Application of Artificial Intelligence to Accounting, Tax, and Audit Services: Research at Brigham Young University

RAYMAN D. MESERVY, ERIC L. DENNA, AND JAMES V. HANSEN
Brigham Young University, Provo, Utah 84602

Abstract—The domains of accounting, tax, and audit services have become much more complex during recent years. This development has generated a need for leveraging professional expertise and judgment processes. In an attempt to meet this need, there are a number of projects in the Marriott School of Management, Brigham Young University (BYU) that apply artificial intelligence techniques and methods to augment professional expertise. Much of the development and application of artificial intelligence in accounting, tax, and audit services has been in the form of expert systems. This paper examines some of the early expert system work by BYU faculty and provides a framework to guide current and future efforts.

1. INTRODUCTION

In recent years there has been a growing interest in the development of expert systems (ES) for various accounting and financial problems. Several of the large accounting firms have actively pursued research and development of such systems to augment the professional's knowledge and judgment. Current efforts are also exploring more powerful methods of knowledge acquisition and knowledge representation, as well as performance evaluation (Denna et al., 1990). At Brigham Young University (BYU), there are a number of projects aimed at applying artificial intelligence (AI) techniques and methods to augment professional expertise. Some projects have resulted in commercial successes, some have furthered domain knowledge, while others have examined the applicability of AI methods to accounting, tax, and auditing problems.

This paper will first present some of the early expert systems developed by faculty members, then provide a framework that is guiding current efforts, and conclude with a discussion of planned activities.

2. EARLY SOFTWARE AND EXPERT SYSTEMS

Brigham Young University faculty have been noted for their applications' research, entrepreneurial spirit, and commercial software application spinoffs that have become commercial successes. Among these are WordPerfect and Novell, both internationally headquartered near BYU. WordPerfect, the word processing giant, was formed by two BYU faculty, Bruce Bastian and Alan Ashton. Bruce, the director of the BYU marching band, contacted Alan, a computer science professor, to computerize marching drills for the band. Later they wrote a word processor to fulfill practical needs for Orem City and BYU, and WordPerfect was born from their garage. Another success story resulted from three graduating students who recognized and solved some networking problems, resulting in today's LAN giant, Novell. Although these and other commercial software successes were not based on ES, they fostered an atmosphere supportive of new technology applications, such as ES. One related project resulted in PLANMAN.

2.1. PLANMAN

PLANMAN is an expert/decision support system for personal financial planning developed by Lynn J. McKell, a systems' professor in the School of Accountancy, and James W. Jenkins, a finance professor in the Marriott School of Management. PLANMAN was designed as a decision support system to be used by qualified professional planners. Using quantitative and qualitative input from both the individual and the planning professional, PLANMAN does diagnostic analysis and makes recommendations in the following areas:
1. income tax planning,
2. cash flow planning,
3. portfolio and debt management,
4. wealth growth management (i.e., planning for re-
2.2 EDP-XPERT

EDP-XPERT (Hansen & Messier, 1986b) is an expert system which is intended to assist computer audit specialists (CASs) in making judgments of the reliability of controls in advanced computer environments. The initial knowledge base was refined with the help of a senior CAS resulting in a system containing 133 rules structured into four goals (the reliability of supervisory, processing, graphics, and databases; and (5) practical solutions to complex expert system concerns such as: 1. legal liability, 2. how much control over ES logic to trust to users, 3. how to limit control over ES logic, 4. what input/output interfacing methods are practical.

PLANMAN is a commercial success as a product, and has also been adopted by several financial, insurance, and consulting firms such as Arthur Anderson.

Other early ES experiments and developments by current BYU faculty members were meant to increase domain knowledge and investigate the methodologies of expert systems. Two of these rule-based systems are EDP-XPERT by James V. Hansen (BYU) and Bill Messier and ARISC by Rayman D. Meservy (BYU).

2.3. ARISC

ARISC (Audit Risk of Identified System Controls) was developed as a simulation model, simulating auditor's evaluation of internal controls, a requisite of every audit performed by CPAs. 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 result of the task (and the output from the expert system) includes (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 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. Some of the more important aspects of expertise were discovered during the tuning, 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 guide their thinking about each case. Another important issue addressed by this study is verification, discussed later.

The next section presents a framework classifying our current work and suggesting future research.

3. A FRAMEWORK FOR EVALUATING ES WORK

For the purpose of classifying our current work in accounting, tax, and audit services and guiding future efforts, we propose a three-dimensional framework as shown in Fig. 1 (see Denna et al., 1991). The framework consists of three axes which represent three main areas of ES research at BYU. 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 knowledge elicitation from experts, a problem pervading ES development in all domains of specialization. Wood (BYU) and Ford define the goal of knowledge acquisition as, "knowledge should be elicited, organized, and documented with minimal concern for how it will be implemented in a working system." Similar objectives have been proposed by Alexander et al. (1987), and Johnson (1989).

The axis labeled KR represents work dealing with the problem of knowledge organization and representation. 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 (Riesbeck and
Schank, 1989) and generic tasks (Chandrasekaran and Mittal, 1982).

The axis labeled KV represents work dealing with the problem of knowledge validation. Knowledge validation focuses on evaluating various computational models of judgment in an effort to determine the validity of the model. This is an area which is crucial to ES, but has not been fully explored. Most work in this area has focused on the simple evaluation of output from an ES. Only one study in auditing has provided an in-depth analysis of the reasoning processes and explored 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 (1986) 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's applicable to the real problem." Complexity in our framework is represented by efforts which either extend understanding along one of the axes, or which 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.

3.1. Knowledge Acquisition

Wood (BYU Psychology Department) and Ford (1990) have categorized knowledge elicitation methods into four classes: descriptive elicitation, structured expansion, scripting, and validation. Descriptive elicitation focuses on helping an expert identify his abstract declarative knowledge. The methodology attends to the categories, objects, models, and other conceptualizations used in problem solving. Structured expansion attempts to expand and integrate fragmented portions of the expert’s knowledge derived from descriptive elicitation. Such methods have emanated for the work of ethnographers and are not widely known to knowledge engineers in the accounting, tax, and auditing domains. Scripting requires capturing the thought processes of the expert while solving a problem and is most often accomplished through methods of protocol analysis. Validation refers to building a model and tuning that model with an expert. Previous experiments in knowledge acquisition are included in the work of Hansen and Messier (1986a) and Meservy, Bailey, and Johnson (1986). Denna, Hansen, Meservy, and Wood are currently conducting various experiments aimed at enhancing knowledge acquisition from experts in the accounting, tax, and auditing domains.

Automated methods of knowledge acquisition are also being explored. As a follow-on to their work on EDP-XPERT, Messier and Hansen (1988) 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 ES, particularly those exhibiting consistent structure in the archival data.

Greene, Smith, and Meservy (1989) have conducted experiments comparing genetic algorithms with other machine learning techniques and are doing further experiments using the sales tax audit selections data from the Commonwealth of Pennsylvania. The results appear to be quite promising. In addition, Hansen and Meservy are comparing various neural net algorithms with the data. We believe that a combination of several
of these activities and techniques may be useful in future ES acquisition.

3.2. Knowledge Representation

This section represents what we consider to be "second generation" work on the issue of representing expert reasoning. We characterize "first generation" ES development projects as those which use a production rule paradigm exclusively.

Recent projects have taken advantage of more advanced representation techniques such as frames and object oriented programming. These studies focus on providing more robust and useful ES applications by developing more accurate representations of both the reasoning process and the underlying domain knowledge of the expert. Denna (1989), with the support of a grant from Coopers and 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 auditor determines the likelihood of material error (LME) for each financial statement assertion, formulates a preliminary audit approach for gathering evidence, and aggregates these into an overall audit program. APE (Audit Planning and Evidence) is currently limited to planning the audit of a large retail grocery client. As illustrated in Figs. 2 and 3, 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. Figure 2 shows the first two levels of abstraction while Fig. 3 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.

Meservy is also currently working with former colleagues at Carnegie Mellon University (Amiri et al., 1990) to combine process modeling concepts from TICOM (see Bailey et al., 1985) with deontic reasoning (authorization and company policy) and ARISC (Meservy, Bailey, & Johnson, 1986), an ES which evaluates internal control. The underlying system knows a lot about how a business functions in addition to the instantiation of the system being modeled and evaluated.

Denna, Hansen, Meservy, and Wood (all of BYU) have applied Case-Based Reasoning (CBR) algorithms to an audit planning problem. CBR applies complex reasoning processes, such that the CBR system can learn from past examples and reason by analogy. They are currently testing other accounting domains which require experience and learning from previous case histories such as is common in tax research.

3.3. System Validation

Initial knowledge validation is closely intertwined with knowledge acquisition activities. During development, experts are often asked to validate individual bits of information and the way in which the information is organized. The expert then provides feedback on what might be done to improve the system. Yet at some point, the knowledge engineer decides that the system has reached completion of the initial developmental phase and is ready for testing. Unfortunately, developers have not been able to formulate any single critical experiment to validate systems, but rather rely on a variety of tests which address various aspects of the knowledge. These tests may be organized into three categories: external, internal, and advanced experiments.

Expert system validation to date has relied mainly on having the initial expert or other professionals watch the system as it solves problems, and then expressing an opinion as to how well they thought the system performed. We refer to such a process as a type of external validation of system outcomes, but generally feel that this is not sufficient. Meservy (1985) suggests that external validation may also involve a form of Turing test, with blind peer review of the ES along with solutions by experts on new real-world problems that the ES has never encountered. Such peer reviews may take several different forms, judging for completeness, effectiveness, and whether they personally agreed with the results. (see Meservy, Bailey, & Johnson, 1986, for examples.)

Internal validation, or tests of process quality and data usage, is established by determining that the inferences made by the system are not only legal, but are of the variety experts make. For example, a chess-playing ES would not only make legal moves, but the moves should be similar to what a chess expert would have done. Methods that have been used include a top-down global analysis and bottom-up knowledge state/process analysis. The top-down analysis evaluates major hypothesis generation and the processing scripts or views, while the bottom-up analysis examines similarities of knowledge states and types of processes. (See both Hansen and Messier, 1986b; and Meservy, Bailey, & Johnson, 1986.)

Advanced experiments in knowledge validation 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 expert may fail (Denna, Hansen, & Meservy, 1990). Once identified, additional knowledge may then be added which would make it possible to avoid such errors.

4. SUMMARY

This article has presented some of the research that is being done in artificial intelligence and particularly in expert systems at the Marriott School of Management along with other faculty at Brigham Young University. Our approach has been to apply AI methodologies to decision problems in the accounting, tax, and auditing domains. There is a synthesis of productivity in having several management and nonmanagement faculty working together in these more narrowly defined domains.
REFERENCES


