Implementation of an Experimental, Microcomputer-Based Medical Diagnosis System

Robert B. Trelease, Jr.

UCLA Sleep Disorders Clinic
1184 Reed Neurological Research Center
Los Angeles, California 90024

Abstract

An experimental sleep disorders diagnosis system is demonstrated, illustrating the use of FORTH and the EXPERT-2 expert system toolkit in implementing medical diagnostic functions on a microcomputer.

Introduction

In the application of artificial intelligence (AI) programming techniques, perhaps the greatest practical success has been obtained with the development of expert systems. Functional expert systems have been created for emulating human reasoning in solving a number of specialized problems (4,12) including those involved in computer system configuration (XCON), geological prospecting (PROSPECTOR), and specialized medical diagnosis (MYCIN).

In the medical domain, expert system development has stimulated research yielding clinically useful tools (2,11). Functional systems can act as technical assistants or consultants to physicians, by providing expeditious analyses of complex laboratory data or specialized "second opinions". The methodical process of "knowledge engineering" (creating machine usable rules/knowledge bases) has complemented scientific investigations of diagnostic techniques, promoting new insights about diagnosis (6). AI researchers have themselves borrowed liberally from the work of neuroscientists and cognitive psychologists (6). Ultimately, in experimenting with AI programming techniques, medical researchers have created and tested different models for decision making, mental function, and brain activity.

Currently, there are many tools for the development of diagnostic expert systems (12). Ranging from specialized AI programming languages (LISP and PROLOG) to "skeletal" expert systems (EMYCIN), they provide the designer with the necessary environments for encoding working rule bases, inference engines, and functional user interfaces. Although specialized AI programming tools may be well-suited for many types of expert system development, other high-level language systems—notably FORTH—can be particularly useful. FORTH's special advantages—compactness of code, rapid execution, native extensibility, and ease of interfacing with intrumentation—have been employed in implementing an exemplary LISP-originated expert system for microcomputer-based diagnosis of diesel-electric locomotive problems (5). In the medical domain, the particular utility of FORTH-based tools has been recently demonstrated in a PC-based "expert operator," which performs on-line real-time analysis of electrophysiologic data and identifies human sleep states (10).

The purpose of this article is to demonstrate an experimental system for specialized medical diagnosis, developed with the EXPERT-2 FORTH expert system toolkit. This FORTH expert system performs provisional differential diagnosis of human sleep disorders and provides a research tool for evolving more powerful diagnostic criteria. Examples will be given of EXPERT-2's capacity for calling "analytic subroutines" and for integrating with conventional FORTH vocabularies.

System Software and Instrumentation

EXPERT-2 consists of about 40 FORTH blocks of extension words (8,9). Obtainable from a number of FORTH vendors, EXPERT-2 will run on a variety of computers, from Apple II to the Digital Equipment VAX. The original FORTH-79 compliant source code has been published, along with an extensive review of system functions and examples (8), and readers are urged to consult this source for a full description. For the purposes of the present article, it should be noted that EXPERT-2 implements a simple, backward-chaining inference engine operating under the guidance of a knowledge base composed of production rules laid out in FORTH blocks. Rule blocks can be created with a standard FORTH screen editor. In addition to "string assertion" rule elements consisting of text strings, EXPERT-2 also employs "analytical subroutine" rules which execute precompiled words. This latter category of rule elements allows access to previously defined FORTH functions, such as hardware calls or data entry prompts, which leave truth flags on the stack for use by the inference engine. "Analytic subroutine" words are called by rule elements ("**-RUN**") beginning with **IFRUN**, **ANDRUN**, etc.

The examples presented in this paper were generated using EXPERT-2 running with Unified Software Systems UNIFORTH (MS-DOS and CP/M-86 versions) on the author's DEC Rainbow 100 personal computer. The original experimental system (FORTH-79 version) also ran acceptably along with other tasks under Concurrent CP/M-86. Minus the pull-down menu utility words, the MS-DOS version of this expert system was also run on an IBM PC-XT.

The initial UNIFORTH 2.18 (FORTH-79 compliant) version of EXPERT-2 differed from the published MVP-FORTH version (8) primarily in the use of different block loading schemes and different double precision operator words. Including MS-DOS, FORTH, a video editor utility, the EXPERT-2 system, and the sleep diagnosis knowledge base demonstrated below, 128 kilobytes of RAM were required for proper function (full 64 Kb task space).

The original demonstration knowledge base code (FORTH-79) was also run unmodified with a version of EXPERT-2 under UNIFORTH 3.01 (FORTH 83 compliant) with MS-DOS. Because of splitting of header/stack and data segments in this particular version of UNIFORTH, the EXPERT-2 inference engine and compiler code was modified for proper CFA and name field calls. Normal FORTH-83 **TRUE** and **FALSE** constants were also allowed. FORTH-79 rule source code was thus compatible with an updated FORTH-83 EXPERT-2 inference engine and rule compiler. With the split segment design, about 22 Kb of additional dictionary space were available. The figures shown in this paper were generated with the FORTH-83 version of the system.

Medical Diagnosis and Sleep Disorders

Hayes-Roth and others (4,11) have observed that contemporary expert systems function best within a relatively restricted domain of specialized knowledge. Thus, separate medical expert systems are devoted to the problems of diagnosis and treatment of infectious disease (MYCIN), analysis of lung function (PUFF), and internal medicine diagnosis (INTERNIST-I/CADUCEUS). This specialization of computer-based expert systems is comparable to that of different types of medical practice (e.g., cardiology, neurology, and obstetrics).

Like other specialized medical problem domains, sleep disorders medicine has its own distinctive characteristics (3,7). Overall, sleep disorders are very common, including a range of problems from sleep apnea to narcolepsy. Sleep apnea, characterized by periodic failure to breathe, can occur in association with a variety of other conditions, including high blood pressure, ulcerative disease, heart and lung diseases, and diabetes. Narcolepsy is relatively rare and is characterized by uncontrollable sleepiness along with a variety of other symptoms. Other, less common disorders include night terrors, nocturnal epilepsy, somnambulism, bruxism, and enuresis.

Consideration of the specifics of sleep disorders diagnosis also provides some insights into the logical intricacies of medical and general diagnostic problem-solving. As formulated by the Association of Sleep Disorders Centers (1), disturbances of sleep can be differentiated into four main categories: a) disorders of initiating and maintaining sleep (DIMS; e.g., psychophysiologic insomnia); b) disorders of excessive sleepiness (DOES; e.g., hypersomnia sleep apnea syndrome); c) disorders of the sleep-wake schedule (e.g., "jet-lag'); d) dysfunctions associated with sleep, sleep stages, or partial arousals (parasomnias; e.g., night terrors). Subdivisions of these major categories are distinguished by the principal presenting sleep complaint, as well as by the nature of the underlying medical, behavioral, or pharmacological problem. Effective diagnosis thus requires proper categorization of the patient's presenting symptoms and complaints, with treatment directed at relieving the underlying problem. The complexity of the diagnostic problem is increased by the fact that underlying pathophysiologic conditions, such as sleep apnea, may produce both complaints (symptoms) of excessive daytime sleepiness (DOES) or of difficulty in sleeping (DIMS). Diagnosis must thus distinguish different disease processes with similar presenting symptoms, and similar disease processes with different presenting symptoms.

Sleep disorders are assessed following consideration of the medical history, the patient's report of symptoms and experiences, and pertinent laboratory data (1,3,7). Diagnostic support is provided by laboratory data, including polygraphic records from all-night studies of patient sleep. Analysis of polygraphic records is provided by technical sleep experts (polysomnographers) who quantify periods of disturbed sleep and cardiorespiratory problems. Scoring of sleep states is based on EEG ('brainwave'), eye movement, muscle, respiration, and cardiac activities recorded on paper. Automation of such laboratory data analysis is the goal of the previously mentioned expert operator – FORTES (10).

The system demonstrated in this article was intended to investigate the practicality of a machine-consultant to provide diagnosis-at-interview for human sleep disorders. For assessment of experimental system performance, real-world data was available from an operational sleep disorders clinic, including patient medical history reviews, physical examination and laboratory data, as well as human expert-analyzed sleep study data. In order to evolve a reliable diagnostic system, successive cycles of testing and refinement were performed on different different rule systems.

In constructing a usable knowledge base for sleep disorders diagnosis, several key elements were considered. As shown by a recent survey of clinic patients, sleep apnea DOES syndrome was the most common disorder encountered (>80% of all cases). Given the common occurence of apnea, and the occurrence of DOES sleepiness with narcolepsy, initial rules were used for differentiating DOES problems. Depicted in Figure 1 is part of a hypothetical "decision tree", with rules designed to make such an initial differentiation between DIMS and DOES sleep disorders. Additional rules were employed to differentiate between other sources of excessive somnolence, parasomnias, and sleep-waking schedule disorders. Initial diagnostic rule criteria were derived from published medical source references (1, 3, 7), from clinical consultations, and from personal research experience.

The Experimental System

The complete experimental system for sleep disorders diagnosis was constructed of several components. The underlying FORTH system was compiled with video editor blocks, which permitted the writing of rule source code and conventional FORTH extension words. Several blocks of diagnostic system-specific extensions and `-RUN` words were also compiled prior to the `RULE` blocks. The `RULE` blocks used in the first prototype knowledge base consisted primarily of string assertion rule elements, as typified by the screens shown in Figure 2. Following the knowledge base, several additional blocks of extension words provided starting screens, test utilities, a pull-down menu utility (ANSI standard VT-100), and a simple report generator for the system. `DIAGNOSE` (invoking the functional EXPERT-2 inference engine utility) was thus called from and in turn called conventional FORTH vocabulary words.

Figure 1. Part of a hypothetical "decision tree", with rules designed to make an initial differentiation between DIMS and DOES sleep disorders.

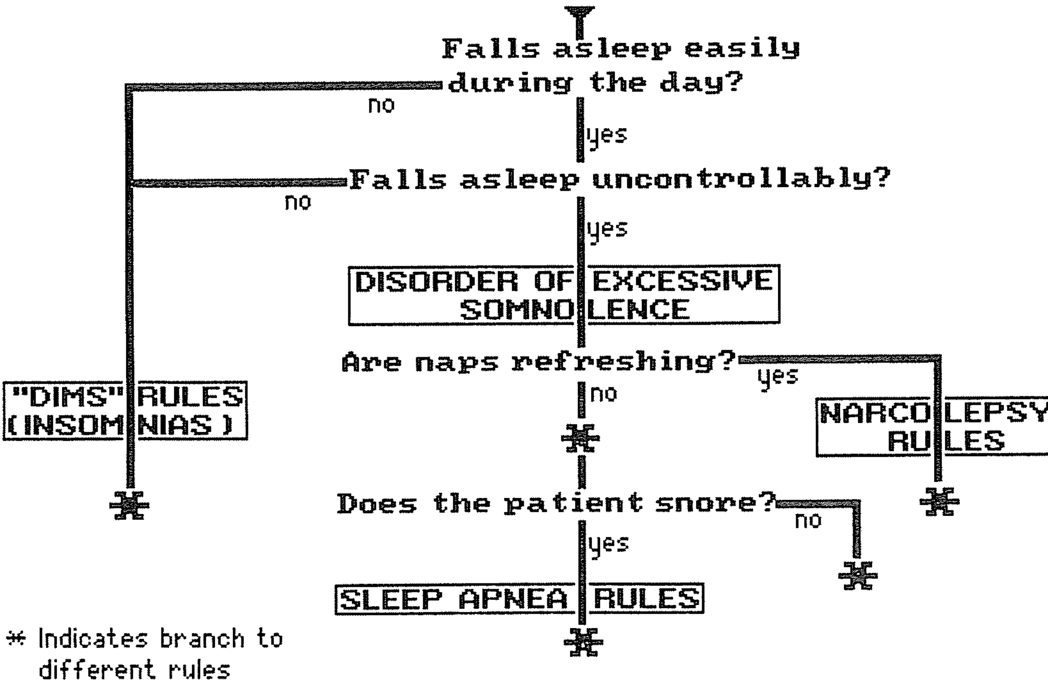

Figure 2. FORTH listing showing block sof "sleep disorders" rules composing a portion of the prototype diagnostic knowledge base. These rules implement part of the decision tree depicted in Figure 1.

```
SCR # 66
\ XPERTSYS.DAT:  Report generator and consultant    rbt 062385



IF Patient has physical complaints
AND Patient falls asleep easily during the day
AND Patient falls asleep uncontrollably or at odd times
THEN Patient has a disorder of excessive somnolence (DOES)

IF Patient has a disorder of excessive somnolence (DOES)
AND Patient finds naps very refreshing
AND Patient has vivid or unreal dreams while falling asleep
THEN Patient has narcolepsy
-->
```

```
SCR # 67
\ XPERTSYS.DAT:  Sleep Disorders Diagnoser           rbt-062385


IF Patient has narcolepsy
AND Patient gets weak in the legs or can't stand at odd times
AND Patient awakens and can't move or feels paralyzed
THENHYP Patient has narcolepsy with cataplexy

IF Patient has narcolepsy
ANDNOT Patient gets weak in the legs or can't stand at odd times
ANDNOT Patient awakens and can't move or feels paralyzed
THENHYP Patient has narcolepsy without cataplexy

-->
```

```
SCR # 68
\ XPERTSYS.DAT:  Sleep Disorders Diagnoser           rbt-062385

IF Patient has a disorder of excessive somnolence (DOES)
AND Patient sleeps poorly
AND Patient snores
THEN Patient has a sleep-induced respiratory impairment

IF Patient has a sleep-induced respiratory impairment
AND Patient has sleeping partner
AND Partner notes patient's breathing pauses during sleep
THEN Patient has obstructive sleep apnea  (c.f.>0.9)
-->
```

```
SCR # 69
\ XPERTSYS.DAT:  Sleep Disorders Diagnoser           rbt-052586
IF Patient has a sleep-induced respiratory impairment
ANDNOT Patient has sleeping partner
AND Patient has been told s/he snores very loudly
THEN Patient has obstructive sleep apnea  (c.f.>0.9)

IF Patient has obstructive sleep apnea  (c.f.>0.9)
AND Patient has heartburn or ulcers
THENHYP Patient has moderate to severe obstructive sleep apnea

IF Patient has obstructive sleep apnea  (c.f.>0.9)
ANDIFRUN HTNCHECK
THENHYP Patient has severe sleep apnea
-->
```

The remaining figures demonstrate the performance of the diagnostic system using the first prototype sleep disorders rules. Figure 3 shows the starting screen for the system, with the pull-down menu and execution of a conventional FORTH word (PHYSDATA) which acquires starting patient data. Following initial data entry and screen output, DIAGNOSE invokes the inference engine, as shown at the top of Figure 4. The system chains backward to the first antecedent of the first hypothesis in the rule base, and then asks about its truth—in this case, asking whether the patient has a particular symptom. The system continues to inquire about the presence of certain symptoms, ruling out unsupported hypotheses and branching until a final conclusion is drawn.

Figure 4 also demonstrates the traversal of a decision tree, similar to that shown in Figure 1, using some of the production rules shown in Figure 2. The explanatory function of EXPERT-2, invoked by a W response to the "statement true" prompt, shows the logic tested by asking whether the patient snores. The ultimate diagnosis of severe sleep apnea further requires that the patient have hypertension, as confirmed by an IFRUN rule element which acquired current blood pressure data. Defined in the following way,

```
: HTNCHECK CVDATA? @ 0=
     IF CVRDATA CR CR DROP
     THEN BPSYS @ BPDIA @ - 3 / BPDIA @ + 110 <
          IF 1 ELSE 0
          THEN  ;
```

HTNCHECK determines whether blood pressure data was previously acquired during pre-DIAGNOSE patient data entry. If not, it prompts/acquires the data (with CVRDATA) and calculates whether the mean blood pressure is greater than a hypertensive "threshold" value (i.e., > 110 mmHg). If hypertension is confirmed, a logic flag value of 1 is left on the stack, causing the IFRUN HTNCHECK statement to evaluate as true. Although HTNCHECK and CVRDATA were written to acquire blood pressure data through keyboard entry, these words could have also been defined to handle instrumentation such as an automatic arterial pressure gauge system.

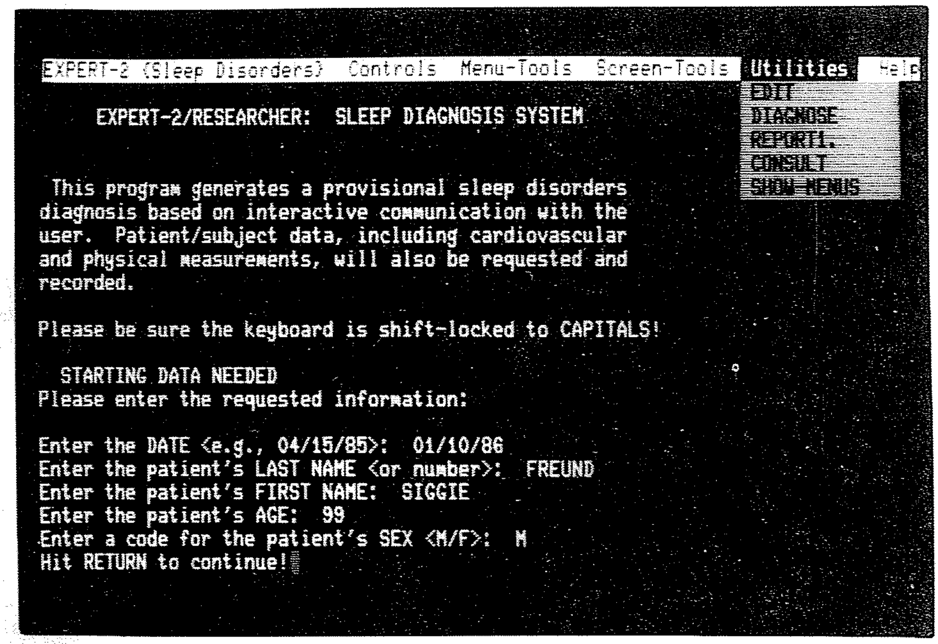

Figure 3. Photograph of the "sleep disorders diagnoser" starting screen, demonstrating a pull-down menu interface and data acquisition sequence implemented with conventional FORTH words.

Figure 4. Output from a "sleep diagnosis" session which shows callup of the inference engine (DIAGNOSE) after printout of previously acquired patient data by conventional FORTH words. The explanation feature of the inference engine (invoked by WHY?) is also demonstrated. Prior to the final inference, a data input and analysis word (HTNCECK) is called by an ANDIFRUN rule element.

```
01/10/86

SIGGIE FREUND
Current weight=  154 lbs


Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  We can proceed  Y

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  You want to use the machine consultant  Y

I deduce
  We can go on with the consultation

I deduce
  Patient has physical complaints

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient falls asleep easily during the day  Y

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient falls asleep uncontrollably or at odd times  Y

I deduce
  Patient has a disorder of excessive somnolence (DOES)

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient finds naps very refreshing  N

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient sleeps poorly  Y

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient snores  W


We are trying to prove
  Patient has moderate to severe obstructive sleep apnea

We are testing rule# 8
IF Patient has a disorder of excessive somnolence (DOES)
IF Patient sleeps poorly
IF Patient snores
THEN Patient has a sleep-induced respiratory impairment

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient snores  Y

I deduce
  Patient has a sleep-induced respiratory impairment
```

```
Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient has sleeping partner  Y

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Partner notes patient's breathing pauses during sleep  Y

I deduce
  Patient has obstructive sleep apnea  (c.f.>0.9)

Is this statement TRUE?  (Y=YES, N=NO, W=WHY)
  Patient has heartburn or ulcers  N

 PLEASE enter the requested information:

Enter the patient's current HEART RATE <##/min>:  74
Enter the patient's current SYSTOLIC BP:  165
Enter the patient's current  DIASTOLIC BP:  95
Enter the patient's current RESPIRATORY RATE <##/min>:  18


I deduce
  Patient has severe sleep apnea

I conclude
  Patient has severe sleep apnea
```

Figure 5. A simple report generated by the sleep diagnostic system, summarizing some patient data, as well as the logic employed in making the final diagnosis.

```
01/10/86

SIGGIE FREUND
Current weight=  154 lbs

Systolic blood pressure=  165  mm Hg
Diastolic blood pressure=  95 mm Hg
Mean blood pressure=  118  mm Hg
Heart rate=  74  beats/minute
Respiratory rate=  18  /minute


Given or Inferred facts:

  Patient falls asleep easily during the day
  Patient falls asleep uncontrollably or at odd times
  Patient has a disorder of excessive somnolence (DOES)
  Patient sleeps poorly
  Patient snores
  Patient has a sleep-induced respiratory impairment
  Patient has sleeping partner
  Partner notes patient's breathing pauses during sleep
  Patient has obstructive sleep apnea  (c.f.>0.9)

Final Hypothesis examined:

  Patient has severe sleep apnea
```

Figure 5 shows a brief report of the diagnostic session, summarizing some patient data and listing confirmed facts which lead to the ultimate diagnosis. This simple report generator was defined using EXPERT-2 words for accumulating proven rules and for printing string assertion text given a rule element address.

Discussion

As demonstrated here, FORTH and the EXPERT-2 toolkit can readily be used to develop experimental expert systems for medical diagnosis. Of course, EXPERT-2 need not be restricted to medical applications, and the original source code includes knowledge bases and support words for stock market speculation, digital circuit analysis, and the ANIMALS guessing game (8). Furthermore, although these sample knowledge bases employ primarily string assertion rules, greater use can be made of "analytical subroutine" rules. At the other extreme, for example, systems can thus employ primarily "analytical subroutine" rules for control and data aquisition with laboratory intrumentation.

The EXPERT-2 toolkit employs "categorical reasoning" rather than probabilistic ("Bayesian") computations to arrive at deductions (see Szolovitz and Pauker in reference 2). Boolean logic is used with string assertions and analytical subroutine elements, and confidence factors or "fuzzy flags" are not employed in evaluating the truth of the rules. Such categorical reasoning may be appropriate for many specialized diagnostic tasks, and if desired with EXPERT-2, confidence limits can be assigned by the developer to the text of different hypotheses and subhypotheses. Probabilities or confidence limits for specific categorical diagnoses can be determined experimentally by using real-world data (e.g., patient records) as input for successive trials of different knowledge bases.

One of the distinct advantages of EXPERT-2 is the simple clarity of its knowledge base structures. With the use of English string assertion rules, human experts can easily comprehend the representation scheme for knowledge which they have contributed. The use of the basic extension word set requires little special FORTH expertise, other than familiarity with block LOADing and with a FORTH editor for creating rules. In addition to facilitating feedback and interaction between consulting domain experts and the developing knowledge engineer, the design of EXPERT-2 makes it possible for the experts themselves to experiment with the knowledge engineering process.

As a "well-tempered" extension of FORTH, EXPERT-2 can integrate with existing vocabularies. This allows FORTH developers to experiment with hardware and software configurations which may not be practically available to users of more conventional AI programming environments like LISP. The compact microcomputer-based system demonstrated here, for example, shows how numerical data obtained by conventional FORTH input words can be used in assessing the truth of diagnostic expert system rules. Although FORTH may not be an "ideal" environment for general-purpose AI development, EXPERT-2 demonstrates its utility for experimentation and production of specialized expert systems.

Acknowledgements

This work was supported in part by NIMH MH-17140-02 and the Basic Research Support Grant of the UCLA Neuropsychiatric Institute. The author is grateful for expert consultation provided by Drs. Donna Arand, Emery Zimmermann, Arne Henden, and Jack Park.

References

1. Association of Sleep Disorders Centers. Diagnostic classification of sleep and arousal disorders. *Sleep* 2: 1-137, 1979.

2. Clancey, W.J. and Shortliffe, E.H. *Readings in Medical Artificial Intelligence—The First Decade*. Reading, MA: Addison-Wesley, 1984.

3. Hauri, P. *Current Concepts: The Sleep Disorders*. Kalamazoo, MI: The Upjohn Company, 1982.

4. Hayes-Roth, F., Waterman, D.A. and Lenat, D.B. *Building Expert Systems*. Reading, MA: Addison-Wesley, 1983.

5. Johnson, H.E. and Bonissone, P.P. Expert system for diesel electric locomotive repair. *Journal of FORTH Application and Research* 1(1):7-16, 1983.

6. Kleinmuntz, B. Diagnostic problem solving by computer: A historical review and the current state of the science. *Computers in Biology and Medicine* 14(3): 255-270, 1984.

7. Kupfer, D.J., and Crook, T. Physicians' Guide to the Recognition and Treatment of Sleep Disorders in the Elderly. New Canaan, CN: Mark Powley Associates, 1984.

8. Park, J. *MVP-FORTH Expert System Toolkit*. Mountain View, CA: Mountain View Press, Inc., 1983.

9. Park, J. A consequent reasoning inference engine for microcomputers. In: *1984 FORML Conference Proceedings*. San Carlos, CA: FORTH Interest Group, 1984.

10. Redington, D. A Forth oriented real-time expert system for sleep staging: A FORTES Polysomnographer. *Journal of FORTH Application and Research*, in Press, 1986.

11. Szolovits, P. *Artificial Intelligence in Medicine*. Boulder, Co: Westview Press—American Association for the Advancement of Science, 1982.

12. Waterman, Donald A., *A Guide to Expert Systems*. Reading, MA: Addison-Wesley, 1986.

R. B. Trelease is a biomedical research scientist who received his Ph.D. in Anatomy from the UCLA School of Medicine. Current research interests include studies of state-related control of the autonomic nervous system, as well as computer modeling of cardiovascular function and brain mechanisms. The work described in this paper was performed during the course of an NIMH Postdoctoral Fellowship (MH-17140-02) at the UCLA Neuropsychiatric Institute.

Manuscript received February 1986.