A Forth Oriented Real-Time Expert System for Sleep Staging: a FORTES Polysomnographer

Dana Redington [1]

Sleep Research Center
Department of Psychiatry and Behavioral Sciences
Stanford University School of Medicine
Stanford, California 94305

Abstract

The established programming perspectives that encode knowledge directly into software are currently being challenged by alternative approaches. Standard software practices work well for simple and small tasks. However, more effective computer methods are required as the size of problems escalate in terms of complexity and real-time demands. A recent approach has been to apply the techniques of Knowledge Engineering and develop Expert Systems in Forth. Most Expert Systems previously implemented in Forth act as consultants; they do not collect data directly. The alternative is to develop real-time capability in Forth so that knowledge-based computer systems can act as Expert Operators. This paper introduces, for the first time, a Forth based Expert Operator aimed at real-time sleep staging.

Introduction

A paradigm exists in Computer Science which is in the process of being overthrown. In science this process has been referred to as a paradigm shift [KUHN70]: a paradigm is a unique perspective that enables problems to be solved better than from other perspectives within the same field and a shift occurs when an alternative perspective provides better solutions than the established point of view. The established programming paradigm is to encode knowledge directly into computer software via mainstream languages, such as Basic, Pascal, and C. Here, dedicated applications programs are written in the computer languages. Examples of dedicated applications software include word processing programs such as Wordstar and Word, and dedicated spread sheet programs such as VisiCalc and 1-2-3. Even examples of the current generation of more "integrated software", such as Excel and Jazz, use a dedicated programming paradigm.

However, this established paradigm began to break down when it was discovered that certain classes of "real-world" problems just could not be solved in a timely fashion with the ordinary programming paradigm [MOOR74; BARR81]. Enter Lisp (and its derivatives) and Forth, although for different reasons (see Review Article on Expert Systems this volume).

The alternative was to develop computer programs that are knowledge-based, have motors of machine intelligence, and can utilize the knowledge-base to act with human-like intelligence. This class of programs are written in languages such as Lisp, Prolog, C, eventually Smalltalk, and Forth which is an object/action oriented environment. An example of programs with machine intelligence is the so-called class of computer software called Expert Systems, i.e., programs that have a knowledge-base comparable to a specialist in a particular field and are able to perform at or beyond

---

[1] Address correspondence to: D. Redington, P.O. Box 620479, Woodside, California 94062.

the level of an expert consultant. (A consultant processes information in non-real-time whereas an operator processes data in real-time.) Examples of Expert System consultants that have achieved super-human performance in an extremely limited domain are DENDRAL, superceded by CONGEN, a heuristic program that determines the possible molecular structure based upon known spectroscopic analysis [BARR82] and MYCIN which diagnoses and suggests therapy to physicians for blood infections and meningitis [BUCH83]. Clearly, the class of problems like molecular identification and bacterial diagnosis and therapy go beyond what is capable in the established programming paradigm and have required a newer, more viable perspective.

There is another problem that is equally representative of this problem class: the computer scoring of sleep/wake states in contemporary Sleep Research . Over the last twenty-nine years this problem has remained unsolved. Even though standard computer methods have been applied with increasing emphasis, only partial solutions have been developed: most attempts at computer staging work with restricted data, i.e., data that is relatively noise free and where the subject is a healthy young adult. Various computer techniques have been tried, including use of fast fourier transforms and power spectral analyses, time series modeling, and various statistical approaches, such as clustering techniques, principle component analysis, discriminant analysis, and adaptive segmentation [GEVI80; JOHN77; LARS70; SMIT78]. Yet, the most reliable and convenient form of sleep staging has remained the expert human interpretation of polysomnograms.

An alternative approach incorporating the newer programming paradigm may provide steps towards a solution to the problem of automated sleep staging [SMIT86]. The application of Expert Systems to sleep staging has not been previously described in the literature although expert system approaches have been suggested for rapid eye movement analysis [LEE84], EEG analysis [BOUR83], and evaluation of sleep disorders (see Trelease article this volume). The focus of this paper is to give a brief overview of standardized sleep staging and to present the initial implementation of an Expert System sleep stager in Forth.

A Brief Overview of Sleep Staging

The first convenient method of sleep staging was described in 1957 [DEME57B]. The method identified four major categories of sleep: Stages 1, 2, 3, and 4. The method also defined two discrete periods within sleep [DEME57A] by using the profound changes in rapid eye motility originally identified by Aserinsky and Kleitman [ASER53]. These two periods were based on the presence or absence of rapid eye motility and were observed to occur in an alternating cyclical pattern within a period of 90-100 minutes. The new period, similar to the EEG of Stage 1, was defined by the novel appearance of rapid eye movements (REMs) – Stage 1-REM; the other period was labeled non-REM (NREM). Stage 1-REM was found to be associated with subjects' reports of dream activity.

A standardized sleep staging system was established in 1968 [RECH68] . It did not experimentally refine sleep staging. Although, another objective indicator, electromyographic activity (EMG, or muscle atonia) was incorporated in an attempt to help distinguish REM sleep from wake; but changes in eye motility still remained superior in defining the "second state of sleep". The standardized scoring procedure is outlined here. The scoring record was divided into 30 second epochs (conveniently, one sheet [30 cm] of polygraph paper at 10 mm/sec). Each epoch was scored as representing one of seven stages when at least 50 per cent of the epoch was a certain stage (see Table 1 for definitions of the sleep stages).

There are inherent limitations with standard sleep staging [STKU82]: sleep staging is an imperfect human pattern recognition process; imprecise heuristics are involved; more than seven wake/sleep states may exist; individual response stereotypy accounts for non-ideal sleep records [JOHN77; LARS70]. When sleep disorders are considered, additional physiological variables must be measured and revised scoring procedures must be applied [GUIL78].

Table 1: Standardized Sleep Stages

Stage	Abbreviation	Brief Description
Movement Time	MT	movement artifact in the polygraphic record
Stage Wake	W	EEG activity in the 8-12 hertz region and/or low voltage, mixed frequency activity
Rapid Eye Movement Sleep	S1-REM	"a relatively low voltage, mixed frequency EEG in conjunction with episodic REMS and low amplitude" EMG
Stage 1	S1	a relatively low voltage, mixed frequency EEG without REMs
Stage 2	S2	"12-14 hertz sleep spindles and K complexes on a background of relatively low voltage, mixed frequency EEG activity"
Stage 3	S3	"moderate amounts of high amplitude, slow wave activity" and in the delta bandwidth
Stage 4	S4	"large amounts of high amplitude, slow wave activity"

To summarize, sleep staging is usually performed by a trained human scorer, an expert, who evaluates a polysomnogram by visual inspection. A polysomnogram is a linear time varying graph of the EEG, EOG, and EMG: a record of a subject's sleep. The record represents an average of eight hours of sleep, about 960 pages or approximately 288 meters of paper. In the process of scoring a night's record each page or epoch is evaluated. Scoring a normal record consumes about 90 minutes on the average. The scoring process progresses in three steps: (1) identifying significant tonic features like changes in EEG frequency and amplitude and degree of muscle atonia, and phasic features like presence of spindles, K complexes, slow and rapid eye movements; (2) evaluating and averaging those features into a 30 (or 60) second epoch; and then (3) deciding on an appropriate stage for a given epoch as prescribed by a standardized criterion, [RECH68 or WILL74].

A Description of a FORTES Polysomnographer

An expert system constructed in Forth and oriented towards real-time interaction with the world is referred to as a Forth oriented real-time expert system (FORTES; pronounced "for-tees") [REDI84A]. The description of a FORTES operator that is specialized in automated sleep staging can be divided into five parts: operating constraints, hardware environment, language environment, knowledge architecture, and performance. Each of these parts is described briefly below.

Operating Constraints

There are several operating constraints for a personal expert operator capable of performing real-time sleep staging. The system should be microcomputer (personal computer, [PC]) based to satisfy a variety of uses (monitoring, scoring, record keeping, and report generation); it must be easily moved between laboratories and be somewhat "transparent" in the recording environment. The computer must process data in real-time, i.e., the major sleep physiology (EEG, EOG, EMG) must be captured directly from the polygraph amplifiers' outputs; this means no specialized hardware to pre-process the physiological activity [BAAS84]. The system must keep compressed records of the data. A complete record (one night of polysomnographic data) should be stored on one standard diskette. The system should also be able to generate a summary of nocturnal behavior, e.g., a hypnogram, total sleep time, sleep latency, number of arousals, etc.

Hardware Environment

The hardware environment is influenced by the operating constraints. One feasible system configuration is to use a 16-bit microprocessor based computer system. A standard IBM PC was used as a preliminary test system with 512K of dynamic memory, two floppy disk drives, and a color graphics card. A numeric co-processor (an 8087) was added to the personal computer to enhance its "number crunching" capability. A final piece of hardware was installed in the system; a standard "off-the-shelf" multi-channel 12-bit analog to digital and digital to analog converter with digital input/output (A/D- D/A converter)[2]. The A/D-D/A converter provided the means to connect the personal computer system to the polygraph.

Language Environment

The language environment for developing an expert operator was a 32-bit implementation of Forth (PC/Forth+; Laboratory Microsystems, Inc., Marina Del Rey). There were several reasons for choosing a 32-bit implementation over the standard 16-bit versions. A 16 bit microprocessor like the 8086 or the 68000 can directly address well over 64K bytes of memory. Contiguous data collection arrays of larger than one 16-bit memory page and the use of a numeric co-processor greatly influenced the selection of a "larger" Forth.

Knowledge Architecture

The architecture of a FORTES polysomnographer follows the general structure of an expert system found in Knowledge Engineering. Three primary components are found in the current test implementation: a small knowledge-base, a micro-slate, and knowledge engines (see Figure 1).

Knowledge-base. The knowledge domain for sleep staging can be divided into three segments: selection of what is to be measured; selection of what features should be extracted from the data; and selection of sleep staging rules.

The first segment involves the selection of primary sleep variables. Classically, only three measures have been used to assess sleep in normal humans: (central and occipital) EEG, EOG, and EMG. There is a qualification that other physiological variables are essential for Sleep Disorders Medicine: respiration variables for measuring hypopneas and apneas; cardiovascular variables for measuring heart function, e.g., heart rate [GUIL82]. Thus, the knowledge-base for sleep staging may be extended to include additional variables. Initially, only three primary variables were used; additional variables could be added at a later date.

The second segment of a knowledge domain for sleep staging involves extraction of the meaningful features in a sleep "epoch". Adequate feature extraction of sleep physiology has only become possible on micro-computer based systems within the last few years. This is due primarily to the introduction of 16/32 bit microprocessors, floating point co-processors (8087/80287), and the development of specialized digital signal processors, such as the Texas Instruments' TMS 32010 [REDI84B]. The structure in the EEG can be directly represented through digital signal processing methods, for example, frequency domain techniques, such as, the fast fourier transform (FFT, and power spectral density), Walsh functions [DZWO84], or by variants of time series modeling [BOX76]. The structure in EOG activity can be also represented by these techniques. EMG activity can be represented by integration and averaging.

The third segment of a knowledge domain for sleep staging involves explicitly defining classification rules. This means translating the inferential rules used by humans into explicit propositional constructs "understood" by machine intelligence. A complete listing of the knowledge rule-base is clearly beyond the scope of the current discussion. However, two simplistic examples help illustrate the rule constructs. A description representing MT in Table 1 might be: an epoch that appears to have extremely large EMG artifact, the EOG and EEG are unreadable. The translation into the rule-base is shown in the screen listing (Screen 28). Likewise, a description representing

---

[2] The A/D converter was a Data Translation (Malborough, MA) DT-2801.

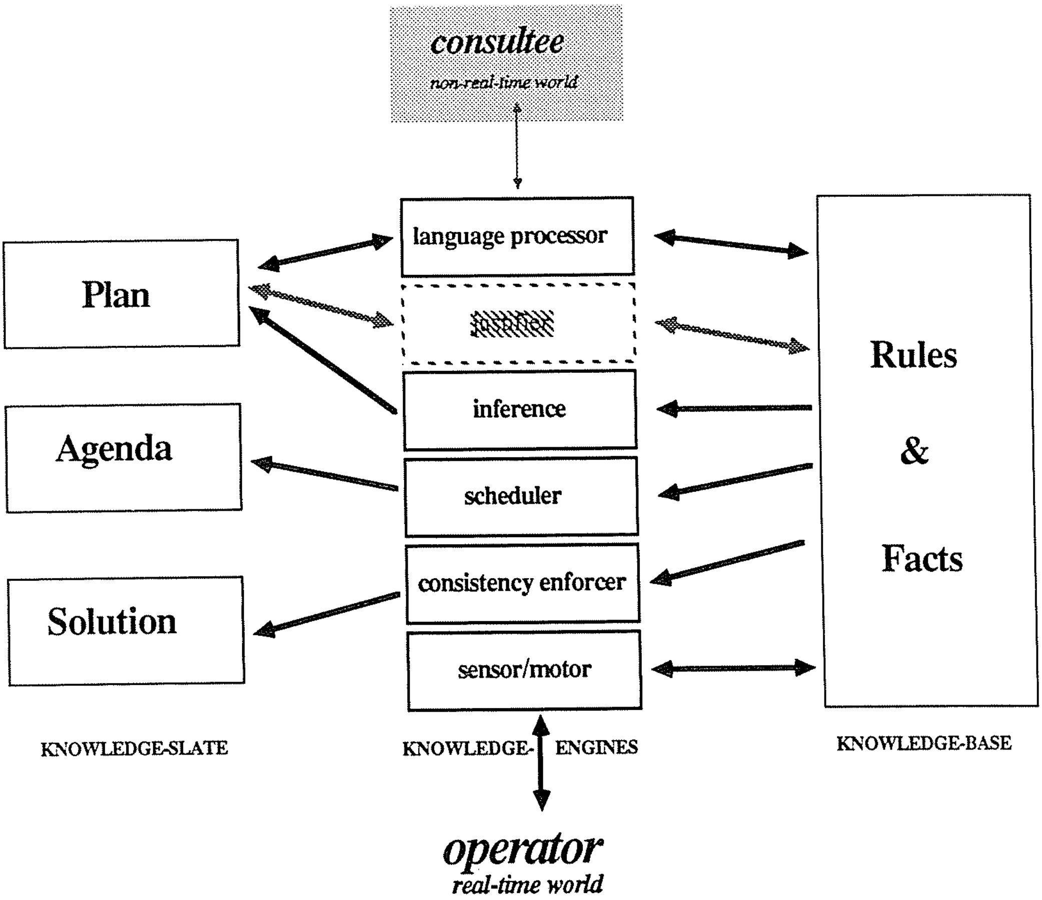

Figure 1. FORTES Knowledge Structure

1- REM in Exhibit 1 might be: an epoch that appears to have substantially reduced EMG, low voltage fast EEG, and rapid EOG activity; the translation of this description into a rule is shown in the second screen (Screen 29).

Micro-slate. The micro-slate is a reduced blackboard with a plan, an agenda, and a solution-slate. The plan is divided into several types only one of which is active at any given moment; for example, the analyze plan is the dominant behavior of a machine polysomnographer. During analyze, the machine operator initializes the black-board, activates the data collection engine, and establishes a "stage epochs until" plan. The agenda is a reduced slate, actually an execution stack, which has a list of the strategies to be tested on the current epoch. The solution-slate quite simply holds the status of the current epoch as the epoch is evaluated.

Knowledge Engines. The knowledge engines are the motors of a machine polysomnographer. Three distinctly different engines are used; each engine operates separately although time sequenced to the other engines. The first knowledge motor is a sensor/motor or data collection/feature extraction engine. The incoming data is sampled periodically and the pre-determined features for the EEG, EOG, and EMG are extracted via digital signal processing methods. The inference/scheduler engine is a simple motor that processes elementary stack structured propositions contained on the plan and agenda-slate. Finally, a consistency enforcing engine maintains a check on the scoring of each epoch; if time runs out this engine forces a solution, i.e., `EPOCH UNSTAGED` (see Glossary).

```
SCR # 28           File: FORTES
( RULE-BASE: stagable                              djr 02 sep 84)

    RULE STAGABLE
         EMG ARTIFACT?         ( check if EMG too high)
          -> EPOCH UNREADABLE  ( if so then skip this epoch.)
   END-RULE
```

```
SCR # 29        File: FORTES
( RULE-BASE: la-rem                              djr 02 sep 84)

RULE 1A-REM
            EEG-1 HI-FREQ?
            EEG-1 LOW-VOLT? AND
            EMG   LOW-AMPL? AND
            EOG   REMS?     AND
            ->  EPOCH S1-REM
END-RULE
```

Performance

The performance of a FORTES polysomnographer has been partially evaluated. The data acquisition/feature extraction engine was designed for pure speed and requires 87% of the processing time in an epoch. This leaves approximately 14% of processing available to the other knowledge engines. The inference/scheduler engine operates at approximately 1095 LIPS (1.1 KLIPS; this was tested by iterating a simple test proposition). With less than 15 percent of the processing time available not more than 148 primitive logical inferences can be processed during each epoch. This limits the amount of evaluation that can be performed on each epoch. Future performance reliability and accuracy comparisons between the machine and human polysomnographer will have to be evaluated via clinical trials and are beyond the scope of this paper.

Discussion and Conclusion

It is expected that a FORTES operator will match or out-perform a human expert in sleep staging. The anticipated performance is based on the accuracy of assessing the primitive features found in the physiological variables and in the reduced duration of a scoring epoch; the task of scoring 14,400 epochs for an average night of sleep requires more effort on the part of a human stager. A precursor to the FORTES operator with eight rules has demonstrated that automated scoring methods compare favorably with trained scorers in estimating sleep latency in humans in a pilot study[3]. However, more careful development of Expert Systems in the FORTES approach and refinement of the sleep staging knowledge-base is a prerequisite before further clinical studies can be used to compare machine versus human polysomnographers. The availability of more computational speed in the form of 16/32 bit personal computers, e.g., the AT and compatibles, also provides additional processing power for building Expert Systems capable of operating in real-time.

A FORTES polysomnographer could be applied to other areas of physiological/behavioral monitoring. Some important areas are real-time assessment of depth of anesthesia, cardiac function [PRES84], process control [SPEC84], monitoring of operator alertness, e.g., piloting an aircraft or other vehicle, or controlling air traffic. The latter examples become feasible with the development of real-time expert systems that run in microprocessor based instruments. In conclusion, the widest application of Knowledge Engineering and use of Expert Systems will be found in the personal computer arena; personal expert consultants may dominate at first. If the objective of real-time machine thinking can be accomplished then personal expert operators may find widespread application in a variety of environments. One step in the evolution of Knowledge Engineering is the development of real-time expert systems in Forth.

Acknowledgements

I would like to thank the Sleep Disorders Center Computer Sleep Staging Group: Boyd Hayes, Russ van Gelder and Drs. William Dement, Christian Guilleminault, Mary Carskadon, German Nino-Murcia, and Kelvin Lim for discussions on computerized sleep staging. I would also like to thank John Bumgarner, Dave Boulton, and Kim Harris in the Forth community.

References

[ASER53] Aserinsky, E., & Kleitman, N. Regularly occurring periods of eye motility and concomitant phenomena during sleep. *Science,* **118**, 273-274, 1953.

[BAAS84] Baas, L., & Bourne, J. R. A rule-based microcomputer system for electroencephalogram evaluation. *IEEE Transactions on Biomedical Engineering,* **BME-31**(10), 660-664, 1984.

[BARR81] Barr, A., & Feigenbaum, E. A. *The Handbook of Artificial Intelligence*. Stanford, California: HeurisTech Press. 1981.

[BARR82] Barr, A., & Feigenbaum, E. A. *The Handbook of Artificial Intelligence.* Volume 2. Los Altos, California: William Kaufmann, Inc. 1982.

[BOX76] Box, G. E. P., & Jenkins, G. M. *Time Series Analysis Forecasting and Control.* San Francisco: Holden-Day. 1976.

[BOUR83] Bourne, J. R., Matousek, M., Friberg, S., & Arvidsson, A. SEER-I: The semantic EEG evaluation regime. *IEEE Transactions on Biomedical Engineering,* **BME-30**(4), 1983, 239-244.

[BUCH83] Buchanan, B. G., Barstow, D., Bechtel, R., Bennett, J., Clancey, W., Kulikowski, C., Mitchell, T., & Waterman, D. A. Constructing an expert system. In: F. Hayes-Roth, D. A. Waterman, & D. B. Lenat (Eds.), *Building Expert Systems*. Reading, Massachusetts: Addison-Wesley Publishing Company, Inc., 1983, pp. 127-167.

---

[3] Data were collected from four subjects participating in an insomnia study during sixteen multiple sleep latency tests. Computer estimates of sleep latency based on single channel EEG spectral properties were not statistically dissimilar in comparison to trained human scorers estimates.

[DEME57A] Dement, W., & Kleitman, N. The relation of eye movements during sleep to dream activity: an objective method for the study of dreaming. *Journal of Experimental Psychology,* **53**(5), 1957, 339-346.

[DEME57B] Dement, W., & Kleitman, N. Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming. *Electroencephalography & Clinical Neurophysiology,* **9**, 1957, 673-690.

[DZWO84] Dzwonczyk, R., Howie, M. B., & McDonald, J. S. A comparison between Walsh and Fourier Analysis of the electroencephalogram for tracking the effects of anesthesia. *IEEE Transactions on Biomedical Engineering,* **BME-31(8)**, 1984, 551-556.

[GEVI80] Gevins, A. S. Pattern recognition of human brain electrical potentials. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **PAMI-2**(5), 1980, 383-404.

[GUIL82] Guilleminault, C. (Ed). *Sleeping and Waking Disorders: Indications and Techniques.* Reading, Mass: Addison-Wesley Publishing Company. 1982.

[GUIL78] Guilleminault, C., Hill, M. W., Simmons, F. B., & Dement, W. C. Obstructive sleep apnea: electromyographic and fiberoptic studies. *Experimental Neurology,* **62**, 1978, 48-67.

[JOHN77] Johnson, L. C. The EEG during sleep as viewed by computer. In: A. Remond (Ed.), *EEG Informatics. A Didactic Review of Methods and Applications of EEG Data Processing.* New York: Elsevier/North-Holland Biomedical Press, pp. 385-406. 1977.

[KUHN70] Kuhn, T. S. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press. 1970.

[LARS70] Larsen, L. E., & Walter, D. O. On automatic methods of sleep staging by EEG spectra. *Electroencephalography and Clinical Neurophysiology,* **28**, 1970, 459-467.

[LEE84] Lee, C.N., & Smith, J.R. REM detection for sleep EEG expert systems. *Proceedings of the 29th International Symposium in Medicine and Bioengineering.* New York, 1984.

[MOOR74] Moore, C. H. Forth: A new way to program a mini- computer. *Astronomy and Astrophysics Supplement,* **15**, 1974, 487- 511.

[PRES84] Preston, K., Fagan, L. M., Huang, H. K. & Pryor, T. A. Computing in medicine. *Computer,* **17**(10), 1984, 294-313.

[RECH68] Rechtschaffen, A., & Kales, A. (Eds.). *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects.* Public Health Service, U.S. Government Printing Office, Washington, D.C. 1968.

[REDI84A] Redington, D. Outline of a Forth Oriented Real-Time Expert System for Sleep Staging: a Fortes Polysomnographer. Paper presented at the Sixth FORML Conference, Asilomar, California, November 23-25, 1984.

[REDI84B] Redington, D. Towards real-time identification of sleep states. *Wescon '84 Conference Record* **36**/4/1-6. Los Angeles: Electron. Conventions. 1984, 1272pp.

[SMIT78] Smith, J. R. Computers in sleep research. In: T. C. Pilkington (Ed.), *Critical Reviews in Bioengineering.* West Palm Beach, Florida: CRC Press, Inc. 1978, pp 93-148.

[SMIT86] Smith, J. R. Automated analysis of sleep EEG data. In: F. J. Lopes da Silva, and W. P. Storm van Lewen (Eds.), *Part B: Application and Clinical Neurophysiology; Handbook of EEG and Clinical Neurophysiology, Volume 3: Computer Analysis of the EEG and other Neurophysical Signals.* Editor in chief: A. Rémond). Amsterdam: Elseivier. 1986; pp 129-145.

[SPEC84] Spector, A. Z. Computer software for process control. *Scientific American* **251**(3), 1984, 175-186.

[STKU82] St. Kubicki, Herrmann, W. M., Holler, L., & Scheuler, W. "Kritische Bemerkungen zu den Regeln von Rechtschaffen und Kalesuber die visuelle uswertung von EEG-Schlafableitungen." [Comments on the rules by Rechtschaffen and Kales about the visual scoring of sleep EEG recordings]. *Z. EEG-EMG,* **13**, 1982, 51-60.

[WILL74] Williams, R. L., Karacan, I., & Hursch, C. J. *Electroencephalography (EEG) of Human Sleep: Clinical Applications*. New York: John Wiley & Sons. 1974.

Manuscript received June 1985.

Glossary

WORD	STACK	DESCRIPTION
`->`	(fuzzy-flag -> ---)	Test the pseudo fuzzy flag on the top of the stack and branches around the consequence if "false"; `IF` by another name.
`1A-REM`		A rule in the Knowledge-base (see SCR #29).
`AND`	(fflag\fflag -> fflag)	Performs a logical bit-wise "and".
`ARTIFACT?`	(addr -> fuzzy-flag)	Evaluates whether the EMG value at `addr` is beyond a pre-defined artifact level and returns a flag.
`EEG-1`	(--- -> addr)	Places the address of EEG channel-1 within the solution-slate onto the top of the data stack.
`END-RULE`		Terminates a `RULE` definition; `THEN ;` by another name.
`EMG`	(--- -> addr)	Places the address of the EMG channel within the solution-slate onto the top of the data stack.
`EOG`	(--- -> addr)	Places the address of the EOG channel within the solution-slate onto the top of the data stack.
`EPOCH`	(--- -> addr)	Places the address of the current epoch in the solution-slate onto the top of the data stack.
`HI-FREQ?`	(addr -> fuzzy-flag)	Evaluates whether the EEG frequency at `addr` is above Beta (7Hz) and returns a flag.
`LOW-AMPL?`	(addr -> fuzzy-flag)	Evaluates whether the EMG amplitude at `addr` is below a pre-defined value and returns a flag.
`LOW-VOLT?`	(addr -> fuzzy-flag)	Evaluates whether the EEG amplitude at `addr` is below a pre-defined value and returns a flag.
`REMS?`	(addr -> fuzzy-flag)	Evaluates whether the EOG activity at `addr` is above a pre-defined value and returns a flag indicating the presence of rapid eye movements.
`RULE`		A Knowledge-base defining word add rules to the rule-base; `:` by another name.

`S1-REM`	(addr -> ---)	Places a solution value in the micro-slate at `addr`.
`STAGABLE`		A rule in the Knowledge-base (see SCR #28).
`UNREADABLE`	(addr -> ---)	Places an "MT" solution value in the micro-slate at `addr`, saves the current epoch, initializes the solution-slate and clears the agenda-slate
`UNSTAGED`	(addr -> ---)	Places an "unstaged" solution value in the micro-slate at `addr`, saves the current epoch, initializes the solution-slate and clears the agenda-slate.