PREDICTING STOCK MARKET BEHAVIOR THROUGH RULE INDUCTION: AN APPLICATION OF THE LEARNING-FROM-EXAMPLE APPROACH

Helmut Braun
Fisher School of Accountancy, University of Florida, Gainesville, FL 32611
John S. Chandler
Department of Accountancy, University of Illinois at Urbana-Champaign, Champaign, IL 61820

ABSTRACT

An artificial intelligence-based rule-induction approach to the analysis of stock market prediction is presented. A single investment analyst was used as the expert for this study. Predicting intermediate fluctuations in the movement of the market for nonconservative investors was selected as the decision to analyze. Commercially available rule-induction software was used to generate rules that predicted the market calls of the market analyst and the actual movements of the market. Rules predicting actual market movement performed better than rules predicting the analyst's calls and better than the analyst himself. Such an approach may prove useful in designing a decision support system for market analysts or in improving the decision-making processes of such analysts. The dynamic nature of the stock market represents a substantially different decision environment than those previously analyzed by learning-from-example (LFE) techniques. This study provides insights into the limits and applications of LFE approaches.

Subject Areas: Decision Processes, Decision Support Systems, and Financial Planning and Modeling.

INTRODUCTION

Stock market analysts use many techniques in making their predictions, but most use some form of technical analysis such as trend analysis, cycle analysis, charting techniques, or one of many other types of historical data analysis. Each of these tools provides the analyst with information about how the market will move. However, the reliability of each tool is low. As a result, analysts must use many techniques and then integrate the evidence from all of them to make market predictions. Although many market analysts follow a rigorous program of data analysis and evaluation, they find it difficult to explain how individual data elements are combined to arrive at specific market predictions.

Constructing a model of stock market predictions is very difficult. The various technical analysis procedures produce a mix of categorical and quantitative measures. Some are specific financial ratios while others are general indicators of trends. Statistical techniques such as discriminant analysis and regression analysis can analyze the quantitative measures. Techniques such as logit and probit regression analysis can address the categorical measures. A market analyst, however, employs both in practice.

---

*The authors thank Donald Michie for his encouragement and advice on this study and the expert market analyst who participated in the study.

Each factor reviewed by the analyst is important. Some (e.g., Dow Jones indices) are used in every instance and therefore are of greater relative importance. No factor, however, has zero importance. In fact, the singular anomaly in an obscure measure may provide an analyst with the small edge to beat the competition—it may yield large financial rewards or prevent a major disaster. Modeling approaches that focus on factors that best explain the aggregate behavior of the market may be missing just those factors an analyst wants to see. Thus, a modeling approach that can integrate categorical and quantitative measures and can recognize the importance of individual measures is desirable.

One possible approach to analyzing and modeling the market prediction decision is a process called rule induction. This method is an artificial intelligence (AI)-based technique under the general rubric of learning-from-example (LFE) approaches. In this method, the rule-induction system is presented with examples of a decision (its inputs and outcomes) and attempts to induce a decision model. This is in contrast to learning-by-being-told methods that try to extract the decision model explicitly from the decision maker using extensive interviews.

The goal of this study is to investigate one market prediction situation through the rule-induction approach. A commercially available software system, ACLS (Analog Concept Learning System), is used to analyze past examples and formulate decision rules. Rules are generated to predict not only an expert market analyst's prediction of the market but also to predict the actual market's movement. Validation exercises are performed to demonstrate the effectiveness of the rules generated. Analysis of the rules provides interesting insights into the process of applying an LFE approach to a business environment.

OVERVIEW OF THE LFE APPROACH

Several steps are involved in applying the LFE approach. First, a source of expertise must be identified (an expert or a panel of experts). Just what constitutes a "most appropriate source" still is debated in the AI literature. As with any AI-based system, the most crucial step in the process is identifying the decision to be analyzed; some decisions are too easily supported while others are too poorly defined. The expert must determine which decision to support and the appropriate set of outcomes for that decision. This may not be as easy as it sounds. Some decisions have a continuum of possible outcomes—such as those in medicine (alternative treatments) or in business (levels of profit)—and these must be quantized into a finite set of outcome ranges for most LFE approaches.

Once a decision environment has been chosen, the expert must identify the relevant cues. This is similar to what takes place in other AI-based approaches (e.g., production systems like MYCIN [17]). In these other approaches, however, the expert also must identify the relative importance of and interrelationships among these cues. Because decision makers, even experts, may not be able to reveal how they integrate evidence [8], AI-based approaches that depend on this task may have difficulty in building their systems. In the particular approach used in this study, these cues can be a mixture of quantitative and categorical variables.

The third step in an LFE approach is to establish the example data base (EXDB). The cues and outcomes identified above set up the structure of the EXDB. What remains is to fill in the EXDB with examples. These may come from many sources. If this decision has been made many times in the past, there may be documentated cases to refer to. If there are "holes" in the set of examples, cases may be simulated by assigning values to the cues and then asking the expert what the probable outcomes would be. In either case, knowing the line of reasoning employed by the expert, implicitly or explicitly, is not required.

PREVIOUS RESEARCH

Rule induction techniques have been successful in many domains. One of the earliest techniques was Winston's [18] [19] system designed to teach toy block constructions. Lenat's [5] AM system developed proofs in elementary mathematics. A more practical application was the Meta-DENDRAL system's [1] [2] analysis of mass spectroscopy data.

Of particular note have been studies on the identification of soybean disease [6] and chess endgames [12] [16]. Michalski and Chilausky [6], using the program AQ11, induced a rule from a set of 290 examples and tested it on 340 cases. Their rule provided the correct diagnosis of soybean disease in every case; its first choice was correct 97.6 percent of the time. The rule then was compared to one generated from interviews with soybean experts. This second rule was tested on the same 340 cases and proved to be correct only 96.9 percent of the time; its first choice was correct only 71.8 percent of the time.

Experiments by Quinlan [12] and Shapiro and Niblett [16] showed that rule-induction methods can be very successful at inducing rules from very small subsets of examples from large data bases. Both experiments involved rules for chess endgames. A complete table lookup was available to verify the correctness of the induced rule. Quinlan showed that a decision rule induced from a set of examples that represented .07 percent of the entire population was 99.67 percent correct. When the example set was expanded to .36 percent of the population, the induced rule was 99.92 percent correct. These results were confirmed by Shapiro and Niblett. Using ACLS, they induced rules using .2 percent of their data base and found these rules to be over 99 percent correct. O'Rorke [10] compared the rule-induction methodologies of AQ11 and ACLS on several criteria but determined no clear advantage of one over the other.

APPLICATION OF THE LFE APPROACH

The Mechanics of ACLS

ACLS was used in this research as the vehicle for rule induction. The program was written by Paterson, Blake, and Shapiro under the direction of Michie (see [11] for details). It was developed from Quinlan's ID3 (Iterative Dichotomizer 3) program [12] [13] which in turn was developed from Hunt, Marin, and Stone's [4] original work on CLS (Concept Learning System). The induction algorithm of ACLS creates a decision tree that partitions the examples in the EXDB with respect to combinations of cue values. The method for determining the specific branches in the decision tree is based on information theory. Details of the algorithm and an example are given in the Appendix.

Selection and Description of an Expert

This study used a single expert as its source of expertise. The expert was an investment analyst from a medium-sized midwestern city.[1] He owned his own investment firm and had been analyzing the stock market for a period of 12 years prior to this study. For three years prior to this study he had been writing a newsletter in which he gave biweekly recommendations on the stock market. The performance of his recommended stocks was very good over this period (1979 to 1982). Average annual returns exceeded 40 percent for the overall period with returns for 1982 at 41.8 percent. Although his newsletter was not at the top of those ranked by *Barrons*, it was in the upper quartile.

In our initial meeting with the expert the nature of the project and the information that would be required were discussed. The expert was skeptical of computer-based systems. His only previous involvement with computers had been for graphing purposes and the results had been poor. He was not favorably impressed with computers in general. However, he was willing to discuss the details of the methods he used to make his predictions.

Most of the techniques used by the expert would be considered technical analysis. These included using trends and various other charting techniques to determine future market trends. Economic and political factors were considered, but not in any formal manner. The expert's primary tools were charts and graphs of various data for past periods. Some of these had been developed by the expert himself; others were from other analysts such as Joseph Granville, Stan Weinstein, and Robert Farrell.

Definition of Cues and EXDB

The expert explained that the stock market is a complex system with many subsystems operating simultaneously. There are many cycles in the market and some cues work well in the prediction of certain cycles but not well for others. According to the expert, three types of cycles could be selected for study, based on fluctuations in the Dow Jones industrial average (DJI). These were small fluctuations (<10 percent change in DJI), intermediate fluctuations (10 percent–20 percent change in DJI), or large fluctuations (>20 percent change in DJI). We decided to study intermediate fluctuations in the DJI because these probably were the ones most traders used in timing their buy and sell decisions. The DJI, the most commonly quoted indicator of market activity, was used as a reference point in this study.

[1]For reasons of confidentiality the identity of the expert cannot be revealed.

At the initial meeting with the expert the tools and cues used in making a forecast of market behavior were discussed. A list of these is shown in Table 1. From this initial list of approximately 50 variables, a final list of 20 was selected to be used in the induction of prediction rules by ACLS. We reduced the initial list for several reasons. Some of the cues were not relevant to the intermediate market. Others were described insufficiently by the expert as useful signals of market movement. Several were eliminated because of lack of data. Table 1 indicates these reductions and identifies those cues used to build rules. It should be noted that the expert could use all these cues in making his own predictions.

To construct the example data base a specific time period was determined. Initially this time period was from 20 March 1981 to 24 September 1982. Later it was extended to 9 April 1983 to produce a larger EXDB. Data at closing time on Friday for each of the 108 weeks were accumulated.[2] Most of the data were collected from *Wall Street Journal* microfiche by one of the authors. Some of the cues were interpretations of trend-charting techniques. These presented special problems for the authors in drawing and in interpreting the trend lines. More skill is required to draw the trend lines than the authors could develop. Procedures for interpreting these charts after the lines had been drawn, however, were stated explicitly and were followed easily. Thus, for this study, the expert drew the trend lines; to avoid bias in interpretation, the authors applied the formal procedures for evaluation.

Definition of Outcomes

The expert's weekly newsletter contained recommendations, concerning the market as a whole and particular stocks, for specific types of investors (conservative, intermediate, and aggressive). Conservative investors keep their holdings in cash until a long-term bull market is forecast. They never sell stocks short and take advantage of only about one-half of the movement predictions (i.e., those for a bull market). We therefore decided to study only those recommendations made for intermediate and aggressive investors. Data also were collected from a call-in service available to investors during weeks when no newsletter was published.

Three types of outcomes (predictions by the expert) were used to categorize these weekly recommendations: bullish (forecasting an upward trend), bearish (forecasting a downward trend), and neutral (indicating that either call was too risky). These predictions were interpreted for each of the 108 weeks in the EXDB. Limiting the number of outcomes to three is a simplifying assumption, but may be realistic in that more-accurate calls are correct considerably less often. Data also were collected on the actual movements of the market for each of the 108 weeks. Only the outcomes bullish and bearish applied here. Thus, essentially there were two EXDBs. Both had the same values for the 20 selected cues for each of the 108 weeks, but one had the expert's predictions for each week as the outcomes while the other had actual market movements for each week as the outcomes.

[2]Attribute values for all 108 weeks are available on request from the authors.

Table 1: List of potential cues.

1. Put-call ratio, Chicago Board of Options Exchange (PCCBOE)[a]
2. Put-call ratio, American Options Exchange (PCAMEX)[a]
3. Granville Cumulative Climax Indicator (GCCI), nonconfirmation cumulative climax[c]
4. GCCI, retrogress (GRANRET)[a]
5. GCCI, trend (GRANTRN)[a]
6. Weinstein Last-Hour Activity, volume NYSE index nonconfirmation[c]
7. Weinstein Last-Hour Activity, volume NYSE index trend[c]
8. Weinstein Last-Hour Activity, price DJI nonconfirmation[c]
9. Weinstein Last-Hour Activity, price DJI trend[c]
10. Dow Jones moving average, 10-day cycle (DJI10)[a]
11. Dow Jones moving average, 30-day cycle (DJI30)[a]
12. Dow Jones moving average, conjointly (when DJI10=DJI30)[b]
13. On-balance volume DJI, trend (OBVDOW)[a]
14. On-balance volume DJI, nonconfirmation[c]
15. Cash of DJI, trend (CASHDOW)[a]
16. Cash of DJI, nonconfirmation (NETCDOW)[a]
17. Specialist short sales, ratio vs. odd-lot sales (SSOLS)[a]
18. Specialist short sales, 4-week moving vs. total shorts (SSTS)[a]
19. Market pressure index, 1-day moving average (MPI)[a]
20. Intensity DJI, trend (expert's trend model)[b]
21. Intensity DJI, nonconfirmation (expert's trend model)[b]
22. Dow theory, compare transportation to industrials[b]
23. Dow theory, over-bought over-sold oscillator (OBOSOS)[a]
24. NYSE composite index, trend and field trend (NYSECI)[a]
25. NYSE composite index, nonconfirmation vs. DJI price index[c]
26. Dow Jones Industrial, trend and field trend (DJIFT)[a]
27. Dow Jones Transportation, trends and field trend (DJTT)[a]
28. Dow Jones Transportation, trend breaks (DJTTB)[a]
29. S&P front spread, trend[c]
30. S&P front spread, nonconfirmation[c]
31. Cash of DJI Weekly, trend[b]
32. Cash of DJI Weekly, nonconfirmation[b]
33. Dow Jones figure point objective, 5 points[c]
34. Dow Jones figure point objective, 10 points[c]
35. Optimism-pessimism index, trend[c]
36. Optimism-pessimism index, nonconfirmation[c]
37. Optimism-pessimism index, 10-point figure chart[c]
38. Optimism-pessimism index, 25-point figure chart[c]
39. Wycoff Wave, trend (WWTRN)[a]
40. Wycoff Wave, nonconfirmation (WWREV)[a]
41. Trend barometer, momentum[b]
42. Trend barometer, force[b]
43. Trend barometer, technometer[b]
44. Ratio of ratios, trend of 6-day ratio[b]
45. Ratio of ratios, value of 6-day ratio (RATRAT)[a]
46. Ratio of ratios, trend of 10-day ratio[b]
47. Ratio of ratios, value of 10-day ratio[b]

[a]Variable used as cue to develop rule
[b]Variable primarily for long-term fluctuations
[c]Unable to use because of insufficient data

Phase I—Initial Pilot Results

In phase 1 of this study, we used an initial data base of 80 examples of our expert's predictions dating from 20 March 1981 to 24 September 1982. The first test split the examples into two groups of 40 on a random basis. One group was used to induce a rule using ALCS; the other was used to test the induced rule. The induction produced a 29-node rule that used 10 cues from the data base. This rule was then applied to the remaining 40 examples, and the degree of consensus between rule-generated predictions and actual predictions was compared. The rule correctly matched the predictions of our expert in 23 of 40 cases (a hit rate of 57.5 percent).

In an effort to improve these results, the number of examples used to construct the rule was increased to 60. This left only 20 examples on which to test the induced rule. This second induced rule matched the expert's predictions on 13 of the 20 examples (a hit rate of 65 percent).

The last rule induced during phase 1 split the actual-market-movement data base into two groups of 40 and constructed a rule that would predict actual movements of the market. This induced rule had 27 nodes and used 8 of the 20 cues. It had a hit rate of 62.5 percent.

We discussed these results with our expert. He examined the decision rules that had been generated by ACLS and was favorably impressed by the general structure of the decision tree and also by the cues that had been included. Our expert considered most of those cues included in the trees to have higher predictive power than cues that had been excluded. He also was favorably impressed that a rule could be generated that would predict actual market movements correctly more than 60 percent of the time.

Phase II—Validation Exercises

Because the EXDBs contained both quantitative and categorical cues, external validity for the entire decision modeled could not be established by comparing our technique to a standard statistical technique such as discriminant analysis. Therefore, it was decided to establish internal validity through repeated applications of the holdout technique used in phase 1. Each EXDB was expanded to 108 examples. Two situations were replicated: one predicting the expert's predictions and one predicting actual stock market movements. The EXDB was split into two equal subsets five times; five rules were induced for each situation. Table 2 presents the results of these exercises.

Our main question was, "How good is the average 64.4 percent correct prediction of actual market movements obtained using ACLS?" In comparison, for the complete set of 108 examples the expert was correct 60.2 percent of the time. Thus, the induced rules performed slightly better than our expert. (A note is needed here to explain that if the analyst or the rule predicting the analyst called for a neutral condition and the market continued in the same direction as the last call made by the analyst, we considered this a correct call. This is because investors would hold their current positions and thus profit from the decision.)

Table 2: Results of validation exercise.

Trial	Nodes	Cues	Number Correct	Percent Correct
A. Predicting the expert's calls				
1	36	8	35	64.8
2	37	9	24	44.4
3	40	11	27	50.0
4	38	7	24	44.4
5	43	9	28	53.8[a]
Average	39	8.8	27.6	51.5
B. Predicting the actual market				
1	20	6	36	66.7
2	23	7	38	70.4
3	22	6	36	66.7
4	20	6	33	61.1
5	26	9	31	57.4
Average	22.2	6.8	34.8	64.4

[a]Two situations not predicted (28/52=53.8 percent)

It also is interesting to note that it was more difficult to predict calls made by the expert than it was to predict actual market movements. This may indicate that our expert was not consistent in making his market calls. Table 2 also shows that predictions of the actual market were less variable. It was learned in later discussions with the expert that some changes had been made in the types of data collected. It also appears that the rules used to classify data to predict the expert's recommendations were more complex than those used to predict the actual market; the average number of nodes was higher (39.0 to 22.2) and the average number of cues was higher (8.8 to 6.8). These measures are not absolute indicators of greater complexity, but provide general rules of thumb for rule-induction techniques.

ANALYSIS OF RESULTS

Predicting Market Behavior

One of our most interesting findings was that predictions of actual market movements were more accurate than predictions of the expert's recommendations. As discussed above, this may have been the result of the expert changing some of his analytical techniques during the data collection period. Over the two-year period of study, certain cues and their associated values may have taken on different meanings for the expert and this may have made a difference in the consistency of his recorded predictions. However, such changes probably take place continually; markets are dynamic and a decision rule effective in one time period may not be effective in other time periods. The implications this has for LFE techniques are discussed later.

It also was interesting to note that the ACLS-generated rule predicted actual market movements better than our expert. This may be somewhat deceiving, however, because the decisions a successful investor needs to make are not as simple as this study made them appear. For example, individual stocks do not always move exactly with the market—they may move ahead or behind the rest of the market by short periods of time. As a result, our expert sometimes is cautious and instructs investors to begin selling before an actual market peak is reached or to begin buying before a trough is reached. Sometimes, when a market move is missed, he advises investors to keep their holdings in cash to protect their capital. These calls were counted against him in this study. The expert might well have been able to call the last two weeks of a three-week rise correctly, but if 60 percent of the move already had occurred during the first week, it would have been too late for investors to make profits. Thus, our expert would have recommended a neutral position and these weeks also would have been counted against him when measuring the accuracy of his predictions.

Furthermore, many of this study's interpretations are somewhat subjective. It is difficult to make absolute statements about market predictions without more-detailed descriptions of market performance. The interpretations made in this study also may be idiosyncratic to our particular expert. An interesting extension of this study would be to analyze the decisions of other market analysts and compare the results.

Rule-Induction and LFE Approaches

The effectiveness of any LFE approach depends on the examples in the EXDB. One crucial aspect concerns the size of the EXDB [3]. One rule of thumb applied by AI researchers is that more examples always are better. However, another perspective is that not the number but the particular content of the EXDB (i.e., positive, negative, and "near misses") is important [18]. The problem is that defining a "near miss" for a stock market prediction problem is not as easy as it was in Winston's [19] toy block problem. Finally, researchers such as Quinlan [12] and Shapiro and Niblett [16] demonstrated outstanding accuracy with a relatively minute EXDB of chess positions. As reported in the section on phase 1, in this study a marginal improvement in accuracy was obtained by increasing the size of the EXDB. It may not, in fact, be valid to compare a highly dynamic decision problem such as stock market prediction to a chess decision problem that can be completely enumerated.

The importance of individual examples to LFE approaches is shown in Table 2. For each of the ten rules induced, a different set of cues resulted. If these results were obtained using a statistical technique, we would question the stability and robustness of the approach. For ACLS, each of these ten cases represents a different set of 54 individual examples. Whereas a statistical approach attempts to build a "straight wall" between outcome groups that minimizes misclassifications, ACLS attempts to build a "meandering wall" customized to each individual example so that none are misclassified. This implies that great care should be taken to include the correct set of examples in the EXDB (i.e., to select a set of examples that covers the spectrum of possible situations). Significant situations not represented in a sample EXDB should be entered with a simulated outcome obtained from interviews with the expert. Anomalies and "holes," whose importance can be diluted in statistical approaches, on the other hand, may receive unwarranted importance in LFE approaches.

This biasing effect can be discussed with respect to the data base of examples in this study. Both positive and negative examples (bull and bear) are needed. If, for example, only cases that produced a bull outcome were analyzed, a simplistic prediction rule would be "the market is always bull" which would be false. In this case, negative (bear) examples are needed to bound the rule for predicting the positive (bull) outcome.

One aspect of EXDB composition peculiar to this study is temporal effectiveness. Decisions on chess endgames and toy blocks do not change over time, but stock market predictions may change daily (or even more frequently). New economic news, a political crisis, or a change in financial status can alter an expert's decisions. To demonstrate this, we split the actual-market EXDB into two groups (one containing first-year data, the other containing second-year data) and examined it sequentially. A rule was generated from the first group to predict actual market movements. We then tested this rule against the second group on a quarterly basis. The percentage of correct predictions per 13 calls per quarter (1st quarter: 77 percent, 2nd quarter: 62 percent, 3rd quarter: 62 percent, 4th quarter: 54 percent) clearly shows how the effectiveness of the rule wanes as the time from rule generation to rule application widens. This suggests that such rules should be monitored constantly and updated to maintain their accuracy.

The question of predicting next year's movements based on this year's activity also was analyzed using discriminant analysis. Of the 20 cues, 7 were quantitative (PCCBOE, PCAMEX, OBOSOS, SSOLS, SSTS, MPI, and RATRAT). A separate data base was constructed for years 1 and 2 with just these cues. The SAS DISCRIM procedure [14] was applied to year 1 to derive a discriminant function which then was tested against year 2. The accuracy results per quarter (1st quarter: 77 percent, 2nd quarter: 46 percent, 3rd quarter: 38 percent, 4th quarter: 54 percent) indicate a generally lower level of accuracy than was obtained by ACLS and an inconsistent trend.

We also compared the characterization accuracy of each method (i.e., the accuracy of the discriminant function when applied to the base examples, in this case, those in year 1). The discriminant analysis method was correct only 70 percent of the time, while ACLS was correct 100 percent of the time. This again emphasizes the importance of individual observations to ACLS. ACLS attempts to build a rule that covers all cases equally, allowing no errors, whereas statistical approaches attempt to minimize the aggregate deviation, allowing for individual errors. This may force ACLS at times to produce an overly complex and idiosyncratic rule that may not be an effective predictor. ACLS (or LFE approaches in general), however, may produce better characterization rules.

One final point concerns what we learn about the decision under scrutiny through an LFE approach. Analysis of the induced decision rule may help a decision maker improve his or her own decision processes. In this study, for example, a seventh and final random halving of EXDB was made and two rules were generated: one predicting the expert's call and the other predicting actual market movement. The former correctly predicted the test group 50 percent of the time, the latter 59 percent of the time (significant at the .05 level). This difference may partially be explained by examining the cues in each rule.

Market-rule cues: PCCBOE, PCAMEX, MPI, GRANRET, RATRAT

Expert-rule cues: PCCBOE, PCAMEX, MPI, GRANRET, DJTTB, NETCDOW, OBOSOS, DJTT, SSTS, GRANTRN

Notice that the market rule cues form almost a subset of those in the expert rule. This suggests that our expert may have overwhelmed himself with information. It might be better to concentrate on a reduced set of cues.

CONCLUSION

A rule-induction approach to analyzing the stock market prediction decision has been presented. A particular segment of the market (intermediate fluctuations) and a particular type of investor (nonconservative) were selected for study. The results demonstrate that for this decision environment, a rule-induction technique can produce predictions as good as those of an expert market analyst. Such a tool can be beneficial in developing a decision support system for market analysts or in improving market analysts' own decision-making processes. The dynamics of the stock market make the particular decision environment studied quite different from those previously analyzed using LFE approaches. It should provide a rich test bed for further analysis of the effectiveness of LFE and rule-induction techniques. [Received: March 3, 1986. Accepted: November 10, 1986.]

REFERENCES

[1] Buchanan, B. G., & Feigenbaum, E. A. DENDRAL and meta-DENDRAL: Their applications dimension. *Artificial Intelligence*, 1978, *11*(1,2), 5-24.

[2] Buchanan, B. G., Smith, D. H., White, W. C., Gritter, R. J., Feigenbaum, E. A., Lederberg, J., & Djerassi, C. Applications of artificial intelligence for chemical inference XXII. Automatic rule formation in mass spectrometry by means of the meta-DENDRAL program. *Journal of the American Chemical Society*, 1976, *98*, 6168-6178.

[3] Chandler, J. S. *Studies on the limits of rule inductance techniques.* Working paper, University of Illinois, Department of Accountancy, 1986.

[4] Hunt, E. B., Marin, J., & Stone, P. T. *Experiments in induction*. New York: Academic Press, 1966.

[5] Lenat, D. B. *AM: An artificial intelligence approach to discovery in mathematics as heuristic search.* Unpublished Ph.D. dissertation, Stanford University, 1976.

[6] Michalski, R. S., & Chilausky, R. L. Learning by being told and learning from examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis. *Journal of Policy Analysis and Information Systems,* 1980, *4*(2), 125-160.

[7] Michalski, R. S., & Dietterich, T. G. A comparative review of selected methods for learning from examples. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning.* Palo Alto, CA: Tioga Publishing, 1983.

[8] Michie, D. The state of the art in machine learning. In D. Michie (Ed.), *Introductory readings in expert systems.* New York: Gordon & Breach, 1982.

[9] Mozetic, I. *NEWGEM: Program for learning from examples. Program documentation and user's guide.* Urbana: University of Illinois, Department of Computer Science, Artificial Intelligence Laboratory, 1985.
[10] O'Rorke, P. *A comparative study of inductive learning systems AQ11 and ID3* (Intelligent Systems Group Report 82-2). University of Illinois, Department of Computer Science, 1982.
[11] Paterson, A., & Niblett, T. *ACLS user manual.* Glasgow, Scotland: Intelligent Terminal Ltd., 1982.
[12] Quinlan, J. R. Discovering rules from large collections of examples: A case study. In D. Michie (Ed.), *Expert systems in the micro electronic age.* Edinburgh, Scotland: Edinburgh University Press, 1979.
[13] Quinlan, J. R. Semi-autonomous acquisition of pattern-based knowledge. In D. Michie (Ed.), *Introductory readings in expert systems.* New York: Gordon & Breach, 1982.
[14] SAS Institute. *SAS user's guide: Statistics.* Cary, NC: 1982.
[15] Shannon, C. E., & Weaver, W. *The mathematical theory of communication.* Urbana: University of Illinois Press, 1975.
[16] Shapiro, A., & Niblett, T. Automatic induction of classification rules for a chess endgame. In M. R. B. Clarke (Ed.), *Advances in computer chess* (Vol. 3). Edinburgh, Scotland: Edinburgh University Press, 1982.
[17] Shortliffe, E. *Computer based medical consultations: MYCIN.* New York: Elsevier, 1976.
[18] Winston, P. H. *Learning structural descriptions from examples* (Technical report AI-TR-231). Cambridge: Massachusetts Institute of Technology, 1970.
[19] Winston, P. H. Learning structural descriptions from examples. In P. H. Winston (Ed.), *The psychology of computer vision.* New York: McGraw-Hill, 1975.

APPENDIX

Description of ACLS Induction Algorithm

This description draws heavily on the reasoning behind the conceptual designs of Hunt et al. [4] and Quinlan [12] [13] and the practical design of Paterson and Niblett [11]. A comparison of induction approaches can be found in Michalski and Dietterich [7] and O'Rorke [10].

ACLS begins with a set of examples, EXDB, to be classified according to a set of classes or categories. These examples ultimately are divided into a decision tree where each leaf node contains examples of the same class or category. The following sequence of steps is used to create this decision tree:

1. Create the root node and associate all examples with it.
2. If all examples at the current node are of the same class then stop.
3. Choose the best attribute to split the examples at the current node. This is done according to an information theoretic measure described below. The attribute chosen may be either logical [categorical] or integer [quantitative]. If [it is] logical then the tree splits at this point with one branch for each value of the chosen attribute. If [it is] integer then a binary split is performed by partitioning the integer range about a threshold value. The threshold value is chosen to split the examples at the current node according to an optimality criterion (see below).
4. The current node now has two or more branches. The examples at the current node are then associated with the branch nodes according to the value of the split attribute. Examples which have a "don't care" are associated with all branch nodes.
5. For each branch node go to step 2. [11, p. 7, reprinted with permission]

The splitting algorithm is based on information theory. Its fundamental premise is that a decision tree can be thought of as a source of information. The decision tree provides information about a collection of examples by classifying them into groups based on attributes or cues [13]. The attribute selection part of ACLS is based on the assumption that the cognitive complexity of the decision tree is strongly related to the information content of the message conveyed by the tree [11]. For example, say a decision tree will be induced to classify two classes of stock market data on several attributes. Let the classes be bearish (B) and bullish (b). If the probabilities of these are $P(B)$ and $P(b)$ respectively, then, based on the communication theory view of information developed by Shannon and Weaver [15], the expected information content conveyed by the tree (ICT) would be

$$ICT = -[P(B) \cdot \log 2(P(B))] - [P(b) \cdot \log 2(P(b))].$$

Given that ICT is the information content of the entire tree, the objective at each node is to choose the attribute that provides the maximum information about the classification being considered. In order to determine this, ACLS evaluates the information content of each of the immediate subtrees formed when each attribute is chosen. The sum of the information content from each of the subtrees, multiplied by the probability of taking the branch attributes that lead to those subtrees, yields the remaining information necessary to complete the tree ($RIC(i)$) at the ith attribute of the current node. Thus for each attribute i, ACLS determines $IC(i)$ as $IC(i) = ICT - RIC(i)$, where $IC(i)$ = information content of attribute i at current node, ICT = information content of the entire tree, and $RIC(i)$ = remaining information content to complete tree if attribute i is chosen for the split. ACLS then selects the attribute i that maximizes $IC(i)$ as the attribute on which to branch.

The above splitting algorithm works as is for logical or categorical attributes. In the case of integer or quantitative cues, however, the value of $IC(i)$ and the integer value on which to split the attribute must be calculated. Assume that attribute i has N given attribute values labeled $V(1), \ldots, V(j), \ldots, V(N)$. Further assume that all $V(j)$ are ordered in increasing value. For each j, $1 < j < N$, we can split the values into two subsets: $\{V(1), \ldots, V(j)\}$ and $\{V(J+1), \ldots, V(N)\}$. These subsets define a value for $IC(i)$. ACLS chooses $V(j)$ such that $IC(i)$ is maximized. If $V(j)+1 = V(j+1)$, the attribute is split on $V(j)$; otherwise, the split value is $(V(j)+V(j+1))/2$. This ability to handle integer-valued classes separates ACLS from other rule-induction techniques such as NEWGEM [9] which can handle only categorical attributes.

An Example of the ACLS Algorithm

Possible Outcomes: bullish or bearish

Cues:

1. Put-call ratio on the Chicago Board of Options (PCCBOE)
   Values: positive (pos) or negative (neg)
2. On-balance volume of the Dow (OBVDOW)
   Values: positive (pos), neutral (neu), or negative (neg)
3. Granville Cumulative Climax Indicator (GCCI)
   Values: positive (pos) or negative (neg)

Set of Examples:

PCCBOE	*OBVDOW*	*GCCI*	*Outcome*
pos	neg	pos	bullish
neg	neg	neg	bearish
neg	neu	pos	bullish
pos	pos	pos	bearish
neg	pos	pos	bearish
neg	neg	pos	bullish
neg	pos	neg	bearish
pos	neg	neg	bearish

This set of examples contains three situations in the class bullish and five in the class bearish. With a known set of examples, the probabilities can be estimated using the relative frequencies. We can calculate *ICT* as $ICT = -[3/8 \log 2\ 3/8] - [5/8 \log 2\ 5/8] = .954$.

If the first attribute is tested, the tree structure begins:

PCCBOE:
Negative: neg,neg,neg,bear
neg,neu,pos,bull
neg,pos,pos,bear
neg,neg,pos,bull
neg,pos,neg,bear
Positive:
pos,neg,pos,bull
pos,pos,pos,bear
pos,neg,neg,bear.

In the positive branch, one-third of the examples are bullish and two-thirds of the examples are bearish. Thus the information still needed for the positive branch would be $-[1/3 \log 2\ 1/3] - [2/3 \log 2\ 2/3] = .918$. In the negative branch, two-fifths of the examples are bullish and three-fifths of the examples are bearish. Thus the information still needed for the negative branch would be $-[2/5 \log 2\ 2/5] - [3/5 \log 2\ 3/5] = .971$. The total information content of the remaining subtrees after splitting on the cue PCCBOE, given that three of the eight examples are in the positive branch and the remaining five examples are in the negative branch, would be $[3/8 \cdot .918] + [5/8 \cdot .971] = .951$. Thus the information gained by using PCCBOE as a branch would be $.954 - .951 = .003$.

If the second attribute, OBVDOW, were chosen for the root node, the tree would begin:

OBVDOW:
Positive:
pos,pos,pos,bear
neg,pos,pos,bear
neg,pos,neg,bear
Neutral:
neg,neu,pos,bull
Negative:
pos,neg,pos,bull
neg,neg,neg,bear
neg,neg,pos,bull
pos,neg,neg,bear.

At this juncture, two of the branches (the neutral and positive branches) are complete since all examples at the end nodes are of the same class. Therefore, the remaining information necessary to complete these branches is 0. The other branch (negative) is not complete and, because it has two bullish and two bearish outcomes, the information to complete it would be: $-[2/4 \log 2\ 2/4] - [2/4 \log 2\ 2/4] = 1.00$. The total remaining information needed to complete the tree if it were split on the attribute OBVDOW would be $[4/8 \cdot 1.00] + [1/8 \cdot 0] + [3/8 \cdot 0] = .500$. Thus the information gained by testing the attribute OBVDOW would be $.954 - .500 = .454$.

Performing these same calculations for the attribute GCCI would show that the information to be gained by testing the attribute would be .347. Based on this data, the algorithm would select OBVDOW to become the root node of the tree. It then would continue to follow the algorithm and test nodes at the ends of all branches where more than one class of examples is present (in this case, only the OBVDOW negative branch). The resulting decision tree for this limited example is the following:

OBVDOW:

$$\text{Negative:}\quad \text{GCII}\begin{cases}\text{Neg}\rightarrow\text{Bearish}\\ \text{Pos}\rightarrow\text{Bullish}\end{cases}$$

$$\text{Neutral:}\quad \rightarrow \quad \text{Bullish}$$

$$\text{Positive:}\quad \rightarrow \quad \text{Bearish}$$

Helmut Braun is currently provisional Assistant Professor in the Fisher School of Accountancy at the University of Florida. He is nearing completion of his doctoral work in the Department of Accountancy at the University of Illinois at Urbana-Champaign. For the past several years, Mr. Braun has been actively involved in the construction and application of expert systems for business.

John S. Chandler is Associate Professor of Accountancy and Director of the Office for Information Management at the University of Illinois at Urbana-Champaign. He received his doctorate in computer and information science from The Ohio State University. Dr. Chandler is involved in constructing expert systems in accounting and auditing and also has interests in international information systems, the performance evaluation of information systems, and the strategic applications of information systems.