Archives of Business Research – Vol.7, No.7
Publication Date: July. 25, 2019
DOI: 10.14738/abr.77.6777.

Zhang, X., Moynihan, G., Ernest, A., & Gutenson, J. (2019). Application of Expert System Technology for the Decontamination of Water Distribution Networks. *Archives of Business Research, 7(7),* 269-279.

Application of Expert System Technology for the Decontamination of Water Distribution Networks

Xiaoyin Zhang
Department of Civil, Construction & Environmental Engineering,
The University of Alabama, Tuscaloosa, Alabama, United States

Gary Moynihan
Department of Civil, Construction & Environmental Engineering,
The University of Alabama, Tuscaloosa, Alabama, United States

Andrew Ernest
Department of Civil Engineering, The University of Texas –
Rio Grande Valley, Brownsville, Texas, United States

Joseph Gutenson
Hydrologic Systems Branch Coastal and Hydraulics Laboratory,
U.S. Army Engineer Research and Development Center, Vicksburg, Mississippi, United States

ABSTRACT

Decontamination of water distribution networks (WDNs) is a difficult process to conduct. Creation of an effective approach necessitates integrating rules and requirements from diverse knowledge domains in such a way that the operational goals are achieved with minimally available situational information. To date, there has been a limited amount of work in applying expert systems in this problem domain. This research 1) identifies and assimilates the knowledge necessary for WDN decontamination; and 2) evaluates the relative benefits of forward and backward chaining inferential logic in WDN decontamination. Based on the results of this analysis, we developed a backward-chaining prototype expert system for WDN decontamination. The system provides reasoning routines and recommendations on the type of event and consequences on the water operator's clients, the public in general, the environment, and the potential hazards from the resulting chemical interactions.

Keywords: Backward chaining; Expert system; Decontamination; Water distribution network; Drinking water; Python.

INTRODUCTION

The United Nations International Strategy for Disaster Reduction defines a disaster as: "a serious disruption of the functioning of a community or a society involving widespread human, material, economic or environmental losses and impacts, which exceeds the ability of the affected community or society to cope using its own resources" [1]. Disasters can be categorized into three main groups as summarized by Shaluf [2]: natural disasters, manmade disasters, and hybrid disasters (e.g. deforestation causing soil erosion leading to landslides). The safety of water distribution networks (WDNs) has long been a concern since water can be polluted by a variety of these natural or manmade disasters. Decontamination of these WDNs is difficult due to both their complexity and the characteristics of the specific disaster event. After a contamination event takes place, human experts and decision makers must process knowledge from a variety of different fields to find an efficient and effective solution in a short time as well as assimilate a large volume of information regarding the impacted WDN. Issues that impact assimilation of WDN information include: 1) lack of a well-defined policy framework; 2) difficulty in information collection; 3) variability of WDN topology; 4) aging of WDN; 5) shortage of human experts with sufficient empirical knowledge; 6) loss of access to the tacit and undocumented expertise once the specialists leave or retire; 7) inconsistency of heuristic experience from different human experts; 8) limitation of expertise in the distinct areas [3]. Frequently, the domain expert may only rely on personal heuristics, intuition, and experience.

Decision making during a WDN decontamination event is a potentially complex problem where expertise is highly valued. These decision making procedures usually begin with limited initial information, and must be able to adapt and evolve as further information becomes available. A logical next step is to develop a computer-based tool to assist in this decision making process.

BACKGROUND STUDIES

The Critical Infrastructure Partnership Advisory Council (CIPAC) Water Sector Decontamination Working Group identified and prioritize 16 key decontamination issues and 35 corresponding recommendations in their 2008 report "Recommendations and Proposed Strategic Plan: Water Sector Decontamination Priorities" (referred to as the CIPAC 2008 report). Utilizing these recommendations, we investigated: Priority Issue 4: Decision-making frameworks for decontamination; Priority Issue 7: Utility communications to public officials, responders, the public and others on decontamination; and Priority Issue 9: Treatment procedures of contaminated drinking water and wastewater [4] to define the project scope and problem domain.

An expert system is a computer-based system that can emulate a human problem solver by applying knowledge and reasoning normally known and used by experts in that specific field. They represent an increasingly practical tool for a variety of functions [e.g. 5]. The knowledgebase, inference engine, interface, and support environment represent the components of a traditional expert system [6]. The knowledgebase is the repository for the problem-specific heuristics. These heuristics are frequently obtained from a human domain expert, then structured and input by the knowledge engineer through the system's interface and support environment [6]. Feigenbaum [7], often considered to be the "Father of Expert Systems", asserted that expert systems gain their power from the specific knowledge they process, rather than from any one particular scheme or formalism.

Recently, a considerable amount of research utilizing AI or expert system technology has been applied to the field of water management. Examples include: 1) irrigation system network management [8]; 2) minimization of water leakage in the pipe networks [9]; and 3) prediction of drinking water distribution system pipe breaks [10] However, within the water management research community, there is lack of expert system research specifically targeting WDN decontamination.

RESEARCH OBJECTIVE AND METHODOLOGY

The objective of this project was to investigate the application of expert system technology to the WDN decontamination problem domain through the development of a prototype system. In addition, this study will: 1) identify and assimilate the knowledge necessary to WDN decontamination; 2) evaluate the relative benefits of forward and backward chaining inferential logic relative to WDN decontamination.

Accepted expert system methodology identifies three primary steps: knowledge engineering, system development, and system verification and validation [5]. During the initial knowledge engineering phase, the key concepts, relationships and heuristics are identified. An extensive literature review was carried out via both internet and manual searches. Decontaminating water distribution networks requires the application of knowledge from a diverse range of knowledge domains including health impact, utility operations, chemical characteristics, and hydraulics. These include identifying possible contaminants and taxonomies [11], EPA drinking water regulations [12], pipe material characteristics [13], pipe material – contaminant interaction [14], and treatment technologies [15].

During the development phase, the acquired knowledge was organized. After identifying the knowledge necessary for WDN decontamination from the literature review, we pursued a process of assimilation into machine readable formats in order to complete the remaining project elements. The selection of a rule-based system was felt appropriate after considering other schemas. Inference rules and objects were constructed and formalized.

Goal Representation

Based on the information obtained from the CIPAC 2008 report [4], we identified our system goals as: 1) Goal Exceedances (G_E), to determine whether the concentration of contaminant exceeds the permissible limit; 2) Goal Warnings (G_W), to identify whether public health and/or environment is in danger; 3) Goal Interaction (G_I), to indicate whether the contaminant harms the WDN infrastructure, e.g. degrading the pipe; 4) Goal Treatment Technologies (G_T), to suggest the potential treatment technologies for decontamination. These four goals are specific components of the general goal, to respond to a WDN contamination event. These component goals are not comprehensive and do not address every issue that users might face. New demands or goals can be incorporated to accommodate an expanding set of issues.

Knowledge Representation

In our systems, the knowledge covering the health impact, utility operations, chemical characteristics and hydraulics of the WDN decontamination area, can be broken down into two primary categories: facts and rules. Facts are simple statements containing data values that represent, and show relationships among entities; Rules are declarative knowledge linking sets of premises and conclusions [6].

With the intention of stability, ease of maintenance, and flexibility of our system, we classify facts as static global facts and dynamic case facts. Like a dictionary, global facts represent those general and common facts applicable to all scenarios. On the other hand, case facts, like a single word or only a few words in the dictionary, record the specific information of each particular case. Case facts can be represented in the knowledge base by a series of questions that function as placeholders for dynamic provided information. Currently, four types of questions would be obtained: 1) Contaminant and its associated concentration; 2) Taxonomies of each unknown contaminant; 3) Materials found in the System; 4) Expected effectiveness of the treatment technologies, if needed. Besides the four essential series of questions mentioned, a unique case ID will be requested to specify different scenarios after the introductory screen as well. The reasoning rules are named correspondingly to the goals they prove. For example, R_E, R_W, R_I, and R_T are four sets of rules to prove G_E, G_W, G_I, and G_T, respectively. How the knowledge primitives are incorporated/programmed into the system is dependent on the software platform selected.

Copyright © Society for Science and Education, United Kingdom

Development Approach

During the programming phase, these formal knowledge representations were translated into computer code. A series of commercial off-the-shelf software products were reviewed and considered. These included: Visual Basic [16], JESS (Java Expert System Shell) [17], CLIPS [18], MATLAB or NETLAB [19], Visual Rule Studio [20], and PyKE (Python Knowledge Engine) [21].

A functional system design (FSD) document was developed to further identify the overall system architecture, processing considerations, and definition of displays and reports [5]. The functional system design carries the design to a sufficiently detailed level to support the actual programming of the system. By reviewing the FSD, recommendation of a specific software was then based on an investigation of functionality, installation, and integration characteristics, as well as compatibility with existing hardware, software, and communications investments.

Within the expert system, the inference engine provides the control mechanism for the expert system by identifying the heuristics to be activated, as well as the sequence of activation [6]. Inference engines work primarily in either a forward chaining or backward chaining mode. Forward chaining starts with the known facts and then works top-down to assert conclusions or new facts. Backward chaining begins with goals and then works bottom-up to determine what facts must be asserted so that the goals can be achieved [6].

We selected an expert system shell that was capable of both backward and forward chaining inferential logic, in order to support the subsequent comparative analysis. We selected a combination of Python and its Knowledge Engine PyKE [21]. Use of PyKE facilitated system construction. It can use Microsoft Windows to provide a flexible, intuitive and expandable environment for delivering knowledge-based systems [21]. Using PyKE, we developed two preliminary systems for further investigation: a forward chaining expert system framework, and a backward chaining expert system framework.

COMPARISON AND RESULTS

The architectures of the two preliminary systems are shown in Figure 1 and Figure 2, respectively. In forward chaining and backward chaining expert systems, operation structures are compiled in the respective inference engines separated from the knowledge base containing questions, case facts, global facts, and rules [6]. The forward chaining inference engine begins with the collection of all available information. However, the backward chaining inference engine starts from the goal selection. Similar in structure to other reasoning rules, the strategy for asking questions is controlled by certain information query rules in the backward chaining system. According to the query rules for each selected goal, the backward chaining inference engine collects necessary information from the existing case facts or previous analyses first, then conducts conversations between users and the system to collect the rest of the essential data (if there is any). After new case facts are asserted, the engine proves a certain goal with all related case facts, global facts, and reasoning rules. The following discussion illustrates and analyzes the effectiveness of these inferencing approaches to the decontamination problem domain. Simplified forward and backward chaining logic is shown in Figures 3 and 4, respectively.

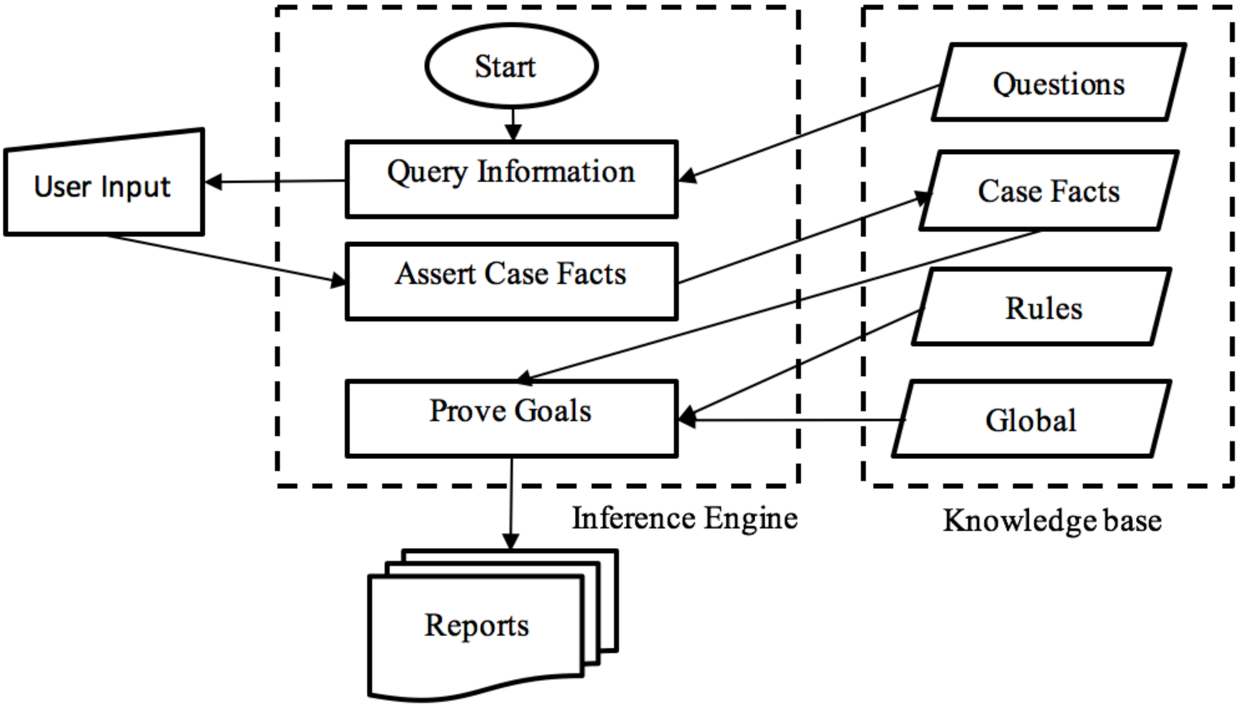

Figure 1. Architecture of the forward chaining expert system.

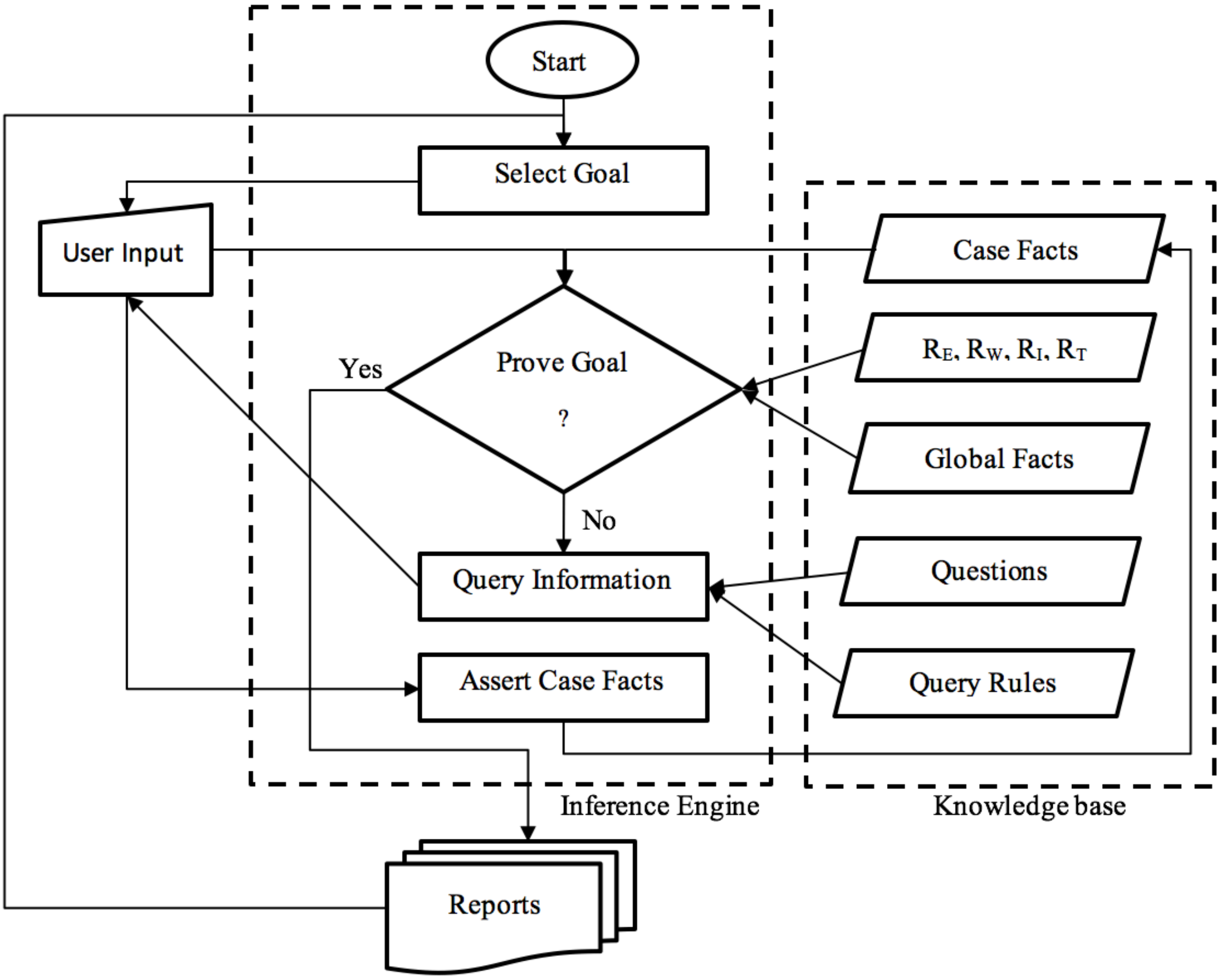

Figure 2. Architecture of the backward chaining expert system

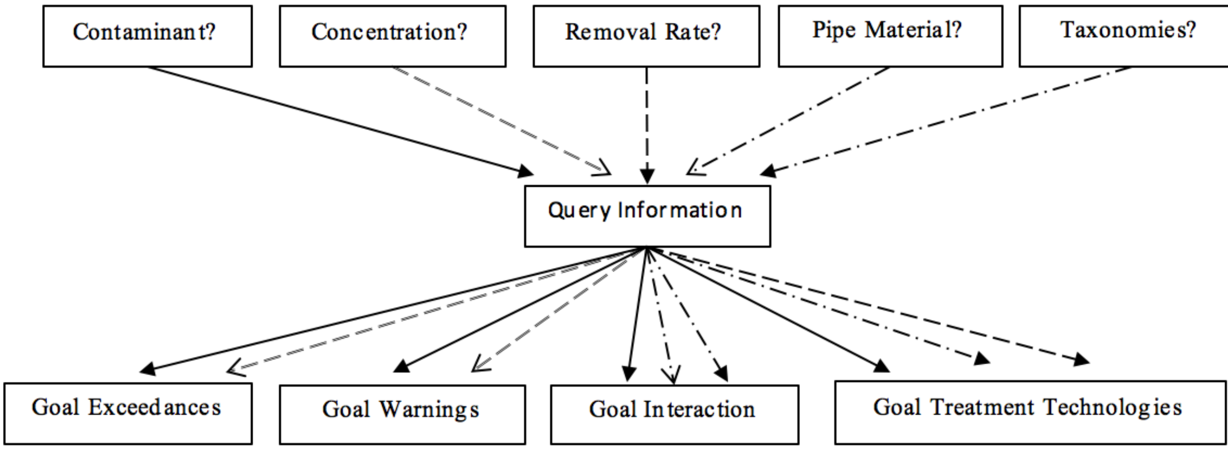

Figure 3. Simplified forward chaing logic to prove goal.

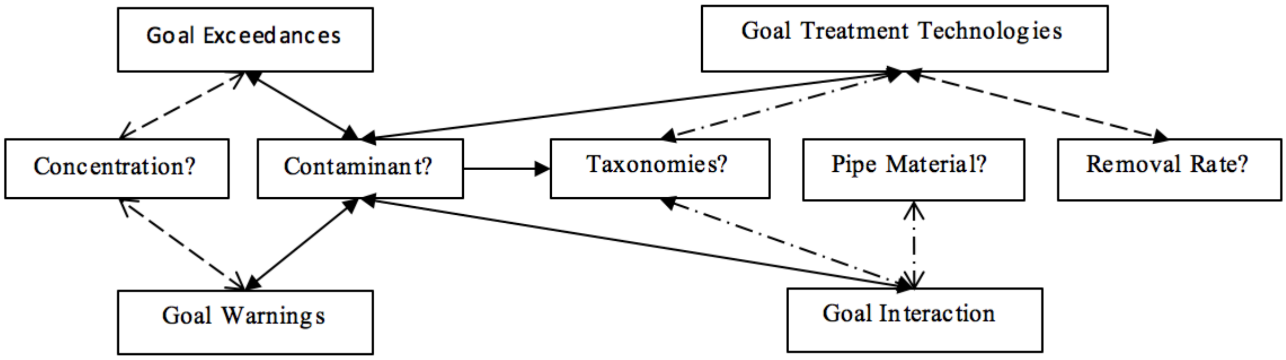

Figure 4. Simplified backward chaining logic to prove goal.

With regard to proving Goal Exceedences, only the information about the species and concentration of contaminant is essential. Therefore, the backward chaining inference engine only collects those two facts; while the forward chaining inference engine also collects extraneous information, e.g. pipe material, target removal rate, and taxonomies of the contaminant. The performance of a predecessor forward chaining expert system [22] indicated that users have to provide the complete query sessions for all incorporated goals in order for the expert system to complete processing, even when some of this information is not available or is not of interest to the uses.

The identification of pipe material, target removal rate, and taxonomies of the contaminant are unnecessary if contaminant concentration is not greater than regulatory exceedance levels. Thus, user input is not always necessary. An expert system can infer much information from the analyses of previous results and/or from other user input. For example, the expert system can create another new fact, e.g. "Benzene is a VOC" (volatile organic compound), to prove G_T if needed. (See Figure 5.) Only when the species of contaminant is unknown and the users need to prove G_I or G_T, should the expert system require user input on taxonomies. Consequently, direct user input should be required only as needed.

User input: Contaminate is benzene.
Knowledge primitive 1: Benzene is a hydrocarbon.
Knowledge primitive 2: Hydrocarbons can degrade PVC (polyvinyl chloride).
➔ Infer new knowledge primitive: Benzene can degrade PVC.

Figure 5. Example of systemic creation of new fact.

The efficiency of the user query session, and subsequent processing of the knowledge base, were considered critical to the success of this expert system. Based on the comparison and analysis of these preliminary systems, we continued to develop a more detailed backward chaining prototype expert system for WDN decontamination, referred to as "Decon". The prototype system was developed and delivered on a stand-alone IBM compatible microcomputer. As noted in Figure 6, Decon's extended knowledge base encompasses the health impact, utility operations, chemical characteristics and hydraulics of a WDN decontamination event.

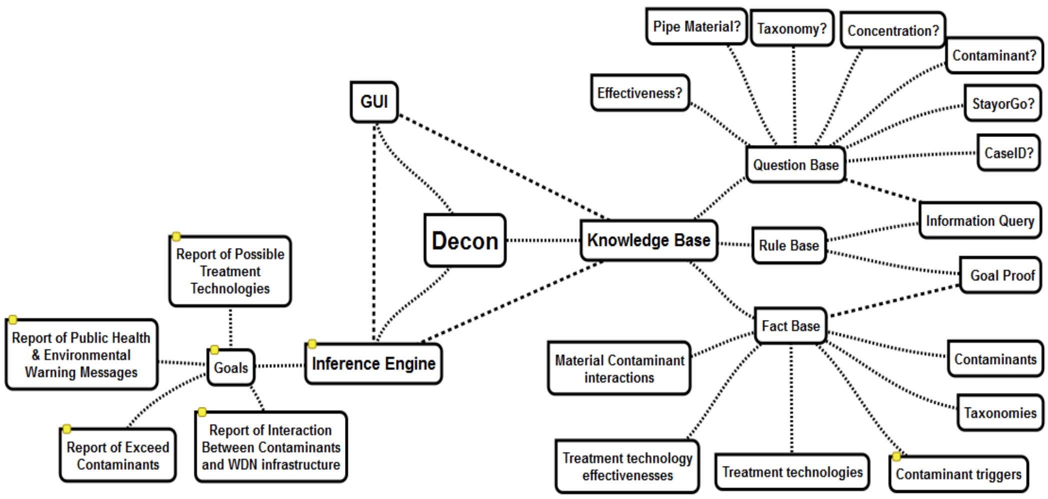

Figure 6. Knowledge contents of Decon.

USE OF THE DECON EXPERT SYSTEM

System initiation activates the main module, from where the four primary modules can be accessed. These subordinate modules are based on the system sub-goals previously discussed: Goal Excedences (G_E), Goal Warnings (G_W), Goal Interaction (G_I), and Goal Treatment Technologies (G_T). These four sub-goals are specific components of the general goal of response to a WDN contamination event.

Consider a scenario in which a large fuel tank is damaged in an accident. Benzene, the most common component of gasoline, permeates into the soil, and may enter into the underlying water pipe through cracks [23, 24]. In this hypothetical disaster, the water utility managers may first want to know if they need to be concerned regarding contamination of the WDN due to the spill. Upon system initiation, Decon instantiates the G_E module logic to determine whether the benzene exceeds local drinking water standards. In this scenario, assume the concentration of benzene is 0.05mg/L. As indicated in Table 1, three permissible limits are exceeded: the maximum contaminant level (MCL) of 0.005mg/L, maximum contaminant level goal (MCLG) of 0mg/L, and the minimal risk level (MRL) of 0.0005mg/kg/day [15]. At this point, the water sector managers will want to know "How does the contaminant impact the public health and/or environment?" (referring to G_W). A warning report is then generated by the system to address these questions. A portion of this report is depicted in Table 2.

Table 1. Example exceedences report.

Contaminant	Concentration	Unit	Trigger	Limit	Unit
Benzene	0.05	mg/L	MCL	0.005	mg/L
Benzene	0.05	mg/L	MCLG	0	mg/L
Benzene	0.05	mg/kg/day	MRL COI	0.0005	mg/kg/day

Table 2. Example warnings report

Contaminant	Concentration	Unit	Alert Type	Action Needed	Health or Environment
Benzene	0.05	mg/L	Public Health	Concentration is sufficiently high to cause a public health concern. Please notify your consumers and your public health agency. Potential health impacts include:	Anemia; decrease in blood platelets; increased risk of cancer

After the managers read the warning report, they may want to know: "Which technologies can be used in response?" Processing for G_T is initiated. Assume that the target removal rate is 80%. Based on the fact that benzene is a volatile organic compound (VOC), possible technologies with the efficiencies greater than or equal to 80% are listed in the output technologies report. In this scenario, nine potential technologies (i.e. activated carbon, activated alumina, air stripping, chloramination, chlorine dioxide, direct filtration, ozonation, ultraviolet, and advanced oxidation processes) and their brief descriptions are provided in the output [15].

A subsequent user inquiry to the expert system might be G_I: "Does the contaminant harm the water distribution pipe?" In the case of the previously discussed benzene spill, assume that the pipe material for this WDN is polyvinyl chloride (PVC). As indicated in Table 3, the Decon system then displays the message "Prolonged exposure to hydrocarbons causes PVC to degrade." This output indicates that benzene is a hydrocarbon, and warns the managers of the interaction between the contaminant and the WDN pipes [14].

Table 3. Example interactions report.

Benzene	Hydrocarbon	PVC	Prolonged exposure to hydrocarbons causes PVC to degrade

VERIFICATION AND VALIDATION

Verification and validation are important requirements to ensure quality and reliability of the system [25]. The verification of the system ensures that the system is error-free and robust. All components of an expert system are subject to verification: i.e. the knowledge base, the inference engine, and the user interface [26]. The knowledge base is verified by examining the consistency, correctness, and completeness of the knowledge. It involves detecting errors such as redundancy, contradiction, and circular dependency. The inference engine is verified by examining the reasoning process of the system. It assesses not only if the expert system is producing correct intermediate and final results but also if the expert system is using the correct reasoning process. The user interface is verified by examining the functionality of every component in the appropriate screens and reports.

Verification is performed to determine if the expert system is producing the correct answers and using the correct reasoning process. Consistent with accepted software engineering practices, the verification of DECON involved conducting unit tests, module tests, and system tests [27]. Unit tests were conducted to verify individual functions of the various modules of the system. Module tests were conducted to verify the outputs given by these modules. Additionally, a system test was also required to ensure that all the modules of the system have been seamlessly integrated and that they perform well as a single system [27]. All components of the prototype system were verified after the development was completed. Each module was tested using pre-designed cases. Pre-designed test cases were also utilized to verify the inference engine while conducting system testing. Upon indication of any errors, a structured problem resolution procedure was conducted. Problem resolution included resetting improper parameter values, and the correction of control code.

The validation process establishes that the system's functionality will address the original engineering problem and the user's needs. According to the literature, there are two validation strategies available for consideration: validation against expert performance, and validation by field testing [28]. They can be used alone or in combination, depending upon the application. Validation against expert performance can be achieved by: comparing the expert system against operational expert judgments; comparing test case results; or comparing against project expert judgments [28]. Validation by field testing involves running the program under actual or equivalent operational conditions. The field trial may be conducted either during initial production use or in parallel. A validation against expert performance was conducted to evaluate the prototype system developed during this research. This included utilization of pre-designed test cases. These compared to the outcomes of the system with the expected results from the literature. A validation against expert performance was conducted by nine water decontamination practitioners. The outcomes of the system were compared to their judgments. The knowledgebase was subsequently modified and extended in response to the practitioners' feedback. System validation confirmed that a correct system was developed according to user's requirements.

CONCLUSION

Decision making regarding WDN decontamination events is a complex problem requiring highly specialized expertise. This decision making process normally begins with incomplete and vague information, and must be able to adapt and evolve as new information becomes available. Consistent with the established capabilities of expert systems [e.g. 29], WDN utility managers need a computer-based tool that can: 1) replicate the logic and reasoning of human experts; 2) recognize and collect data related to progressive goals; 3) determine possible solutions in a stable and fast manner; 4) explain basis for the reasoning; and 5) can easily adapt to meet new standards or methods of decision making. This research identified the knowledge required to conduct WDN decontamination. Upon assimilation, we developed a knowledge base incorporating both dynamic and static knowledge primitives collected from official documents and refereed journal papers. We also explain how the knowledge is compiled, and how the compiled knowledge is stored, shared and driven by inferential logic. The resulting prototype system applies the collected knowledge to provide analysis of both contaminant exceedance and public health threats in order to guide water sector managers. Further, information is provided by the system to guide these decision makers in selecting potential technologies for WDN decontamination.

Further, to meet the water utility managers' evolving demands (referred to as goals) and growing new information on WDN decontamination, the backward chaining method is proved to work more effectively than forward chaining because it enables the expert system to search for the necessary data from the result of the previous analyses, and thus limit the requirements for direct user input.

Although the existing version of Decon works with limited goals for WDN decontamination, more goals (related to other key issues identified in the CIPAC 2008 report) and their corresponding knowledge can be added to the system. For example, interactions between multiple contaminants could be taken into consideration, as well. Moreover, external sources such as EPANET could be linked to the expert system to conduct a more comprehensive analysis for the entire WDN, rather than one isolated node.

References

UNISDR, UNISDR Terminology on disaster risk reduction, 2009, Geneva: UNISD.

Shaluf, I. M., An overview on disasters, Disaster Prevention and Management, 2007, 16(5): p. 687-703.

Kulshrestha, S. and Khosa, R., Expert system for management of water distribution network, International Journal of Engineering, Science and Technology, 2010, 2(12): p. 7401-7412.

Critical Infrastructure Partnership Advisory Council, Recommendations and proposed strategic plan: water sector decontamination priorities, 2008. Downloaded 3December 2018 from http://www.nawc.org/uploads/documents-and-publications/document_ca70ed5-40ed-afcl-a92a8beb3988.pdf

Moynihan, G.P., Expert systems in foundry operations, in Leondes, C. (Ed.): Knowledge-based systems: technologies and applications, 2000, Cambridge, MA: Academic Press, p. 1273-1292

Coppin, B, Artificial intelligence revisited, 2004, Sudbury, MA: Jones and Bertlett Publishing.

Feigenbaum, E., The art of artificial intelligence: themes and case studies of knowledge engineering, 1977, Stanford, CA: Stanford University Computer Science Department,.

Kenov, K. N. and Ramos, H. M, Water and energy sustainable management in irrigation systems network, International Journal of Energy and Environment, 2012, 3 (6): p. 833-860.

Ladopoulos, E. G, Non-linear pipe networks water management real-time expert telematics system, Journal of Water Resource and Hydraulic Engineering, 2013, 2(1): p. 13-20.

Francis, R., Guikema, S. and Henneman, L., Bayesian belief networks for predicting drinking water distribution system pipe breaks, Reliability Engineering and System Safety, 2014, 130(1): p. 1-11.

Agency for Toxic Substances and Disease Registry, ATSDR toxic substances portal. 2012. Downloaded 3 December 2018, from http://www.atsdr.cdc.gov/ substances/index.asp

U.S. Environmental Protection Agency, National primary drinking water regulations, 2013, Washington, D.C.: U.S. Environmental Protection Agency.

Knuuttila, H. et al., Advanced polyethylene technologies – controlled material properties in long-term properties of polyolefins, 2004, Berlin: Springer.

Ong, S. et al., Impact of hydrocarbons on PE/PVC pipes and pipe gaskets, 2008, Denver, CO: American Water Works Association Research Foundation.

U.S. Environmental Protection Agency, Module 6: remediation and recovery guide, 2004, Washington, D.C.: U.S. Environmental Protection Agency.

Spyridakos, T., Pierakos, G., Metaxas, V. and Logotheti, S., Supporting the management of measurement network with an expert system: the NeMO system, Operational Research: An International Journal, 2005, 5(2): p. 273-288.

Robindro, K. and Sarma, S. K, JESS based expert system architecture for diagnosis of rice plant diseases: Design and prototype development, in Proceedings of the 4th International Conference on Intelligent Systems, Modelling and Simulation, 2013.

Ooshaksaraie, L. and Basri, N. E., An expert system applied in constuction water quality monitoring, American Journal of Environmental Sciences, 2011, 7(11): p. 75-81.

Mounce, S., Boxall, J. and Machell, J., Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows, Journal of Water Resources Planning and Management, 2010, 136, pp. 309-318.

Chau, K. W. and Phil, M., Knowledge-based system on water resource management in coastal waters, Water and Environmental Journal, 2004, 18(1): p, 25-28.

Frederiksen, B., Welcome to Pyke , 2015, 12 December 2018, from http://pyke.sourceforge.net

Gutenson, J. L et al., Water Expert: a conceptualized framework for development of a rule-based decision support system for distribution system decontamination, Drinking Water Engineering and Science, 2015, 8(4): p. 9-24.

Atkin, E., Company that caused historic chemical spill leaks more waste into West Virginia waters. 2014. Downloaded 4 August 2018, from http://thinkprogress.org/ climate/2014/06/13/3448678/freedom-industries-spills-again/

National Oceanic and Atmospheric Administration| National Ocean Service| Office of Response and Restoration, Oil and chemical spills, 2015, Downloaded 1 December 2018, from http://response.restoration.noaa.gov/oil-and-chemical-spills

Alter, S., Information systems: the foundation of e-business, 2002, Upper Saddle River, NJ: Prentice Hall.

Smith, S. and Kandel, A., Verification and validation of rule-based expert systems, 1993, Boca Raton, FL: CRC Press.

Laudon, K.C. and Laudon, J.P., Essentials of management information systems, 4th edition, 2001, Englewood Cliffs, NJ: Prentice Hall.

Prerau, D., Developing and managing expert systems, 1990, Reading, MA: Addison-Wesley.

Cretescu, I., et al., Development of an expert system for surface water quality monitoring in the context of sustainable mangement of water resources, Environmental Engineering and Management Journal, 2013, 12(8): p. 1721-1734.