A MEDICAL DECISION SUPPORT SYSTEM WITH UNCERTAINTY: A CASE STUDY FOR EPILEPSY CLASSIFICATION

Emmanuil Marakakis*, Kostas Vassilakis*, Emmanuil Kalivianakis* and Sifis Micheloyiannis**

*Science Department, Technological Educational Institute of Crete, Heraklion, Crete, Greece
**School of Medicine, University of Crete, Heraklion, Crete, Greece

mmarak@cs.teicrete.gr, kostas@cs.teicrete.gr, mkaliv@venus.cs.teicrete.gr, mixelogj@med.uoc.gr

Abstract: This paper presents a medical decision support system which derives differential diagnosis about epilepsy cases in childhood, according to international classifications. This system simulates the reasoning process of top neurologists specialized in epilepsy. It is intended to be used by neurologists specialized on epilepsy. Meta-rules drive the reasoning process of the system. Uncertainty is handled by using weights and certainty factors. The user of the system, physician specialist, can update the knowledge base of the system through a user friendly interface. The initial evaluation results of the system have given 83.3% successful diagnosis.

Introduction

Epilepsy is a chronic disease characterized from recurrent seizures that cause sudden but revertible changes in the brain operation. Classification of epilepsy cases, according to international classifications, during childhood is difficult because individual laboratory findings and symptoms are often inconclusive.

This paper presents the development of a *Medical Decision Support System (MDSS)* which derives differential diagnosis about epilepsy cases in childhood. The diagnosis of epilepsies by our system follows the classification of epileptic syndromes and epilepsies of International League Against Epilepsy (ILAE) [4]. Our expert system is called HIPPOCRAT-EES (HIPPOCRATes Epilepsy Expert System - Hippocrates had made the first observations on Epilepsy). This system is intended to be used by neurologists specialized on epilepsy domain as a consultation system. The reasoning mechanisms of the system simulate the reasoning process of top neurologists. The system's inference engine applies uncertain reasoning by using certainty factors and weights, suggested by expert physicians. Meta-knowledge guides the reasoning process of the HIPPOCRAT-EES. The interface of the system is window-based and user-friendly. The working environment is proportional to the real clinical environment of neurologist. In addition, the interface for updating the knowledge base takes into consideration the computing capabilities of neurologists. The following operations are supported by this MDSS. First, the physician fills in the symptoms and the lab findings of the patient. HIPPOCRAT-EES responses with a list of possible epilepsy types (differential diagnosis) sorted in increasing order based on certainty factors. Second, the doctor specialist can update the knowledge base from a window-based, pop-up menu driven interface. Finally, the physician can have graphical display of the rules of the system. This feature helps the user to understand the decision process of the system and eventually update the rules.

This paper is organized as follows. First, materials and methods are presented. Next, the results of our system are presented. After that, our system is discussed with respect to related work. Finally, conclusions are discussed.

Materials and Methods

The Epilepsy

Epilepsy is a chronic illness characterized from recurrent seizures that cause sudden but revertible changes in the brain operation [1]. According to ILAE there is a fundamental difference between seizures and epilepsies. Epilepsy is a chronic disorder while epileptic seizure is an acute phenomenon that has a beginning and an end. ILAE publishes periodically different classifications for seizures, epilepsies, and epileptic syndromes [4], [5]. Causation of epilepsy is not always straightforward. Epilepsy in most cases is generated by acquired damages of brain cortex (symptomatic). However, brain predisposition to seizures is also a good reason for the appearance of the disease (idiopathic). The last international classification of ILAE [5] includes four main classes of epilepsies, namely:

1. Localization related (focal, local, partial) epilepsies.
2. Generalized epilepsies.
3. Epilepsies and syndromes undetermined as to whether they are focal or generalized.
4. Special syndromes.

Prevalent rates of epilepsy have been reported from many countries, but there are significant differences among various studies. Nonetheless, most studies point a prevalence of 0.5% in the general population [1]. It is estimated that in England there are more than 300,000 with active epilepsy and over than 1,000,000 people with a history of seizures. The prevalence of epilepsy is bigger in children and elderly people (over 60), while

IFMBE Proc. 2005 11(1)
ISSN: 1727-1983 © 2005 IFMBE

the incidence is much bigger (75%) in children and in those under 20 years old [3]. It has been estimated that 2-4% of children in Europe and United States before 4 years, have the experience of at least one epileptic seizure usually during fever [8]. Proportionally epilepsy affects more male than female population. Epilepsy may be a life-threatening condition, but there no documentary studies about that.

Diagnosis of epilepsy is achieved based on the type of the epileptic seizure observed. Various clinical data such as motor / physic/ somatosensory symptoms, impairment of consciousness, absence etc can help doctors to define the seizure type. On a second stage electroencephalograph's (EEG) findings are also interpreted to clarify the seizure type, mainly by figuring out the focus and often they are very helpful for the diagnosis and the differential diagnosis of epilepsies. The epileptic seizure type and the electroencephalographic findings are the main diagnostic criteria used by the classification of epilepsies according to the international classification. Furthermore, very helpful for the diagnosis of epilepsy are the coexisting miscellaneous clinical data, patient's demographics, as well as laboratory findings such as Computerized Tomography (CT) and Magnetic Resonance Imaging (MRI) [3].

Data Analysis and Knowledge Representation

The proposed MDSS exploits the history, the clinical data, the laboratory tests and the EEG findings of the patient. We identified a set of different factors which support knowledge representation and modeling. These factors have been figured out by taking into consideration the diagnostic categories, the types of epilepsies, the epileptic syndromes and ILAE classifications [6], [5]. We divided these factors into the following four groups, motivated by clinical practice.

Patient demographics: This group includes pregnancy and delivery status and family inheritance. Demographic data are assessed using scales ranging from 1 to 5 (where, 1: normal, 2: suspicious, 3: slight pathological, 4: pathological, 5: sever pathological). Zero (0) indicates lack of information (unknown situation).

Clinical data: The type of epileptic seizure is the primary indication for the diagnosis of epilepsy. The seizure should be estimated according to the descriptions in the classification for epileptic seizures, as proposed by ILAE [4]. In this classification, there are 19 different categories. In addition, various findings related to the epileptic seizure observed, should be estimated. These are: the severity (slight, intermediate, severe, status), the focus (frontal, temporal, parietal, occipital, with multiple focuses, unknown) and the diffusion (focal, diffuse, both) of the epileptic seizure.

Apart from the data related to the seizure, other clinical findings should be considered as well. That is, the behaviour, the neurological estimation, the psychomotor development and (if applicable) the school performance of the patient. These are assessed using scales ranging from 1 to 5. Zero (0) again indicates lack of information (unknown situation).

Laboratory findings: Laboratory findings encompass MRI and/or CT readings. Readings are summarized using a five point's scale in proportion to the seriousness, while localization takes one of the following values: focal, diffuse, both and not traceable.

EEG Finding: EEG findings are very important for the diagnosis of epilepsy and they contain very helpful information for the classification of a patient case. They are sorted in four (4) main categories. Each category is in turn codified in subcategories. These subcategories correspond to specific EEG descriptions related to epilepsy and that could help and guide the doctor in her/his diagnostic procedure. The four main EEG categories are the following:

1. Non-specific abnormal EEG patterns, as widespread intermittent slow abnormalities and bilateral/focal persistent findings (3 sub-categories).
2. Abnormal EEG patterns - epileptic paroxysmal as spikes, polyspikes, sharp waves, spike wave complex, small sharp spikes, polyspikes of multiple spikes etc (12 sub-categories).
3. Specific EEG patterns, as hypsarrhythmia, rolantic spikes, typical neonatal EEGs etc (26 sub-categories in total).
4. Epileptic EEG findings in specific recordings as during sleep, photo-stimulation, hyperventilation, during seizure, long-time recording, video recording etc (9 sub-categories).

Analyzing the above, significant for the task, data we conclude specific rules that govern the knowledge base of the system. The last procedure in this stage was the detailed theoretical examination of these rules by firing them on sample data, in order to check their correctness.

According to our conceptual design each epilepsy type corresponds to one rule of the knowledge base. ILAE classification suggests more than 50 epilepsy types [5]. Each epilepsy type, in our system, is expressed in terms of 28 diagnostic criteria, according to the expert neurologist. These criteria are the following [13], [14]:

Seizure type, seizure focus, seizure severity, electroencephalograph (EEG) type, patient age, pregnancy status, delivery status, family inheritance, school performance, lab findings existence, lab finding focus, behavior estimation, neurological estimation, psychomotor development, vocalization during seizure, fever, number of seizures per day, seizure during/after sleep, metabolic symptoms, toxic poisoning, head injury, existence of disease that affects the nervous system, existence of disease that does not affect the nervous system, acquired aphasia symptoms, primary visual ictal seizure, speech problems during seizure, behavioural problems during seizures.

There are 19 different seizure types [4] and about 41 different EEG types [13], [14]. The general form of the if-then rules is as follows:

ISSN: 1727-1983

If (Seizure_type ϵ ST_set) ∧
(Seizure_focus ϵ SF_set) ∧
(Seizure_severity ϵ SS_set) ∧
(EEG ϵ EEG_set) ∧
(Age ϵ Age_set) ∧
(Pregnancy_status ϵ PS_set) ∧
(Delivery_status ϵ DS_set) ∧
(Family_inheritance ϵ FI_set) ∧
(School_performance ϵ SP_set) ∧
(Lab_findings ϵ LF_set) ∧
(Behavior ϵ Be_set) ∧
(Neurological_estimation ϵ NE_set) ∧
(Phychomotor_development ϵ PD_set) ∧
(Lab_focal ϵ LFoc_set) ∧
(Vocalization ϵ yes_no_set) ∧
(Seizures_Per_Day ϵ SPD_data_set) ∧
(Seizure_During_Sleep ϵ yes_no_set) ∧
(Seizure_After_Sleep ϵ yes_no_set) ∧
(Seizure_Fever ϵ yes_no_set) ∧
(Metavolic_Symptoms ϵ yes_no_set) ∧
(Toxic_Poisoning ϵ yes_no_set) ∧
(Head_Injury ϵ yes_no_set) ∧
(Neurous_System_Disease ϵ yes_no_set) ∧
(NOT_Neurous_System_Disease ϵ yes_no_set) ∧
(Acquired_Aphasia ϵ yes_no_set) ∧
(Primary_Visual_Ictal_Seizure ϵ yes_no_set) ∧
(Speech_Problems ϵ yes_no_set) ∧
(Behavior_Problems_During_Seizure ϵ yes_no_set)
then Epilepsy_type = *ET*

where, ∧ stands for *and*. All the elements of the sets ST_set, SF_set, …, LF_set are pairs of the form (value, weight). The elements of the yes_no_set are pairs of the forms (yes, weight) and (no, weight). The elements of the set SPD_data_set have the form ((minimum_value, maximum_value), weight). *ET* has the form (epilepsy type, certainty factor). The certainty factor is derived from the weights of the diagnostic criteria.

The rules of our system have been implemented in SICStus Prolog [12]. Each rule of HIPPOCRAT-EES has been implemented as a Prolog fact and as a Prolog rule. This implementation allows the dynamic update of the knowledge base of the system by using the meta-programming features of Prolog.

Reasoning under Uncertainty

The term uncertainty means non-availability of accurate information in decision making. The main sources of uncertainty in problem solving are due to imprecise data, incomplete data, and subjective description of knowledge. The existence of uncertain knowledge requires the development of reasoning mechanisms which will handle this type of knowledge. Several techniques have been proposed to handle uncertain knowledge like certainty factors, fuzzy logic and others [2], [7]. Experts use inexact reasoning methods because exact methods either may not be known or may be impractical. Inexact methods of reasoning are important in many expert systems applications. Correct medical diagnosis is possible to be derived from ambiguous symptoms.

Certainty Factors (CF) in our system are arithmetic values in the interval [0,1]. They express the expert's belief for the truth of the derived epilepsy type. Each value represents the degree of truth of the derived epilepsy type. For example, the values 1, 0.5 and 0 stand for absolute certainty, 50% certainty and 0% certainty, i.e. absolute uncertainty, for the truth of the derived epilepsy type. Certainty factors appear only in the conclusions of the rules of HIPOCRAT-EES.

Weights are assigned to each sub expression in the premise of each rule. That is, each of the 28 diagnostic criteria is assigned a weight. The weights are values in the interval [0, 1]. Each weight represents a percentage of the certainty factor of the derived epilepsy type. The weight of each of the 28 diagnostic criteria contributes to the certainty factor, truth, of the conclusion of the rule. Let's assume that a diagnostic criterion has weight *w*. This means that the weight of this criterion for the truth of the epilepsy type is *w*. For example, weight 1 or 0.5 or 0.03 etc in a value of a diagnostic criterion in the premise of a rule means contribution to the certainty factor of the epilepsy type in the conclusion of the rule by 1 or by 0.5 or by 0.03 etc respectively.

The certainty factor CF of the conclusion of each rule is the summation of weights of sub expressions in rule hypothesis. The summation of weights expresses the expert's belief for the truth of the derived epilepsy type. If we see a rule as a tree then the certainty factor of the rule corresponds to the total weight of the tree [11]. The summation of the weights in the premise of each rule must be ≤ 1. This is verified by the system during construction of the rule.

Diagnosis Directed by Meta-rules

The clinical and laboratory values for a specific patient are inserted into the system. The system fires rules in order to derive possible epilepsy types and the corresponding certainty factors. There are four meta-rules which guide the rule selection. These meta-rules drive the diagnosis process by following different diagnostic paths. The design of these meta-rules depends on four classes of important clinical and laboratory data which direct the diagnosis towards

ISSN: 1727-1983

specific classes of epilepsy types. The meta-rules are the following.

1. *Seizure type meta-rule.* The seizure type is an important clinical criterion. This criterion results in the selection and firing of the rules that have in their premises a sub expression with same seizure type as the one of the patient case.
2. *The EEG type and seizure focus meta-rule.* Sometimes EEG focus is not same as the seizure focus. The estimation of seizure focus is derived subjectively by the physician. On the other hand, EEG focus is derived mechanically. When a neurologist has observed a seizure focus, he assumes it as important factor for the diagnosis. If he doesn't observe any focus, then he assumes seizure as generalized. In this case, EEG results are critical for the diagnosis especially if they show focus acceptable by the physician. Therefore, EEG data and the seizure's focus are combined to formulate this meta-rule. This meta-rule specifies the important diagnosis criterion for certain cases of patients. Then, rules are selected and fired based on this criterion.
3. *Important electroencephalograph (EEG) types meta-rule.* There are some specific EEG types which affect very much , i.e. the derived epilepsy type, the final conclusion. For example, the observation of "rolantic spikes" in EEG directs diagnosis towards to "localization-related, idiopathic, benign childhood epilepsy, with centro-temporal spikes". In most such cases no other data are required in the diagnosis process.
4. *Other important symptoms meta-rule.* This meta-rule uses a list of secondary symptoms, e.g. vocalization during seizure, seizures per day, seizure during sleep, seizure after sleep, fever before and during seizure, metabolic symptoms etc, in order to select appropriate rules. One or more of these symptoms may have occurred in a patient case.

An Example

The following notation is used in the presentation of this example. Variables start with capital letter. Values, i.e. constants, either start with a lower case letter or they are in double quotes. The rule for the epilepsy type named "localization-related, idiopathic, benign childhood epilepsy with centro-temporal spike" is as follows:

If (Seizure_type ϵ {("simple partial with motor symptoms", 0.1), ("partial seizures secondary generalized", 0.1), (anything, 0.0)}) ∧
(Seizure_focus ϵ {(temporal, 0.05), (parietal, 0.05), (frontal, 0.05), (anything, 0.0)}) ∧
(Seizure_severity ϵ {(slight, 0.05), (intermediary, 0.05), (anything, 0.0)}) ∧
(EEG ϵ {(" rolantic spikes (typical or atypical)", 0.7), ("rolantic spikes - in sleep only", 0.7), (anything, 0.0)}) ∧
(Age ϵ {(4, 0.04), (5, 0.04), (6, 0.04), (7, 0.04), (8, 0.04)}) ∧
(Family_inheritance ϵ {(indications, 0.03), ("severe indications", 0.03)}) ∧
(Lab_findings ϵ {(normal, 0.03), (anything, 0.0)})
then Epilepsy_type = **(**"localization-related, idiopathic, benign childhood epilepsy with Centro-temporal spike"**, CF)**

Let's assume that the clinical and laboratory data values of a patient case are as follows:

Findings-symptoms	**Values**
Seizure Type	simple partial with motor symptoms
Seizure Focus	temporal
Seizure Severity	intermediary
EEG Type	rolantic spikes - in sleep only
Age	6
Family Inheritance	severe indications
Lab Findings	normal

Based on the above clinical and laboratory data of the patient the system fires among others the rule for epilepsy type "localization-related, idiopathic, benign frontal epilepsy". The instance of this rule that is evaluated is the following.

(Seizure_type = "simple partial with motor symptoms")$^{W1=0.1}$ ∧
(Seizure_focus = temporal) $^{W2=0.05}$ ∧
(Seizure_severity = intermediary) $^{W3=0.05}$ ∧
(EEG = "rolantic spikes - in sleep only") $^{W4=0.7}$ ∧
((Age ≥ 4) ∧ (Age ≤ 8) $^{W5=0.04}$) ∧
(Family Inheritance = "severe indications") $^{W7=0.03}$ ∧
(Lab_findings = normal) $^{W6=0.03}$
→ Epilepsy_type = "localization-related, idiopathic, benign childhood epilepsy with centro-temporal spike"^{CF}

The final certainty factor is the summation of the weights of the diagnostic criteria of the hypothesis. That is,

CF =0.1 + 0.05+0.05+0.7+0.04+0.03+0.03 = **1.0**

The diagnostic category named "localization-related, idiopathic, benign childhood epilepsy with centro-temporal spike" will be in the list of epilepsy types suggested by the system with CF 1.0. Other sugested epilepsy types with smaller certainty factors are "localization-related, idiopathic, partial epilepsy with GSSEP" with certainty factor 0.5, "localization-related, symptomatic, simple partial seizures arising from temporal lobes" with certainty factor 0.35 etc.

Illustration of HIPOCRAT-EES through Screen Snapshots

The screen snapshots in figures 1, 2, 3, and 4 illustrate some features of HIPOCRAT-EES. Note that

ISSN: 1727-1983

the interface of the system has been implemented in Visual Basic. The first and second screen snapshots illustrate the forms that have to be filled in by a neurologist for a patient and the system's response from the processing of these data. The third and the fourth screen snapshots illustrate the list of available rules in the KB and their graphical presentation. The weight of each data value is shown by touching the value with the mouse.

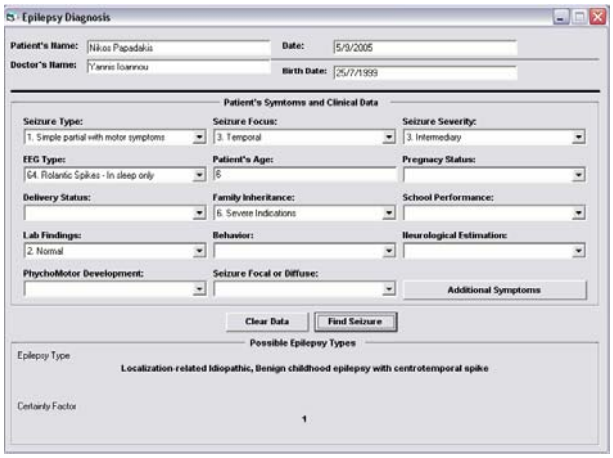

Figure 1: Primary diagnostic criteria and system's response.

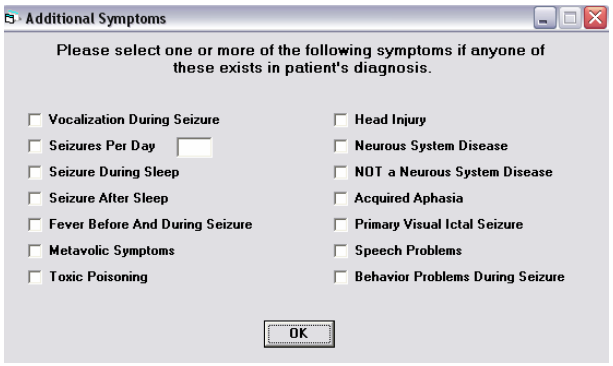

Figure 2: Secondary diagnostic criteria.

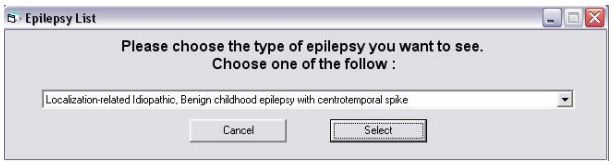

Figure 3: List of available rules in the KB.

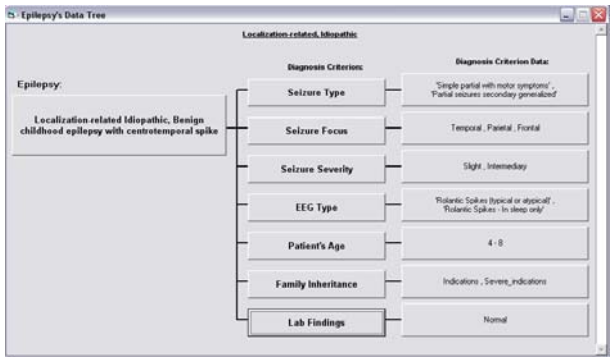

Figure 4: Graphical display of a rule.

Results

The evaluation of our system has shown the following preliminary results. HIPPOCRAT-EES has correctly diagnosed 35 cases out of 42, i.e., 83.3% successful diagnosis. These diagnosed cases had the highest certainty factors in the suggested list of the possible ones. In addition, 3 correctly diagnosed patient cases out of 42, i.e. 7.1%, were in the lists of possible ones. That is, total successful diagnosis 90.4%.

Discussion

The results from the evaluation of the decision support system in [14] are as follows. The system has derived correct diagnosis in 85.2% of patient cases, partial successful diagnosis in 8.2% of patient cases and absolute incorrect diagnosis in 6.6% of patient cases.

The diagnosis of the system in [9] has been compared with the diagnosis of three experts. The evaluation results are 72% correct diagnoses, 8% partially correct diagnoses and 20% incorrect diagnoses. The evaluation of the subsystem that uses electroencephalographs for the diagnosis has given 48% correct diagnoses, 28% partially correct diagnoses and 24% incorrect diagnoses. The subsystem which is based on hypertext has given the best results, i.e. 80% correct diagnoses [10].

HIPPOCRAT-EES preliminary results are satisfactory compared with the results of the other systems in [14] and [9]. These results can be further improved by adjusting the weights in the rule premises. In addition, neurologists can enrich the knowledge base of the system and refine rules which do not produce accurate diagnosis. Another novel feature of our system with respect to the previous expert systems in this domain is the use of uncertainty.

Conclusions

The main new technical features of this system compared to other epilepsy diagnosis expert systems are the following.

1. Use of uncertain reasoning.
2. Graphical presentation of the rule base.
3. The rule base of the system can be updated directly by the doctor specialist and no intervention of the

ISSN: 1727-1983

knowledge engineer is required. This feature makes the system adaptable to new knowledge.
4. The reasoning of the system is directed by meta-rules.

The features of our system that can be further improved and the directions for future research are as follows.

1. The uncertainty reasoning mechanism can be improved by adjusting the weights in some rules. Fuzzy logic can be considered as an alternative technique for uncertainty.
2. A machine learning subsystem can be added into our system. Such a component can derive rules from diagnosed cases of patients.
3. An explanation reasoning subsystem is required as well.

This MDSS is intended to be used as a consultation system by neurologists in order to reach a decision and for differential diagnosis of epilepsy cases.

Acknowledgements

This study is partially supported by:

- BIOPATTERN: Computational Intelligence for biopattern analysis in Support of eHealthcare (EU - 6th FP/IST).
- Archimedes II: Funding of research groups at Technological Educational Institute of Crete (EU-ESF, Operational Programme for Education and Initial Vocational Training).

References

[1] BALLIS T. (1998): 'Epilepsies, in Neurology', LOGOTHETIDES J. (in Greek), (Ed): (University Studio Press, Thessaloniki), pp.577-600.

[2] BUCHANAN B., SHORTLIFFE E. (1984): 'Rule-Based Expert Systems: The Mycin Experiments of the Stanford Heuristics Programming Project', (Addison Wesley).

[3] COCKERELL C. O. , SHORVON S. D. (1996): 'Epilepsy - Current Concepts', (Current Medical Literature).

[4] COMMISSION ON CLASSIFICATION AND TERMINOLOGY OF THE ILAE. (1981): 'Proposal for revised clinical and electroencephalographic classification of epileptic seizures', *Epilepsia*, **22**, pp. 489-501.

[5] COMMISSION ON CLASSIFICATION AND TERMINOLOGY OF THE ILAE. (1989): 'Proposal for revised classification of epilepsies and epileptic syndromes', *Epilepsia*, **30(4)**, pp. 389-399.

[6] DREIFUS F. E. (1989): 'Classification of Epileptic Seizures and the Epilepsies', *Pediatric Clinics of North America*, Vol. 36, No. 2.

[7] GIARRATANO J., RILEY G. (2005): 'Expert Systems: Principles and Programming'. (Thomson Course Technology, 4th Edition).

[8] HAUSER W.A. (1994): 'The prevalence and incidence of convulsive disorders in children', *Epilepsia*, **35** (Suppl.2):S1-S6.

[9] KORPINEN L. (1993): 'Computer-aided decision-making for epilepsy and sleep diagnostics', Doctoral Thesis, Tampere University, *Acta Neurologica Scandinavica*, Vol. 87, No. 144.

[10] KORPINEN L., PIETILA Y., PELTOLA J., NISSILA M., KERANEN T., TOUVINEN T., FALCK B., PETRANEK E., FREY H. (1994): 'Evaluation of Epilepsy Expert a decision support system', *Computer Methods and Programs in Biomedicine*, **45**, pp. 223-231.

[11] ROSS K., WRIGHT C. (1992): 'Discrete Mathematics', (Prentice- Hall, 3rd Edition).

[12] SICSTUS PROLOG. Internet site address: http://www.sics.se/sicstus/.

[13] VASSILAKIS K., MICHELOYANNIS S., E. MARAKAKIS E. (2004): 'Decision support system as a tool to diagnose Epilepsies in children', Proc. of the 1st European Workshop on the Assessment of Diagnostic Performance, Milan, Italy, 7-9 July 2004.

[14] VASSILAKIS K., VORGIA L., S. MICHELOYANNIS S. (2002): 'A Decision Support System for Classification of Epilepsies in Childhood', *Journal of Child Neurology*, **17(5)**, pp. 357-63.