1-1-1986

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

Kirk P. Kelly

Gary S. Ribar

John J. Willingham

Recommended Citation

Auditing Symposium VIII: Proceedings of the 1986 Touche Ross/University of Kansas Symposium on Auditing Problems, pp. 167-181;

7

Interim Report on the Development of an Expert System for the Auditor's Loan Loss Evaluation

Kirk P. Kelly
Gary S. Ribar
John J. Willingham
Peat, Marwick, Mitchell & Co.

Introduction

The Audit Research Group at Peat, Marwick, Mitchell & Co. has been interested in Artificial Intelligence (AI) and Expert Systems for a number of years. Under the auspices of the Research Opportunities in Auditing program, we have funded a number of academic research projects on the application of AI to the audit task. With the growing interest in the field and the advances in technology, it was decided to undertake a project oriented toward the development of an application model. The initial thrust was to build a prototype model for test and evaluation with the implicit intent that the model would eventually be developed into a useful audit tool for field work. This paper reports on that project in terms of the rationale for the project, the current status of the project, and the future directions for this project.

Rationale for Expert Systems

The rapid advances in computer technology and ensuing applications require that those engaged in the accounting and auditing profession be involved in exploring new application opportunities. Artificial intelligence and expert systems are clearly in the forefront of these technologies; however the conventional wisdom of expert system developers suggests that considered applications ought to be limited to environments that exhibit certain characteristics. For example, it is suggested that there should be clearly definable experts in the problem task, that there should be appropriate measures of correct vs incorrect judgments, and problems should be small yet have a high payoff.

The auditing environment has some unique characteristics that tend to make it a less likely candidate for successful deployment of expert systems. For example, many areas of auditing do not have a feedback mechanism that allows for determination of correct vs incorrect decisions. Auditing is more process oriented than results oriented, wherein the quality of work is judged not by results, but by traces of process to be found in the work papers. Moreover, auditors learn acceptance of processes that may diverge significantly from their own as long as they "appear reasonable." A side effect of this is that we do not have a set of clearly defined "experts" whose technical skills find "material errors" in an audit with a significantly higher frequency than other auditors.

While these factors may mitigate against using expert systems, we do not believe they are fatal. The issue surrounding the feedback and correctness of judgments in the audit environment is, we believe, a knowledge representation issue that will clarify itself through the knowledge engineering tasks. We also believe that there is expertise, albeit spread out, and that the professed need for a singular expert is a knowledge engineering problem that can and will be addressed pragmatically as the art of knowledge engineering advances.

We believe that AI technology offers the following significant benefits:

1) *Support of Field Work:* There are any number of applications for the AI technology that, when harnessed, can be used as tools in the support of auditing field work, thereby freeing the auditor from many of the more mundane tasks and making the work of the auditor significantly more interesting. At the same time, the technology can lead to a greater consistency in the quality of field work, and hopefully reduce the time requirements for the field work.

2) *Diffusion of knowledge:* The complexity of modern auditing, as dictated by the complexity of modern business, leads to areas of audit specialization. Expertise relates to certain industries, such as banking or oil and gas, and across industries as in EDP auditing. Even within industries, there are pockets of expertise, e.g., in the banking industry there are those who are expert in auditing community banks, moderate size banks and the extremely large banks. Additionally, many banks themselves perform in specialized industries, e.g., agricultural banks, oil and gas, etc. The data or information available in these varying circumstances require varying types of expertise. It is very difficult if not impossible for one auditor to be an expert in all these areas. By capturing the expertise in specialized areas, however, we can provide knowledge where the expert is not available.

3) *Uniformity of documentation:* Through the proper design of an expert system, the required documentation to support a given judgment can be automatically provided as the output of the judgment exercise and included in the working papers of the audit. The expert system not only provides uniformity of documentation, but also frees the auditor from another time consuming and costly chore.

4) *Staff Training Aids:* Training is an extremely costly investment in a large public accounting firm. Technological advances are providing the potential vehicles for both increasing the effectiveness of training while concurrently reducing the huge costs involved.

5) *Research:* We should not forget the role of research in the design of expert systems. Designing expert systems is research oriented, in that problems chosen are seldom well enough understood to be solved algorithmically. The knowledge engineering process can and should lead us to a greater understanding of the problems, thereby advancing our knowledge.

Based on the above reasoning, a decision was made to embark on the development of an expert system that would at once provide insights into the development process, provide knowledge about resource requirements, and produce a useful audit tool.

Selecting a Project

Since the project to be developed had multiple objectives, it was agreed that the project should be of a very limited scope and nature, yet have the potential for a very high payoff. Additionally, since we were not overly committed to the expert system technology, we wanted to attempt the development at a minimal investment. The decision was therefore made to develop the model in a microcomputer environment using commercially available development shells.

Hoping for the potentially high payoff, we wanted to focus on a problem that was meaningful to our firm's audit practice and yet might be successful given the constraints we were imposing. Since bank audits are a large part of our audit practice, it was decided to focus on a problem in that area. We found that there was significant support from bank audit partners in the form of enthusiasm and willingness to invest expert bank auditors' time and cooperation. This was considered important, since we knew the development work would require a considerable amount of time and effort from bank experts at no small cost.

The next issue was to settle on a specific problem. We were guided by two considerations: 1) the problem had to be small enough to accomplish within a reasonable time, and 2) it had to be sufficiently important within the context of a bank audit. An area of bank audits that filled both of these requirements was the loan loss evaluation, the process of estimating the dollar amount of the reserve for the bank's portfolio of loans. This problem is basically a classification problem, which is a type of problem that has been successfully attacked by rule based systems before. (Most commercially available development tools for microcomputers are rule based.)

Project Description

Since we did not have an in-house AI capability for the development of such a system, we contracted the project to an outside consultant. The consultant's project proposal suggested the following stages of development:

1) Review current literature.
2) Develop a preliminary model of the loan loss evaluation process.
3) Implement the preliminary model as a computer program.
4) Extend knowledge acquisition to include the process of expert loan evaluation.
5) Combine knowledge into a final task expertise model and complete prototype expert system.

The proposal initially indicated that the above stages would require nine months to complete, employing one full-time consultant with the availability of audit experts in the loan loss evaluation task. To date we are somewhere in the fourth stage. What follows is a description of our model and how the system works.

Description of Model

For ease of reference, we have named the model CFILE, for *c*redit *file* analysis. The current working model is based on the conceptual model shown in

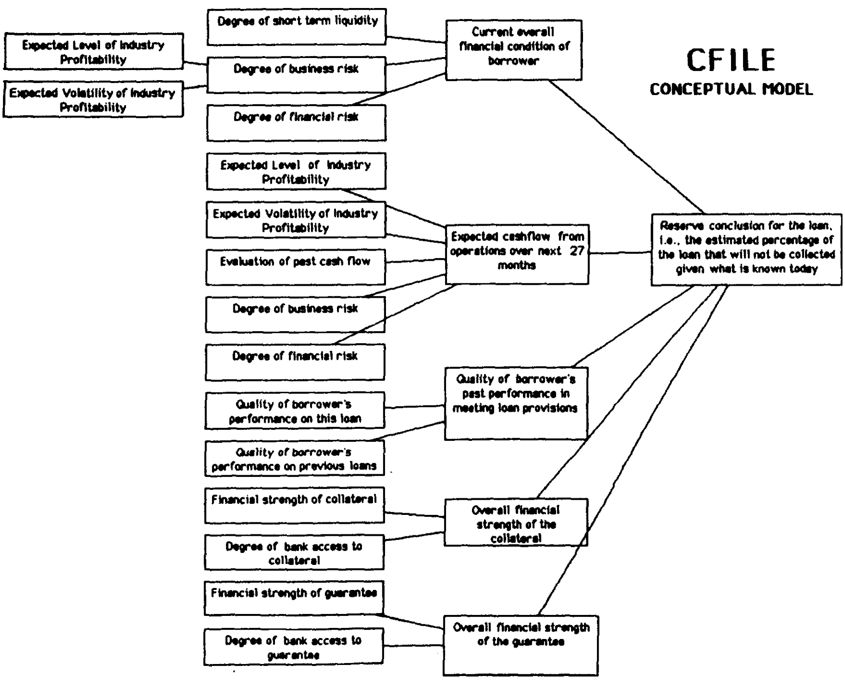
Figure 1. The model is modularized and illustrates the various factors considered when making the reserve judgment. The first column of factors to the left of the reserve conclusions are 'level one subgoals' and the second column of factors are 'level two subgoals' which affect the level one subgoals. For example, the conclusion on the current financial condition of the borrower is based on conclusions concerning the borrower's short term liquidity, finanical risk, and business risk. These judgments are reached internally by the model with the exception of the industry profitability and volatility, which temporarily are user inputs.

The consideration underlying the control structure of CFILE is efficiency. Efficiency is often considered one of the hallmarks of the expert. Like an expert, the model is designed to arrive at a conclusion as soon as possible with the minimum amount of information.

A session with CFILE begins with screens explaining the purpose of CFILE and what it will do. Immediately following this explanation, the user is asked for some basic information about the loan including its size, due date, and what kind of collateral and/or guarantees exist relating to the loan.

What CFILE asks next depends on the answers to the initial questions. If, for example, it is indicated that there are bank deposits pledged as collateral, CFILE will ask a series of questions about those bank deposits. These include questions about both access and financial strength, which are the two 'level two subgoals' relating to collateral. CFILE will want to know whether or not the bank has the legal right to dispose of the collateral in the event of a default. It

Figure 1



might also ask if those bank deposits were pledged as security for another loan. If the model concludes that there is adequate access to those deposits and their strength is sufficient to cover the loan, the analysis would stop with a no-reserve decision.

If the bank deposits were not sufficient, the model would start dealing with the three 'level one subgoals' that are needed to perform an analysis on an unsecured loan: current financial condition, overall loan history, and expected net cash flow. The model would ask the usual questions about hard data such as the current ratio of the borrower and would also ask about soft information, such as whether or not the borrower is planning any major projects that are going to be financed through the use of current assets. Again, how many of these subgoals would be pursued and to what extent would depend on the situation. For example, if the loan were due in the next 12 months and the borrower had a very strong current financial condition, no reserve would be necessary and the system would conclude without asking any questions about loan history or expected cash flow.

The system has some other interesting features. In general the questions are asked in abbreviated form. This is useful for the experienced user who will be familiar with the system. For example, the question about major projects alluded to above would appear as illustrated in Figure 2. However, help screens are available to provide more details and guidance to understand the question. The help screen for the same question as shown previously appears in Figure 3.

Another feature of the system is the ability to do limited sensitivity analysis. It is possible for the user to see how sensitive the conclusion is to a particular question. For example, one might be interested in determining the impact of the loan officer's opinion of the borrower's liquidity (see Figure 4), given an otherwise constant set of input judgments.

Figures 5 and 6 illustrate the conclusion reports provided by the model. Both conclusions came from identical information except for the response to the question noted in Figure 4. One can see that, in this case, the answer to the question had a fairly substantial impact. There is a difference in the evaluation of current financial condition which leads to different conclusions. In one case, we find an evaluation of the current financial condition of the borrower as weak and a conclusion of a 25 to 34 percent reserve before considering collateral. In

Figure 2

PMM—CFILE Preliminary version 2.02 November 25, 1985

Select what describes:

current assets used for new commitments

MMM no

yes

2 UNKNOWN 3 REPORT 4 EXPAND 5 MENU 6 HELP

Figure 3

PMM—CFILE

Based on your judgment, is there a significant chance the borrower will use a substantial amount (i.e., at least 25 percent) of current cash, accounts receivable and marketable securities or incur a significant amount of new short term liabilities for commitments to finance a major new project?

A major new project could be an acquisition, stock repurchase, an expanded advertising campaign or plant expansion program. A yes response would also be appropriate here if the borrower is involved in a continuing problem situation (e.g., a legal dispute) such that it is possible (FASB #5) that a new significant liability will emerge for the borrower.

enter no — if any new commitments will not use significant current assets or generate significant new current liabilities.

enter yes — if new commitments will use significant current assets or generate significant new current liabilities.

2 RESTART 5 GO BACK 6 HELP 7 EXIT

Figure 4

PMM—CFILE

Based on your judgment, if a set of financial statements were to be generated as of today, do the comments provided by the loan officer suggest to you that the loan officer, based on his/her knowledge of the borrower's current financial condition, believes the borrower is in a strong, moderate or weak short term liquidity condition?

enter

strong — if the loan officer believes the short term liquidity condition of the borrower is strong

moderate — if the loan officer believes the short term liquidity condition of the borrower is moderate

weak — if the loan officer believes the short term liquidity condition of the borrower is weak

2 RESTART 5 GO BACK 6 HELP 7 EXIT

Figure 5

PMM—CFILE Conclusions

Client Name: ABC BankCorp
Audit Period: 12-31-85

Borrower: XYZ Company

Analysis prepared by Joe Auditor on 12-1-85

Extent of available information is adequate.
Based on the available information, the following factors are indicated:

Industry prospects: expected profitability = moderate.
expected profit volatility = high.

Intermediate conclusions (scaled from very weak to very strong):
Current financial condition is weak.
Future cash flow potential is weak.
Borrower's past loan performance is moderate.

The amount of the loan is $150,000.

The loan is covered by bank deposits having an accessible value of $100,000. Of this, $90,000 is considered available to cover the loan.

No guarantee is available for this loan.

A reserve of 25 to 34 percent of the loan would appear appropriate, if it were unsecured. After considering the collateral available, no reserve would appear to be required.

I agree with the conclusion suggested by the system and the underlying reasoning.

________________________ preparer.

Figure 6

PMM—CFILE Conclusions

Client Name: ABC BankCorp
Audit Period: 12-31-85

Borrower: XYZ Company

Analysis prepared by Joe Auditor on 12-1-85

Extent of available information is adequate.
Based on the available information, the following factors are indicated:

Industry prospects: expected profitability = moderate.
expected profit volatility = high.

Intermediate conclusions (scaled from very weak to very strong):
Current financial condition is moderate.
Future cash flow potential is weak.
Borrower's past loan performance is moderate.

The amount of the loan is $150,000.

The loan is covered by bank deposits having an accessible value of $100,000.
Of this, $90,000 is considered available to cover the loan.

No guarantee is available for this loan.

No reserve appears to be required.

I agree with the conclusion suggested by the system and the underlying reasoning.

____________________ preparer.

the other we find a moderate evaluation leading to a no-reserve conclusion even before the collateral is considered.

This facility is useful to both user and developer. It gives the user, who is uncertain about the appropriate response, the ability to see the impact of alternatives without repeating a lot of data entry. It gives the developer a tool for testing the reasonableness of the rules in the system.

Perhaps the most important feature in this system is the user's ability to find out why a question is being asked. Through function key, one can look at the rule that has caused a specific question to be asked, and in turn ask about that rule. Figure 7 illustrates the screen that would appear asking about the loan officer's view of the borrower's liquidity. In this way it is always possible for the

Figure 7

PMM—CFILE Preliminary version 2.02 November 25, 1985
The highlighted fields indicate the antecedent
and conclusion being pursued.
The rule currently being pursued is:

RULE 3850
IF
MMM quick ratio is (are) weak
AND current ratio is (are) moderate
AND current ratio trend is (are) decreasing
AND loan officer liquidity judgment is (are) strong
THEN
stliquid is (are) very strong CF 0
AND stliquid is (are) strong CF 0
AND stliquid is (are) moderate CF 100
AND stliquid is (are) weak CF 0
AND stliquid is (are) very weak CF 0

2 ALL RULE 3 OR CLASS 4 FORWARD
5 GO BACK 6 HELP 7 EXIT

user to understand the line of reasoning that the system is using. This not only allows the user to understand the basis for the conclusion the system reached but facilitates review and avoids the blanket acceptance or rejection that is common with algorithmic systems. The model becomes a transparent box which is essential to the audit review process and it places the user in a position to be able to make constructive criticism, which may aid in further system development.

Limitations of Current Model

The current model has limited capabilities that have resulted from design decisions intended to keep the project manageable. CFILE applies only to loans due on demand or within one year and are either unsecured or secured by bank deposits or marketable securities. The model requires two years of audited financial information or three years of unaudited financial information from the borrower and is limited in its ability to perform and integrate cash flow analysis into its decision process. The model is further limited by its inability to deal with situations involving bankruptcy and liquidation analysis.

These limitations resulted from design decisions made early in the project and compose a major portion of the work yet to be performed. Again, our intent was to build a working prototype model that we hoped would be easily expanded to cover situations through the addition of modules to the knowledge base. It is envisioned that the prototype will then be of assistance in future knowledge engineering work.

With the prototype model working, it was decided that we should test the system against the modeled 'expert' to determine how well we captured the experts' decision model. A field test of CFILE was carried out in late February and early March of 1986.

Field Test of CFILE

For a number of reasons dealing with logistics, time constraints and purpose, the field test was not set up as an experimental design but rather as a pilot test to determine if we were on the right path with our model. It provided the opportunity to deal with actual loan files in bank audit environments and to compare how different auditors performed the tasks in process as well as judgment.

The testing was carried out at four of our client banks. Two of the banks are large regional banks and the other two are smaller community banks. A total of 16 cases were chosen either from client's listings of unsecured loans or with the assistance of the local audit team. First priority was given to loans which had a reserve allocated to them either by the audit team or by the bank's internal loan review department.

Each case was reviewed by three subjects, two at the partner level and one at the senior accountant level. The partners chosen were from our bank audit practice. One of the partners was the 'expert' employed in the development of the model. The other partner had only a cursory understanding of the model. The senior accountant had neither bank audit experience nor knowledge of the model to be tested. Our intent here was to see how much the model might assist the novice in the field and the senior accountant level is the appropriate level for performing this task during an actual audit.

Cases were reviewed first without the use of the model and then with the use of the model by each of the three people. Unfortunately, one of our partner subjects, the 'expert', was unable to participate at the first bank setting due to illness and therefore only evaluated ten of the 16 cases.

The results of the test are summarized in Figure 8. By way of explanation, CFILE uses nine reserve classifications expressed in percentage: no reserve, 1 to 10, 11 to 15, 16 to 24, 25 to 34, 35 to 44, 45 to 59, 60 to 74, and 75 to 100 percent. All analyses of the data were made using these ranges. If the reserve suggested by the subject fell into the same range or on the border, the comparison was marked OK. If the reserves fell in different ranges, the number of ranges by which they are different is noted. Starred entries indicate that one party suggested a reserve and the others did not. In addition the cases were analyzed for a comparison of the reserve vs. no reserve decision.

Comparisons were made between individual judgments with and without the use of the model. This comparison allowed us to consider how closely the unaided partner's judgments agreed on the same loan and how closely the non-expert's judgment agreed with the partners. Additional comparisons were made between the partner's judgments without the model and between the senior's judgments with and without the model in order to determine if the system was moving the non-expert judgment closer to the partner judgment. The loans were also analyzed according to whether no reserve or some reserve was required without respect to the reserve amount in order to test how the model did on the reserve vs. no reserve decision.

Figure 8

CFILE TEST RESULTS

Loan	PMM or Bank Eval	Expert Partner w/o CFILE	Expert Partner w/ CFILE	Expert Partner Difference	Second Partner w/o CFILE	Second Partner w/ CFILE	Second Partner Difference	Senior w/o CFILE	Senior w/ CFILE	Senior Difference		Partner Divergence	Average Ptr Judgment (AP) w/o CFILE	AP vs. Snr w/o CFILE		AP vs. Snr w/ CFILE
D2	n/a	n/a	n/a	n/a	0	0	OK	0	0	OK		n/a	0	OK		OK
D3	0	n/a	n/a	n/a	0	0	OK	0	0	OK		n/a	0	OK		OK
D8	n/a	n/a	n/a	n/a	0	0	OK	0	0	OK		n/a	0	OK		OK
D10	0	n/a	n/a	n/a	0	0	OK	10	0	*1		n/a	0	*1		OK
D11	0	n/a	n/a	n/a	0	13	*2	18	20	OK		n/a	0	*2		*2
D13	0	n/a	n/a	n/a	10	0	*1	20	0	*2		n/a	10	1		*1
P1	0	0	13	*2	0	0	OK	25	52	2		OK	0	*3		*4
P3	2	0	0	OK	0	13	*2	0	0	OK		OK	0	OK		OK
P4	66	67	67	OK	75	67	OK	60	40	1		OK	71	OK		1
P6	52	75	87.5	OK	50	87.5	2	60	13	4		1	62.5	OK		5
M-P1	n/a	0	0	OK	0	0	OK	0	0	OK		OK	0	OK		OK
M-P2	n/a	0	0	OK	0	0	OK	0	0	OK		OK	0	OK		OK
M-D4	n/a	0	0	OK	0	0	OK	0	0	OK		OK	0	OK		OK
M-D5	n/a	0	0	OK	0	0	OK	0	0	OK		OK	0	OK		OK
M-D6	n/a	87.5	87.5	OK	90	52	2	70	87.5	1		OK	88.75	1		OK
M-D7	n/a	0	0	OK	0	0	OK	0	0	OK		OK	0	OK		OK

Note: Evaluations are given in percentages of the outstanding loan balance.
Differences are in number of CFILE categories (as discussed on page 3).
Starred differences also indicate a disagreement about the reserve vs. no reserve decision.

SUMMARY STATISTICS	Expert Partner w/ & w/o CFILE		Second Partner w/ & w/o CFILE		Senior w/ & w/o CFILE		Agreement between Partners		Average Partner Judgment vs. Senior w/o CFILE		Average Partner Judgment vs. Senior w/ CFILE	
ALL CASES												
AGREEMENTS	9	90%	11	69%	10	63%	9	90%	11	69%	11	69%
DISAGREEMENT	1	10%	5	31%	6	38%	1	10%	5	31%	5	31%
AVE. CATEGORIES DISAGREEMENT	1		1.8		1.8		1		1.6		2.6	
RESERVE VS. NO RESERVE DECISIONS												
AGREEMENTS	9	90%	13	81%	14	88%	10	100%	13	81%	13	81%
DISAGREEMENT	1	10%	3	19%	2	13%	0	0%	3	19%	3	19%
CASES WITH RESERVES												
AGREEMENTS	3	100%	1	33%	1	20%	2	67%	2	50%	1	33%
DISAGREEMENT	0	0%	2	67%	4	80%	1	33%	2	50%	2	67%
AVE. CATEGORIES DISAGREEMENT	0		2		2		1		1		3	

Summary of Results

The following table summarizes the results of the individuals' judgments compared to the model's judgments when the model is used by that individual.

	All Cases		Res vs. No Res		Reserve Cases	
	Agree	**Disagree**	**Agree**	**Disagree**	**Agree**	**Disagree**
Expert partner	90%	10%	90%	10%	100%	0%
Second partner	69%	31%	81%	19%	33%	67%
Senior	62%	38%	88%	12%	20%	80%

In terms of the test's first objective, i.e., determining whether the system is consistent with the judgments of the designated expert, the results are very positive. On ten loans, the model's judgment is consistent with the expert partner's judgment nine times. Reserve vs. no reserve decisions were consistent in 90 percent or nine out of ten loans. In three cases where the expert and the model both suggested a reserve, the reserve amounts are in agreement. On the one disagreement, the model suggested a reserve of 11–15 percent while the partner suggested no reserve. We interpret these results as very positive and we intend to expand the scope of the model to produce a significant audit tool.

The second partner's percentages do not look quite so good in terms of agreement with the model. The second partner evaluated 16 loans and agreed with the model 11 times while disagreeing on five of those loans. These results become much more positive, however, when viewed in relation to other data. First of all, the percentages improve when looking at the agreement between a reserve vs no-reserve judgment. Here the model disagreed on only three loans. If we then scrutinize the degree of disagreements we note the model was never more than two classifications away from the second partner.

In attempting to explain the disagreement we note that the two partners' judgments, independent of the model, agree in nine of ten or 90 percent of the cases, (with only one classification separating them on the one disagreement). Since the use of the model is the only variable, and we know that the model is constant when given the same inputs, we hypothesize that the problem is not in the model itself, but in the user/model interface. We explain this as follows. The expert partner, who was instrumental in the design of the model, fully understands the questions and the impact of the responses on the model since he essentially wrote the questions. The other users of the system only had the cryptic wording of the questions and the help screens to indicate what the questions intended to ask. To support this hypothesis, we looked at the model's consistency of performance across users. We have 42 runs of the model which consisted of running ten cases three times, once by each subject and six cases two times by the subjects which we designate as partner-2 and senior. This provides us with 36 two-way comparisons. Of these 36, 20 runs involving ten of the 16 cases had complete three-way agreement. All of these agreed on zero reserve. In the additional 16 comparisons, involving only six of the cases, the consistency of the model was significantly different, agreeing with itself only five times or 31 percent of the time when a reserve is indicated.

Based on this it appears that the model performs well on the easy cases that require no reserve, but struggles when the case becomes more difficult and where more user judgment comes into play. While one reason for the degradation may be the user interface, we also suspect that the depth of the knowledge base may be inappropriate, thereby requiring too much user judgment in interpreting what the model is asking for. If the model were sufficiently robust to deal with facts rather than user judgments about, for example, the strength of the current ratio, we would expect that a good deal of the inconsistency would disappear. Yet another cause may be the attempt to be too specific about the amount of the reserve. In attempting to specify the ranges, it is possible that we have overrefined by attempting to be more specific than the experts themselves. While this may be a cause, we tend to discount it somewhat since there was no definable pattern to the disagreements between the model and the users. The model was not consistently higher or lower nor off by one or two classifications. The differences appeared to be more random, leading us to believe that the shallowness of the model's knowledge coupled with the user/model interface are the major problems.

We could apply the same analysis to the figures associated with the senior subject performing the task; however, in this case, we are not primarily interested in whether the model agreed with the senior. Since one of the objectives of the model is to improve the inexperienced decision maker's ability to emulate the partner decision, the more important data deal with how the senior's judgment independent of the model compared to the partners' judgment independent of the model, and then how the model altered the senior's judgment in relation to the partners'.

The data indicate that the senior's unaided judgments agreed with the partners' unaided judgments in only 69 percent of the cases. This, of course, is expected based on experience and knowledge of the senior. Ideally, when using the model, the senior's judgments should be closer to the partners' decisions. The data show that the model did alter the senior's decision in four of the cases; however, the model moved toward the partners' decision on only two loans and moved further away from the partners' decisions on the other two loans. While these results are inconclusive, we again hypothesize that the interface or communication problem cited above is the major culprit. In any event, negative conclusions should not be drawn on the basis of this test. Further testing with improved user interface will provide more insight in this matter.

Summary of Field Test Results

Based on the results obtained from the field test, we conclude that the model performs very well within the stated limitations of the design and when used by the expert who was involved in the design of the model. We must also conclude that the model performs less well in the hands of others.

This problem can be thought of as an interface or communication problem that may be very simple to rectify, or may require a considerable amount of effort. The solution lies in determining how to structure the questions in such a manner that, given a specific loan, user responses to the model's questions will be consistent. To obtain the solution, existing questions may need to be restructured and/or users may need more training in the use of the model. A third and more time consuming solution is to enhance the model's knowledge base to a depth that allows the model to work from more basic information.

Additional Insights from Field Testing

Through observation and recording verbalized protocols of certain cases, we were able to gather additional knowledge that a) lends more support to our hypotheses above and b) provides a focus for the immediate development work that is required. Since the analysis of the protocols is not yet complete, we will informally discuss these in the following paragraphs.

We are pleasantly surprised in finding that our bank partner's unaided judgments agreed in nine of ten loans and disagreed by only one reserve classification on the tenth. We are fortunate that this one case is included in the six cases for which we have protocols, and these protocols provide a plausible explanation for the partner's disagreement.

The second partner made reference in the protocol to having just recently read an article in a leading business journal concerning the borrower's history of problems, actions taken, and forecast for their survival. (In a later discussion we found he had read the article on the airplane in route to the lending bank's city.) The expert partner made no such reference to any additional outside information. The article provided an optimistic outlook for the company's ability to turn its problems around and survive in its market. While both partners recommended a rather high reserve (75 and 50 percent), the second partner was lower, perhaps indicating the impact of the article on the amount of his reserve judgment. This would indicate the need for the model to account for more soft data in greater detail than currently available. This is further supported in other parts of the various transcripts.

While we have not yet completed our analysis of the protocols, they appear to provide clear evidence of a significant weighting differential based on two primary characteristics of data: the recency of the data in relation to the date of evaluation, and the independence of the source of the information. While this is not terribly surprising, it is surprising in that the degree of change in the weighting appears to be significant. While we have not yet drawn any conclusion, it appears at this point that the model will have to account for these information characteristics.

Another fact that is becoming increasingly evident is the need for the model to deal with cash flow. It was originally thought that cash flow projections would not be a significant factor until we expanded the scope of the model to longer time horizons. Our protocols clearly indicate otherwise. In fact, as soon as a loan is considered to be a candidate for a reserve, the cash flow model comes into play. Furthermore, as the loans become increasingly suspect, there is a point when the partners change to a liquidation model, attempting to determine how much the bank may salvage from a liquidation and/or bankruptcy proceeding. These are important considerations even within our limited scope model.

Conclusions

We are basically pleased with the results of our field test not only because they indicate the model provides results consistent with the expert, but also because we believe that the model will provide significant assistance to the senior in the field. While we are aware that in the longer term the model's knowledge base must be expanded depthwise, we also believe that many of the user/model communication problems can be rectified through a restructuring of questions and help screens, as well as training of the intended users.

Our intention is to pursue the development of this model in three directions: a) to improve the interface to the point we can release the model to the bank practice personnel for more extensive field tests, b) to improve the model's current scope by increasing the depth of its knowledge and provide the ability to deal with the cash flow and liquidation requirements, and c) to begin expanding the scope of the model to handle other types of security and time horizons.