Development and Application of Expert Systems in Audit Services

Eric L. Denna
James V. Hansen
Rayman D. Meservy

Brigham Young University, rayman_Meservy@byu.edu

Follow this and additional works at: https://scholarsarchive.byu.edu/facpub

Part of the Accounting Commons

BYU ScholarsArchive Citation
https://scholarsarchive.byu.edu/facpub/3240

This Peer-Reviewed Article is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Faculty Publications by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.
Development and Application of Expert Systems in Audit Services

Eric L. Denna, James V. Hansen, and Rayman D. Meservy

Abstract—Professional auditing has become much more complex over the past decade. This development has engendered a need for leveraging auditor expertise. In an attempt to meet this need, public accounting firms have been actively involved in supporting the development and application of artificial intelligence methods in audit services, which forms the foundation of their business.

Much of the development and application of artificial intelligence in auditing has taken the form of expert systems. This paper examines the current state of the art and provides a framework for analysis of present work, as well as a guide to future efforts.

Index Terms—Auditing, decision support, expert systems, knowledge acquisition, knowledge representation, performance evaluation.

I. INTRODUCTION

During the past decade there has been a growing interest in the development of expert systems (ES) for various financial auditing problems. Research and development has been vigorously pursued by several large public accounting firms. Current efforts have extended into exploration of more powerful methods of knowledge acquisition and knowledge representation, as well as means of performance evaluation.

There are several reasons for the appeal of expert systems in auditing. First, the audit environment is becoming increasingly complex. The past 20 years has resulted in a proliferation of detailed rules and documentation requirements to be followed by auditors on all engagements. The rapid growth in requirements has made it impossible for an individual auditor to maintain a working knowledge of how to assess all eventualities [10].

Second, competitive pressures are motivating auditors to seek new ways to perform professional services efficiently and to reduce redundancies and inconsistencies. Changes in the Code of Ethics pertaining to competitive bidding, solicitation, and advertising have led to increased competition among firms. This competition results in more competitive audit fees and smaller profit margins per engagement. To counter these changes, public accounting firms are searching for ways to become more efficient, while maintaining a high level of effectiveness [11].

Third, despite a plethora of analytical techniques which have been proposed for evaluating audit evidence, some of which are in general use, audit practice needs methods which can systematically and logically aggregate judgments and evidence. Knowledge of the relative merits and performance of various inference systems, or combinations of systems, would be an invaluable contribution to audit decision making.

The central role of judgment in audit decision making places a premium on experience and client knowledge. However, transfer of knowledge from the partner and the manager to the junior auditor is difficult. The auditor is often faced with too much information to fully and completely analyze. The expert auditor with this problem by assuring that all the key issues and events are identified as early in the audit process as possible. This helps to reduce the information that must be analyzed. Further, it is the expert auditor’s ability to transfer knowledge and experience to other auditors in a timely manner that leads to the right questions at the right time that leads to the identification of the risks and issues [10]. These questions are often asked in the field in response to new information revealed during the audit engagement.

The need for systems which can aid the auditor by bearing the best expertise available in the firm is clear. Systems technology offers a powerful way of addressing that issue. Expert systems technology offers a powerful way of addressing that issue. Expert systems technology offers a powerful way of addressing that issue.

The purpose of this paper is to evaluate research and development in the design of expert systems for the audit domain. Section II provides an overview of the domain of audit expertise involved in the audit process. We consider a framework that we use to present and analyze work done as well as to guide future efforts. Section III examines methods of knowledge acquisition being used to develop audit applications. Section IV addresses knowledge representation in the audit domain. Most systems developed for auditing have rule-based systems. Currently, however, other representations are being tested that may be more appropriate for dealing with audit complexity. Section V considers performance evaluation in expert system validation. The paper concludes with a brief summary of the current state of development along with suggested directions for future efforts.

II. A FRAMEWORK FOR EVALUATING ES WORK IN AUDIT SERVICES

Recently, Akresh et al. [1] provided a taxonomy of activities involved in the audit process that require various degrees of audit expertise. This taxonomy is shown in Fig. 1, relatively few of the tasks have been studied using an approach, although interest continues to increase significantly for both academics and practitioners. The taxonomy serves as a useful means of scoping the domain of audit judgment tasks which researchers could investigate.
For purposes of summarizing and evaluating ES research in the domain illustrated in Fig. 1, we propose a three-dimensional framework as shown in Fig. 2. Our interest in presenting this framework is twofold: first, to provide a structure for selecting, reviewing, and analyzing work performed to date; and second, to provide a structure for proposing additional work. The framework consists of three axes which present the three main areas of ES research in auditing. Our paper analyzes developments in each of the three areas, as well as examining how the three interact to affect research salability.

The axis labeled KA represents work dealing with the problem of knowledge acquisition. Efforts along this dimension have attempted to address the traditional bottleneck of eliciting knowledge from experts, a problem pervading ES development in all domains of specialization. The goal of knowledge acquisition, as stated by Wood and Ford [24], is that, “knowledge should be elicited, organized, and documented with minimal concern for how it will be implemented in a working system” [24, p. 1]. This objective is similar to that which has been proposed by Alexander et al. [2] and Johnson [13].

The axis labeled KR represents work dealing with the problem of knowledge representation and organization. Work in this area has ranged from the use of simple production-rule systems, to more sophisticated hybrid models utilizing rules and frames, to the development of generalizable models of
memory organization such as case-based reasoning (cf., Riesbeck and Schank [19]) and generic tasks (cf., Chandrasekaran and Mittal [7]).

The axis labeled KV represents work dealing with the problem of knowledge validation. Knowledge validation focuses on evaluating various computational models of audit judgment in an effort to determine the validity of the model. This is an area which is crucial to audit ES’s, but has not been fully explored. Work in this area has focused on the simple evaluation of output from an ES. Only one study has provided an in-depth analysis of the reasoning process and exploration of the robustness of domain knowledge.

A summary observation regarding the proposed evaluation framework concerns the way in which it addresses the issue of domain and project complexity. By complexity we mean the degree to which a research effort attempts to address the real-world nature of a particular problem domain. As Waterman [23] states, “When gross simplifying assumptions are made about a complex problem, and its data, the resulting solution may not scale up to the point where it is applicable to the real problem” [23, p. 27]. Complexity in our framework is represented by efforts that extend understanding along one of the axes, or that attempt to test the usefulness of the extant knowledge in each of the three areas by developing a tool for actual use in the profession. This aspect of the framework attempts to recognize the contribution of the profession to academic research by testing the scalability of academic theories and ideas.

The sections which follow examine the more recent ES projects in auditing. Our objective in this examination is to abstract general propositions about ES development in the audit domain.

III. EXPERT SYSTEM KNOWLEDGE ACQUISITION

A. General Techniques

Wood and Ford [24] have categorized knowledge elicitation methods in a way that is well suited to our discussion. They group knowledge-acquisition methods into four classes: descriptive elicitation, structured expansion, scripting, and validation.

Descriptive elicitation focuses on helping the expert to provide a large quantity of his abstract declarative knowledge about the problem-solving domain. This methodology attends to the categories, objects, models, and other conceptualizations used by experts in problem-solving.

Structured expansion attempts to expand and integrate fragmented portions of the expert’s knowledge derived from descriptive elicitation. The use of structured expansion is not yet well evidenced in the development of expert systems in the audit domain. One possible reason is that the most recent methods of accomplishing structured expansion elicit hierarchies of concepts based on language cues. These methods have emanated from the work of ethnographers and are not widely known to knowledge engineers in the audit area.

Scripting requires capturing the thought processes of the expert while solving a problem. Formalizing the details of problem-solving behavior and their sequence is the goal of scripting. This approach is most often accomplished through methods of protocol analysis [6].

Validation refers to “checks,” or attempts to validate knowledge in the possession of the knowledge engineer, and “controls,” or methods designed to reduce error and misunderstanding in the knowledge elicitation process. The standard method involves building prototype modules as knowledge is gathered. The overall prototype is then evaluated by the expert as a means of checking the validity of the model. By observing the performance of the prototype the expert is able to assist in refining the system to enhance the validity of the prototype model.

B. Knowledge Acquisition in Developing Expert Systems for Auditing

A common method of knowledge acquisition in the development of expert systems for auditing has been direct elicitation. This method was used in the development of the commercial ES, RISKADVISER [10]. Direct elicitation has, however, been found insufficient in some studies [6] and has been augmented with other knowledge acquisition methods.

Spradley [21] provides a compelling explanation for the use of more than one elicitation method. He notes that experts tend to translate their knowledge into terms that they believe the interviewer will find easier to understand. However, these artificial representations do not faithfully represent the expert’s customary modes of thought.

As shown in Fig. 2, validation methods are often used in concert with direct elicitation methods [12]. A common approach to the development of ES’s for auditing has been to use direct elicitation to develop a small prototype. The expert is then asked to evaluate the prototype, indicating where errors exist, knowledge is incomplete, or other dissonances are observed. The knowledge engineer analyzes this information and modifies the prototype accordingly. Elements of the process are repeated until satisfactory results are achieved (c.f., [9] and [16]).

In a search for more effective methods, some recent studies have applied scripting, as represented by protocol analysis. Biggs et al. [6] used protocol analysis in the knowledge acquisition phase of developing EDP-XPERT. A thorough understanding of the auditor’s information processing and decision-making activity was produced; however, the method did not generate sufficient production rules for a complete prototype. Protocols do, however, provide a framework for structuring a production system [6], [16].

Structured expansion would appear to have a useful role to play in knowledge acquisition for ES’s in auditing. As these methods become better defined they will no doubt be utilized.

C. Automated Methods of Knowledge Acquisition

There has been a recent proliferation of automated methods of assisting in the knowledge acquisition process. Work on knowledge-acquisition workbenches shows promise of being effective in capture and transfer of expertise in complex problem domains [24].
As a follow-on to their work on EDP-EXPERT, Messier and Hansen [17] have explored the use of inductive algorithms as a way of generating production rules from archival data involving problem-domain scenarios and the expert’s resulting decision. While the results were favorable, the method may be limited to developing the production rules for modules within the expert system, particularly those exhibiting consistent structure in the archival data.

On balance, auditing remains a rich opportunity for the testing and development of knowledge acquisition methods. Methods based on current theories and software-assisted methods have not been widely used. We believe that this is because most research and development of expert systems for auditing has been driven by the need to get something immediately into operation which can be tested. The end results may be less than they otherwise could be.

IV. EXPERT SYSTEM KNOWLEDGE REPRESENTATION

This section describes what we consider to be “second generation” work on the issue of representing expert auditor reasoning. We characterize “first generation” expert system development projects as those which use a production rule paradigm exclusively. Because these initial audit ES projects were developed using only production rules, they offer no real contribution to the issue of knowledge representation. Hence, our focus is on “second generation” projects.

Three recent projects have taken advantage of more advanced representation techniques, thus extending the representation methods for modeling audit expertise. These studies focus on providing more robust and useful ES applications by developing more accurate representations of both the reasoning process and the domain knowledge of the auditor. In addition to modeling the heuristic processes of a specific judgment task, these projects concentrate on modeling the underlying, domain knowledge of the expert. The idea here is that if the domain knowledge of an auditor can be encoded in computational form, developing applications for the various audit functions simply become extensions or uses of the domain knowledge base. This argument is akin to the idea of data modeling for database development. If the model is epistemologically correct then it should anticipate, to some extent, future application development efforts. Although each of these projects is still in the research prototype stage, those involved continue to work toward the development of each as a commercial product for use in audit practice.

The purpose of this section is to review the progress toward developing models of auditor domain knowledge. For each of the three projects considered, we explain the purpose of the project and the method of knowledge representation used.

A. GCX—Going Concern Expert

Selfridge and Biggs [20] focused on enhancing the representation of audit domain knowledge. This effort concentrated on developing an ES to determine whether an audit client is a “going concern.” The going concern judgment centers on determining whether the audit client can be expected to remain in business throughout the foreseeable future. This judgment is typically performed by the engagement partner and can result in an auditor refusing, or withdrawing from, an engagement. When the auditor determines the client’s continued existence is doubtful, a “going concern” opinion is rendered in the auditor’s public report attached to the company’s annual report.

When no “going concern” opinion is rendered and the client thereafter fails (e.g., declares bankruptcy), investors will typically sue the auditor for investment losses, citing their reliance on the auditor’s opinion in the annual report as the reason. Such lawsuits have grown to become very costly to the audit profession. Hence, audit firms are extremely concerned that going concern judgments be made correctly and have contributed substantial financial support toward the work of Selfridge and Biggs [20].

Fig. 3 illustrates the structure of the GCX system. At the core of the system is a detailed model of the client, industry, and general economic events. For example, the figure illustrates auditor knowledge concerning the normal chain of events involved in the shipping of goods and the effect of various events on the financial statements of the client. GCX contains numerous such descriptions of both the normal and actual operating events of a client. This event knowledge provides the basis for an extensive causal network involving it and eight other types of knowledge as shown in Fig. 4.

Implementation of the GCX system is based on the use of frames to represent the event memory, as explained in Charniak and McDermott [8]. The reasoning processes are represented using relatively simple LISP procedures. Selfridge and Biggs [20] show that by assuring a well-conceived model of auditor memory the procedures utilizing the memory can be rather simple.

B. APE—Audit Planning and Evidence

More recently, Denna [9], with the support of a grant from Coopers & Lybrand, concentrated on representing expert judgment demonstrated during the audit planning process. Specifically, the system focuses on the risk analysis performed by an auditor during the detailed planning stage of the audit.

The detailed audit planning process typically consists of the steps illustrated in Fig. 5. Moving from left to right in the diagram, the planning process begins by identifying a set of potential problems for each financial statement assertion. In turn, the auditor determines the likelihood of material error (LME) for each financial statement assertion. Based upon the LME, the auditor then formulates a preliminary audit approach for gathering evidence regarding the specific audit concern. Finally, the auditor aggregates each preliminary audit approach into one, overall audit program that serves as a workplan for the audit engagement team.

Although each of the steps in Fig. 5 involves expert judgment, the most complex and most critical is that of assessing the LME for each audit concern. Due to the importance of the LME judgment, Denna [9] developed a system capable of performing the LME judgment process, which has been named APE (Audit Planning and Evidence). The domain of application for APE at this point is that of planning the audit for a large retail grocery client.
As with GCX, APE extends the representation of audit judgment expertise beyond the use of production rules. APE includes the auditor’s knowledge of company function, actual and normal operating events, financial knowledge, and the reasoning processes manifested during the audit planning process.

APE differs from GCX in that it involves various levels of domain knowledge abstraction as illustrated in Figs. 6 and 7. During the knowledge acquisition phase, the experts demonstrated a working knowledge of the client operations at three levels of abstraction. Fig. 6 shows the first two levels of abstraction while Fig. 7 illustrates the detail of the “Order Inventory” operation of the client. APE possesses the ability to work at any of the three levels of abstraction and can jump to another level as the reasoning process requires.

A key feature of the APE system is its ability to utilize its knowledge of direct operations during the planning process. As shown in Fig. 6, during the analytical review process the auditor accesses the domain knowledge of company operations by utilizing the causal links (labeled as arrows in Fig. 8) between financial statement items (e.g., the amount of warehouse inventory) and an event (e.g., increase purchase quantities). With this knowledge, the auditor is able to determine the cause of a financial statement item fluctuation and determine whether the item behavior should be of concern to the auditor.
APE consists of over 100 instances of a frame hierarchy, over 50 rules, and over a dozen special purpose LISP procedures. As mentioned earlier, APE’s domain is limited to the planning of the inventory portion of a retail grocer audit. It is currently being extended to include the remaining operating functions of a retail grocer.

Rather than pursue the traditional method of external evaluation procedures, the APE evaluation process concentrates on an internal validation (discussed in Section IV) of the computational model. Specifically, Denna [9] demonstrates the model’s ability to explain what it knew about the client and its ability to reason through the effects of previously unencountered events, on the audit planning process.

For example, as is illustrated in Fig. 8, APE contains knowledge of the effort of the client to become the discount grocer for the area during the year. APE utilizes this knowledge to reason through its effects on both company operations and resources (the English equivalent of the APE output for such a session is shown in Table I). This process was used to demonstrate APE’s ability to handle previously unencountered events and reason through their effects on the client.

C. IRE—Inherent Risk Evaluator

Closely related to the APE and GCX projects is the work of Peters [18] on the problem of the auditor’s assessment of overall inherent risk. Preceding the detailed planning of the audit engagement discussed earlier (see Fig. 5), auditors make a determination as to the overall level of risk of the audit engagement. Again, due to the increasingly litigious nature of the audit engagement, the audit profession is interested in formalizing and disseminating the firm’s expertise to avoid costly mistakes in the audit practice.

Peters’ [18] work concentrates on representing the domain knowledge of the auditor and then utilizing this knowledge
the development of a specific application (e.g., inherent risk assessment). The structure of the system development by steps is shown in Fig. 9. As with both GCX and APE, the process of assessing the inherent risk of the audit engagement draws heavily from the domain knowledge of the auditor.

V. Contribution of Representation-Focused Studies

Studies concentrating on extending the representation of auditor knowledge have focused on answering at least two of the three criticisms often leveled against expert system work in auditing. These criticisms of traditional (rule-based) representations are summarized by Selfridge and Biggs [20] as follows:

1) They break down or perform poorly when confronted with situations outside the boundaries of their domain.
2) They lack human-level explanation and learning capabilities that characterize expert behavior.
3) They often do not outperform human experts.

The three studies summarized in this section do not address any of the criticisms in their entirety. However, they represent efforts to work toward answering these criticisms. From these studies we can conclude the following regarding the application of the expert system approach to modeling the audit judgment process:

1) Auditor domain knowledge appears to rest upon a detailed understanding of client operations and the surrounding environment.
2) Auditor knowledge appears to be segmented into various levels of abstraction starting with a very simplified abstraction to a detailed identification of the events, resources, and agents associated with the client.
3) Model explanation and deformation capabilities seem more feasible using frame-based representations rather than simple rule-based systems.
4) Development of deep models of auditor knowledge for commercial uses will require significant strides in both the technology needed to deliver the models as well as the further refinement of the models to make them generalizable.

With the support of the audit profession, continued progress is anticipated in the development of the three systems presented here, as well as other projects using expert systems to study the audit judgment process.

V. Expert System Knowledge Validation

The validation of the knowledge contained in an expert system is often closely intertwined with the knowledge acquisition activities. During the developmental phase, experts are often asked to validate individual bits of information and the way in which they are organized in the knowledge model or expert system. While tuning the system at the close of the development phase, various problems are often given to the expert system while an expert observes the system's lines of questions, tentative conclusions, and final decisions. The expert then provides feedback on what might be done to improve the system. Tuning, along with other development activities, is often an iterative process.
Fig. 8. Hook to client knowledge during deep reasoning.

Fig. 9. Structure of IRE.
It is important that systems not only be built and tuned, but that they also be tested and verified. At some point, the knowledge engineer decides that the system has reached the completion of the initial developmental phase and is ready for testing. Typically, developers have not been able to formulate single critical experiments which validate systems, but rather rely on a variety of tests which address various aspects of the knowledge base. We have organized these tests into three classifications: external, internal, and advanced experiments.

We focus on an expert system developed by Meservy, Bailey, and Johnson [16] to discuss knowledge validation. ARISC (Auditor Response to Identified System Controls) investigated the strategies and processes by which experienced auditors evaluate systems of internal accounting controls. Its development involved the extraction and simulation of auditor expertise, expert system building and tuning, and extensive validation of the resulting expert system.

A. Conceptualization of the Task

The auditor's evaluation of internal accounting controls is a problem requiring the expertise of well-trained auditors and is thus a requisite of every audit performed by CPA's. The strengths and weaknesses of an internal accounting control system are evaluated by determining control objectives, identifying controls and faults from a description of the system, and then combining the controls and faults into an overall evaluation of the sufficiency with which each control objective has been met. The results of the task (and the output from the expert system) consist of 1) a suggested list of controls for the compliance testing phase of the audit and 2) a list of control weaknesses: significant problems that will need to be discussed with management.

The task is difficult because no unique set of normative controls exists; rather, accounting systems can be configured using a wide variety of acceptable combinations. Bailey et al. have shown this task to be an NP-hard problem [5].

B. Discovery Phase

The discovery phase consisted of two steps: 1) knowledge acquisition and system development which included "thinking aloud" protocols, interviews, and structured descriptions, and 2) tuning. Significant findings include a) the various "views" used to evaluate data and b) that uncertainty in this domain is represented symbolically. Never once in this protocol, as in [6], was uncertainty represented numerically. This may be due to the fineness of the decision differences, along with the low level of expected error. A particular expert system shell, Galen [22], was selected which has the ability to partition the knowledge base, to search for a hierarchical set of goals, to apply forward and backward chaining, and to revise decisions.

Tuning involved running several prototype cases through the evaluation process and, in collaboration with the expert, making adjustments in the rules. Some of the more important aspects of expertise were discovered in this process, i.e., knowing when to discontinue the current line of reasoning or begin another. Aspects of expertise incorporated during the tuning phase included rules about the use of other rules, known as meta-rules, which were used by experts to their thinking about each case.

C. Verification Phase

There is no single critical experiment which can be used to validate an expert system. Furthermore, due to small sample sizes, statistical evaluations of experimental results are generally not available and researchers are constrained to graphical techniques. Meservy et al. [16] included several empirical tests, each of which addressed different types of data and different aspects of the system's behavior. The framework used to evaluate the system's performance included major features: 1) tests of sufficiency or adequacy of system outcomes, and 2) tests of the quality of system process of data usage.

Because accounting information systems differ, each represents a new challenge to the expert system. Therefore, the test the expert system, the cases chosen should not recombine recombination of portions of previous cases used, but new cases representing a wide range in risk, reliance, and auditor documentation. Three such cases were obtained from the accounting firm.

Since the system was fine tuned around the expert one auditor (the primary subject, S1), the expert system initially validated against this individual using the three cases. It was then cross-validated against three additional auditors (subjects 4, 5, and 6). For each case, the auditors were asked to evaluate the actual engagement document and to list 1) recommendations for specific controls compliance tested, and 2) weaknesses identified from the system description. The expert system also received each made similar evaluations while providing a trace of data analyzed and rules fired.

D. External Validation—Adequacy of System Outcomes

External validation involves the use of domain expertise by the same experts used in developing the system, or by not previously exposed to the system. These experts evaluated the performance of the ES a) when given developmental b) when given new cases, or during c) field testing.

Adequacy of ARISC results was evaluated using peer reviews. Three new auditors were given the decisions by the expert system and the four auditors for each of the three cases, along with the case materials, one case at a time. The results were retyped so that the review was performed by the auditors. The reviewer was asked to judge the decisions in each of three categories, completeness, effectiveness, and whether they personally agreed with the results. Table II shows that out of the nine reviews, the system decisions received five first place rankings in each of the three categories evaluated (tie results are each awarded a first place ranking), thus the systems overall ranking is higher than for any auditor.

The results of a good simulation should be compared to the expert modeled. The task required the peer reviewers to separate the five decisions for each case into first a) two groups and then b) three groups based on perceived similarity results showing how often the expert system was placed...
same group with each of the auditors are presented in Tables II and III. Analyzing the above two tables over all cases and all subjects, the expert system appears to be most like subject 1, the modeler. However, it is interesting to note that for the 2 and 3 groupings, the system was placed in the same group with each auditor at least once.

E. Internal Validation—Tests of Process Quality and Data Usage

The quality of expert system rules is established by determining that the rules are not only legal, but are of the variety experts make. A chess playing expert system should not only make legal moves, but the moves should be similar to what a chess expert would have done. The analytical approach used, synthesized a top-down global analysis with a bottom-up knowledge state/process analysis.

One top-down analysis performed was the evaluation of major hypothesis generation. Fig. 10 is a hypotheses graph of the major hypotheses generated by case for both system and subjects. The “+” represents a positive decision; a “−” represents a weakness, problem, or question, and a “?” represents an assumption or issue about the hypothesis which needs to be resolved. System-generated hypotheses are more similar to auditor hypotheses for any given case than to system generated responses for the other three cases. This would be expected if it is truly data driven. Also, subject 1’s hypotheses agree with the system generated hypotheses 96% of the time, which was much higher than any of the other subjects. This result would be expected if the system were simulating subject 1. Overall agreement with the auditors was about 72%. In other internal control-related studies, typical auditor consensus was found to be 70% [3] and 67% [4]. Thus, the consensus fell within an acceptable range.

The quality of system processes was also established by six other tests studying data usage, the amount of time spent in various process and episodic categories, and the order of processing on both low and high levels, which are presented elsewhere [15]. The resulting graphs of these studies led to conclusions regarding various lines-of-reasoning and views or focus used by the system and subjects. These studies confirmed that the quality of system processes was acceptably similar to all auditors tested, although more comparable to the simulated subject.

From the external and internal validation studies it was concluded that ARISC appears to simulate the processes of expert auditors, particularly the auditor after whom it was modeled.

F. Advanced Experiments in Knowledge Validation

Advanced experiments are designed to test the generalizability of the knowledge base. Such experiments systematically analyze not only the system knowledge base, but also the underlying knowledge of the domain for critical paths and errors, where even the domain experts may fail. Due to the detailed knowledge base, we should be able to anticipate expert subject behavior when conditions in the task are altered [14].

Expert system behavior often varies systematically as a function of the presence or absence of a small number of critical cues. It turns out the behavior of experts varies in a similar way (see Johnson et al. [14] for a further description of work done in the medical area). When early strong cues suggest one solution and later weak cues suggest a different (and correct) alternative, the task is labeled as a Garden Path case. Garden Path cases often lead to suboptimal decisions by experts [15]. Once identified, additional knowledge may then be added which would make it possible to avoid Garden Path solutions.

Advanced experiments systematically perturb the system and manipulate the data to determine not only Garden Path errors, but also discover the limits of the knowledge base. Such limits may then be encoded as additional meta-knowledge about the generalizability of the ES. Unfortunately, few appropriately designed studies have been carried out in auditing.

In summary, developers have not been able to formulate single critical experiments which validate systems, but rather rely on a variety of tests, each test addressing various aspects of the knowledge base. While there is no single sufficient test, each additional external, internal, and advanced test adds to the validation claim of an expert system.
I. SUMMARY AND RECOMMENDATIONS

In the conclusion, we offer a brief summary and provide some recommendations regarding future efforts in ES research in auditing. In this paper, we have reviewed ES research in auditing using a framework that positions research along three axes: knowledge acquisition, knowledge representation, and knowledge validation. Along each of these axes, we have characterized the work we have mentioned as either "first," or "second," generation work. The primary features of these first two generations of work are shown in the center area of our three-dimensional framework.

ES research in auditing is still relatively new, even by general ES research standards. Nevertheless, significant progress is being made to appropriately address the study of expert judgment using the ES approach. Particularly encouraging are efforts to use a variety of more robust methods of knowledge acquisition, knowledge representation, and knowledge validation.

The following is a list of what we would consider to be attractive and potentially rich areas of focus in each of the three areas of our proposed framework. These suggestions are not intended to be exhaustive, but will hopefully provide some guidance for the individual desiring to participate in this area of study.

1) Knowledge Acquisition—Working toward the noninvasive elicitation of knowledge, we suggest future efforts could focus on the use of several knowledge acquisition methods as proposed by Wood and Ford [24] and on the use of automated tools. As with knowledge acquisition in all judgment areas, work is needed to enhance the effectiveness and efficiency of the knowledge acquisition process. The use of a variety of knowledge...
acquisition approaches and the use of automated tools could provide significant contributions toward resolving this important development bottleneck.

2) Knowledge Representation—Although current ES audit research has shifted away from simple rule-based representations of knowledge, we are just beginning to explore generalizable models of auditor memory and reasoning. A potentially rich area of research might deal with the testing of various memory models such as Riesbeck’s and Schank’s [19] case-based reasoning, or Chandrasekaren and Mittal’s [7] generic tasks. Many who have already captured extensive expert protocols might consider using these recent advances in memory modeling to test the adequacy of the models as an extension of their earlier research. Additionally, such models of memory might allow researchers to be more aggressive in terms of addressing complex instances of judgment of compare auditor judgment.

3) Knowledge Validation—This is possibly the least developed, yet most critical area of ES research. Unless we are confident an ES accurately represents the expertise, it will be difficult to convince the audit profession to extensively utilize such systems. To date, most studies have simply asked other experts to confirm that the final result of the system appears reliable, even though there are a variety of validation techniques in existence that could be applied. No single test is sufficient. Little has been done, beyond the work of Meservy, Bailey, and Johnson [16] to test the internal validity of an ES in auditing. Future research using advanced experiments such as described may lead not only to better ES’s that avoid Garden Path errors, but may also improve the underlying domain knowledge.

4) Tests of Scalability—Although progress in each of the three areas listed above is critical, until these ideas are exposed to the complexities of real audit engagements they will likely remain academic theories. As such, they will simply be bantered about in research journals having little real value to the profession and remaining unproven as to their real world validity and usefulness. This will require a close cooperation between professionals and academicians—something which has occurred in only a few situations to date. Although a great deal can be done at the university, until practice and academics develop a stronger relationship, both practice and academic efforts will be lacking.

As with many other expert judgment tasks, the study of audit expertise using the ES approach has provided valuable insights into the expert judgment involved in financial auditing. Nonetheless, a great deal of work is needed to provide useful tools to leverage audit expertise and provide useful theories of audit judgment by which to train future auditors. We see this as an exciting opportunity for both academics and audit practitioners.

REFERENCES


Eric L. Denna received the B.S. degree in accounting and the M.S. degree in accounting information systems from Brigham Young University, Provo, UT, and the Ph.D. degree in accounting and in formation systems from Michigan State University East Lansing, where he was the Coopers & Lybrand Doctoral Scholar.

He is currently an Assistant Professor in the Marriott School of Management’s School of Accounting and Information Systems. He has worked for the Decision Support Group of Coopers & Lybrand in New York City, with the Micro Systems Group of Ernst.