can be contacted at:
Risk Reduction Engineering Laboratory
U.S. Environmental Protection Agency
Cincinnati, OH 45268

United States
Environmental Protection Agency
Center for Environmental Research Information
Cincinnati, OH 45268

BULK RATE
POSTAGE & FEES PAID
EPA
PERMIT No. G-35

Official Business
Penalty for Private Use
$300

EPA/600/SR-92/218