

## DISCUSSION OF

# A Computational Model of Auditor Knowledge and Reasoning Processes in the Going-Concern Judgment

Daniel E. O'Leary

### INTRODUCTION

Since its presentation at the Audit Judgment Conference in 1984 and the American Accounting Association meeting in New York (Biggs and Selfridge 1986), the GCX research summarized in Biggs et al. (1992) has had substantial impact on research in auditing expert systems. GCX set a standard for subsequent researchers.

#### Contributions of Biggs et al. (1992)

The GCX model developed in Biggs et al. has a number of substantial contributions. First, the model is rich in domain knowledge about the going-concern process. Knowledge in the model takes many forms: financial knowledge, company finances, event knowledge (both actual and normal), company operation knowledge, and procedural knowledge. In many ways the model is as rich in knowledge as a case study might be.

Second, a cognitive model of expert auditor behavior is captured in the transcripts, schematic diagrams, and prototype expert systems. That model represents a theory that can be tested or used for future research in going-concern problems.

Third, rather than just using financial data to investigate the going-concern problem, Biggs et al. (1992) provides the reasoning to understand what events may have led to a going-concern situation. Fourth, the system can then rea-

son about management's plans to mitigate the impact of those events.

Fifth, much of the model's knowledge is episodic. Thus, the research indicates the importance of so-called case-based reasoning to auditor reasoning.

Sixth, the original research having been done in 1986 (and much of it probably prior to that) indicates that the use of a computational model in research is not fleeting. After six years the research is still interesting and fresh.

The reader can examine the GCX research in Biggs et al. (1993) and Selfridge and Biggs (1990) to study the importance of the paper. However, the purpose of a commentary of this type is, to a large extent, critical.

#### Themes of this Commentary

There are three critical themes in this paper. First, it is argued that "computational models" (or qualitative models or expert systems or intelligent systems—whatever we choose to call them) are useful and unique research artifacts. Computational models are particularly well suited to those situations where theory generation or theory testing is required or substantial domain knowledge is desired in the model. However, computational models also have a number of limitations to their use in research.

---

*Daniel E. O'Leary is an Associate Professor at the University of Southern California.*

Second, although GCX is a sophisticated and intelligent system, there are some important limitations in the current model approach. The model is evaluated primarily from a qualitative modeling perspective to derive those criticisms.

Third, the model is evaluated from the perspective of the auditing and accounting domain. Unfortunately, GCX seems to make some very limiting assumptions about auditors. These naive assumptions are characterized as "Green Eye Shade Accountant/Auditor" assumptions.

This analysis of the potential limitations of GCX is intended to facilitate or generate additional research, not to suggest that the current GCX should accommodate these issues. All research has scope limitations. In addition, many of the criticisms offered here may be extended to other computational models and other types of audit research.

#### COMPUTATIONAL MODELS AS RESEARCH ARTIFACTS

This section examines different research artifacts for their contribution to the accounting and auditing research process. In addition, it discusses the advantages and disadvantages of computer prototype models of accountant and auditor behavior.

##### Research Artifacts

An artifact is "a usual object (such as a tool or ornament) showing human workmanship or modification as distinguished from a natural object" (Webster's unabridged dictionary, 1971). A research artifact will be defined as an object that results from the research process or is the object of study of a research process. Such artifacts may include the data from an experiment, records of the analysis of a set of data, research papers, and a variety of other artifacts including prototype computational models.

Different artifacts result from different research approaches. For example, in an experiment, data are generated, while in a case study, case-specific data are generated. Different artifacts have different characteristics. Some artifacts are entirely numeric, some entirely interpretative.

##### Artifacts in the Development of GCX

The development of GCX by Biggs et al. (1993) generated a number of different research artifacts. First, the transcripts with the expert auditors form an artifact, one of the earliest in the process. Second, those transcripts were distilled into schematic diagrams that provided the primary basis for the development of the computational program. Third, a rule-based expert system was developed to make judgments. Fourth, a large frame-based system, developed in LISP, was used as the final format.

Each of these artifacts was designed to capture substantial amounts of domain information. However, one difficulty with limiting the research to the transcripts and schematic diagrams would have been difficulty in validating the model. This has led researchers to use tools and methods to generate artifacts that can be more readily validated, using such paradigms as artificial intelligence (AI).

Each of the two computer programs provides statements of auditor problem-solving models. Both of those artifacts apparently can be manipulated for experimentation purposes.

##### Computer Program Prototypes

Computer program prototypes, like those developed in GCX, provide research artifacts with different objectives and attributes than other types of auditing research. Newell and Simon (1976, 114) discussed the use of computer program prototypes in research:

Computer science is an empirical discipline. ... Each new program that is built is an experiment. It poses a question to nature and its behavior offers clues to an answer. ... [A]s basic scientists we build machines and programs as a way of discovering new phenomena and analyzing phenomena we already know about. Society often becomes confused about this believing that computers and programs are to be constructed only for the economic use that can be made of them (or as intermediate items in a developmental sequence leading to such use). ... It needs to understand that, as in any science, the gains that accrue from such experimentation and understanding pay off in the permanent acquisition of new techniques; and

that it is these techniques that will create instruments to help society in achieving its goals.

In a discussion of research in artificial intelligence, Cohen and Howe (1988, 19) noted that "Whereas ... much research in behavioral sciences is concerned with teasing apart the components of behavior and their causal interrelationships, empirical AI is concerned with putting those components together in one box to produce behavior." Cohen and Howe (1988, 19) also observe that the job of empirical AI researchers is "not to find out [by statistical induction] how the average human organism (or organization) works; but rather to build artificial systems that work in particular ways."

Discussions such as these have led McCarthy et al. (1992, 53) to note that the criticisms that the development of *any* knowledge-based system is not research "are unfounded and are in fact more attributable to a critic's lack of schooling in computer science than to any conceptual shortcomings in the actual systems research method."

#### **Advantages of Prototype Artifacts**

Prototype artifacts can be used for theory testing and theory generation. GCX is proposed as a computer program representation of a theory of going-concern judgments. Structure is brought to a problem that was previously less structured or used a more naive model. Winston (1984, 2) observed:

Computer models force precision. Implementing a theory uncovers conceptual mistakes and oversights that ordinarily escape even the most meticulous researchers.

As a prototype, GCX can be used as the basis of experimentation. In a research environment where experimentation is receiving increasing attention, prototypes offer a number of advantages over other research artifacts. As noted by Winston (1984, 2):

Computer programs exhibit unlimited patience, they require no feeding and do not bite. Moreover, it is usually simple to deprive a computer program of some piece of knowledge in order to test how important that piece really is.

#### **Disadvantages of Prototype Artifacts**

Although prototypes allow the capture and testing of rich models of auditor behavior, there are some limitations. Replication and extension of prototypes are difficult and lack standardization. Prototypes are difficult to transfer from one computing environment to another, for a number of reasons: special fixes in the program to accommodate nuances are likely to consist of up to ten percent of the code (e.g., Fox 1979); research prototypes often are not well-documented; the purposes of computer code may be difficult to understand; researchers may use different operating systems. Thus, when someone other than the developer of the prototype attempts to use or extend the system, subtle knowledge contained in the program is lost. Thus, we are not likely to see different researchers directly extend or generalize GCX or when they do, the results may be very limited.

As is the case with virtually all other research artifacts, prototypes are not theory independent. In particular, various causal and programming structures are indicative of different assumptions. For example, the use of memory organization packets (MOPs) by Biggs et al. (1993) indicates a certain school of cognitive science and artificial intelligence.

Unfortunately, unlike statistical methods, there are no well-defined evaluation rules or stopping rules for the models. In general, measures of fit, etc., are not of use in prototypes. Thus, it is not clear why Biggs et al. (1993) stopped the research process with the version they report on.

Further, it is difficult to measure the fineness of the decision process. Although prototypes capture different levels of knowledge, no measures exist to determine if a desired level of fineness has been attained. Some other advantages and limitations of prototypes are discussed and summarized in O'Leary (1988).

#### **GCX AS A COMPUTATIONAL MODEL**

The area of artificial intelligence referred to as qualitative reasoning (e.g., de Kleer and Brown 1984; Hayes 1985), has established implicit criteria for qualitative models. This section draws on some of those criteria in order to

determine some of the limitations of GCX.

### Scalability

Hayes (1985, 3), following the lead of other researchers in AI, identified scalability as a critical issue for the analysis of qualitative systems:

It is perilously easy to conclude that, because one has a program which works (in some sense), its representation of its knowledge must be more or less correct (in some sense). Now this is true in some sense. But representation may be adequate to support a limited kind of inference, and completely unable to be extended to support a slightly more general kind of behavior. It may be wholly limited by scale factors, and therefore tell us nothing about thinking about realistically complicated worlds.

GCX is very case-specific to the domain of going-concern decisions and it is also very company specific. As a result, there is some question of GCX's scalability to more general company settings. Although auditors may employ very company-specific information in making going-concern decisions, they have much broader decision-making capabilities. Unlike GCX, auditors solve going-concern problems for Merrytime and other companies. Thus, it is unclear that a study of GCX can answer questions such as what knowledge the auditor possesses or how those structures are represented in memory. GCX's representations are likely to be far more situation-specific than auditors' representation.

GCX offers a starting point, but there is need to extend the GCX approach to more general representations that would facilitate more general problem solving and use of knowledge structures. Unfortunately, from Biggs et al. (1992) it is unclear if the current approach is scalable. If it is not scalable, then it is unclear from the paper what approaches would be used to attain that scalability.

### Basic Decision Process

The decision process used in GCX appears to be unidirectional. Analysis of the financial activity leads to the analysis of events that might have manifested themselves in the financial records. This approach is the reverse of the clas-

sic qualitative model approach and the classic learning theory approach, but apparently is consistent with SAS No. 59.

In one of the classic papers in qualitative reasoning, de Kleer (1977) proposed a reasoning process that successively progressed as "envisioning-qualitative-quantitative-mathematical." This approach is consistent with Piaget (1973), who suggested that mathematical ideas begin with a qualitative construction or mathematical formulation, before they acquire a metrical character.

In addition, de Kleer (1977, 299) argued that "as well as being able to solve the difficult problems, the expert problem solver should be able to solve simpler versions of a problem with qualitatively simpler techniques." Thus, we would expect a model of experts to facilitate multiple approaches or directions.

This does not mean that the approach in Biggs et al. (1992) is inappropriate. However, it does mean that it is likely that not all going-concern judgments are initiated from financial measures. It is easy to anticipate a number of situations where the auditor would envision an event or set of events as resulting in a going-concern problem.

### Validation of the Model

An important issue in qualitative modeling is system validation (e.g., Hayes 1985; O'Keefe and O'Leary 1993). In GCX, model validation could focus on at least two levels: the conceptual design level and the computer program level. It appears that most of the validation was done on the conceptual design. For example, the expert auditors reviewed and evaluated schematic representations and revisions were made.

Unfortunately, there is no report of validation of the computer model. Biggs et al. (1993) note that additional work could well be done in the experimentation with this model that would mitigate this issue. However, as noted by Winston (1984), a working prototype is not *prima facie* evidence of a successful representation of an expert.

Such validation may have been limited by the nature of the model. The system is very specific to the Merrytime company. As a result,

it may be difficult to do generic validation of the model as was done by either, for example, Meservy et al. (1986) or Messier and Hansen (1992). If that is the situation, then that may also indicate additional scalability difficulties.

#### **Quantitative Capabilities of the Model**

The quantitative recognition of the existence of a going-concern issue in GCX seems naive when compared to alternative models. Although it is difficult to determine exactly what is done, a brief description is contained in the paper. As Biggs et al. (1993) note, "The first rule obtains or calculates the values of financial measures relating to liquidity and operating performance in the financial reasoning network. The second recognizes as problems all financial measures whose values fall outside acceptable ranges."

In summary, it appears that some ratios are developed and compared to industry ratios. Although individual ratios are important, the work of Johnson and Jamal (1988) indicated that we should be most concerned with patterns of ratios. This is further exemplified in the research of Bell and Ribar (1991), where neural nets are used to establish patterns of ratios in bankruptcy analysis.

Other models of the going-concern process (e.g., Altman and McGough 1974; Mutchler 1985) apparently employ more sophisticated statistical methods. Since auditors have access to those more sophisticated analytical review methods, it seems curious that Biggs et al. did not use a similar approach in the models.

In addition, since one of the primary validation approaches is to compare the performance of the model to other models, it would seem that without the more sophisticated approaches there would be situations where many other models would out-perform GCX. Thus, the use of a naive ratio analysis could interfere with the validation efforts and cast doubt on the overall quality of the model.

#### **GREEN EYE SHADE AUDITOR?**

A number of assumptions are used by accounting and auditing researchers that generally will be referred to here as the "Green Eye

Shade Accountant (Auditor)" assumptions. This label derives from the public's common misconception that (1) accountants or auditors work independent of groups, (2) there are no political issues in the decisions faced by accountants or auditors, (3) accountants follow an absolute set of rules promulgated by some unerring source, (4) accountants' and auditors' primary concern is the generation and use of a set of well-defined numbers, and (5) accountants and auditors use only the full capabilities brought to them by pencil and paper.

Analysis of the so-called "Economic Man" (Simon 1955) has provided substantial insight into the limitations inherent in some economics research. Thus, there is reason to assume that a similar analysis of "Green Eye Shade Man" may yield insights into accounting and auditing research.

Although GCX is probably more sophisticated and domain-specific than any other model built to investigate the issue of going-concern, it still seems to make the assumption of the "Green Eye Shade Accountant." First, the model appears to assume away any audit team or group aspects. Throughout, the concern is with cognitive science processes of individual expert auditors. There seems to be little impact of team, decision hierarchy, and other aspects so apparent in engagements performed by public accountants.

Thus, the focus in Biggs et al. (1993) is on the base disciplines of cognitive science and artificial intelligence. Additional research might focus on organizational theory, group theory, or sociology. In addition, from an artificial intelligence perspective, multiple agent models would be used to capture some of the work group efforts in a computer program.

Second, political aspects of the going-concern decision may be obscured. Kida (1980) found that auditors sometimes recognize problems but do not qualify the opinion. The richness of detail in GCX provides little insight into this issue. For example, different compensation results for audit team members ultimately associated with the issuance of either a going-concern or not a going-concern judgment would conceivably be an important issue. However,

like virtually all expert systems, compensation and other important variables are ignored. Overlooking compensation may be a function of the basic assumptions about expertise. Expertise is assumed to be "pure" and "untouched" by differing reward schemes.

Third, SAS No. 59 provides some processing structure to the model (e.g., start with the financial statement numbers). SAS's are the result of political processes, developed in an increasingly litigatory environment. Thus, the literature guiding auditors is not necessarily aimed at only audit problem-solving, but also posturing for legal purposes. As a result, it is unlikely that such a political document as a SAS would provide a reasonable basis for a computational model.

Tying models to SAS pronouncements also seems to ignore the findings of qualitative models (e.g., de Kleer 1977) which indicate that more expert problem solvers should be able to use simpler techniques, and that experts initiate the task with envisioning (generating a sequence of snapshots of what might or could happen).

Fourth, if the financial numbers are the starting point, then it is assumed that events will manifest themselves in the numbers in time for the auditor to be concerned. It seems that this makes an assumption about the behavior of accounting numbers. Consider the impact of a set of events on accounting numbers. We could anticipate that those numbers would take different functional forms across time, e.g., stepwise, linear, or exponential. In the first case, a major portion of the impact could occur very rapidly. However, in the last situation, the majority of the impact would be delayed, possibly camouflaging the need for a going-concern judgment.

Fifth, the actual and potential effects of computational models used by auditors have been ignored. For example, one of the concerns of Biggs et al. (1993) is "how do auditors recognize trends in financial data?" The focus on qualitative models of auditor behavior seems to obscure the use and capabilities of computational models for tools. In 1992, I do not think that we can divorce audit judgment from the technology. However, the authors are to be excused from this since there is little in the literature about the use, diffusion, or impact of technology on audit judgments.

### CONCLUSION

Biggs et al. (1993) present a sophisticated system that provides a number of research artifacts, with different characteristics. For example, computer program prototype systems are amenable to systematic validation.

As a qualitative model, it is not clear that GCX can be scaled upward to meet the demands of situations beyond a single company, or even to a more extensive analysis of Merrytime. The basic model seems to ignore methods available to auditors, e.g., analytical review. In addition, the basic approach seems to ignore a model of the qualitative reasoning process (envisioning-qualitative-quantitative-mathematical).

Although GCX provides a research tool that has had a substantial impact on related research efforts since its introduction in 1984, it still adopts a limiting set of assumptions about auditors. The relaxation of those assumptions, labeled here as the "green eye shade auditor," provides a basis for a number of opportunities for future research.

## REFERENCES

- Altman, E., and T. McGough. 1974. Evaluation of a company as a going-concern. *Journal of Accountancy* (December): 59-67.
- Bell, T., and G. Ribar. 1991. Neural nets versus logistic regression in the prediction of bankruptcy. USC International Symposium on Expert Systems in Business, Finance and Accounting.
- Biggs, S., and M. Selfridge. 1986. GCX: An expert system for the auditor's going concern judgment. Unpublished presentation, American Accounting Association, New York (August).
- , ———, and G. Krupka. 1993. A computational model of auditor knowledge and reasoning processes in the going-concern judgment. *Auditing: A Journal of Practice & Theory* (Supplement).
- Cohen, P., and A. Howe. 1988. Towards AI research methodology: Three cases in evaluation. COINS Technical Report 88-31, Department of Computer and Information Science, University of Massachusetts.
- de Kleer, J. 1977. Multiple representations of knowledge in a mechanics problem solver. *Proceedings of IJCAI-1977*: 299-304.
- , and J. Brown. 1984. Qualitative physics based on confluences. *Artificial Intelligence* (December): 7-83.
- Fox, M. 1979. On inheritance in knowledge representation. In *Proceedings of the Sixth International Joint Conference on Artificial Intelligence 1*. Menlo Park, CA: American Association for Artificial Intelligence.
- Hayes, P.. 1985. The second naive physics manifesto. In *Formal Theories of the Commonsense World*, ed. J. Hobbs and B. Moore, 1-36. Ablex Publishing.
- Johnson, P., and K. Jamal. 1988. Human judgment: Limitations and opportunities for research. *Auditor Productivity in the Year 2000*. Reston, VA: Arthur Young.
- Kida, T. 1980. An investigation into auditor's continuity and related qualification judgments. *Journal of Accounting Research* (Autumn): 506-523.
- McCarthy, W., E. Denna, G. Gal, and S. Rockwell. 1992. Expert systems and AI-based decision support in auditing: Progress and perspectives. *International Journal of Intelligent Systems in Accounting, Finance and Management* (Vol. 1, No. 1): 53-63.
- Meservy, R., A. Bailey, and P. Johnson. 1986. Internal control evaluation: A computational model of the review process. *Auditing: A Journal of Practice & Theory* 6 (Fall): 44-74.
- Messier, W., and J. Hansen. 1992. A case study and field evaluation of EDP-Xpert. *International Journal of Intelligent Systems in Accounting, Finance and Management* (Vol. 1, No. 3): 173-186.
- Mutchler, J. 1985. A multivariate analysis of the auditor's going concern opinion decision. *Journal of Accounting Research* 23 (Autumn): 668-682.
- Newell, A., and H. Simon. 1976. Computer science as empirical inquiry: Symbols and search. *Communications of the ACM* (March): 113-126.
- O'Keefe, R., and D. O'Leary. 1993. Expert system verification and validation. *Artificial Intelligence Review* 7: 3-42.
- O'Leary, D. 1988. Expert system prototyping as a research tool. In *Applied Expert Systems*, ed. E. Turban and P. Watkins. Amsterdam, Holland. North Holland Studies in Management Science #17, Presented at the Audit Judgment Conference, University of Southern California, 1987.
- Piaget, J. 1973. *To Understand is to Invent*. New York, NY: Grossman.
- Selfridge, M., and S. Biggs. 1990. The architecture of expertise: The auditor's going concern judgment. *Expert Systems Review*: 3-18.
- Simon, H. 1955. A behavioral model of rational choice. *Quarterly Journal of Economics* 59: 99-118.
- Winston, P. 1984. *Artificial Intelligence*. Reading, MA: Addison Wesley.