4th WSEAS Int. Conf. on COMPUTATIONAL INTELLIGENCE, MAN-MACHINE SYSTEMS and CYBERNETICS Miami, Florida, USA, November 17-19, 2005 (pp224-232)

# Classification of EEG signals by radial neuro-fuzzy system

David Coufal
Institute of Computer Science AS CR
Pod Vodárenskou věží 2, 182 07 Prague 8
Czech Republic

*Abstract:* We present a hybrid system for classification of EEG signals into the three classes of mentation, relaxation and micro-sleep. The classifier is based on an neural representation of a radial conjunctive fuzzy system. Conjunctive fuzzy systems are the fuzzy systems which employ fuzzy conjunctions for representation of IF-THEN structure of their rules. Radial fuzzy systems have, in addition, IF parts represented by radial functions which helps to simplify their computation model. GUHA data mining method and genetic algorithms are used for learning the classifier.

*Key-Words:* Neuro-Fuzzy Systems, Radial Fuzzy Systems, Data Mining, Hybrid Systems

## 1. Introduction

Analysis and classification of EEG signals is nowadays well recognized area of interest not only with respect to a medical research [1]. The other domains connected with the area are biocybernetics and the area of study of man-machine interactions where the special role is played by design and manufacturing automatic systems for indicating operators attention decrease episodes.

Compared by frequency, the most common accidents related to the operators attention decrease phenomenon are connected with transportation, especially, with trucks driving. Every year substantial loses on lives and property are reported in connection with accidents caused by drivers attention decrease episodes. Of course, other areas of man-machine interactions are even more important with respect to the damage extent of possible accidents. Let us just mention airport traffic and power stations control.

The research presented in the paper is the contribution to a design of algorithms for an automatic indication of emergency situations caused by attention decrease episodes of human operators. In our case, an operator's attention level is determined on the basis of measurement, mathematical transformation and classification of EEG signals. More precisely, raw EEG signals which are measured on operator's head, are transformed to EEG spectrograms which are classified into three classes of *mentation, relaxation* and *micro-sleep*.

The classifier is build in the form of a neuro-fuzzy system. The fuzzy part of the classifier is treated in the form of a radial conjunctive fuzzy system. The neural part, i.e., the neural representation, corresponds to a radial neural basis network concept. The representation of the system allows the data mining and genetic algorithms to be used for its structure and parameter learning, respectively.

The organization of the paper is as follows: In the next section we review the concept of a radial conjunctive fuzzy system (radial C-FS), its properties, computational model and the neural representation of such a system. In the third section we discuss learning of introduced neuro-fuzzy system. Section four consists of description of EEG data, i.e., how they are obtained, what is its structure and how they are pre-processed so that they can by classified by the presented classifier. The fifth section presents the results obtained with respect to the classification and few remarks on the future research.

## 2. Radial conjunctive fuzzy systems

In the area of fuzzy computing, there are recognized two approaches to the representation of IF-THEN rules and their groups - rule bases [3]. These are so called the conjunctive and the implicative approach. In conjunctive approach IF part of a rule (the antecedent) is combined with the THEN part (the consequent) by a fuzzy conjunction and individual rules are combined by a fuzzy disjunction. In the implicative approach an antecedent is combined with the consequent by a fuzzy implication and individual rules are combined by a fuzzy conjunction. In this paper we are interested in the first approach.

Mathematically, the rule base of a conjunctive fuzzy system is written as

$$RB(\boldsymbol{x}, y) = \bigvee_{j=1}^{m} A_j(\boldsymbol{x}) \star B_j(y), \qquad (1)$$

where $A_j$, $B_j$ are fuzzy sets representing antecedents and consequents of IF-THEN rules , $j = 1, \ldots, m$ ($m \in \mathcal{N}$, $m \geq 2$, is the number of rules in the rule base). IF-THEN structure of individual rules $R_j$ is represented a fuzzy conjunction, i.e., $R_j(\boldsymbol{x}, y) = A_j \star B_j$. Typically, this fuzzy conjunction is the same as it is used for building antecedents. Single rules are then combined by a fuzzy disjunction $\bigvee$ (usually max is used) to obtain the representation of whole rule base $RB$, as it is presented in formula (1).

Fuzzy sets representing antecedents are generally multidimensional, defined on $\mathcal{R}^n$ space, $n \in \mathcal{N} = 1, 2 \ldots$, and composed in the standard way from one-dimensional fuzzy sets $A_{ji}$, employing a $t$-norm $\star$ as a fuzzy conjunction. Formally we have

$$A_j(\boldsymbol{x}) = A_{j1}(x_1) \star \cdots \star A_{jn}(x_n), \qquad (2)$$

where $\boldsymbol{x} \in \mathcal{R}^n$, $\boldsymbol{x} = (x_1, \ldots, x_n)$.

A fuzzy system has nominal consequents if consequent fuzzy sets $B_j$s are defined on finite, generally unordered, universe of discourse $Y$ of so called actions $y_1, \ldots, y_l$, i.e., $Y = \{y_1, \ldots, y_l\}$, $l \in \mathcal{N}$. Particular $y_k$s, $k = 1, \ldots, l$, are treated as possible actions, e.g., *go right*, *go left* or classes *mentation*, *relaxation*, *micro-sleep* without no ordering assumed. The value $B_j(y_k)$ then corresponds to the membership degree $\mu_{kj}$ (indices are switched) of action $y_k$ into the consequent set $B_j$ of the $j$-th rule.

A fuzzy system is radial if antecedents of its IF-THEN rules exhibit the radial property. The property refers to a shape preservation of one-dimensional fuzzy sets in antecedents after their combination by a $t$-norm. The formal definition of radial conjunctive fuzzy systems (radial C-FSs) with nominal consequents follows:

**Definition 1** *A conjunctive fuzzy system with nominal consequents is radial if*:

(i) *There exists a continuous function* $act : [0, +\infty) \to [0, 1]$, $act(0) = 1$ *as follows:* (a) *either there exists* $z_0 \in (0, +\infty)$ *such that* $act$ *is strictly decreasing on* $[0, z_0]$ *and* $act(z) = 0$ *for* $z \in [z_0, +\infty)$ *or* (b) $act$ *is strictly decreasing on* $[0, +\infty)$ *and* $\lim_{z \to +\infty} act(z) = 0$.

(ii) *Fuzzy sets in antecedent and consequent parts of the $j$-th rule are specified as*

$$A_{ji}(x_i) = act\left(\left|\frac{x_i - a_{ji}}{b_{ji}}\right|\right), \qquad (3)$$

$$B_j(y_k) = \mu_{kj}, \qquad (4)$$

*where* $n, m, l \in \mathcal{N}$; $i, j, l = 1, \ldots, n, m, l$, *respectively*; $\boldsymbol{x} \in \mathcal{R}^n$, $\boldsymbol{x} = (x_1, \ldots, x_n)$; $y_k \in Y = \{y_1, \ldots, y_l\}$; $\boldsymbol{a}_j \in \mathcal{R}^n$, $\boldsymbol{a}_j = (a_{j1}, \ldots, a_{jn})$; $\boldsymbol{b}_j \in \mathcal{R}^n_+$, $\boldsymbol{b}_j = (b_{j1}, \ldots, b_{jn})$, (*i.e.,* $b_{ji} > 0$); $\mu_{jk} \in [0, 1]$

(iii) *For each* $\boldsymbol{x} \in \mathcal{R}^n$ *the radial property holds, i.e.,*

$$A_j(\boldsymbol{x}) = act(\,||\boldsymbol{x} - \boldsymbol{a}_j||_{\boldsymbol{b}_j}), \qquad (5)$$

*where* $||\cdot||_{\boldsymbol{b}_j}$ *is a scaled version of some norm in* $\mathcal{R}^n$. *This norm is common to all rules of the fuzzy system.*

Let us comment on the definition. According to the definition a conjunctive fuzzy system is radial if it satisfies three requirements:

The first and the second require membership functions of one-dimensional fuzzy sets forming antecedents to they have a radial shape. That is, to they be formed by an application of a decreasing function $act$ on a distance of argument $x$ from a central point $a$. Examples of act functions which lead to triangular or Gaussian fuzzy sets are $act(z) = \max\{0, 1 - z\}$, $act(z) = \exp(-z^2)$, respectively. Examples of these fuzzy sets are graphically presented presented in Fig. 1.

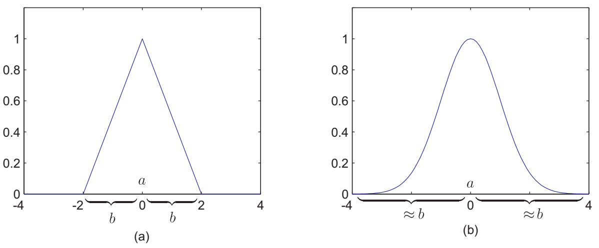


Figure 1: (a) Triangular fuzzy set; (b) Gaussian fuzzy set

The radial property is specified by the third requirement. As indicated, it refers to a radial shape preservation property in antecedents of rules. More specifically, it requires the membership function representing antecedents to be radial and to have the same shape ($act$ function) as one-dimensional fuzzy sets have. The radial character of $A_j$ is formalized by application of $act$ function on the distance of input argument from a central point. The distance is computed employing a scaled version of a norm in multidimensional space $\mathcal{R}^n$; and it is required to scaling parameters correspond to scaling parameters $b_{ji}$ of one-dimensional sets $A_{ji}$. A scaled version of some norm $||\cdot||$ in $\mathcal{R}^n$, with scaling parameter $\boldsymbol{b} = (b_1, \dots b_n)$, $b_i \geq 0$ is defined as $||\boldsymbol{u}||_{\boldsymbol{b}} = ||u_1/b_1, \dots, u_n/b_n||$ for $\boldsymbol{u} \in \mathcal{R}^n$. Moreover, central point $\boldsymbol{a}_j \in \mathcal{R}^n$ has to be composed from central points $a_{ji}$ of $A_{ji}$ sets, i.e., $\boldsymbol{a}_j = (a_{j1}, \dots, a_{jn})$.

The property is not trivial. It means not all, radial fuzzy sets can be combined with the all the $t$-norms. For example triangular fuzzy sets cannot be combined by the product $t$-norm in such a way that the radial property holds. The question which combination of $t$-norms and $act$ functions are allowed in order to the radial property hold, are answered by the representation theorem, see [4].

### 2.1. Computational model of a radial C-FS with nominal consequents

In order to build the computational model of a radial C-FS with nominal consequents we adopt the assumption of standard architecture of whole fuzzy system [5], we assume that the system is composed of four building blocks: 1) singleton fuzzifier, 2) conjunctive rule base, 3) CRI inference engine and 4) some defuzzification block.

Under the assumption of the singleton fuzzifier to be used in the system the general CRI formula simplifies to the form of

$$B(y_k) = RB(\boldsymbol{x}, y_k) \tag{6}$$

for $k = 1, \dots, l$ and $\boldsymbol{x}$ being the actual input to the system. Combining this fact with the representation of implicative rule base (1) we get

$$B(y_k) = \bigvee_{j=1}^{m} A_j(\boldsymbol{x}) \star B_j(y_k). \tag{7}$$

Introducing sets $B'_j$ as those issuing from individual rules we have $B'_j(y_k) = A_j(\boldsymbol{x}) \star B_j(y_k)$; and representing fuzzy disjunction by maximum we have the above in form

$$B(y_k) = \max_j\{B'_j(y_k)\} = \max_j\{A_j(\boldsymbol{x}) \star B_j(y_k)\}. \tag{8}$$

This computational model is standard well known computational model of conjunctive fuzzy systems. However, in radial systems it is simplified as $A_j(\boldsymbol{x})$ values can be easily computed employing the radial character of antecedents. Moreover, the radial fuzzy systems can be easily transformed into the form of a hybrid neuro-fuzzy system.

### 2.2. Neural representation

The neural representation of a radial C-FS with nominal consequents is based on the concept of radial basis functions (RBF) networks. The corresponding network is three-layered, feed-forward network. The first layer is the input layer with only identity transfer function, i.e.,

$$u_i = x_i. \tag{9}$$

The second, hidden layer, consists of $m$ hidden nodes $h_j$ which corresponds to antecedents of individual rules $A_j$. Due to the radial property hidden nodes form radial functions in $\mathcal{R}^n$ space. Weights from the input layer to the hidden layer correspond to values $1/b_{ji}$, i.e., to reciprocals of scaling parameters of norm used in representation of antecedents according to formula (5). Incorporating weights into the computation of hidden nodes we get the following specification of $h_j$s, $j = 1, \dots, m$ computation:

$$h_j(\boldsymbol{x}) = A_j(\boldsymbol{x}) = act(\ ||\boldsymbol{x} - \boldsymbol{a}_j||_{\boldsymbol{b}_j}). \tag{10}$$

The third, the output layer, consists of $l$ output nodes corresponding to the cardinality of $Y$ set. The computation of $o_k$ nodes, $k = 1, \dots, l$ is given as

$$o_k = \max_j\{h_j(\boldsymbol{x}) \star \mu_{kj}\} = \max\{A_j(\boldsymbol{x}) \star B_j(y_k)\}. \tag{11}$$

The output of an $o_k$ node can be seen as possibility degree of taking action $y_k$ as overall input of the system. If only single action is considered as the output, then the winner takes it all strategy is adopted to determine the output action or class. This can be seen as a variant of MOM (mean of maxima) defuzzification method.

The main advantage of the presented neural representation is the possibility of bringing learning algorithms known from neural computing into the area of conjunctive radial fuzzy systems. The next section is devoted to the learning of presented neuro-fuzzy systems.

## 3. Learning of neuro-fuzzy classifier

Learning of a radial fuzzy C-FS in neural representation consists from two subtasks of structure and parameter learning. In this paper we present an employment of GUHA data mining method for the first task and genetic learning for the second subtask. In the following two sections we present basic ideas of this approach. The details will be presented with respect to the concrete task in the fourth section.

### 3.1. GUHA method

Structure learning consists in a specification of a number of hidden nodes (rules) and initial setting of parameters of the network. In this work we have employed GUHA data mining method to perform this task. A limited introduction to the method follows.

The GUHA data mining method is the method of exploratory data analysis based on logical and statistical principles. Its origins fall into to the mid-60 sixties of the last century and from this times it is under continual developement at both levels - theoretical and practical (software implementations).

The basic idea of the method is to mechanically construct at the syntactical level (exploratory data analysis) relational patterns over data an test effectively if these relations hold in the data at semantic level. Data analyzed by GUHA method has form of a table, where rows corresponds to objects and columns to variables observed on these objects. A cell in the $i$-th row and the $j$-th column contains a value of $j$-th variable for $i$-th object.

Relational patterns are called hypotheses in GUHA method. A hypothesis has form $A \approx S$ and consists of three parts - antecedent $A$, succedent (consequent) $S$, and a generalized quantifier $\approx$. Antecedent and succedent (together called as *cedents*) are in fact Boolean conjunctions of a flexible length representing Boolean properties of data. These conjunctions are composed from elementary literals which are called categories. Categories are set by user of the method and they are devised from variables as subsets of their ranges. As an example consider variable *sex* and its two categories *sex*[male] and *sex*[female]; or variable *age* and the category *age*[20,30], i.e., to be between 20 and 30. An example of compound cedent of length two is *sex*[male] & *age*[20,30].

Each cedent for each object can be evaluated as true or false, simply by checking if the corresponding values of variables fall into the respective categories or not. If a value falls into the category then the evaluation of this category is 1 otherwise 0. The evaluation of whole cedent is obtained according to laws of Boolean logic for elementary conjunction. Thus for pair $sex$=female and $age$=25 the evaluations of above categories is false and true, respectively; and the evaluation of the compound cedent is therefore false.

For a given pair of cedents, based on evaluation for every object, we can construct contingency table (four fold table, ff-table) consisting of four integers $a, b, c, d$, where $a$ is the number of object satisfying (evaluation is 1) simultaneously $A$ and $S$, $b$ is the number of object simultaneously satisfying $A$ and not satisfying $S$ (evaluation is 0). Similarly $c$ for non(A)&S, and $d$ for non(A)&non(S).

Based on a concrete ff-table for a certain pair of cedents $A$, $S$ and the chosen quantifier, a hypothesis is evaluated as valid or invalid in GUHA sense. A quantifier is mathematically represented by its associated function which maps ff-tables (quadruples $a, b, c, d$) to set $\{0, 1\}$. The form of this function depends on the type of relations we are looking for in the data. Most common are associative relations such as "many objects satisfying (having the property) $A$ satisfy (have) also (the property) $S$"; or "to have the property $A$ is statistically dependent with to have property $S$". As an example, let us show explicitly the associative function for the first case. The quantifier is called founded implication (FIMPL) and its associated function has two parameters $cp \in [0, 1]$ driving the specificity and $base \in \mathcal{N}$ driving the support of hypothesis. The function depends only on values $a$ and $b$ of a corresponding ff-table, and mathematically is written as

$$\mathrm{FIMPL}(a, b) = \\ = \begin{cases} 1 & \text{if } a \geq base \text{ and } \frac{a}{a+b} \geq cp, \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

If, for a given hypothesis $A \approx S$, the evaluation of respective associated function (corresponding to $\approx$) is 1 then hypothesis is taken as valid otherwise as unvalid. The main power of GUHA method lies in the fact that it tests enormous number of hypotheses for their validity. The tested hypotheses, called relevant hypotheses, are built on the basis of syntactical patterns specified by user. The patterns consists in the specification of maximal lengths of both antecedent and succedent, variables and categories used for building cedents and type of the quantifier with setting its parameters. A hypothesis satisfying these patterns is relevant hypothesis and each relevant hypothesis is tested for its validity. Valid hypotheses then form the output of the GUHA method.

More details about the method can be found in [7, 8, 9, 10].

### 3.2. Structure learning by GUHA method

As we have announced, the GUHA method can be used for structure learning of our neuro-fuzzy system. At the general level, the GUHA method is used to identify a set of valid (with the high significance) hypotheses using FIMPL quantifier. Antecedents are build from categories of input variables. The specification of categories is up to user, but few semi-automatic procedures are available in the GUHA method implementing software such as equifrequent and equidistant splits of variables ranges. Succedents are formed from categories of $Y$ variable where single actions $y_k$ correspond to different categories.

The number of hypotheses found and their strengths can be interactively driven by user by setting values of $cp$ and $base$ parameters. Revealed valid hypotheses are then used for structure learning as follows: Each hypothesis corresponds to a single IF-THEN rule. Antecedents of GUHA hypotheses correspond to antecedents of IF-THEN rules, i.e., to hidden nodes $h_j$. Parameters $a_{ji}$ correspond to midpoints of respective categories and parameters $b_{ji}$ to halves of theirs widths (lengths of intervals). Finally, values $\mu_{kj}$ are set to be one if the succedent of the $j$-th hypothesis corresponds to $y_k$, otherwise it is set to zero. The other specific details depend on concrete applications.

### 3.3. Parameter learning

Parameters of the presented neuro-fuzzy system are of three types. Central points of antecedents $\boldsymbol{a}_j$, scaling parameters $\boldsymbol{b}_j$ and memberships degrees of actions $\mu_{kj}$. Elements of central points $a_{ji}$ are real numbers, the scaling parameters $b_{ji}$ are non-negative numbers and $\mu_{kj}$ values are from the unit interval. In order to have unified approach during genetic learning we transformed $a_{ji}$ and $b_{ji}$ values on unit scale employing the sigmoid function.

The search for suitable readjusting of initally set parameters is performed by the following genetic algorithm. The number of bits $nb$ for representing one parameter is chosen. Using this number of bits the values from unit intervals are linearly transformed on interval $[0, 2^{nb} - 1]$ and then its binary representation is used to form one gene. Single gene consists of binary representation of $m(2n + l)$ parameters, so the final length of single gene is $nb{\cdot}m(2n+l)$ bits.

After coding the initially set parameters from structure learning, the initial population is created and the standard genetic learning is performed consisting of crossover and mutation operations and selecting the best evaluated genes. The evaluation is done in the following way: A gene corresponds to a certain neuro-fuzzy system. Based on a given set of training inputs, it determines the set of outputs, in fact set of actions $y_k$ these are then compared with desired outputs. If two single outputs matches then de difference is set to zero otherwise to 1. Differences are then summed and the obtained value form the evaluation of gene. Clearly, in the ideal case this evaluation is zero. So during the learning the genes with lower evaluation are the better.

In the next section we demonstrate an employment of presented framework in a concrete example of classification of EEG signals and their spectrograms.

## 4. EEG signals and their spectrograms

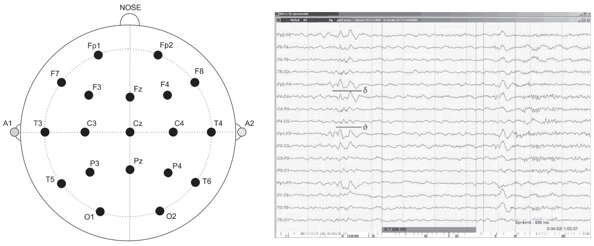
The measurement of EEG (electroencephalographic) signals is now well established area. Our research is based on the measurements performed at Joint Laboratory of System Reliability of Czech Technical University, Prague, which is equipped by relevant hardware and software tools; and cooperates with the Institute of Computer Science AS CR, in the project aimed at micro-sleeps detection [1, 2].

The EEG signals we have analyzed were obtained from measurements on approximately 60 volunteers (probands) who were students and professional drivers. The probands were asked to attend the measurement after experiencing an as long as possible sleep deprivation period in order to micro-sleeps events occur. During measurement sessions probands had to solve several psychological tests (Raven, addition of single and double-cipher integer numbers). They were driven to exhibit mentation phase in their brain activity, relaxation phase and also, in successful sessions, micro-sleeps events were detected.

At the hardware level, the EEG signals measurement was performed by using a special hat consisting of 19 electrodes spread over the proband's head, see Fig. 2(a) The raw signals were recorded using a special hardware card installed in PC and loaded into viewing and processing software. An example of the raw time record is presented in Fig. 2(b).

In the related software, preprocessing of raw EEG signals is performed together with Gabor analysis yielding spectrograms for 3 min moving windows. An example of spectrogram is presented in Fig. 3. We can see that it consists of signal intensity in 14 frequencies (1-14 Hz).

According to the medical terminology, individual frequencies are grouped into the 4 bands,




(a) (b)

Figure 2: (a) The distribution of electrodes over the proband's head; (b) An example of raw EEG records

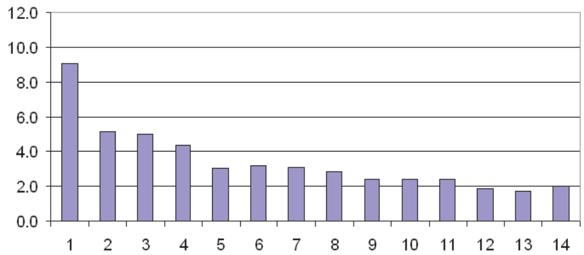



Figure 3: An example of spectrogram

called delta, theta, alpha, beta; and the first three bands are further split into the so called slow and fast sub-bands. The separation of frequencies is presented in Table 1.

| delta - d 1 - 3 Hz | | theta - t 4 - 7 Hz | | alpha - a 8 - 13 Hz | | beta - b 14 - 30 Hz |
|---|---|---|---|---|---|---|
| d1 1 - 2 | d2 3 | t1 4 - 5 | t2 6 - 7 | a1 8 - 10 | a2 11 - 13 | b no split |

Table 1: Separation of frequencies into bands

The aim of the research presented in the paper is to classify spectrograms into one of three classes - *mentation, relaxation* and *micro-sleep*. During the recent research [1, 2] it was verified that instead of using directly spectrograms (signal intensities in individual bands) for classification it is advantageous to use their ratios in order to eliminate individual biases of probands. Especially, it was shown that a/d ratios is tightly related to the classification of EEG signals to the mentioned classes.

Ratios of bands' intensities were also employed in our approach to classification. The classificator was designed in the form of a radial neuro-fuzzy system. The process of building (learning) of this neuro-fuzzy system/classifier is described in the next two sections.

### 4.1. EEG spectrograms - structure learning

Structure learning of neuro-fuzzy classifier was performed using the GUHA method as it was presented in previous sections. The analyzed data consisted of 79 spectrograms (objects) with 24+1 variables computed/observed on them. The first 24 variables corresponded to 24 ratios t/d, a/d, b/d, d/t, a/t, b/t, d/a, t/a, b/a, d/b, t/b, a/b, t1/d1, a1/d1, d1/t1, a1/t1, d1/a1, t1/a1, t2/d2, a2/d2, d2/t2, a2/t2, d2/a2, t2/a2; and the last variable to the classification into one of three class of mentation, relaxation and micro-sleep.

For the ratio variables, the categories were created on the basis of equifrequent splits of variables' ranges. For each variable three categories were created. Variable *class* had also three categories corresponding to the individual classes.

During the GUHA analysis we have searched for FIMPL based hypotheses with parameters to be set as $cp = 0.9$ and $base$ varying from $10$ to $5$. Antecedents were built from categories derived from ratio variables. Succedents were built from categories corresponding to the class variable. Maximal length of antecedent was set to be 3 and succedent to be 1.

In Table 2 there are summarized the hypotheses we have found by the analysis:

1. a/d$[1.3, 4.7]$ & a/t$[1.5, 4.9] \rightarrow_{0.9,\,10}$ *relaxation*
2. b/t$[0, 0.1]$ & d1/a1$[1.5, 4.9]$ & t1/a1$[1.5, 4.9] \rightarrow$ *ment.*
3. a/d[1.3,4.7] & b/d[1.5,4.9] $\rightarrow_{0.9,\,10}$ *relaxation*
4. a/d[1.3,4.7] & b/a[,4.9] & $\rightarrow_{0.9,\,10}$ *micro-sleep*

Table 2: The results of GUHA analysis

The categories in antecedents, actually the intervals forming these categories were used for initial setting of parameters of the neuro-fuzzy classifier. Mid-points of intervals were used to determine central points $a_{ji}$ and halves of theirs lengths to determine initial values of $b_{ji}$ parameters. These initially set parameters were then readjusted by parameter learning.

### 4.2. EEG scpectrograms - parameter learnig

To perform parameter learning we had to chose a $t$-norm and activating function to get the full specification of classifier. In our case we have used the product $t$-norm and exponential activating function, i.e., $act(z) = \exp(-z^2)$. Thus antecedents of rules/hidden nodes were represented by multidimensional Gaussians.

A genetic learning algorithm was used for parameter learning of the classifier. In every step of the algorithm we have worked with a population of 40 genes, each of them coded parameters of the classifier. Coding was performed in two steps. In the first step, real valued parameters were transformed on unit interval range employing the sigmoid function, and then, in the second step, each parameter was coded employing 10 bits.

After the coding, the standard learning procedure was employed. We separated population of genes into two halves and performed crossover between genes of two groups. Mutation process with the probability of mutation 0.05 was incorporated into the learning procedure as well. After each epoch, 2/3 of best genes were retained and 1/3 was replaced by new genes. These genes were generated by mutations from initial population of genes which issued from structure learning. We performed 1000 epochs of learning. The results are presented in the next section.

## 5. Results and conclusions

After the parameter learning we have obtained the final adjustment of parameters of neuro-fuzzy classifier. The accuracy of classification for analyzed spectrograms is summarized in Table 3.

The neuro-fuzzy classifier achieved accuracy of 69% (54/78) which is not high at the first glance. However, it was the best what we were able to get with the presented model. For example C4.5 algorithm achieved accuracy of 61% what is the comparable result. The important thing is that with respect to the correct classification of micro-sleep events we have achieved accuracy 73% (19/26) which is good performance.

| | predicted class | | |
|---|---|---|---|
| real class | mentation | relaxation | micro-sleep |
| mentation | 15 | 3 | 8 |
| relaxation | 3 | 20 | 3 |
| micro-sleep | 5 | 2 | 19 |

Table 3: Result of classification

The main portion of misclassification was caused by the fact that the spectrograms for mentation and micro-sleep classes are very similar.

The aim of the research presented was to introduce a neuro-fuzzy system for classification of EEG signals. The roots of the system are in the theory of radial fuzzy systems and radial basis neural networks. We have introduced a hybrid approach to the learning of presented neuro-fuzzy system which is based on the GUHA data-mining method and genetic learning.

In the future research we will further develop the presented classifier, especially on the basis of experimental data obtained currently from ongoing measurements.

### Acknowledgement

The research presented in the paper was supported by grant ME 701 of Ministry of Education, Youth and Sports of the Czech Republic.

*References:*

[1] Novák M., Faber J., Votruba Z. 2004. *Problem of Reliability in Interactions between Human Subjects and Artificial Systems (First Book on Micro-Sleeps)*, Neural Network World - monographs edition.

[2] Faber J., Novák M., Svoboda P., Tatarinov V. 2003. Electrical Brain Wave Analysis during

Hypnagogium. *Neural Network World*, Vol. 13 (2003), No. 1, 45-54.

[3] D. Dubois, H. Prade. What are fuzzy rules and how to use them? *Fuzzy Sets and Systems*, Vol. 84 (1996), No. 2, 169-185.

[4] D. Coufal. Representation of Continuous Archimedean Radial Fuzzy Systems. In Proceedings of IFSA 2005 World Congress, Beijing, China, 2005.

[5] L. X. Wang. *A course in Fuzzy Systems and Control*. Prentice Hall, 1997

[6] Hájek P., Feglar T., Rauch J., Coufal D. 2002. The GUHA Method, Data Preprocessing and Mining. Position Paper. In: Database Technologies for Data Mining. Working Notes. - Prague, - 2002, pp. 29-34, A Workshop of the International Conference EDBT 2002, Prague, CZ, 02.03.24

[7] Hájek, P., Havránek, T. 1978. *Mechanizing Hypothesis Formation - Mathematical Foundations for a General Theory*. Berlin - Heidelberg -New York, Springer -Verlag, 396 p. The Internet free version is available at http://www.cs.cas.cz/ hajek/guhabook

[8] Harmancová D., Holena M., Sochorová A. 1999. Overview of the GUHA Method for Automating Knowledge Discovery in Data Sets. *In: Proceedings of the Conference of Knowledge Extraction and Symbolic Data Analysis: KESDA 98*, Luxembourg, 1999, pp. 65-77, LU, 98.04.27-98.04.28

[9] GUHA +– project:
http://www.cs.cas.cz/research/ sotware.shtml

[10] Rauch J., Mining for Scientific Hypotheses. 2002. *In Meij, (Editor): Dealing with the data flood. Mining Data, Text and Multimedia*. STT/Beweton, The Hague. pp. 73-84. Internet version of software is available at http://lispminer.vse.cz