AI Magazine Volume 16 Number 1 (1995) (© AAAI)

Countrywide Loan-Underwriting Expert System

Houman Talebzadeh, Sanda Mandutianu, and Christian F. Winner

■ Countrywide loan-underwriting expert system (CLUES) is an advanced, automated mortgage-underwriting rule-based expert system. The system was developed to increase the production capacity and productivity of Countrywide branches, improve the consistency of underwriting, and reduce the cost of originating a loan. The system receives selected information from the loan application, credit report, and appraisal. It then decides whether the loan should be approved or whether it requires further review by a human underwriter. If the system approves the loan, no further review is required, and the application is funded. CLUES has been in operation since February 1993 and is currently processing more than 8500 loans each month in over 300 decentralized branches around the country.

Countrywide loan-underwriting expert system (CLUES) is one of the largest implementations of a decentralized AI system. It is located in 300 branch locations and is used by over a thousand users. The model used in CLUES is an elegant solution to a complex problem. CLUES has an abstract decision-making model that evaluates a loan file in the same way that a human would, without a combinatory explosion of all possible cases. The decision-making model has had a tremendous impact on the ability to enhance and maintain the system. In addition, it has heuristics to provide extensive justification for the decisions made at each intermediate level and the final step of the evaluation. Finally, the design allows the system to recognize when it has reached the limits of its knowledge to make a sound decision, at which time, it refers the final decision to a human underwriter.

CLUES is also a successful proof of the tremendous value of data-driven, rule-based AI techniques in solving difficult problems. The small team that developed CLUES was able to take the system from concept to production release in less than 12 months.

The Life of a Loan

Loan underwriting is the process of evaluating a loan application to determine whether the loan should be funded. The process often starts with a potential borrower walking into a branch office and requesting a loan to purchase or refinance a home. A processor asks the borrower to fill out an application, setting in motion a lengthy information-gathering process in which as many as 1500 data-element pieces will eventually be collected. This loan information includes items about the borrower's employment, income, assets, liabilities, and monthly expenses. During the process, a credit report and appraisal will be ordered from a third-party vendor. When these data are electronically received, it is uploaded into the loan-origination software.

When the necessary information is collected, the stack of documents is an inch thick. At this point, the loan file is turned over to an underwriter, who evaluates the information, verifies the information against the supporting documents, and decides whether to approve the loan or deny the application.

If the application is approved, the final documents are signed, and the funds are sent to a closing agent. Subsequently, the loan is pooled with other loans, converted into securities, and sold in the secondary market. The average loan takes 45 days to close, and an additional 30 days to be converted into securities and settled in the secondary market.

The life of the loan ends when it is ultimately paid off, or if the borrowers fail to make the payments, the property is foreclosed.

The Problem

Countrywide is the largest originator and servicer of mortgages in the country. It is currently funding over $5 billion a month and is servicing a portfolio of approximately $80 billion.

Underwriting is difficult, expensive, and time consuming. It takes approximately 50 minutes for an underwriter to review a typical loan. Nontypical loans, such as those for FHA and VA mortgages, second homes, investment properties, and small apartment buildings, can take even longer.

The typical underwriter is a highly skilled individual with many years of experience in mortgage banking. The underwriter has to be knowledgeable about hundreds of specific product and investor underwriting guidelines to determine if the loan profile satisfies the requirements of the lender and the ultimate investors. The sheer volume of information that has to be considered, along with the interdependence of this information, makes underwriting a challenging decision-making process.

The low interest rates of the last two years unleashed a substantial increase in the number of people applying for home loans, many to refinance under better terms. This increase has led to a serious shortage of qualified underwriters, creating major bottlenecks for all lenders, including Countrywide.

To cope with the increasing number of loans and the shortage of underwriter personnel, in late 1991, Countrywide began looking at the possibility of creating a system that would automate the underwriting process. Such a system would alleviate the critical shortage of underwriters; better use the current personnel; and, most importantly, allow Countrywide to increase the number of loans processed by each employee. Underwriters, freed from the need to process simple loans, could then concentrate on difficult loans, study the local markets more fully, and spend more time with customers.

Objectives

In a step toward developing such a system, a list of key objectives was compiled:

First was to develop a system that would underwrite a loan and determine whether it should be funded or whether a second review should be made by a human underwriter. If the loan were approved, it would be sent directly to funding, and no further human involvement would be required. If the system referred the loan to an underwriter, the system should be specific about the key reasons that caused the referral, allowing the underwriter to concentrate on the problematic areas.

Second was to focus on the quality of the decisions, even at the cost of making the system conservative. With volumes as high as $5 billion a month going through the branches, a single error could result in millions of dollars in bad loans. It was acceptable if the system referred a questionable loan that was later approved; however, it was unacceptable for the system to approve a loan that would typically be denied.

Third was to make the system capable of justifying its decisions. It was not acceptable to assign a meaningless score of 1 to 100, as is commonly done by some scoring systems, and let the underwriter figure out why 71 is a good loan, and 69 is not. A score says nothing about the individual strengths or weaknesses of a loan.

If a loan was referred with simply a score, the underwriter would need to reunderwrite the entire loan to make a final decision. However, if the system could explain the strengths and weaknesses of a loan file and could point to the specific reasons why it referred the loan, the underwriter could focus on the problematic areas and make a quicker decision.

Fourth was to make the system easy to maintain and enhance. Because market conditions change on a regional basis, underwriting guidelines are updated to reflect these changes. The technology used to develop the system should allow the developers to make overnight changes in the underwriting guidelines.

Fifth was to make the system portable and easy to integrate with other lenders' loan-origination software. If the system could be made independent from the front-end interface, the company would be able to change the interface without affecting the integrity of the underwriting system. Making the system independent would also allow the company to make the system available to other lenders.

Selecting the Technology

Technologies other than AI were considered only briefly. Underwriting is a complex process that is completely data driven. Several

hundred data elements can be included in a loan. However, a typical loan file contains only a small subset, about 40 percent of all the possible values. The data present depend on the profile of the borrowers, for example, the types of assets that they own or the sources of their income; their credit history (for example, whether they have any foreclosures or bankruptcies and the status of the bankruptcy); and the type of the subject property, that is, single-family residence versus condominium. As a result, the data present in the file dictate how the system should process and underwrite the loan. In addition, there is a strong interdependence between the data fields. The value of one field can directly influence how another value is used by the system. For example, if the length of time at the current job or in the profession is too short, the income from the job cannot be included in the total income of the borrower. Thus, nonintelligent, procedural-driven systems were determined to be incapable of easily handling the varying amount and the interrelatedness of the data for this application. Four alternate AI technologies were evaluated: (1) rule-based expert systems, (2) neural networks, (3) mentoring systems, and (4) case-based reasoning.

The case-based technology proved unwieldy. Because each loan is unique, a large sample of loans would be required for the case-based approach. Maintaining such a large case base would be difficult, especially because underwriting guidelines are apt to change, and a loan that was acceptable three months ago might not be acceptable today.

The neural network and mentoring approaches initially appeared attractive but had to be ruled out when they could not meet some key project requirements. Because both systems require a large sample of loans to retrain themselves, the ability to update the system and rapidly send new changes to the branches was limited. These systems could also not justify or explain their decision to the extent necessitated by the objectives. Finally, they are, in fact, a black box and difficult to test.

Unlike the other approaches, a ruled-based expert system lends itself well to solving underwriting problems. A rule within a rule-based system can naturally express an underwriting guideline. With a rule-based system, rules can be added or modified easily and quickly, allowing new releases to be made on demand, sometimes overnight—an important consideration in an industry where conditions change rapidly. The system is also predictable, and regression testing can be performed at the rule level, module level, and system level to verify that the decisions are consistent with the requirements.

The potential for a rule-based expert system to underwrite loans was promising; however, no known rule-based systems had been deployed for loan underwriting on such a large scale. It would also be the first use of AI within the company. Although the project had strong advocates in upper management among those who understood the technology's potential, others were understandably unsure of deploying a new, untried technology for the critical task of underwriting; their support was conditional upon early proof that the system worked.

Development

With the selection of the technology and a clear set of objectives, the project could begin. In January 1992, an Artificial Intelligence Department was created and a small development team put together: two knowledge engineers and a project manager for the expert system, two database developers for the database and analysis tools, and one quality control member.

ART-IM from Inference Corporation was chosen as the development tool, and Microsoft WINDOWS and RS6000 were chosen as the development platforms by the developers (WINDOWS was chosen for the facilities offered by that version of ART-IM). When in production, the system would need to be deployed in the OS/2 environment because it was the platform of the existing loan-origination software. To test and validate the system during development, the UNIX version on the RS6000 platform was chosen for its speed.

Thus, the application environment was spread over several platforms: (1) Microsoft WINDOWS on the PC, UNIX on the RS6000 (development); (2) UNIX on the RS6000 (testing and validation); and (3) OS/2 on the PC (production).

System Requirements

The scope of the project was ambitious: to develop, in as short a time as possible, a system capable of underwriting loans as well as expert underwriters could. In addition, the developers would need to keep several requirements in mind when designing the system.

The system had to be integrated with existing corporate software: The overall computing environment of the branches into

which the system was to be deployed is based on a decentralized approach. Each branch is equipped with a server and several PC-based workstations, with all the software required to process the loan files. The branches are connected to the central facilities by modem lines. Several times a day, prices and updates of the software are transmitted to branches, and information about the loan pipeline and funding is transmitted back to the central facilities.

The company had developed two software systems for processing loans: (1) loan-origination software and (2) data-communications software for electronic links to business partners. Both software systems run under OS/2 on the PC.

The loan-origination software is used by the loan processor to enter all the information regarding the borrower, the credit history, and the appraisal report. The loan information thus collected is stored in a database where it can be recalled by the software to allow the processor to track loans and to print any of the more than 1300 different types of loan documents that might be required in a loan file.

Electronic communications software allows the branches to establish direct electronic links to Countrywide business partners, send requests, and receive electronic data.

The final AI system had to be able to accept data from the corporate databases as well as the data incoming from credit vendors and other Countrywide business partners.

The system had to be easy to maintain and expand: Easy maintenance was needed both for enhancing the system to accommodate changing underwriting requirements and new loan products and for fine tuning the system to add more underwriting guidelines that would allow the system to approve a larger percentage of loans.

Underwriting guidelines change constantly, in some cases overnight. Changes in the economy of a region and the potential changes to housing prices affect whether a loan can be approved; loans that might have been approved three months ago might no longer be accepted. In other cases, new conditions must be satisfied before a loan is funded in certain geographic areas. For example, if a natural disaster occurred in a localized region, the guidelines for the area could be modified to require drive-by appraisals to make sure the property was still in good condition. The need to respond to local conditions is especially acute for a company such as Countrywide that does business in all 50 states.

In addition to changes in existing underwriting guidelines, new products are added on a regular basis, each with different underwriting requirements. The system would have to be enhanced periodically as new loan programs were added, without affecting the integrity of the system.

The decisions had to be validated at each development stage to ensure high-quality decisions: The decision to fund the loan is critical to mortgage bankers, who survive by their ability to correctly predict which loans are good and which loans have a high probability of being foreclosed. Before an automated underwriting system could be deployed in the field, the system would have to be tested at every stage of development and each time the system was modified. The design itself should make it easy for the system to be tested.

Knowledge Acquisition

The first step in designing the system was for the knowledge engineers, who had no previous experience with mortgage banking, to learn about the underwriting principles and guidelines and become familiar with the vast amount of terminology specific to the banking industry.

As with any introduction to a new domain, being able to understand the terminology and principles is a slow and painstaking process. The knowledge engineers began this knowledge-acquisition process by interviewing the chief underwriters—the experts—in the company and watching them underwrite numerous loans. The knowledge engineers took copious notes and recorded the interview sessions with a tape recorder. Sessions were spaced out over alternate days, with the off days spent listening to the recordings, reviewing the notes, and making lists of questions. In all, more than 100 hours of interviews were conducted over a 2-month period.

One difficulty was the sheer number of guidelines. Another difficulty was understanding how the underwriters applied the principles to the facts of the loan file. In the banking industry, underwriting is considered an art that is taught from one generation of underwriters to the next. It takes years of experience to learn to judge and evaluate the loan information to make an informed decision on the risk of the loan. Underwriters do not follow a systematic, step-by-step approach to evaluate a loan. Instead, the underwriter must look at the strengths and weaknesses of the individual elements in a loan file and evaluate how all the data elements affect one another. The process is intu-

itive rather than scientific.

The challenge for the knowledge engineers was to represent the thought process of an underwriter in a manner that could be implemented as a software system.

CLUES Architecture

By carefully interviewing and analyzing how the underwriters make their decisions, the knowledge engineers made their first tentative steps at breaking down the underwriting thought process into its separate steps. They identified three distinct analyses made by the underwriters: (1) the financial ability of the borrower to make the monthly mortgage payments, (2) the credit history of the borrower, and (3) the appraisal report.

Conceptual Model

The knowledge engineers then began building a conceptual model of the problem in which the system was divided into three separate analyses: ability, credit, and appraisal. For each analysis, the knowledge engineers created a module consisting of rule sets that applied to the specific analysis. Based on the results of each module, the final decision to approve or refer the loan would be made (figure 1).

Two other modules that supported the analyses were also created. One module contained rules for checking the completeness of the data to ensure CLUES had sufficient information to make a good decision; the other contained rules for calculating the derived values, such as the PITI (sum of payment, insurance, and taxes over total income) and debt ratios (total monthly debt over total income), that are needed for the analysis modules.

The modularity of the design had many advantages. It allowed the knowledge engineers to concentrate on a core set of rules that could be used for conventional loan types (such loans accounted for 80 percent of Countrywide's business). Later, the ability to process more specialized and difficult loans, such as FHA and VA loans or loans for second homes and investment properties, could be added by creating other modules.

The design's modularity also enhanced the maintainability of the system. Changes could be made to one module without affecting other modules. Additional modules could later be added for fraud checking or new loan types.

Once the conceptual model was completed and approved by the experts, the knowledge engineers began coding. To keep the number of rules manageable, the knowledge engineers were determined to reduce the number of data items needed to make a correct underwriting decision. Initially, all 1500 data items available in the loan package were looked at. However, by carefully interviewing the experts, the knowledge engineers were able to determine what data items were decisive to the underwriter. The number of data items was thus reduced to 550, although for many loans, the actual number is much lower.

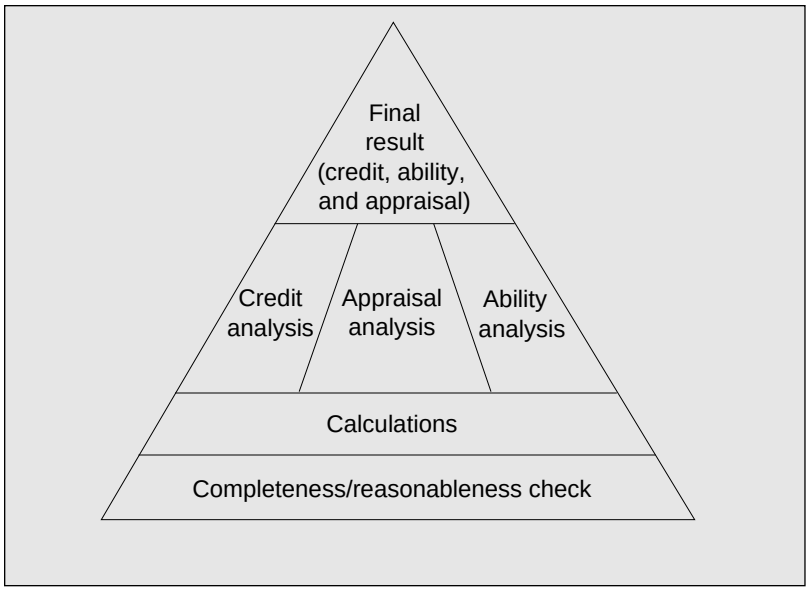
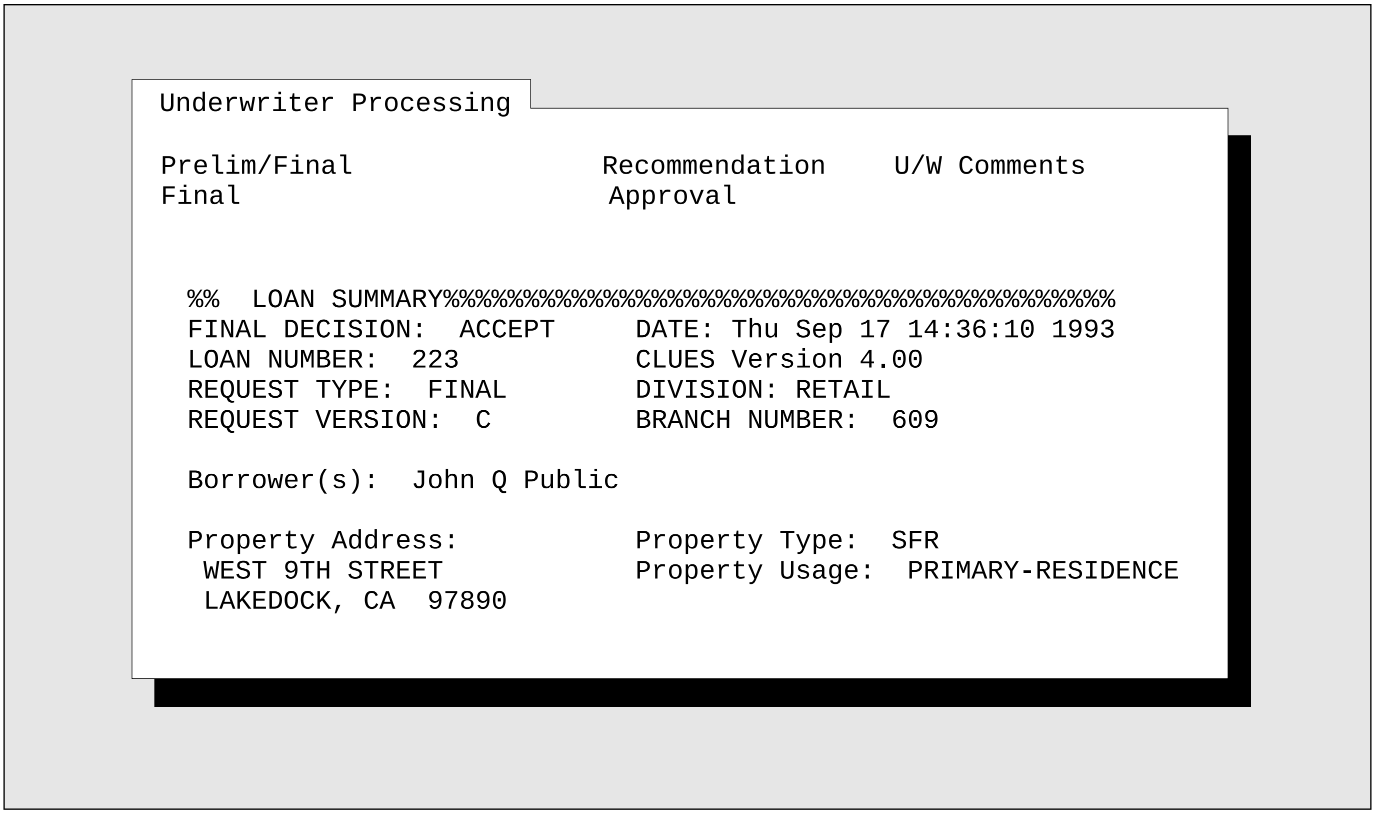


Figure 1. CLUES Decision Process.

By May 1992, the system, which by now had acquired the name CLUES, was essentially complete. Further development would focus on fine tuning the process to discover if specific underwriting guidelines had been omitted or whether they were implemented incorrectly.

Architecture

In its current form, CLUES is an embedded system consisting of approximately 1000 rules, 180 functions, and 120 objects. For each specific portion of the loan information that needed to be entered, the developers defined a rule set. Rule sets were developed for evaluating such data concepts as income, employment, credit worthiness, appraisal, comparable properties, assets, liabilities, and the risk of the loan. Each rule set is included in one of the three main modules: (1) ability, (2) credit, and (3) appraisal. Each rule set can positively or negatively affect the results of another rule set. For example, if the employment of one of the borrowers is determined to be unstable,



```
Underwriter Processing

Prelim/Final                Recommendation    U/W Comments
Final                        Approval


  %%  LOAN SUMMARY%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
  FINAL DECISION:  ACCEPT     DATE: Thu Sep 17 14:36:10 1993
  LOAN NUMBER:  223           CLUES Version 4.00
  REQUEST TYPE:  FINAL        DIVISION: RETAIL
  REQUEST VERSION:  C         BRANCH NUMBER:  609

  Borrower(s):  John Q Public

  Property Address:           Property Type:  SFR
   WEST 9TH STREET            Property Usage:  PRIMARY-RESIDENCE
   LAKEDOCK, CA  97890
```

Figure 2. Portion of the CLUES Final Output Screen.

the income from this job is not included in the income analysis rule set. Conversely, if the income is determined to be less than desirable, but the borrowers have a strong source of liquid or semiliquid assets, then the risk of the loan is significantly reduced.

Data are input to CLUES in the form of a file with a specific syntax that allows it to be read by CLUES. This file is created from loan information taken from the corporate databases; it can contain as many as 550 data elements, although many files have far fewer than this number. When CLUES is executed, the data elements in the file are loaded into the knowledge base. The first step is to check the integrity of the data to ensure it is complete and reasonable. If the data passes the validation step, the system proceeds to make a decision on the loan, and the final result is reported in an output file.

CLUES output, in fact, consists of two files. The first file is a simple text file that is displayed on screen to notify the loan processor of the results of the CLUES evaluation, including the final decision, the calculations used by CLUES, and the justification for the final decision (figure 2). The text file is formatted to be understood easily by a nonprogrammer.

The second output file is similar to the input file except that it contains the intermediate results created during the CLUES evaluation. This file is used for testing the system and archiving the results of CLUES evaluations and is not available to the loan processor.

Within the branches, CLUES fits easily into the work flow. The loan processor uses the company's proprietary loan-origination software to enter loan information. When this process is complete, the processor then chooses a menu item in the loan-origination software to request a CLUES evaluation. The loan information is then passed from the loan-origination software to CLUES, which evaluates the data and returns the final decision to the loan processor, who views it on screen.

Integration

To be integrated with the existing corporate application environment, it was decided to make the AI system an embedded system that would sit in the background until the user initiated the request (figure 3). CLUES would then execute and return the results of its evaluation to the loan-origination system for display to the user.

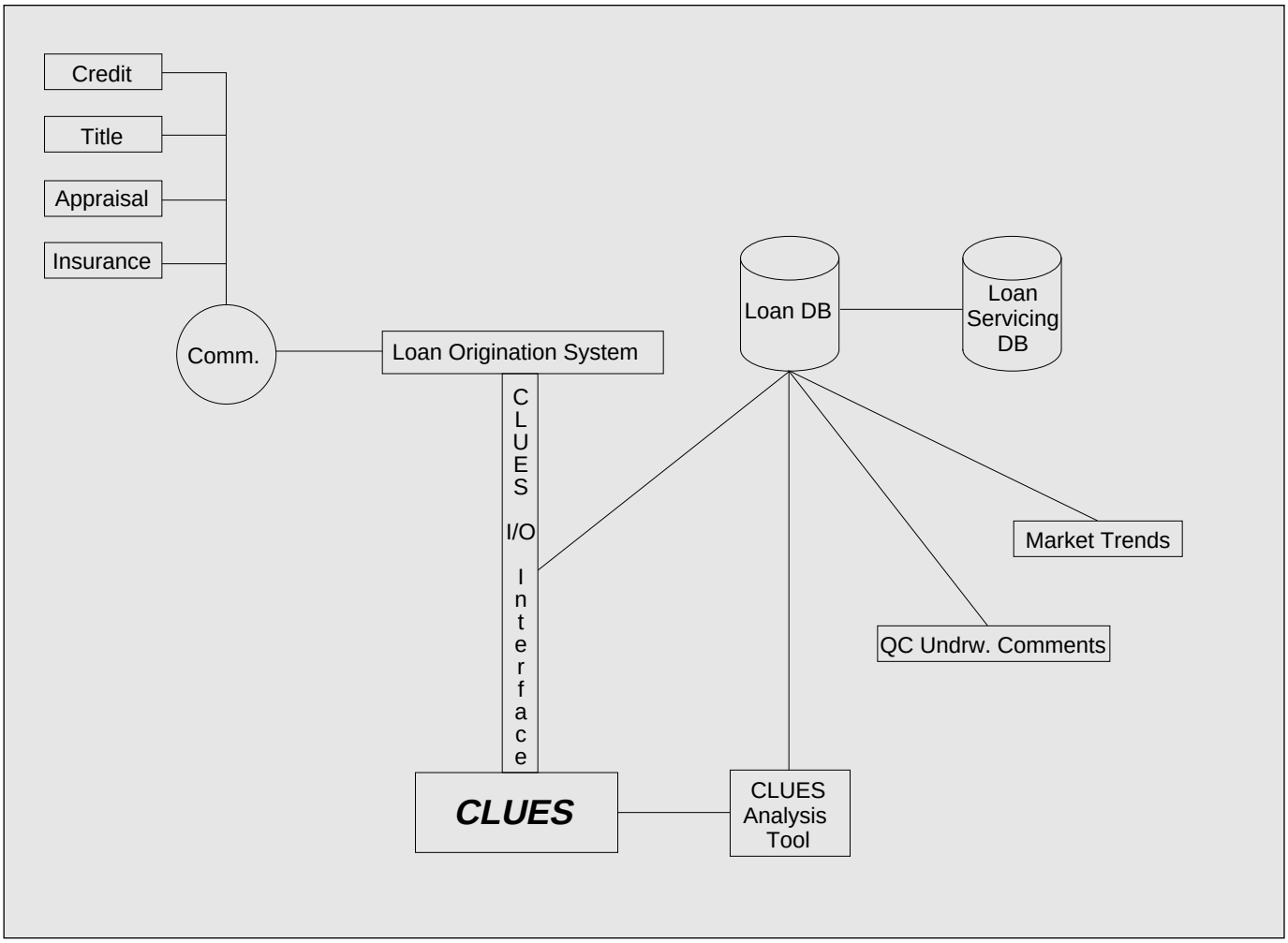


Figure 3. CLUES in the Application Environment.

November 1991	Problem stated. Investigation into possible solutions begins.
January 1992	AI Division formed.
February to April 1992	Knowledge acquisition conducted.
April to May 1992	Modeling and prototyping done.
May to June 1992	Coding conducted.
June to September 1992	Internal and system integration testing performed.
September to December 1992	Alpha test conducted.
December 1992	Beta test begins.
January 1993	Beta test ends.
February 1993	Product released to retail branches.

Table 1. Summary of CLUES Development.

In addition, the knowledge engineers had to consider how to pass data between CLUES and other corporate systems. This task proved to be almost as big as developing the expert system itself. For data to be passed from the corporate databases to CLUES, an intermediate representation of the data was used to make it acceptable to the internal representation of CLUES. Such an approach allows CLUES to be independent from the data sources.

Data translation and mapping programs were designed to achieve the communication between CLUES and various other parts of the corporate data system. Credit and appraisal information received from vendors also had to be represented in a way that allowed CLUES to access the information. Data protocols were developed for CLUES and stand now as the company standards.

Validation

In developing an AI program or any software program, testing and validation of the system is important to ensure that the functional requirements were satisfied and that the system meets the needs of its users. However, in the case of CLUES, testing and validation became especially important. If the system did not perform well and made bad decisions, the result could be millions of dollars of bad loans that would directly affect the company's financial standing. In the 12 months it took to develop the first release of CLUES (table 1), the developers spent approximately 7 months validating the system.

Validation also served a second important purpose. As with any other AI application, CLUES had to be shown to be a viable solution, and the people within the company had to be convinced that a machine could make decisions that were as good as those of an underwriting expert. Many in the company had little experience with, or even knowledge of, AI; their first experience with AI was with CLUES, a system that would take on one of the most important tasks within the company. People needed to be confident that the system would work.

Validation efforts started early. For each change to the system, tests would be run, the results studied, and modifications made as needed. The cycle would then repeat. At each stage, the experts were kept informed of the progress.

Testing the Prototype

The first prototype of CLUES was completed in June 1992. To test the prototype as well as the entire model, a set of real cases was put together for the alpha test. A stand-alone loan-origination interface was also developed to ensure the independence of the prototype expert system from any interface problems. The stand-alone interface not only achieved the desired independence but also served as a streamlined loan-origination system that collected precisely the information needed for the CLUES evaluation. A sample screen from the streamlined origination system is shown in figure 4.

The prototype was demonstrated internally, and comments were fed back to the development team and integrated into the specifications. The original prototype was then expanded several times as comments and suggestions were addressed. Thus, the system was built up using an iterative approach in which the system would be revised, then tested and demonstrated. Feedback from the tests and demonstrations would then be incorporated into the new version. This cycle was repeated until the design was finalized.

Alpha Test

The alpha test was an early attempt to simulate how well the system would handle the real-world branch environment. A set of 1400 loans that had already been processed by human underwriters was gathered. Half of these loans had been approved; half had been declined. The Quality Control Department also reviewed the loans, rating the three major areas of the loan: (1) appraisal, (2) ability, and (3) credit. The loans were then passed to CLUES.

CLUES approved 32 percent of the loans passed to it; because the goal had been between 30 and 35 percent, the knowledge engineers were encouraged by the approval rate. However, CLUES also approved 30 bad loans. This result was alarming; in the real world, CLUES could not accept a single bad loan. The knowledge engineers examined the loans carefully and found that a number of bad loans had been approved because the system did not contain a few of the specific underwriting guidelines that were needed. These guidelines had been missed during the series of interviews with the expert underwriters. The rest of the approvals occurred because some data had been entered incorrectly. Although CLUES had approved some bad loans, the causes were the result of simple mistakes and not to the overall design of the system. The system only needed to be fine tuned and have some underwriting guidelines added.

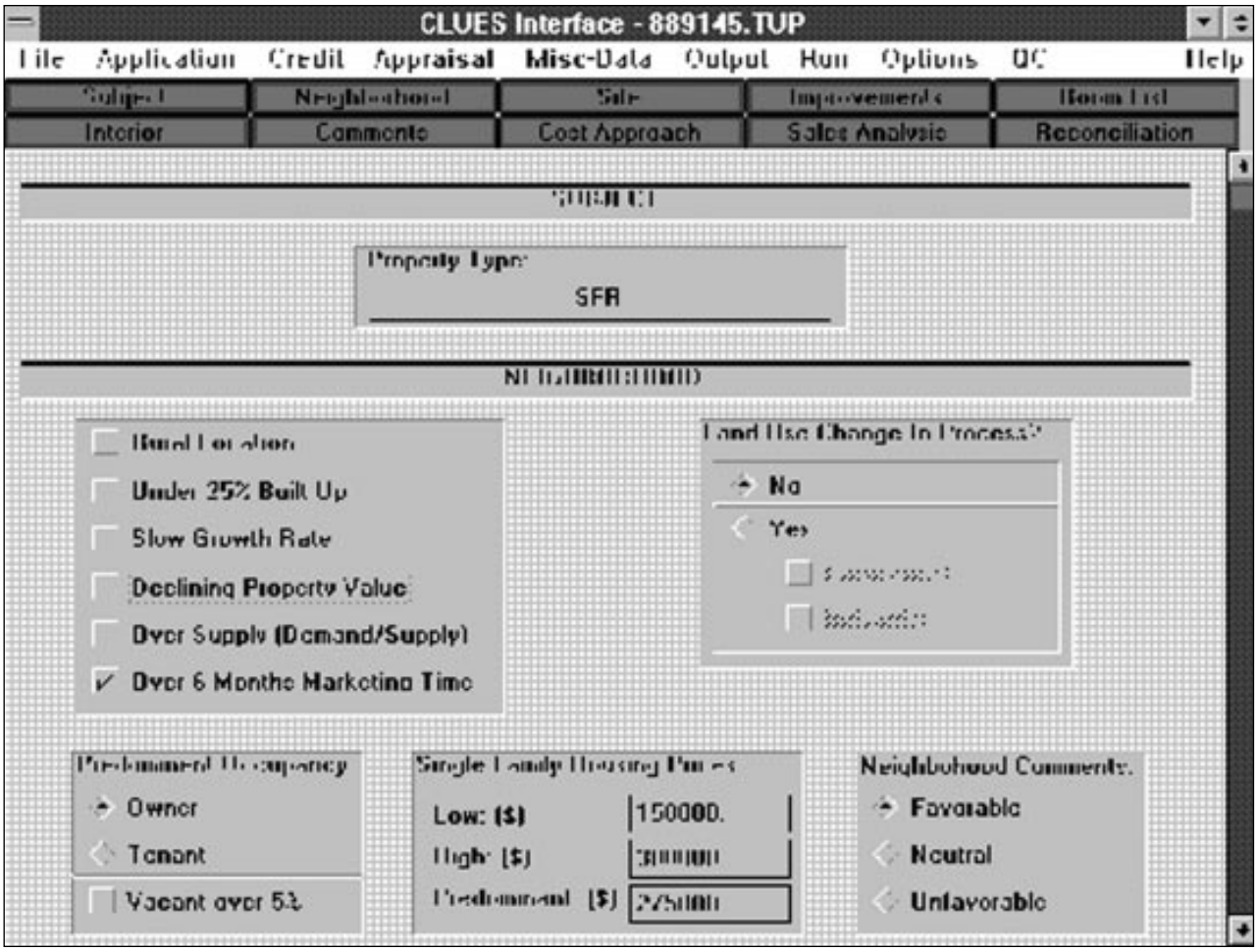


Figure 4. Input Screen of the Stand-Alone Loan-Origination Software.

Beta Test

By the time of the beta test in December 1992, the knowledge engineers were fairly confident that the system would work as planned. The two-month beta test in which the system would be used daily by a selected number of representative branches was the first test of what the system could do in the field. The beta test was of interest not only to the developers but also to upper-level managers and others within the company.

During the beta test, the loans underwritten by CLUES were also manually underwritten. Thus, each loan was underwritten twice. Underwriters were asked to comment on the decisions made by CLUES. The developers could then have a first-hand expert opinion about how CLUES performed. The underwriters, too, received some reassurance that their concerns would be addressed and that their advice and feedback would be able to correct any outstanding issues.

To test the results of the beta phase, the AI team developed the CLUES analysis tool (CAT) on the RS6000. CAT is a dedicated test bed that has a graphic interface to extract the data from the loan database. This database has all the data captured on the loans processed by CLUES, including input, output, and underwriter comments. The developers used CAT to determine what factors caused CLUES to refer a specific loan. If CLUES consistently referred loans with a particular set of factors, but the underwriters approved loans with the same factors, the developers could then adjust the system to make it less conservative in this one area.

Since the beta test, 4600 test cases have exhaustively been created. Each test case represents one aspect of the system that needs to be checked. For example, one case includes a bankruptcy, another includes a property in a flood plain, and another checks for slow mar-

Date	Division	Number of Branches[1]	Percentage of Loan Volume in Dollars
February 1993	Retail	180	20
May 1992	Consumer	6	10
August 1993	Wholesale	70	35
February 1994	Correspondent	4	35

1. The plans for expansion project that by the end of 1995, the retail division will have 400 branches and the wholesale division 100.

Table 2. Clues Deployment to the Branches.

keting rates. By running these test cases, the developers can check that each rule within the system fires when the necessary conditions are present.

The final beta results were impressive. Virtually all the decisions made by CLUES were confirmed by the underwriters themselves. Detailed analyses of the beta results were issued to upper management. For some, the results confirmed their expectations in the potential for the system; for others, the results were the proof needed to convince them that the system would work. The CLUES effort now had the full support of upper management. This support would be necessary for the general release. After the beta test, only minimal adjustments were needed before CLUES was deemed ready for production.

CLUES in Production

In February 1993, CLUES was released to the first branches. Branches within Countrywide belong to one of four divisions: (1) retail, (2) consumer, (3) wholesale, and (4) correspondent. Each branch is responsible for underwriting loans. However, the manner in which the loan is originated differs according to the division. Subsequently, the way in which CLUES is used is slightly different in each division.

The retail and consumer branches start the origination from the time the borrower walks into a branch (retail) or calls a centralized branch (consumer). These two divisions get all the needed information from the borrower, working closely with the borrower at each stage. As problems are discovered, the branch can work with the borrower to investigate and solve problems or suggest another loan type more appropriate for the needs of the borrower.

The wholesale and correspondent divisions, rather than originate loans, receive already packaged loans from either a mortgage broker or another lending institution (correspondent). In both cases, the loan file is complete.

Releasing CLUES to the Branches

Because of the differences in the divisions and the need to build support within the company, CLUES was phased in one division at a time. The first division to receive CLUES was retail, followed by consumer, wholesale, and correspondent. Table 2 shows the dates on which CLUES was introduced to the different divisions. In terms of physical locations, the retail division has branches in 25 states, each with an average of 4 employees; the consumer and correspondent divisions are more centralized, with few locations but as many as 50 to 100 employees; and the wholesale division has numerous medium-sized branches in the Western Pacific area and the Northeast. The entire rollout process was completed in 12 months.

Increasing the Use of CLUES

Simply releasing CLUES to the branches did not ensure a high rate of use. Only if the users accepted CLUES would it be possible to increase use of the system.

Among the hindrances to wide use of the system was the natural tendency of people to distrust new systems. Although many underwriters, especially those accustomed to new advances in technology, immediately took to the system and understood the benefits, other underwriters were more comfortable with manually underwriting loans or were skeptical that CLUES could perform as well as an experienced underwriter. However, by the time CLUES was ready for production, the company was committed to CLUES, and the underwriters were strongly encouraged to employ the new system.

As part of the introductory phase, underwriters could manually underwrite the same loans that were passed to CLUES. In this way, the underwriters could get first-hand experience of how well the CLUES decisions matched

their own. The underwriters, too, developed a good appreciation of how CLUES affected the work flow within their branches. As underwriters began using (and testing) the system, they gained confidence in it and saw their productivity increase. Even those underwriters who might have initially been opposed to the system were won over.

With the growing user acceptance of the system and as more divisions received CLUES, the number of loans processed by CLUES increased monthly. Initially, with only the retail division online, fewer than 1000 loans a month were processed. With the addition of the consumer division in May 1993, the volume increased somewhat—to approximately 1200 loans each month (figure 5). The number of loans processed by CLUES increased substantially with the addition of the wholesale division, which accounts for 35 percent of Countrywide's loans.

Within the first two months of releasing CLUES to the wholesale division, the volume of loans processed through the system more than doubled. At this time, Countrywide experienced a cumulative six-fold increase in total use of CLUES. Currently, CLUES is processing more than 4000 loans a month in the retail division alone and over 8500 loans across all divisions.

Increasing the Acceptance Rate

In addition to increasing the number of loans handled by CLUES, an additional (and ongoing) goal is to gradually increase the acceptance rate of the loans processed through CLUES. The *acceptance rate* is the percentage of loans approved by CLUES. As already mentioned, CLUES was initially conservative to allow the users to gain confidence in the system. The initial acceptance rate was 32 percent.

As CLUES was used in production and as the developers received feedback from the branches, the developers worked to fine tune the system, so it could approve a larger percentage of loans. They did this fine tuning by looking carefully at those loans that had been referred by CLUES but were then later approved by the underwriter. By concentrating on the underwriter comments that explained why the underwriters were approving the loan, the developers were able to better understand what loans should be approved. After this exercise, the knowledge engineers were able in a gradual and controlled manner to increase the acceptance rate to close to 50 percent, which compares with an average of 75 percent for underwriters within Countrywide (figure 6). Over the next year, it is expected that the CLUES acceptance rate will approach 65 percent.

The Effect of CLUES on Work Flow

Any new automated system has an immediate and major impact on the environment into which it is introduced. Adjustments almost always have to be made, some planned and some unplanned. CLUES was no exception.

Before CLUES, the purpose of the loan-origination software was to track loans and print loan documents from the data collected. The data entered in the loan-origination software were not used in evaluating the loan; rather the underwriter referred to the original loan forms. Any data-entry errors introduced in the loan-origination software would typically be caught when the printed forms were compared to the original documents. Data entry then was not all that important and was often delegated to a loan processor or other junior employee with no particular expertise in underwriting.

However, when CLUES was released to the branch, data entry became critical because the data entered are the only information used by CLUES to evaluate the loan. CLUES could not verify the data or compare it against the supporting documents. Unless CLUES referred a loan, there was no need for the underwriter to check the original loan documents.

As a result, the position of loan processor took on added responsibility. Loan processors had to be more qualified than before and fully aware of how data entry affected the final decision. Loan processors also had to be somewhat knowledgeable about underwriting because they were the only ones reviewing the forms for loans that were eventually approved. They needed to do a certain amount of fraud checking, looking for whiteouts or otherwise altered data. They would also have to be knowledgeable about which sources of income to include.

Although the introduction of CLUES actually created more work for the loan processors, forcing them to be more careful, verify, check, and double-check the loan documents, it also meant that they had more responsibility and more direct participation in the important underwriting task. Loan processors could learn first hand about underwriting. Because they entered the data and received back the decisions, they knew the reasons why CLUES approved or referred the loan. The underwriters could thus offload some of their

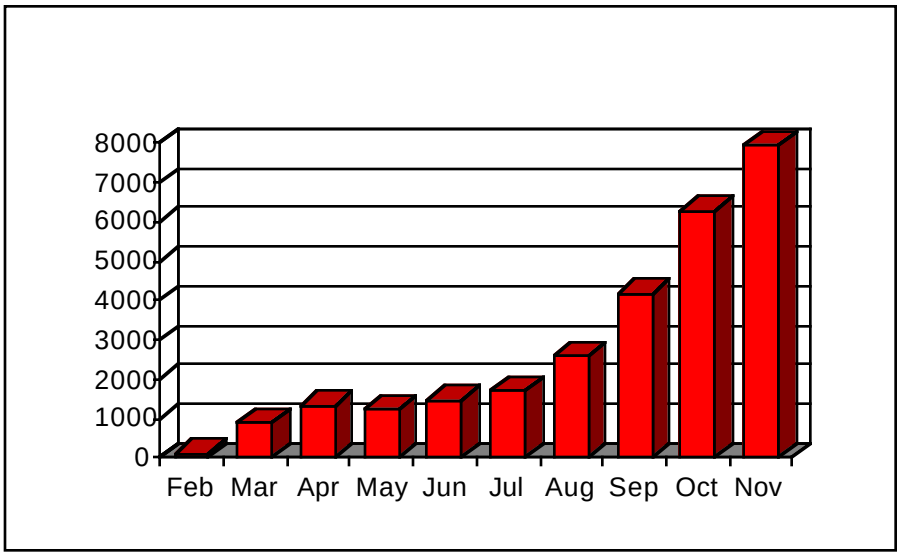


Figure 5. Number of Loans Processed Monthly by CLUES *in 1993.*

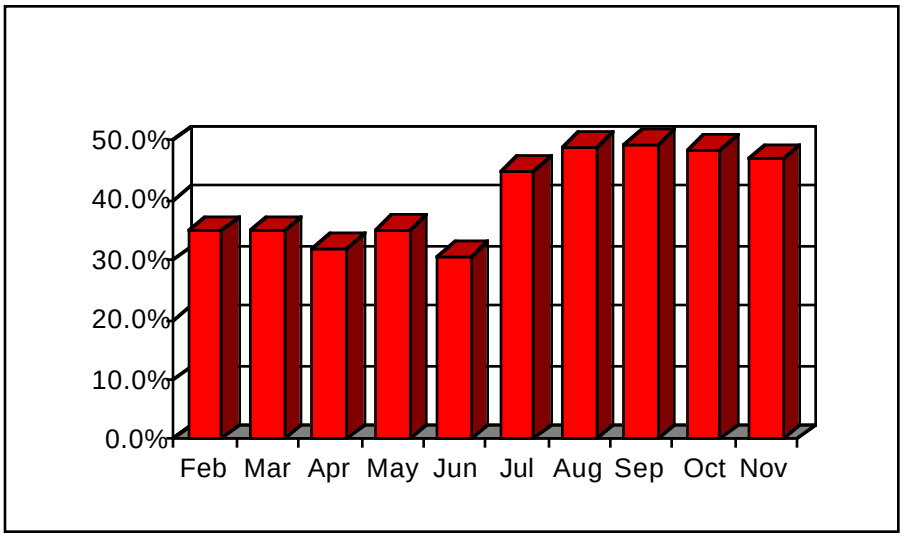


Figure 6. CLUES *Acceptance Rate in 1993.*

work to loan processors and concentrate on more complex loans or other tasks, such as studying the local markets or finding new business opportunities.

In the wholesale and correspondent divisions, the work flow was even more dramatically changed by the addition of CLUES. These two divisions did not originally have the position of loan processor because the loans they bought had already been processed. When CLUES was released to these divisions, a new position had to be created that was responsible for the data entry. Rather than hire an underwriter for this task, the divisions created the new position of junior underwriter.

Benefits to Countrywide

After a year in production, CLUES has already made several significant tangible and intangible contributions.

Cost savings: Underwriting is the most expensive step in the loan-origination process. A typical loan takes about 50 minutes to

review and underwrite. With CLUES, 50 percent of the loans are automatically processed and do not require review by a human underwriter. In the remaining 50 percent of the cases, CLUES provides a detailed review of the loan file and identifies the areas of concern, allowing the underwriter to focus on problematic areas. As a result, a loan processed by CLUES can be underwritten in 10 to 15 minutes.

Currently, the system processes over 8500 loans a month, out of a total of 35,000 loans a month. With the completion and training of the branch personnel, the current use rate of CLUES is contributing $0.91 million a year to company profits. By August 1994, it is expected that the system will process over 30,000 loans a month, resulting in savings over $2.7 million a year.

Consistency of underwriting: Countrywide has been growing at a rate of approximately 15 new branches a month. Several hundred underwriters have been added to increase the capacity of the production divisions from less than $1 billion a month two years ago to over $5 billion a month in December 1993. CLUES has been effective in helping to establish consistency.

Underwriters themselves are not always consistent in the way they look at a loan. As part of the knowledge-acquisition efforts of CLUES, the same loan scenarios were passed to 10 senior underwriters in the company. On a single loan, underwriters would find themselves at opposite extremes. For example, one or more underwriters might decline a specific loan whether or not it was acceptable according to recommended underwriting guidelines. At the other extreme, underwriters were willing to approve the same loan far beyond what the guidelines recommended. This disparity of opinion is primarily the result of each underwriter being more or less sensitive to specific issues in the loan file.

CLUES underwrites the loan the same way every day of the week in every branch—unless the system is intentionally modified to respond to the requirements of a local region. For example, because of the slow housing market in Southern California, underwriters were willing to allow longer marketing times in this region than in other regions. CLUES, too, could be made to respond differently.

Removal of the human bias: Over the last two years, discrimination has been a concern of the U.S. Department of Housing and Urban Development. CLUES does not recognize sex, race, color, or national origin; thus, it cannot discriminate. It makes an unbiased evaluation of the loan file based purely on the quality of the financial profile of the borrowers, their credit worthiness, and the appraised value of the property.

Training tool: An unanticipated benefit of CLUES has been the use of the system as a training tool. Underwriting is a difficult task to teach. It takes a loan processor many years to learn all the guidelines and requirements to advance to the level of underwriter. With CLUES, the loan processors responsible for entering the data are getting immediate feedback on what the strengths and weaknesses of the loan file are and whether the loan should be approved or why it requires further review. This feedback has allowed the loan processors to learn underwriting much faster.

Improved customer service: Because Countrywide underwriters now have to look at only 50 percent of the loans that are funded at their branches, they have a lot more free time to spend with their customers. This availability clearly results in better service to Countrywide borrowers.

Maintenance and Future Enhancement

Initially, the project focused on conventional types of loan. However, once CLUES was in production and had proved that it could successfully handle common loan types, such as conventional single-family residences, the development team could then focus on developing modules capable of underwriting more specialized loan types, such as FHA (Federal Housing Administration) and VA (Veterans Administration) loans or loans for second homes, investment properties, or multiple-family residences.

Knowledge acquisition was once again needed, and more effort was spent in the design and development of tools for the application environment: the test bed, the independent loan-origination interface, and so on. This phase started in January 1993 and is still in progress.

Also, the focus shifted from development to performance. The declared goal was to increase the percentage of loans approved by CLUES but maintain the robustness and the integrity of the system. The system had to become less conservative without a decline in quality.

Future enhancements will allow the system to be modified by the users. Modules are being developed to make it possible for underwriters to add rules within the knowl-

edge base on a limited basis. Such an ability would not only allow users to make immediate changes to the system but would also make CLUES extremely responsive to local needs. For example, the system could be modified to enforce stricter underwriting guidelines in a region that has experienced a large-scale natural disaster.

Summary

After a year of development and another year of production, CLUES has been a success for Countrywide. It more than fulfilled the objectives stated at the beginning, allowing the company to substantially increase the number of loans processed, resulting in a lower processing cost per loan and benefiting the company and its customers alike. The success is even more impressive when one considers that CLUES was developed in less than a year by a small development team. Other departments within the company are now in the process of identifying areas to which AI technologies can be applied.

For the mortgage banking industry as a whole, the success of CLUES has shown that AI has application within the industry. Mortgage banking, which for so long has relied on manual tasks, will likely be reinvigorated as automated loan underwriting becomes widespread. In addition, other applications will be found for AI in secondary marketing and loan servicing.

Finally, the success of CLUES has proved that AI is not simply a set of esoteric methods with little application for real-world problems. On the contrary, AI can be an efficient tool for building sophisticated applications that solve difficult and complex problems. Now that CLUES and other AI solutions have been accepted within the mortgage banking industry, the focus will shift from the design and development of underwriting systems to a search for more applications that will benefit from innovative AI approaches.

Acknowledgments

The success of CLUES is the result of the contributions of many individuals. The authors would like to thank Ralph Mozilo, the project head and champion and our underwriting expert, for his vision and invaluable support throughout the development and deployment of the system. Special thanks to Gene Soda, who tirelessly and cheerfully provided us with his underwriting expertise. We would also like to thank all the following individuals on the AI Team, whose dedicated efforts and contributions turned this vision into a reality: John Edgar, Jennifer Ting, Janis Cole, Thomas Lee, Hoshyar Sayah, Sujatha Marsden, LaDeanne Estrada, and Yongsuk Park. Finally, we would like to thank our users, whose support and feedback have helped us enhance the system since its initial deployment.

Houman Talebzadeh is currently a senior manager at Andersen Consulting. His primary focus is on developing automated systems for financial institutions. Previously, he was the vice-president of AI at Countrywide Funding, where he was responsible for the design, development, and production release of CLUES and several other related systems. Talebzadeh holds a bachelor of science in electrical engineering from Rensselaer Polytechnic Institute and a master of science in computer engineering and a master of business administration from the University of Southern California.

Sanda Mandutianu was a senior knowledge engineer in the Artificial Intelligence Division at Countrywide Funding. She had the lead as a principal developer of the credit and financial risk analysis of CLUES. She holds an MS in physics from the University of Bucharest in Hungary. In the last 14 years, she has been working and leading various projects in AI, including knowledge representation, expert systems, and natural language understanding. She is currently working on an AI application in telecommunications at AGL Systems in Pasadena, California.

Chris Winner is currently a senior member of the Artificial Intelligence Division at Countrywide Funding. He has spent the last five years in the financial industry developing automated underwriting systems. He received his certificate in AI from the University of California at Los Angeles in 1990 and has a graduate degree in industrial engineering from the University of Pittsburgh. His future projects are to expand applications in finance from underwriting to include more automation for instruments in the secondary market.