J.E. Heiss, C.M. Held, P.A. Estévez,
C.A. Perez, C.A. Holzmann,
J.P. Pérez
Department of Electrical Engineering,
Universidad de Chile

Classification of Sleep Stages in Infants: A Neuro Fuzzy Approach

Editor's Note: This article won second-place in the 2001 EMBS Student Paper Competition. Except for minor editing and formatting changes, the article appears as it was submitted for the competition.

The sleep classification process is divided into three steps: data acquisition, pattern identification, and sleep-waking state-stage classification. In the first step, several signals generated by bioelectrical and biomechanical activity of the infant's body are recorded by a polygraph, generating a large number of pages with graphical data. The pattern identification process is performed for each page. The expert determines the background predominant frequency range in the electroencephalogram (EEG) according to [1]-[3]; relevant for this paper are the slow delta (SD) (0.5-2 Hz) and theta (TH) (3-7 Hz) frequency ranges. The EEG is also examined to detect sigma spindles (SS), which are in the 12-14 Hz range. The electrooculogram (EOG) and the electromyogram (EMG) are used to determine the presence of rapid eye movements (REMov) and muscular tone (MT), respectively. The polygraph records additional signals which the expert uses as context information, such as electrocardiogram (ECG), detection of body movements (BM), abdominal ventilatory movements, nostrils airflow, body temperatures, and oxymetry.

The most basic division in sleep classification is to distinguish between wakefulness (WA) and being asleep. There are two sleep states called REM and non-REM (NREM). NREM is subdivided in turn into four stages called NREM-I, NREM-II, NREM-III, and NREM-IV. NREM-III and IV were considered as a single stage called NREM-III&IV in this paper. The difference between the two is the threshold of SD presence.

To determine the sleep state or stage, the experts established certain rules, based on [2] and [3], that are shown in Table 1. However, sleep classification is not completely standardized and usually experts from different research centers have slightly different approaches. Even between expert co-workers there is usually less than 90% agreement in sleep classification [4].

The large amount of data, the complexity of the classification analysis, and the variability among human experts are reasons to develop an automated sleep classification system. An evaluation of the computerized system ALICE 3 using 50 subjects [5] showed substantial differences between automated computer scoring and manually scored paper polysomnographies. A manual edition of the computer scoring enhanced agreement to 75.7% with the paper polysomnography scoring. In [1], a pattern identification system for sleep stage classification that emulates the way the expert searches for each of the five relevant patterns was implemented. A ganglionar lattice system performed the classification,

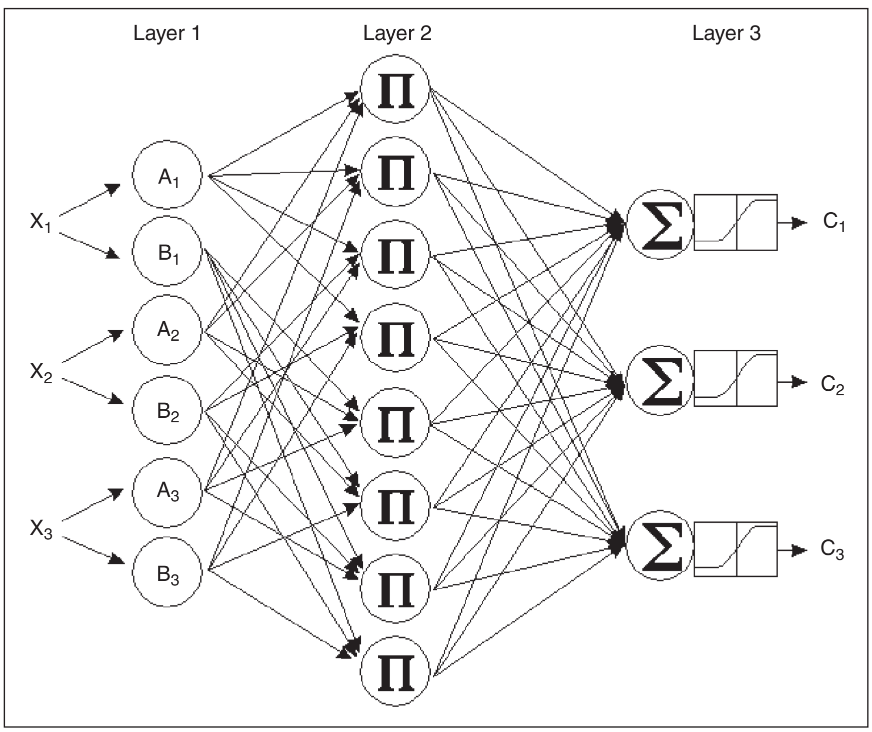
 0739-5175/02/$17.00©2002IEEE

achieving 84.9% of expert agreement, after manually removing several "noisy" pages from the database. Later on, in [6], the pattern detection algorithms were redesigned in order to enhance their robustness and evaluated with an enlarged database using the expert's rules of Table 1. An 86.7% correct classification rate was achieved for the testing set, which had no manual intervention and included "noisy" data.

In order to discover rules that may explain how the classification process should be performed and to find parameters that define the degree of presence or absence for a pattern, a neuro-fuzzy approach was chosen. The weight of each rule and the parameters of the membership functions were determined by supervised learning through an ANFIS [7] based neuro-fuzzy classifier (NFC) [8]. Nonrelevant rules were eliminated by applying a pruning algorithm. The remaining rules were analyzed and compared with the expert's rules.

Table 1. Expert's rules for sleep-waking state-stage classification. A: absent, P: present, X: irrelevant. A particular state or stage has to last at least 1 minute to be assessed as such.

Pattern	Sleep-Waking States and Stages				
	NREM-I	NREM-II	NREM-III&IV	REM	WA
REMov	A	A	A	P	P
TH	P	X	X	P	X
SD	A	A	P	A	A
SS	A	P	X	A	A
MT	X	X	X	A	P



1. Neuro-fuzzy classifier architecture. Layer 1 is the fuzzification layer. Three input variables are shown here (X_1, X_2, and X_3), each with two associated fuzzy concepts (A_i and B_i). Layer 2 generates all the possible rules of the form IF X_1 is A_1 and X_2 is B_2 and X_3 is A_3, with a T-norm operator ( ), considering one fuzzy concept per input variable. The output of layer 2 is a strength parameter for each of the rules. Each node at layer 3 performs a linear combination of the rules and uses a sigmoidal function to determine the degree of belonging of the input pattern to each class (C_1, C_2, C_3).

Methods

Data Acquisition

Eight continuous sleep recordings were obtained from infants between 6 and 13 months of age on a TECA 1A97 18-channel polygraph connected as follows: five EEG channels with electrode placement adapted for infants from the international 10-20 system (FP1-C3, C3-O1, FP2-C4, C4-O2, and C3-C4); EOG for REMov detection; tonic chin and diaphragmatic EMGs; ECG; body movement detection of upper and lower limbs using piezoelectric crystal transducers; abdominal ventilatory movements, using a mercury strain gauge; and nostrils airflow, by means of a thermistor. All data were simultaneously recorded on paper and on digital means at a 250-Hz sampling rate. The digital data were collected on hard disk and then stored in laser media for off-line analysis. Infant behavior was also observed directly and noted on the polygraph paper. Depending on the polygraph settings, a page can last 20 or 30 s. The digital recordings were divided in 20-s frames, which represented one paper page in most cases.

Pattern Identification

The system described in [6] was applied to obtain a level of presence for each of the five relevant patterns. The pattern detection system outputs are either percentages of presence or quality indices of a given pattern per frame. The outputs are in the [0, 1] range.

The data set was divided into four records with 2,067 frames for the training set, two records with 585 frames for the validation set, and two records with 858 frames for the testing set. The training set was used to adjust the parameters with supervised learning in order to achieve over 80% agreement on the validation set for each sleep-waking state-stage when using the expert's rules of Table 1. Two additional recordings were left for testing the system.

ANFIS-Based Neuro-Fuzzy Classifier

A neuro-fuzzy classifier (NFC) based on [7] and [8] was applied on the detected patterns to perform sleep-waking state-stage classification. Each of the five relevant patterns were associated to two fuzzy concepts, present and absent, with sigmoidal fuzzification functions

$$1/1 \quad e^{-s(x-c)}, \tag{1}$$

where s is the slope and c is the center of the sigmoid. The sign of s determines if the concept means present (+) or absent (–). Parameters s and c are determined through a training process, using the delta rule

$$\Delta W = -\mu \frac{\partial \varepsilon}{\partial W}, \quad (2)$$

where ΔW is the adjustment for the parameter W, ε is the sum of the squared error, and μ is the learning rate. The weights of the linear combinations at layer 3 were also determined by supervised learning using the delta rule (2). The NFC architecture allowed us to implement a fuzzy classification system with differentiable fuzzification functions at layer 1 (in our case sigmoidal functions), including parameters that were trained using the delta rule (2) with the squared error as the objective function

$$\varepsilon = \sum_{j=1}^{m} \sum_{i=1}^{n} \left(d_i^j - o_i^j \right)^2, \quad (3)$$

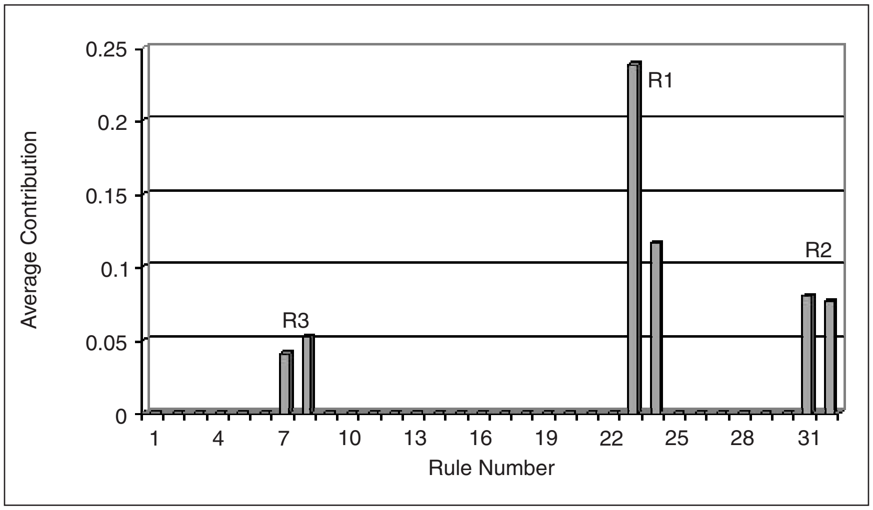
where m is the number of output classes, n is the number of examples, d is the desired output (0 or 1), and o is the node output at layer 3.

The validation set was used to establish when to stop the learning algorithm and to consider the tuned parameters as final. The testing set was used to evaluate the performance of the tuned system with independent data. Nonlinear relations resulted after training between the NFC input and output spaces.

A simplified diagram of the NFC system architecture is shown in Figure 1. A detailed explanation about the training process of an ANFIS network is given in [7]. The actual NFC model applied to the sleep problem had five inputs, each one with two associated fuzzy membership functions, and five output classes (WA, NREM-I, NREM-II, NREM-III&IV, and REM). The combination of the fuzzy concepts of layer 1 produced a total of 32 rules at layer 2. The maximum output at layer 3 determined the class associated to each input vector. The weights at layer 3 were initialized with random values in the [0, 1] interval. The center c was initially determined at half of the maximum input from all the respective training examples and the parameter s was set at ±5.

Postprocessing

In order to reduce the number of rules and thus produce a more expert-like set of fuzzy if-then rules, a pruning algorithm was implemented. For every output class, the average contribution of each rule was evaluated and a threshold of 0.01 was used to eliminate the least significant rules (the observed contribution values were always in the [0, 1] range). The last



2. System Rules for NREM-I. The average contribution of the surviving rule for NREM-I after pruning is shown. To simplify the analysis, the rules have been grouped in pairs (R1, R2, and R3), in accordance with their order of average contribution. The only difference in each pair of rules is the fuzzy concept associated to the MT input (presence or absence).

Table 2. Performances of the implemented NFC, a multilayer perceptron neural network (MLP), and the expert's rules. The results show the overall classification performance on a frame by frame basis except for the last column, which shows results on a 1-minute basis after applying the State Duration Algorithm (SDA).

	Training	Validation	Test	Test with SDA
NFC	86.2 ± 0.1%	87.7 ± 0.2%	83.9 ± 0.4%	88.2± 0.5%
MLP	87.1 ± 0.7%	87.3 ± 0.4%	83.4 ± 0.6%	87.3 ± 0.9%
Expert's Rules	84.1%	87.2%	82.6%	86.7%

Table 3. Fuzzy rules generated to assess stage NREM-I. The letters represent: A: absent, P: present, X: irrelevant. Absent and present are fuzzy concepts defined by sigmoidal functions.

Pattern	Fuzzy Rules Generated		
	R1	R2	R3
REMov	A	A	A
TH	P	A	P
SD	A	A	P
SS	A	A	A
MT	X	X	X

Table 4. Classification performance degradation for stage NREM-I when one of the rules (R1, R2, R3) is supressed.

NREM-I Stage	Classification Performance			
Supressed Rule	None	R1	R2	R3
% of Correct Classification	82.4%	19.7%	65.0%	76.8%

Table 5. Relative activation frequency for rules R1, R2, and R3 In NREM-I.			
Previous State to NREM-I	**R1**	**R2**	**R3**
NREM	40.7%	70.0%	100%
REM OR WA	59.3%	30.0%	0%

step of the classification process took into account that, according to expert criteria, every sleep-waking state-stage had to last at least 1 min [2]. A state duration algorithm (SDA) was developed to ensure this condition [6].

In order to compare the performance of the system with a general classification method, a multilayer perceptron (MLP) neural network with five input nodes and a hidden layer with ten nodes and five output nodes was trained, using the same training, validation, and testing sets as for the NFC.

Results

Ten simulations with the ANFIS-based NFC were performed and the test results were postprocessed by applying the SDA algorithm. The average results for the training, validation, and test sets are summarized in Table 2. This table also shows the results of classifying these sets using an MLP neural network and using the expert's rules of Table 1.

Only a few of the 32 rules survived after applying the pruning algorithm, for each of the 5 possible outputs (classes). As an example, the rules generated to classify NREM-I with the results of one of the ten simulations will be described in what follows. A similar analysis could be performed for all the other output classes. Figure 2 shows the average contribution to the node output for the rules that were not pruned. Table 3 shows the surviving rules (R1, R2, and R3), with their respective fuzzy concepts associated to each rule (absent or present). Only the examples classified as NREM-I by the NFC were considered in the average calculation.

Table 4 shows the system performance for classifying NREM-I after eliminating one of the three rules. Finally, Table 5 shows the relative activation frequency of rules R1, R2, and R3, as a function of the sleep-waking state previous to NREM-I. A rule was considered active when its contribution to the output was above 0.2 (in the [0,1] range).

Discussion

The results of applying MLP and NFC (Table 2) were statistically nondifferent at a level of significance of 0.01 (t-test), for all data sets (training, validation, and test). Both methods show an enhancement over applying the crisp expert's rules of Table 1. The last column of Table 2 shows the results after applying the state duration algorithm, which improved the classification percentages because it eliminated isolated frames with different patterns. The same partition of sets used in [6] was maintained in order to perform meaningful comparations between NFC, the crisp classifier, and the MLP.

To evaluate the pruning algorithm, the results of the NFC applied to the training, validation, and test set with and without pruning were compared, showing no statistically significant differences.

Figure 2 and Tables 3 and 4 show that there is a hierarchy among rules. R1 can be considered as the main rule while R2 and R3 are complementary rules; their combination made the system achieve a performance of over 80%. R1 matches exactly the expert's rule for NREM-I (Table 1). R2 and R3 are newly discovered fuzzy rules.

Table 5 shows relative activation frequency for the surviving rules as a function of the preceding sleep state. It shows that R2 and R3 activate mainly within NREM sleep. These results suggest that NREM-I sleep within NREM may have different characteristics than NREM-I following WA or REM state. The rules R2 and R3 may help to identify these differences.

Conclusions

An ANFIS based neuro-fuzzy classifier with a pruning algorithm was implemented and applied to the classification of sleep-waking states-stages in infants, using the sleep pattern detection system of [6] to generate the inputs. Including artifacted pages, an average of 88.2% of expert agreement was achieved for testing data. As a result of the training process and pruning, rules and parameters that defined a fuzzy classification system were also determined. Analyzing the rules obtained for sleep-stage NREM-I, it was found that the main rule matched the expert rule to classify NREM-I. Additional rules were discovered that complement the classification and may provide additional information about the characteristics of this sleep stage. This is a promissory result, and further research is needed in this topic.

Future work includes implementation of a clustering algorithm to determine the initial parameters of the system, training the system with a different objective function, such as the max-type error function described in [8], and evaluating the performance of different T-norms at layer 2 in Figure 1. The development of a general methodology for rule discovery and interpretation is also of interest.

Acknowledgments

This research project was partly supported by Conicyt-Chile under Grant Fondecyt 1980909 and the Department of Electrical Engineering, Universidad de Chile. We wish to thank the valuable cooperation of Dr. P. Peirano and M. Garrido from the Sleep Laboratory, INTA, Universidad de Chile.

Jaime E. Heiss was born in Santiago, Chile, in 1974. He received the B.S. degree in electrical engineering from Universidad de Chile, Santiago, Chile, in 1998. He is currently pursuing the P.E. degree in electrical engineering and the M.S. in biomedical engineering, at the same university. His research areas are sleep pattern recognition and neurofuzzy systems applied to the classification of physiological data.

Claudio M. Held received the Lic. Eng. Sc. (EE), the Electrical Civil Eng. and the M.S. degree in biomedical engineering from Universidad de Chile, Santiago, Chile, and the Ph.D. degree in biomedical engineering from Rensselaer Polytechnic Institute, Troy, New York, in 1995. He is currently an assistant professor of electrical engineering at Universidad de Chile and is a managing director at Aparatos de Comunicación Interna Ltda. His research interests are applying fuzzy logic and intelligent systems in biomedical engineering and other areas of decision making. He has worked in automated EKG interpretation, multiple drug infusion, classification of sleep-waking states in infants, handwritten digit recognition,

and fraud detection in long-distance telecommunications.

Pablo A. Estévez received the B.S. degree in electrical engineering and the P.E. degree in electrical engineering, both from the Universidad de Chile, in 1978 and 1981, respectively. He received the M.Eng. degree and the Ph.D. degree in information engineering, both from the University of Tokyo, in 1992 and 1995, respectively. Currently he is an assistant professor in the Department of Electrical Engineering, Universidad de Chile. He has been an invited researcher at the Communication Science Laboratory, NTT- Kyoto, Japan, and at the Ecole Normale Superieure de Lyon, France. His research interests include the application of neural networks and genetic algorithms to pattern recognition and biomedical issues. Dr. Estévez is a member of the INNS and Forest Products Society.

Claudio A. Perez received the B.S.(EE) (1980), the P.E.(EE), and the M.S. (BME) in 1985 from Universidad de Chile. He obtained the Ph.D. (BME) at The Ohio State University (OSU) in 1991. He is an associate professor at the Department of Electrical Engineering, U. de Chile. He is the graduate studies chairman at the E.E. Department. His research interests include computational vision including biologically inspired models for pattern recognition and man-machine interfaces. He is a member of the Pattern Recognition Society, Sigma-Xi, BMES, and OSU Alumni Association. He received a Fulbright Fellowship in 1986 to pursue his Ph.D. degree at OSU. In 1990 he was a Presidential Fellow and received a Graduate Student Alumni Research Award from OSU. In 1991 he received a Fellowship for Chilean scientists from Fundacion Andes.

Carlos A. Holzmann received an Engineering diploma (1961) from Universidad Católica de Valparaiso, Chile, and M.S. degree (1963) and Ph.D. degree (1966) from The University of Notre Dame, Indiana, USA. Since 1969 he has been a faculty member of Departamento de Ingeniería Eléctrica of the Universidad de Chile. He was a research associate at the University of Michigan (1971) and a visiting professor at the University of Notre Dame (1978). In 1975 he formed the Biomedical Engineering Research Group and initiated the Graduate Program in Biomedical Engineering at Universidad de Chile, serving as its chairman to this date. He was vice-dean of the Facultad de Ciencias Físicas y Matemáticas (School of Engineering) at Universidad de Chile (1975-1978). He presently serves as professor of biomedical engineering and electrical engineering.

Juan P. Pérez was born in Santiago, Chile, in 1975. He received the B.S. degree in electrical engineering and the P.E. degree in electrical engineering (with honors), from the Universidad de Chile, in 2000 and 2002, respectively. He is working in the Department of Electrical Engineering, U. de Chile, as a research assistant. His current research interests are neural networks, fuzzy logic, pattern recognition, and medical and meteorological applications.

Address for Correspondence: Jaime Heiss, Área de Automática y Bioingeniería, Departamento de Ingeniería Eléctrica, Universidad de Chile, Av. Tupper 2007, Piso 3, Casilla 412-3, Santiago, Chile. E-mail: jheiss@cec.uchile.cl.

References

[1] C.A. Holzmann, C.A. Perez, C.M. Held, M. San Martín, F. Pizarro, J.P. Pérez, M. Garrido, and P. Peirano, "Expert system classification of sleep-waking states in infants," *Med. Biol. Eng. Comput.,* vol. 37, pp. 466-476, 1999.

[2] A. Rechtschaffen and A. Kales, *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects.* Los Angeles, CA: BIS/BRI, Univ. of California, 1968.

[3] C. Guilleminault and M. Souquet, "Sleep states and related pathology," in *Advances in Perinatal Neurology,* vol. 1, R. Korobkin and C. Guilleminault, Eds. New York: Spectrum, 1979.

[4] R. Norman, I. Pal, C. Stewart, J. Walsleben, and D. Rappaport, "Interobserver agreement among sleep scorers from different centers in a large dataset," *Sleep,* vol. 23, no. 7, pp. 901-908, 2000.

[5] D.P. White and T.J. Gibb, "Evaluation of a computerized polysomnographic system," *Sleep,* vol. 21, no. 2, pp. 188-196, 1998.

[6] P.A. Estévez, C.M. Held, C.A. Holzmann, C.A. Perez, J.P. Pérez, J. Heiss, M. Garrido, and P. Peirano, "Polysomnographic pattern recognition for automated classification of sleep-waking states in infants," *Med. Biol. Eng. Comput.,* vol. 40, pp. 105-113, Jan. 2002.

[7] J.S.R. Jang, "ANFIS: Adaptive-network-based fuzzy inference systems," *IEEE Trans. Syst., Man, Cybern.,* vol. 23, pp. 665-685, May 1993.

[8] J.S.R. Jang and C.T. Sun, "A neuro-fuzzy classifier and its applications," in *Proc. 2nd IEEE Int. Conf. Fuzzy Systems,* San Francisco, CA, 28 Mar. –1 Apr. 1993, pp. 94-98.