1-1-1986

Discussant's response to "Interim report on the development of an expert system for the auditor's loan loss evaluation"

William F. Messier

Recommended Citation

Auditing Symposium VIII: Proceedings of the 1986 Touche Ross/University of Kansas Symposium on Auditing Problems, pp. 182-188;

Discussant's Response to "Interim Report on the Development of An Expert System for the Auditor's Loan Loss Evaluation"

William F. Messier, Jr.
University of Florida

It is a pleasure for me to comment on the paper by Kelly, Ribar and Willingham. As someone who has spent a major part of the last five years in expert systems research, it is good to see this technology begin to impact audit practice. I will make one caveat before I proceed. There are times in my discussion where I raise questions about or criticize this work. In those instances, please recognize that I am fully cognizant of the difficulties of doing this type of research and, more importantly, sympathetic with those difficulties.

Before discussing the specifics of the paper, I would like to make a few overall comments on this work. The research by Kelly, et al. is noteworthy for three reasons. First, it demonstrates the application of expert systems technology to an important audit problem, the assessment of loan loss reserves. Application of this technology to auditing is important because public accounting firms are facing a more competitive environment that will require audits to be conducted with the same level of effectiveness but with increased efficiency. Relatedly, the types of decisions auditors face today are more complex (e.g., EDP auditing) and require more expertise. Expert systems are intended to assist with such complex decisions.

Second, since this is a proprietary system, it is especially noteworthy that Peat Marwick is willing to share the details of the system with academics and practitioners. Until recently, many public accounting firms were unwilling to share these types of developments with the public. I make a point of this because I believe that it is important to our profession to disseminate research and that it should be a two-way street.

Finally, this paper shows that public accounting firms are willing and able to build expert systems. This realization was also brought home to me at the recent expert systems conference at the University of Southern California where Coopers & Lybrand demonstrated an expert system for deferred taxes [Shpilberg and Graham, 1986]. For that project, Coopers & Lybrand hired a full-time computer scientist to assist in developing their system. By bringing their enormous resources to bear on the problem, they were able to construct the system in approximately a year. Peat Marwick has been able to do a similar sort of thing with CFILE. In spite of the generous support from Peat Marwick's Research Opportunities in Auditing grants, most academics who are developing expert systems have faced much longer development times. This raises a question about whether academics any longer have a competitive advantage in developing these systems or whether some type of joint collaboration is necessary. If the second alternative is the most appropriate, then academics have to ask themselves whether this work is research or consulting. My thoughts on this question are that as long as sound research issues are addressed and no limitations are placed on the dissemination of those results, the work qualifies as research.

My specific comments on the paper will center on three topics: the rationale for expert systems, the CFILE model, and the field testing.

Rationale for Expert Systems

Successful development of expert systems requires that certain characteristics be present in the problem domain. These include acknowledged experts, an ability to extract their knowledge, some measure of the correctness of the decision, and manageable problems with high payoffs. Kelly, et al. properly point out that auditing does not have all of these characteristics. For example, in auditing it is very difficult to state *specific* criteria by which to label someone an expert. This is unlike domains such as chess or certain specialities in medicine. Thus, there is some difficulty with identifying an expert(s) to assist in designing a particular system. This is further compounded by the fact that auditing is "process oriented" and two experts may solve the problem differently. Hansen and I [Hansen and Messier, 1986a, b] have encountered some of these problems in developing and testing EDP-XPERT.

In addition, the ability to extract the necessary knowledge from the expert is perhaps the most difficult and time consuming part of constructing an expert system. Given that in auditing we have difficulty identifying an expert and the fact that two experts may solve the problem in different ways, knowledge acquisition poses a major hurdle for constructing expert systems.

Finally, the fact that many audit judgments do not have outcome feedback about the correctness of the decision is an important characteristic for expert system development in auditing. Kelly, et al. argue that this is "a knowledge representation issue that will clarify itself through knowledge engineering tasks." I am not convinced that this is true in all instances. In the loan loss reserve situation, the auditor will get feedback (not immediately, of course) about the collectibility of this short term loan. This situation is probably not true for areas such as the reliability of internal controls (manual or EDP), evaluation of inherent risk, analytical review, or similar areas where expert systems are currently being developed. The absence of a true criterion value by which to evaluate the goodness of the expert system's decision poses the greatest difficulty in validating expert systems in auditing. Note that in validating CFILE, the system's judgments were compared against the expert or the user's unaided judgment rather than against the true outcome of the loan.

The comments just made should not be interpreted as an indictment of the use of expert systems in auditing. They are intended to point out that construction of an expert system is not an easy task. Individuals and firms who decide to build such systems must recognize that this process is long and costly in terms of both time and money[1]. However, I agree with Kelly, et al. that expert systems technology does offer some significant benefits for public accounting firms.

In that vein, I would like to discuss the benefits that Kelly, et al. believe will result from the application of expert systems. They suggest five potential benefits: support of field work, diffusion of knowledge, uniformity of documentation, staff training aids, and research. I will limit my comments to diffusion of knowledge and research.

Expertise in any discipline is usually a scarce commodity. The complexity of auditing today requires individuals within firms to develop expertise in specific areas. Kelly, et al. mention this in the area of banking. Most of my expert systems research has involved the work of computer audit specialists. I am sure we could identify a long list of audit areas where experts exist, and I am fairly certain that future development in auditing will only increase this trend. My point here is that perhaps the greatest potential benefit of expert systems is the ability to provide the expert's knowledge to novices. In auditing, most of the expert's knowledge is not textbook knowledge. Most of it is experiential knowledge accumulated over many years. If firms are able to capture this type of knowledge and make it available throughout the firm, there may be cost savings and improvements in audit effectiveness and efficiency.

Kelly, et al. underplay the role of research in designing expert systems. As they point out, the problems chosen for expert systems development are generally not well understood and the knowledge engineering process can contribute to our understanding. My experiences indicate that the process of developing the knowledge base can provide a major contribution to our understanding of the specific problem and auditor decision-making in general. For example, a number of audit researchers have used a Bayesian formulation for modeling auditor judgment. However, two recent studies by Biggs, Messier and Hansen [1987] and Biggs, Mock and Watkins [1986] that were conducted to develop a knowledge base for expert systems seem to indicate that expert auditors do not follow a Bayesian revision process. Instead they seem to use "reasoned assumptions" and "analogies" to arrive at decisions. This finding not only has implications for modeling auditor judgment but also the type of model used in the inference engines of expert systems. So from a research perspective, I think that construction of an expert system for a complex problem will contribute immensely to our understanding.

The CFILE Model

The section of the paper which describes the CFILE model leaves a number of important questions unanswered. For example, what expert system shell was used to develop CFILE? On what basis was this shell chosen? What type of evidence accumulation model is contained in the inference engine? Is this evidence accumulation model appropriate for auditing? How many rules are contained in the system? Additionally, there is little discussion of how the knowledge was captured from the expert.

The answers to these questions would be helpful to our understanding of the system. For example, the answer to the question concerning the type of model used in the inference engine. It is not clear in the expert systems literature [Gordon and Shortliffe, 1985; Shafer and Srivastava, 1986] which type of model should be used to accumulate evidence in problem domains where some degree of uncertainty exists. Information about the model would provide insight into the reliability of the system's reasoning process. Based on the presence of "CF" factors in the rule shown in Figure 7, I can speculate that the inference engine contains some type of Bayesian process. This is exactly the model that has come under recent criticism in the expert systems literature. However, I am sympathetic with the authors. When expert system developers decide to use an expert system shell, their choice of models is severely limited.

Similarly, if we knew how many rules were present in the current prototype we would have an idea about the number of questions the system asks the user. This question is important because if the system contains a large number of rules there is always some question concerning the consistency of the rule base. The rule contained in Figure 7 is numbered 3850. I am quite certain that the system does not contain that many rules. Obviously, there is some numbering convention within the system. However, a close examination of the rule contained in Figure 7 would suggest that the system does contain a large number of rules. Rule 3850 contains four antecendent conditions and there appears to be five possible categories (e.g., very strong, strong, etc.) for each antecedent. This would suggest that there are 625 possible combinations of this rule.

The system does appear to have some important capabilities. The questions posed by the system are asked in an abbreviated form for users familiar with the system. The less experienced user is assisted by help screens which provide more information on the question. This feature should improve usage of the system. It should also increase consensus in the way the questions are answered since there will be less chance that two users will misinterpret the question and respond differently even though the circumstances are similar.

Two other features appear quite interesting. The ability to do limited sensitivity analysis should prove very useful. Since expert systems are intended to support rather than replace experts, the ability to do this type of analysis should lead to improvements in decision making. I also found the final report generated by CFILE to be very comprehensive. The report not only contains the conclusions about the reserve but it also contains important information on the variables that led to that conclusion. Thus, the report can be used for audit documentation.

I am a little disappointed with the system's explanation capability. Early research demonstrated that experts were interested not only in a system's conclusion but how the system arrived at the conclusion. I suspect that expert auditors will require a similar capability.[2] CFILE's ability to respond to why a question is asked is typical of most expert system shells. The response is a limited parse (see Figure 7) of the rule that led to the question. It would be more helpful to the user if the system could provide an explanation in a more user-friendly manner.

Field Test of CFILE

As I mentioned earlier, the validation of expert systems in auditing will represent one of the major challenges for implementation of such systems. Before expert systems will be adopted for use in the field, public accounting firms will have to be sure of the system's reliability (i.e., ability to yield a correct answer a high percentage of the time). The difficulties with validating expert systems were alluded to earlier. In many auditing areas, the outcome to a particular problem is not immediately known with certainty or may not be known for some time in the future (usually after the audit report has been issued). As a result, it is not possible in many audit settings to test the correctness of the expert system's decision. The alternative in these situations is to compare the system's conclusions with those of the expert. Note that this testing is similar to earlier behavioral research on consensus.

I will not take exception with the fact that the field test of CFILE did not use a formal experimental design. Buchanan and Shortliffe [1984] have suggested that the validation process must be undertaken throughout the life of the system and that the evaluations should get more formal as the expert system is developed further. The authors admit that the system is still in the early stages of development.

My comments are first directed at some relatively simple changes that could have been made or added to the testing. First, it would have been interesting to compare the results from the three subjects (expert partner, second partner, and senior) with the conclusions reached by the audit teams on each of these clients. Second, it should have been possible to use loans from previous years where the client had already determined the amount collectible. In this instance, the subjects' aided judgments could have been compared to a known criterion. Both of these extensions would have provided increased external validity for CFILE's performance.

In terms of the results of the field testing, the system does an excellent job of replicating the expert's judgments. CFILE agreed with the expert in nine of the ten cases. However, performance decreased with the second partner (69 percent) and the senior accountant (62 percent). Additionally, most of the favorable performance is found on cases where no reserve is the suggested answer. Seven of the ten cases evaluated by the expert without CFILE result in a no reserve answer and the second partner agreed with all of those cases.

I am not sure how valid it is to look at the reserve versus no reserve results. It seems to me that differences in the size of the reserve is an important criterion to measure because it relates to materiality. Certainly, the results that examine only the cases with reserves are not very encouraging. However, these results are quite limited since they only include three cases for the partners.

Kelly, et al. contend that the differences in the performance of CFILE can be attributed to two possible causes: (1) interface or communication problems and/or (2) the depth of the knowledge base. The first cause is correctable but may be more difficult than the authors speculate. Hansen and I have encountered this problem in some of the recent field testing of EDP-XPERT. Sometimes the wording of the question (and its explanation) can cause the user to misinterpret what is being asked or cause the user to make an incorrect assumption. We might expect this type of problem in an area such as auditing where there are no "natural laws." I do not know if there is an easy solution to this problem. Adequate training with the system may be one alternative.

The second cause, depth of the knowledge base, is an even more difficult problem. The authors acknowledge that the current prototype has a number of limitations (e.g., cash flow and bankruptcy analysis) and that there is a need to refine the knowledge base to handle more basic information. However, the results indicate that the system handles easy cases (i.e., no reserve) quite well. The difficulty occurs when the system encounters a case where more judgment is required (i.e., the situation where a reserve for the loan is required). It is such loans that are of real interest to the auditor. From my perspective, it appears that it will be necessary to do some detailed refinements to handle the more difficult cases. As a suggestion for future testing, it would seem appropriate to add more cases where a reserve is necessary. It is important to test the boundaries of the system's capabilities.

It is unfortunate that we do not have more detail on the protocol data. It would be interesting to compare the decision processes of the two partners, both with and without the use of CFILE. Such an analysis might provide important insights into expert auditor decision-making.

This last comment raises an important area for future research: auditor expertise. While a lot of effort has been devoted to developing expert systems, relatively little research has examined expertise. There are a number of questions that we are unable to answer at this time. For example, how does an expert become an expert? We know very little about this process. How do expert auditors categorize their specialized knowledge? Do experts use different types of memory structures than novices? Answers to these types of questions will improve our understanding of expert decision-making and may contribute to building better expert systems[3].

Summary

Kelly, Riber and Willingham should be commended on this work. Construction of an expert system is a long process full of many ups and downs. I look forward to seeing the results of the ongoing development of CFILE.

End Notes

1. It should be recognized that there will also be ongoing maintenance costs for updating the knowledge base after the system is introduced into the field.

2. In the questionnaire used in Hansen and Messier [1986b], the question "An expert system when fully developed *should* be able to explain decisions to auditors" received the second highest agreement score: 1.59 on a -2 (disagree) to $+2$ (agree) scale, out of 13 questions.

3. There is a growing recognition in the expert systems literature [Buchanan and Shortliffe, 1984] that a better understanding of *how* experts solve problems may be necessary before expert systems achieve expert level performance.

References

Biggs, S.F., W.F. Messier, Jr., and J.V. Hansen, "A Descriptive Analysis of Computer Audit Specialists' Decision-Making Behavior in Advanced EDP Environments" *Auditing: A Journal of Practice & Theory* (Spring 1987), in press.

Biggs, S.F., T.J. Mock, and P.R. Watkins, "A Descriptive Study of Auditors' Use of Analytical Review in Audit Program Design" working paper, University of Connecticut (1986).

Buchanan, B.G. and E.H. Shortliffe, *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA: Addison Wesley (1984).

Gordon, J. and E.H. Shortliffe, "A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space" *Artificial Intelligence* (July 1985), pp. 323-58.

Hansen, J.V. and W.F. Messier, Jr., "A Knowledge Based, Expert System for Auditing Advanced Computer Systems" *European Journal of Operational Research* (1986a), pp. 371-79.

Hansen, J.V. and W.F. Messier, Jr., "A Preliminary Investigation of EDP-XPERT" *Auditing: A Journal of Practice & Theory* (Autumn 1986b), pp. 109-23.

Shafer, G. and R. Srivastava, "Managing Uncertainty in Expert Systems: A Perspective for Auditing" working paper, University of Kansas (January 1986).

Shpilberg, D. and L.E. Graham, "Developing ExperTAX: An Expert System for Corporate Tax Accrual and Planning" *Auditing: A Journal of Practice & Theory* (Fall 1986), pp. 75-94.