FOREWORD
Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project by Bruce G. Buchanan and Edward H. Shortliffe

Allen Newell

March 1984

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Foreword

The last seven years have seen the field of artificial intelligence (AI) transformed. This transformation is not simple, nor has it yet run its course. The transformation has been generated by the emergence of *expert systems.* Whatever exactly these are or turn out to be, they first arose during the 1970s, with a triple claim: to be AI systems that used large bodies of heuristic knowledge, to be AI systems that could be applied, and to be the wave of the future. The exact status of these claims (or even whether my statement of them is anywhere close to the mark) is not important. The thrust of these systems was strong enough and the surface evidence impressive enough to initiate the transformation. This transformation has at least two components. One comes from the resulting societal interest in AI, expressed in the widespread entrepreneurial efforts to capitalize on AI research and in the Japanese Fifth-Generation plans with their subsequent worldwide ripples. The other component comes from the need to redraw the intellectual map of AI to assimilate this new class of systems — to declare it a coherent subarea, or to fragment it into intellectual subparts that fit the existing map, or whatever.

A side note is important. Even if the evidence from politics is not persuasive, science has surely taught us that more than one revolution can go on simultaneously. Taken as a whole, science is currently running at least a score of revolutions — not a small number. AI is being transformed by more than expert systems. In particular, robotics, under the press of industrial productivity, is producing a revolution in AI in its own right. Although progressing somewhat more slowly than expert systems at the moment, robotics in the end will produce an effect at least as large, not just on the applied side, but on the intellectual structure of the field as well. Even more, both AI and robotics are to some degree parts of an overarching revolution in microelectronics. In any event, to focus on one revolution, namely expert systems, as I will do here for good reason, is not to deny the importance of the others.

The book at whose threshold this foreword stands has (also) a triple claim on the attention of someone interested in expert systems and AI. First, it provides a detailed look at a particular expert system, MYCIN. Second, it is of historical interest, for this is not just any old expert system, but the granddaddy of them all — the one that launched the field. Third, it is an attempt to advance the science of AI, not just to report on a system or project. Each of these deserves a moment's comment, for those readers who will tarry at a foreword before getting on with the real story.

MYCIN as Example. It is sometimes noted that the term *expert system* is a pun. It designates a system that is expert in some existing human art, and thus that operates at human scale — not on some trifling, though perhaps illustrative task, not on some *toy* task, to use the somewhat pejorative term popular in the field. But it also designates a system that plays the role of a consultant, i.e., an expert who gives advice to someone who has a task. Such a dual picture cannot last long. The population of so-called expert systems is rapidly becoming mongrelized to include any system that is applied, has some vague connection with AI systems and has pretentions of success. Such is the fate of terms that attain (if only briefly) a positive halo, when advantage lies in shoehorning a system under its protective and productive cover.

MYCIN provides a pure case of the original pun. It is expert in an existing art of human scale (diagnosing bacterial infections and prescribing treatment for them) and it operates as a consultant (a physician describes his patient to MYCIN and the latter then returns advice to the physician). The considerations that came to the fore because of the consultant mode — in particular, explanation to the user — play a strong role throughout all of the work. Indeed, MYCIN makes explicit most of the issues with which any group who would engineer an expert system must deal. It also lays out some of the solutions, making clear their adequacies and inadequacies. Because the MYCIN story is essentially complete by now and the book tells it all, the record of initial work and response gives a perspective on the development of a system over time. This adds substantially to the time-sliced picture that constitutes the typical system description. It is a good case to study, even though, if we learn our lessons from it and the other early expert systems, we will not have to recapitulate exactly this history again.

One striking feature of the MYCIN story, as told in this book, is its eclecticism. Those outside a system's project tend to build brief, trenchant descriptions of a system. MYCIN is an example of approach X leading to a system of type Y. Designers themselves often characterize their own systems in such abbreviated terms, seeking to make particular properties stand out. And, of course, critics do also, although the properties they choose to highlight are not usually the same ones. Indeed, I myself use such simplified views in this very foreword. But if this book makes anything clear, it is that the MYCIN gang (as they called themselves) continually explored, often with experimental variants, the full range of ideas in the AI armamentarium. We would undoubtedly see that this is true of many projects if we were to follow their histories carefully. However, it seems to have been particularly true of the effort described here.

MYCIN as History. MYCIN comes out of the Stanford Heuristic Programming Project (HPP), the laboratory that without doubt has had the most impact in setting the expert-system transformation in motion and determining its initial character. I said that MYCIN is the granddaddy of expert systems. I do not think it is so viewed in HPP. They prefer to talk about DENDRAL, the system for identifying chemical structures from mass spectrograms (Lindsay, Buchanan, Feigenbaum & Lederberg, 1980), as the original expert system (Feigenbaum, 1977). True, DENDRAL was the original system built by the group that became HPP, and its origins go back into the mid-1960s. Also true is that many basic design decisions that contributed to MYCIN came from lessons learned in DENDRAL. For instance, the basic production-system representation had been tried out in DENDRAL for modeling the mass spectrometer, and it proved highly serviceable, as seen in all the work on Meta-DENDRAL which learned production rules. And certainly true, as well, is that the explicit focus on the role of expertise in AI systems predates MYCIN by a long stretch. I trace the focus back to Joel Moses's dissertation at M.I.T. in symbolic integration (Moses, 1967), which led to the MACSYMA project on symbolic mathematics (Mathlab Group, 1977), a system often included in the roster of early expert systems.

Even so, there are grounds for taking DENDRAL and MACSYMA as precursors. DENDRAL has strong links to classical problem-solving programs, with a heuristically shaped combinatorial search in a space of all isomers at its heart and a representation (the chemical valence model) that provided the clean space within which to search. DENDRAL started out as an investigation into scientific induction (on real tasks, to be sure) and only ended up becoming an expert system when that view gradually emerged. MYCIN, on the other hand, was a pure rule-based system that worked in an area unsupported by a clean scientifically powerful representation. Its search was limited enough (being nongenerative in an important sense) to be relegated to the background and MYCIN viewed purely as a body of knowledge. MYCIN embodied all the features that have (it must be admitted) become the cliches of what expert systems are. MACSYMA also wears the mantle of original expert system somewhat awkwardly. It has never been an AI system in any central way. It has been regarded by those who created it, and now nurture it, as not belonging to the world of AI at all, but rather to the world of symbolic mathematics. Only its roots lie in AI — though they certainly include the attitude that computer systems should embody as much expertise as possible (which may or may not imply a large amount of knowledge).

My position here is as an outsider, for I did not witness the day-to-day development of MYCIN in the research environment within which (in the early 1970s) DENDRAL was the reigning success and paradigm. But I still like my view that MYCIN is the original expert system that made it evident to all the rest of the world that a new niche had opened up. Indeed, an outsider's view may have a validity of its own. It is, at least, certain that in the efflorescence of medical diagnostic expert systems in the 1970s (CASNET, INTERNIST, and the DIGITALIS THERAPY ADVISOR, see Szolovits, 1982), MYCIN epitomized the new path that had been created. Thus, gathering together the full record of this system and the internal history of its development serves to record an important event in the history of AI.

MYCIN as Science. The first words of this foreword put forth the image of a development within AI of uncertain character, one that needed to be assimilated. Whatever effects are being generated on the social organization of the field by the development of an applied wing of AI, the more important requirement for assimilation, as far as I am concerned, comes from the scientific side. Certainly, there is nothing very natural about expert systems as a category, although the term is useful for the cluster of systems that is causing the transformation.

AI is both an empirical discipline and an engineering discipline. This has many consequences for its course as a science. It progresses by building systems and demonstrating their performance. From a scientific point of view, these systems are the data points out of which a cumulative body of knowledge is to develop. However, an AI system is a complex join of many mechanisms, some new, most familiar. Of necessity, on the edge of the art, systems are messy and inelegant joins — that's the nature of frontiers. It is difficult to extract from these data points the scientific increments that should be added to the cumulation. Thus, AI is case-study science with a vengeance. But if that were not enough of a problem, the payoff structure of AI permits the extraction to be put off, even to be avoided permanently. If a system performs well and breaks new ground — which can often be verified by global output measures and direct qualitative assessment — then it has justified its construction. Global conclusions, packaged as the discursive views of its designers, are often the only increments to be added to the cumulated scientific base.

Of course, such a judgment is too harsh by half. The system itself constitutes a body of engineering know-how. Through direct study and emulation, the next generation of similar systems benefits. However, the entire history of science shows no alternative to the formation of explicit theories, with their rounds of testing and modification, as the path to genuine understanding and control of any domain, whether natural or technological. In the present state of AI, it is all too easy to move on to the next system without devoting sufficient energies to trying to understand what has already been wrought and to doing so in a way that adds to the explicit body of science. An explosive development, such as that of expert systems, is just the place where engineering progress can be expected to occur pell-mell, with little attention to obtaining other than global scientific lessons.

This situation is not to be condemned out of hand, but accepted as a basic condition of our field. For the difficulties mentioned above stem from the sources that generate our progress. Informal and experiential techniques work well because programmed systems are so open to direct inspection and assessment, and because the loop to incremental change and improvement is so short with interactive creation and modification. AI, like any other scientific field, must find its own particular way to science, building on its own structure and strengths. But the field is young and that way is not yet clear. We must continue to struggle to find out how to extract scientific knowledge from our data points. The situation is hardly unappreciated, and many people in the field are trying their hands at varying approaches, from formal theory to more controlled system experimentation. There has been exhortation as well. Indeed, I seem to have done my share of exhortation, especially with respect to expert systems. The authors of the present volume, in inviting me to provide a foreword to it, explicitly noted that the book was (in small part) an attempt to meet the calls I had made for more science from our expert-systems experiments. And recently, Harry Pople asserted that his attempt at articulating the task domain of medical diagnosis for INTERNIST was (again, in small part) a response to exhortation (he called it criticism) of mine (Pople, 1982). I am not totally comfortable with the role of exhorter — I prefer to be in the trenches. However, if comments of mine have helped move anyone to devote energy to extracting the science from our growing experience with expert systems, I can only rejoice.

The third claim of this book, then, is to extract and document the scientific lessons from the experience with MYCIN. This extraction and documentation occurs at two levels. First, there has been a very substantial exploration in the last decade of many of the questions that were raised by MYCIN. Indeed, there are some 26 contributors to this book, even though the number of people devoted to MYCIN proper at any one time was never very large. Rather, the large number of contributors reflects the large number of follow-on and alternative-path studies that have been undertaken. This book documents this work. It does so by collecting the papers and reports of the original researchers that did the work, but the present authors have made substantial editorial revisions to smooth the whole into a coherent story. This story lays to rest the simplified view that MYCIN was a single system that was designed, built, demonstrated and refined; or even that it was only a two-stage affair — MYCIN, the original task-specific system, followed by a single stage of generalization into EMYCIN, a kernel system that could be used in other tasks. The network of studies was much more ramified, and the approaches considered were more diverse.

The step to EMYCIN does have general significance. It represents a major way we have found of distilling our knowledge and making it available to the future. It is used rather widely (e.g., the system called EXPERT (Kulikowski & Weiss, 1982) bears the same relation to the CASNET system as EMYCIN does to MYCIN). It is of a piece with the strategy of building special purpose problem-oriented programming languages to capture a body of experience about how to solve a class of problems, a strategy common throughout computer science. The interesting aspect of this step, from the perspective of this foreword, is its attempt to capitalize on the strong procedural aspects of the field. The scientific abstraction is embodied in the streamlined and clean structure of the kernel system (or programming language). The scientific advance is communicated by direct study of the new artifact and, importantly, by its use. Such kernel systems still leave much to be desired as a vehicle for science. For example, evaluation still consists in global discussion of features and direct experience, and assessment of its use. (Witness the difficulty that computer science has in assessing programming languages, an entirely analogous situation.) Still, the strategy represented by EMYCIN is an important and novel response by AI to producing science.

The second level at which this book addresses the question of science is in surveying the entire enterprise and attempting to draw the major lessons (see especially the last chapter). Here the authors have faced a hard task. Of necessity, they have had to deal with all the complexity of a case study (more properly, of a collection of them). Thus, they have had to settle for reflecting on the enterprise and its various products and experiences, and to encapsulate these in what I referred to above as qualitative discussion. But they have a long perspective available to them, and there is a lot of substance in the individual studies. Thus, the lessons that they draw are indeed a contribution to our understanding of expert systems.

In sum, for all these reasons I've enumerated, I commend to you a volume that is an important addition to the literature on AI expert systems. It is noteworthy that the Stanford Heuristic Programming Project previously produced an analogous book describing the DENDRAL effort and summarizing their experience with it (Lindsay, Buchanan, Feigenbaum & Lederberg, 1980). Thus, HPP has done its bit twice. It is well ahead of many of the rest of us in providing valuable increments to the accumulation of knowledge about expert systems.

References

Feigenbaum, E. A. The art of artificial intelligence: Themes and case studies in knowledge engineering. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence.* Pittsburgh, PA: Computer Science Department, Carnegie-Mellon University, 1977.

Kulikowski, C. A. & Weiss, S. M. Representation of expert knowledge for consultation: The CASNET and EXPERT projects. In Szolovits, P. (Ed.), *Artificial Intelligence in Medicine.* Boulder, CO: Westview, 1982.

Lindsay, R. K., Buchanan, B. G., Feigenbaum, E. A. & Lederberg, J. *Applications of Artificial Intelligence to Chemistry: The DENDRAL project.* New York: McGraw-Hill, 1980.

Mathlab Group. *MACSYMA Reference Manual* (Tech. Rep.). Computer Science Laboratory, MIT, 1977.

Moses, J. *Symbolic Integration.* Doctoral dissertation, MIT, 1967.

Pople, H. E. Jr. Heuristic methods for imposing structure on ill-structured problems: The structuring of medical diagnosis. In Szolovits, P. (Ed.), *Artificial Intelligence in Medicine.* Boulder, CO: Westview Press, 1982.

Szolovits, P. *Artificial Intelligence in Medicine.* Boulder, CO: Westview Press, 1982.