AL^{2}X: An Expert System for Off-Site Evaluation of Commercial Bank Financial Stability

John R. Segerstrom

Doctors spend their lives treating patients, who come to them in a never-ending stream, and they inevitably see the answer to any disease in terms of treatment rather than prevention. Nonetheless, historically, progress against disease has come from prevention and vaccines, not cures.

—John Cairns
Department of Cancer Biology
Harvard School of Public Health

By substituting "bank regulators" for "doctors" and "troubled banks" for "patients," Cairns's insight can be applied with equal generality to the U.S. banking system. The regulatory establishment has traditionally sought to fulfill its role as watchdog of the public's interest in the banking system by reacting to problems in individual banks and treating them to restore the bank to health. Unfortunately, in the 1980s, the bank mortality rate soared. Traditional "treatment" failed to protect the public interest, and taxpayer dollars were allocated through bad investments by financial institutions rather than through public policy.

These diversions of public resources are almost unbelievably huge. The mistakes of about 1,000 savings and loans diverted $300 billion, or $1200 for every man, woman, and child in the country. In 1990 dollars, this amount is roughly 6 times the cost of the Marshall Plan, which rebuilt Europe after World War II. Conservative estimates of the cost of commercial bank errors run about $5 billion each year for at least the next 2 years; some estimates are much higher. The Federal Deposit Insurance System is generally agreed to be insolvent in its current form, with some form of subsidy required soon. It is clear that the bank treatment system has been overwhelmed and is no longer adequate for the task of protecting the public interest. An effective prevention mechanism must be found and implemented if the banking industry is to return to stability.

The treatment paradigm on which bank regulation has been based focuses on the emergence of troublesome symptoms, particularly deterioration in loan repayments. The underlying critical assumption is that loans are made conscientiously: "They're all good when they're made," the banker's saying goes. A bad bank loan, in this view, is considered an anomaly.

Until about 1980, this assumption and the resulting model were adequate, and bank failures averaged about four each year. Since this time, however, the bank failure rate has skyrocketed, and there is considerable evidence that failures today are resulting not from good loans gone bad, as before, but from loans, actually bad when made, gone undiscovered by regulators. Rather than being an aberration, bad loans appear to have become the rule in some banks. Because bad loans might not develop symptoms for as long as two years, a fatal volume of such loans can already be in place when traditional regulatory treatment is triggered by symptom presence. By this time, failure of the bank is virtually inevitable.

Genesis of the AL²X Prevention Paradigm

AL²X (pronounced "Alex") was originally intended to assist private-sector auditors (CPAs) in making recommendations to clients. It began as a rudimentary expert system with 150 rules and a 20,000-word structured vocabulary. It produced an essay-style analysis of a bank's financial reports based on the early-warning research available at the time (Koch and Cox 1983). All methods of distinguishing one group from another are subject to two kinds of error: type 1 errors, in which a member of the target group (in this case, a future problem bank) is identified as a nonmember (a strong bank), and type 2 errors, in which a nontarget member is included in the target group. To adjust AL^{2}X to achieve low type 1 error levels, identifying a high percentage of future problem banks, it was necessary to accept the identification of a large number of strong banks as possible future problems, creating an excessive type 2 error rate. Users complained that the original AL^{2}X "cried wolf" too often to be credible.

More seriously, no actionable recommendations were available because the identification of the troubled banks came too late. A typical suggestion might have been, "To help this bank, make good loans starting two years ago and continue to do so until the present." This second failing led to the withdrawal of the original model. AL^{2}X Development Corporation was formed in 1984 to identify and market a more effective bank analysis paradigm. The objective was to find the root cause of an individual bank's abandonment of industry norms and consequent acceptance of excessively risky loans.

First Hypothesis: Risk Sought to Offset Weak Earnings

In 1984, research identified several financial ratios for identifying troubled banks: loan losses as a percentage of total loans, nonperforming loans as a percentage of total loans, and actual losses as a percentage of expected losses, to name three. The second AL^{2}X model retained some of these ratios from the original AL^{2}X design.

Sophisticated cost measures were added to these straightforward calculations. These measures are used to identify banks that require a large amount of income from loans to cover high expenses and also make a profit. The hypothesis is that excessive loan risk can be expected in banks configured to require high loan income; cost pressures can force the bank to make loans that, although riskier over time, produce immediate income.

The cost rules in AL^{2}X cover four distinct areas of expense: overhead, interest, yield curve placement, and nonloan asset yields. Each expense level is estimated, relative to industry norms, by an algorithm in the model and assigned a relative level. All the levels are then combined using a weighted factor analysis system that produces a *loan income pressure score* (LIP score), which is compared to a standard in the model's knowledge base. In general, higher LIP scores create more concern in AL^{2}X about the quality of the loan decisions being made.

Field tests of AL^{2}X were encouraging. Banks with high LIP scores did, in fact, have a tendency toward loan problems after a one- to two-year lag. The correlation was not perfect, however; banks with identical LIP scores apparently respond differently. Robert Long, a well-known bank observer and futurist, suggested that the source of the differences in response is cultural; that is, each bank has a unique predisposition to take risk under pressure (Long 1988). To test this suggestion, a subfactor analysis system that measures risk taking was added to the LIP scoring system. In this subsystem, noncredit risks that surface promptly in the bank's financial statement (funding risk, investment risk, interest rate risk, and regulatory risk) are measured with additional algorithms and aggregated using factor analysis. The resulting *risk propensity score* (RP score) is added as a factor in determining the LIP score. By adjusting factor weights and knowledge base standards, the model can now be tuned to provide a measure of pressure for loan income relative to the bank's cultural resistance to loan risk rather than to a fixed threshold.

The addition of the RP score dramatically improved the performance of AL^{2}X in identifying banks at high risk. The model is now able to identify situations in advance of actual loan deterioration and make recommendations for current action to relieve the pressure for loan income. The system also identifies cultures that are most likely to accept risk and, therefore, are most in need of monitoring by regulators.

A significant retrospective field test of the AL^{2}X first-hypothesis system was conducted in the state of Michigan in early 1988. In the 20 years preceding 1987, Michigan had had 1 bank failure. Using 1985 data, AL^{2}X identified 4 of 250 Michigan banks as seriously unstable. The actual 1987 failure was 1 of the 4. The result was cause for optimism for both the hypothesis and the actual paradigm underlying the AL^{2}X model.

Second Hypothesis: Risk for Its Own Sake

The effectiveness of the first hypothesis was tarnished in other field tests by the appearance of banks that either failed or experienced significant problems but maintained a relatively low LIP score. Every effort was made to eliminate the effects of dishonesty and economic fluctuations, but the inconsistency remained.

In discussions with regulators and insurance companies writing bank insurance, it was suggested that small risks in banks have no natural enemies and rapidly grow into large risks with no apparent bounds. The experts felt that this potential for risk explosion was especially true for risks with which the banker had little or no experience, such as those new risks created by the deregulation of the banking industry beginning in 1979 and 1980.

The process begins with the banker assuming a small risk. If s/he "wins," s/he becomes overconfident and plays again for higher stakes. If s/he "loses," s/he raises his(her) bet, either assuming that the law of averages is now in his(her) favor (the infamous "gambler's fallacy") or simply acting in desperation to recover his(her) loss before it is discovered. In either case, win or lose, the size of the risk balloons until it exceeds the capacity of the bank.

To test this second hypothesis, the new banking risks from deregulation were listed, and measurement algorithms were developed. A separate risk progression model was then added to the LIP model in AL^{2}X. Rerunning the field tests and tuning this segment of the program resulted in a two-screen process with impressive accuracy. Furthermore, the second hypothesis lent itself to practical recommendations for breaking the chain of risk progression. The resulting paradigm was dubbed *stress mechanics* (Segerstrom 1987). The enhanced AL^{2}X system was released in early 1988.

The Question of Risk Capacity

Because banking risk is not symmetrical—that is, because a bad decision involves a loss as much as 50 times greater than the gain from a good decision—and because some losses are inevitable in a risky environment, it was necessary to consider risk capacity in the AL^{2}X model. After further consultation with bankers and other experts, certain levels of capacity were established as minimums, and additional algorithms were added to provide base measures of a bank's ability to accept normal banking risk.

These algorithms were developed heuristically. The capacity to accept risk translates into the ability to incur loss and continue in business, and the capacity measures in AL^{2}X are ranked accordingly. The most serious conditions are those that can result in the immediate demise of the bank (low liquidity, for example: if a bank runs out of cash its demise is immediate). The capacity measures range from liquidity (cash) to equity level to potential profitability and are ranked in descending order of severity. The minimum levels of each are determined by AL^{2}X on the basis of heuristic rules using demographic factors for each individual bank, including size, type of market, and legal environment.

AL^{2}X System Structure

The AL^{2}X model applies the two hypotheses and the heuristic capacity rules to bank data and produces an analysis at one of five levels:

1. Is capacity adequate?
    a. IF yes, proceed with analysis.
    b. IF no, stop and report.

2. Is trouble evident now?
    a. IF yes, revise minimum capacity and go to 1.
    b. IF no, proceed with analysis.
3. Is loan pressure high?
    a. IF yes, stop and report.
    b. IF no, proceed with analysis.
4. Is risk progression evident?
    a. IF yes, revise LP max. and go to 3., ELSE stop and report.
    b. IF no, proceed with analysis.
5. Are strategic improvements available?
    a. IF yes, stop and report.
    b. IF no, summarize and stop.

The system reports on the most serious observations; that is, it does not make recommendations for high-level strategic improvements if the individual bank lacks the basic capacity to survive in a normal banking environment.

The appropriate level at which to report is first estimated by applying broad rules of the form "IF (data) THEN (estimate)" to the results of preliminary algorithms. This estimate is then confirmed by secondary, more sophisticated algorithms and rules of the form "IF (estimate) THEN (data)ELSE (fail)"; failure of the confirmation process generates an adjustment to the preliminary and secondary algorithms and the reprocessing of the data, followed by reapplication of the broad rules and reconfirmation. The process continues until a confirmed estimate (conclusion) is identified, allowing the system to generate a report.

The AL^{2}X report is written from the system's vocabulary on the basis of the final output from the primary and secondary algorithms, using syntax rules in a report assembler in the AL^{2}X program. The vocabulary is stored in phraselike parcels, which are available to the report assembler by coded reference.

The user has the option of bypassing the written AL^{2}X report and directly viewing the results of the system's algorithms on the computer screen. In this mode, the system operates as a specialized calculator, performing complex mathematics for the user who is responsible for interpreting the results. This approach allows the system to support the expert, as well as the novice, user.

AL^{2}X normally accepts data from bank financial statements through the computer keyboard. Data are updated quarterly by the user, and the system retains data for 1 year for trend analysis. Fewer than 50 pieces of data are required each quarter, and data entry requires less than 5 minutes, as does processing. Large-scale users, such as regulators, have the option of acquiring preformatted data about a large number of banks in computer-compatible form, accessible by key code, which eliminates manual data entry altogether.

The AL^{2}X system contains 55 separate analytic algorithms and 3,500 decision rules. It draws its output report from a 60,000-word vocabulary and supports its conclusions with statistical exhibits comparing the actual results of the bank with the standards in the model. AL^{2}X draws initial data from public financial reports prepared by each commercial bank in the United States and allows the user to inquire about the potential impact of changes in the bank's activity on the AL^{2}X analysis.

The AL^{2}X system is written in Borland, Inc.'s TURBO BASIC compiler for MS-DOS–compatible personal computers. The software is designed to be acceptable to the broadest possible range of computer and peripheral device configurations: The minimum requirements are an MS-DOS computer with 256K of random-access memory and a single floppy diskette drive.

The current evolutionary system represents 6 years of development and a combined development and deployment investment of $450,000. The AL^{2}X software is currently used nationwide by over 300 banks; 6 state bank regulatory departments; and a sampling of attorneys, accountants, investment advisers, and other professionals.

Automatic Data Acquisition and Mass Testing

In 1988, just following the public release of the current AL^{2}X system, data from 31 December 1987 (the most current then available) from federal bank reports was purchased on magnetic tape and decoded. AL^{2}X was asked to analyze each bank in the country within its domain (banks with less than $1 billion in assets). The results of the AL^{2}X analyses were summarized for each bank in the form of an index, called the *<STAR> index* (Segerstrom and Meadows 1989). The number of high-risk banks as a percentage of total banks was computed for each state, and the states were then divided into quartiles by the resulting percentages.

The southwestern states, already troubled, figured prominently in the lowest quartile. This result could have been anticipated by the inclusion of traditional measures of loan problems in AL^{2}X. However, Massachusetts, Connecticut, and Rhode Island were also in the lowest quartile at a time when the "Massachusetts Miracle" was being touted as the potential economic salvation model for the country. Skeptics of the AL^{2}X model were easy to find, including then-Governor of Massachusetts and presidential candidate Michael Dukakis. By late 1988, however, the *American Banker* and the *New York Times* featured stories about the decline of the New England economy (Bartlett 1988;

Matthews 1989). The prescience of the AL^{2}X prediction for New England has since been proven beyond debate.

The AL^{2}X system has correctly predicted problems in the Mid-Atlantic states, particularly North Carolina, and beginning in June 1989, the model began predicting serious problems in California that are, in fact, currently unfolding.

Since 1987, the year of the first nationwide application of the AL^{2}X system, no serious deterioration in a state banking system has taken place that was not predicted, and no prediction has proven to be groundless. In each case, the AL^{2}X system identified specific banks requiring attention and made practical recommendations for action in each of these banks 18 months to 2 years in advance of the appearance of traditional symptoms of trouble in the bank.

AL^{2}X System Deployment: Real-World Lessons

There are approximately 14,000 banks in the United States, about 12,000 of which are within the domain of the AL^{2}X system. The federal and state bank regulatory agencies monitor the stability of each of these banks to protect the public interest. Because of the previously mentioned failure of traditional off-site analysis methods, emphasis is currently being placed on increasing the frequency of on-site examinations in all banks at tremendous expense. A bank examination typically costs $25,000 or more; so, the cost of examining each bank in the AL^{2}X domain once each year would be at least $300 million. Based on the results of the AL^{2}X stability study of 30 June 1990, approximately 70 percent of this cost ($210 million each year) could be avoided by examining only those banks too large or too complex to be included in the domain and those banks within the domain that AL^{2}X indicates are unstable. The necessary antecedent to realizing these savings is confidence on the part of the regulatory establishment in the reliability of the AL^{2}X analysis and conclusions.

Although it was not the original mission of the developers of the AL^{2}X system, it is clear now that individual banks can also benefit from the expertise embodied in AL^{2}X. Empirical observation led to the conclusion that reversing the focus of the hypotheses in AL^{2}X (from risk measurement to risk avoidance) defines a management discipline that produces both stability and high profitability in individual banks. William M. Reid, president of a $650 million (assets) bank in Richmond, California, observes:

> The major impact of AL^{2}X on the investment or structural decisions of the bank came from understanding the powerful financial concepts that are embedded in the model. Most important among these are the long-

> term predictable relationships between the various rates that affect the bank. Not all loan, investment, and deposit interest rates change by the same amount or at the same time when interest rates in general rise or fall. By understanding and being able to study these relationships, we realized that the bank was actually positioned to earn substantially less money if interest rates fell. We saw the need to make longer term, fixed rate loans and purchase longer term, fixed rate investment securities. We also saw the trade-offs between various types of risk our bank faces: credit, liquidity and market valuation.
>
> By our acting on these convictions, we saw bank earnings rise immediately as longer term investments and loans yielded higher returns than the shorter term instruments we were previously using. More important, however, is the fact that as interest rates have fallen and as a positively sloped yield curve has been reestablished, the bank's interest spread on a ratio basis has been maintained. This was precisely the bank's objective—to maintain strong and stable income levels regardless of the interest rate environment in which we happened to be operating. Had we used traditionally accepted methods, our income in 1991 would almost certainly have suffered by as much as a million dollars.

Because AL^{2}X is not based on a previous hypothesis, its methods and conclusions are unfamiliar to bankers and regulators alike. AL^{2}X Development Corporation, with significant assistance from the Office of the Comptroller of the Currency, which regulates the 4,200 federally chartered banks in the United States, undertook an extensive study to compare AL^{2}X's results with those produced with the theory being used by this office. The results were substantially different: There was virtually no correlation. The regulator's conclusion was that AL^{2}X failed the test. Ironically, the standard used—the regulator's existing theory—had already itself failed in the real world. As the challenger, AL^{2}X faced several barriers in the study's design, many of which could not be overcome because of legal protections applying to information about banks. No barrier, however, was greater than the fact that no accepted theories or results are appropriate to test the AL^{2}X methodology.

AL^{2}X is an *orphan* expert system: There is no one expert or group of experts whose knowledge is embodied in the system. Lacking this credibility, the AL^{2}X system relies on empirical results to demonstrate its utility. Mention has been made of the demonstrated success of the AL^{2}X methodology in New England, the Mid-Atlantic states, and California since 1987; empirical credibility is high and still growing. However, acceptance remains slow. It is now believed that the system is hindered by two aspects of its design, neither of which bear on the performance of the expert system itself: (1) the potency of the system and (2) the interface between the system and the user.

The word potency describes the power of the output. AL^{2}X is forceful in its reporting; the output is conclusion oriented and concise. The AL²X bank report emulates a consultant's report, usually producing 10 to 12 pages of text—expert opinion—with supporting statistical exhibits. However, this opinion is often contrary to prevalent theory and, thus, seems threatening to a user whose knowledge and experience are based on this theory. The empirical success of AL²X only adds to user discomfort because there is no ready means for discounting its results.

The acceptance of an expert system with new capabilities seems to require that the system's output use implication and suggestion rather than conclusion. To succeed, the output must rely on skillful design to guide the user to the proper conclusion on his(her) own. This theory is the current thrust of AL²X's evolutionary development, and it is proving as difficult as, if not more difficult than, the development of the knowledge base itself. The AL²X knowledge base is maintained by the developers, and the redesign of the output routines can be integrated into the maintenance effort as new knowledge is gained. If AL²X were a learning system, the output design would certainly be much more complex than the development of the knowledge base.

At a less abstract level, the use of a written report creates comprehension problems for the user with this kind of expert system. The written word is assimilated serially, one word at a time; the AL²X algorithms work in parallel. A high percentage of these algorithms determine the structure of other algorithms, in addition to developing conclusions that are often based, in turn, on the results of yet others. The AL²X report is carefully crafted to describe this parallel analysis, but in the process of reading, users almost inevitably make an effort to imply simple causality in the conclusions. This approach can appear successful in an isolated case; however, users become frustrated when they attempt to generalize the same inferences from one analysis to another because they rarely succeed.

Interestingly, the transition from serial to parallel processing in human cognition is synonymous with becoming an expert in a complex field (Buffington 1987). It is probable, then, that written output from a complex expert system can only be used by an expert in the field. The newest challenge in redesigning the AL²X system is to find a way to display, in quasi-real time, all the algorithms or syntheses at the same time on the computer screen but provide the user with access to financial variables through the keyboard (or perhaps a mouse-type interface). This banking "flight simulator" would provide a vehicle through which the non-expert AL²X user could communicate with the AL²X parallel system on a trial-and-error basis until, like a pilot, the user becomes more comfortable with simultaneous interactions. This sort of display module, combined with less intimidating written output, might bridge the existing gap between accepted theory, discredited but entrenched, and the AL^{2}X methodology, proven but unfamiliar.

The AL^{2}X Model: Prospects and Perspectives

A representative of the Xerox advanced research facility in Palo Alto, California, once said that a new technology cannot be integrated into society in less than 10 years and, then, only if it attracts a significant competitor. The AL^{2}X paradigms, if viewed as new technology, are 4 years and 1 competitor away from acceptance. As a product, the AL^{2}X system is a growing commercial success; as a useful technology, the potential benefit has been denied to the general public, to some extent because of the design of the system itself.

The knowledge base in which the details of the AL^{2}X paradigm exist has not been copyrighted or offered for license, nor have the details been published. They are maintained as trade secrets. It is the particular and circular dilemma of small-company research that wide use requires acceptance; acceptance requires competition; and competition, in turn, endangers the small developer. The development of AL^{2}X provided a tremendous amount of insight into the workings of the banking industry in the newly deregulated environment and could have tremendous value to bankers, regulators, and legislators in the immediate future. AL^{2}X Development Corporation now seeks a vehicle that allows the dissemination of the details of AL^{2}X's knowledge without committing commercial suicide. Success in this quest will, it appears, finally determine AL^{2}X's contribution to the prevention of further banking disasters in this country.

References

Bartlett, S. 1988. Bad Real Estate Loans Hurt Northeast Banks. *The New York Times*, 15 November.

Buffington, P. 1987. Experts: Who Says? *Sky*, June 1987.

Koch, D. L., and Cox, W. N., eds. Early Warning Systems and Financial Analysis in Bank Monitoring. *Economic Review* 68(11): 6–48.

Long, R. H. 1988. Secrets of High-Performing Bankers, Thinking Technology Associates, Phoenix, Arizona.

Matthews, G. 1989. Economic Downturn Seen for Northeast. *American Banker*, 31 October.

Segerstrom, J. R. 1987. Stress Mechanics: Understanding How Deregulated Independent Banks Work. *The Supervisor* 24(4): 2–7.

Segerstrom, J. R., and Meadows, G. D. 1989. Why Is AL2X Saying That about My Bank?" AL^{2}X Development Corporation, Portland, Oregon.