

Research Report AI-1992-02

**Design and Evaluation of a Tutoring Module for  
Computerized Reference System *Maclib* to Assist  
in Differential Diagnosis of Primary Brain Tumors**

Katarzyna J. Macura, Robert T. Macura, Donald E. Nute

Artificial Intelligence Programs

The University of Georgia

Athens, Georgia 30602 U.S.A.

**Design and evaluation of a tutoring module  
for Computerized Reference System *MacLib*  
to assist in differential diagnosis of Primary Brain Tumors.**

Katarzyna J. Macura, Robert T. Macura\*, Donald E. Nute  
Artificial Intelligence Programs      \*Department of Radiology  
The University of Georgia      The Medical College of Georgia  
Athens, GA, 30602      Augusta, GA, 30909

## **Abstract**

*MacLib* is a Computerized Reference System to assist pathologists in the differentiation of brain tumors. The system provides a consulting module with collection of relevant images linked to a reasoning system for a close comparison with the analyzed material, and a reference system with several modules: correlated image and text databases; descriptions of particular diagnoses; differential diagnosis facility for comparing two different diagnostic hypotheses; references to relevant literature on particular diagnoses. This article presents our approach to designing a consulting system and adapting it to a tutoring environment. We concentrate on the knowledge acquisition stage and identification of different classes of users likely to use a system with analysis of their requirements and the types of knowledge they bring to bear on a problem. We explain the background and approach that we have taken to design the inferencing mechanism operating like an automated theorem prover.

# 1 Introduction

Pathology is one of those medical domain where diagnosis involves the integration of visual (macroscopic and microscopic), verbal (discussion with colleagues), and written (patients records, textbooks, articles etc.) information. Reference to the prior experience of other pathologist is an essential part of the diagnostic process. But the specialized knowledge of experts is available only for a limited number of diagnostic problems and retrieving necessary information from prior experience encoded in the literature is impaired by one-dimensional access. To practice pathology effectively, the pathologist needs easy and flexible (i.e., via several entries) access to both textual and pictorial information from different sources, a large collection of well documented illustrated cases, and criteria for differentiation among diagnostic categories. Although for several types of tissue abnormalities (especially tumors) the application of quantitative techniques (morphometry, flowcytometry, and image analysis) which are based on measurements of objective and reproducible features of individual cell components (e.g., the number of nuclei or mitoses, the area of nuclei, the amount of DNA in the nuclei) can be a valuable contribution to the diagnostic process, these techniques offer only part of the data necessary to arrive at a final diagnosis (Baak, 1987). The rest is based on a variety of qualitative criteria and depends on a human eye. A high level of diagnostic accuracy requires an educated and continuously practiced eye. Diagnostic categories usually have unclear boundaries and features have fuzzy distinctions that make the judgement highly subjective (Nathwani et al., 1990). Pathologist must be a broad and comprehensive spectrum of diseases in any disease class to sort out the histologic patterns correctly. Most pathologist in a community hospitals have the opportunity to review a limited number of cases, far fewer than pathologists in academic institution and pathologic centers.

Computers have found many applications in pathology not only in quantitative pathology but also in assisting in diagnosis based on qualitative criteria. At present, much attention is focused on the development of expert systems which contain explicit knowledge with reasoning capabilities for diagnostic decision support. Expert systems are intended for use in situations involving professional judgement and uncertainty. Most of the research effort has gone into developing different knowledge representations and inferencing procedures. Rule-based systems have been developed for such small, well-defined domains as diagnosing diseases of the liver (Chang et al., 1984) and diagnosing bone marrow aspirates (Gyde, 1988). In Pathfinder, a diagnostic support system in the field of lymphomas, the computational architecture of the system is based on a hypothetico-deductive approach to diagnosis (a method of sequential diagnosis) (Horvitz et al., 1984; Horvitz et al., 1989). The most recent implementation of Pathfinder is based on probability theory; the previous version used a quasi-probabilistic method similar to the approach used in Internist-1 (Miller et al., 1982) where features had indicated specificity, sensitivity and importance, and these values were used by the inference strategy to generate an ordered list of possible diagnoses. The move to a probabilistic theory was motivated by the precise definition of probability and by significant increases in the diagnostic accuracy of the system (Horvitz et al., 1989). Hence, the quantitative approach of probabilistic inference is inconsistent with the informal, qualitative nature of human reasoning. Building any expert system requires answering some basic

questions about knowledge: what kind of knowledge is involved; how does an expert explore a given problem; how does expert problem solving differ from that of novice; how should the knowledge be represented; how are choices made in an uncertain environment. Diagnostic pathology is a particularly appropriate domain for the study of decision-making because the decisions are made within narrowly defined settings (Nathwani et al., 1990). However, a logically demanding computer system is a hard task-master in the systematic organization of any collection of information, and in representing and manipulating the uncertain relations between features and diseases.

Expert system designers have been criticized for their attempts to simulate the decision skills of human expert (Dreyfus and Dreyfus, 1986) and for their theoretical, technology-oriented approach (Schank, 1991). It has been claimed that since expert systems cannot reach the stage of an advanced beginner, it is unlikely that expert systems will ever be able to deliver expert performance; so they should rather be called *competent systems* (Dreyfus and Dreyfus, 1986).

We have found a promising use for expert systems in the creation of a learning environment. The first attempt to adopt a pre-existing expert system for use in teaching was GUIDON (Clansey, 1987). The underlying expert system was MYCIN (Shortliffe, 1979) the most well-known rule-based expert system. MYCIN's knowledge base is interpreted by GUIDON to provide feedback as the student gathers information about a patient and makes a diagnosis in the domain of infectious diseases (Clansey, 1987). We have designed and implemented a program to assist in learning diagnostic procedures in the pathology of brain tumors. We want the user to figure out how to solve problems in the micro-world and so to learn the conceptual structure of the domain. The micro-world has failed so far as a step toward modeling real-world understanding; but it is useful as a simplified environment in which a beginner can more easily pick out the features he/she needs to recognize and the step-by-step procedures which would eventually become incorporated into subconscious decision making. Primary brain tumors pathology has been used as the prototype domain for this research. The current system reasons about 31 tumors, constructing differential diagnoses through the consideration of evidence of 90 features present in brain tumors pathology.

Besides an expert system which provides a collection of relevant images linked to a reasoning system for a close comparison with the analyzed material, the learning environment includes a reference system for pathology with several modules:

1. correlated image and text databases;
2. descriptions of particular diagnoses;
3. differential diagnosis facility for comparing two different diagnostic hypotheses;
4. references to relevant literature on particular diagnoses.

We wanted to make available at the pathologist's desk, in the form of a Computerized Reference System (CRS), all sources of external knowledge which the pathologist uses in daily

practice. In designing the CRS for pathology we built a modular system composed of six units:

**CONSULTING SYSTEM** - an advisory module which can function as a training system and as a second opinion system with access to help files in the form of microscopic slides and textual descriptions and with the possibility to ask why a certain question is being asked and what is the importance of the questions to the final diagnosis.

**LEXICON** - a textual module providing detailed descriptions of particular tumors (type, occurrence, clinical features, macroscopic features, microscopic features, ultrastructural features, immunohistochemical data, differential diagnosis — other brain tumors which can be potential sources of misdiagnosis, grading, and prognosis).

**ATLAS** - a computerized atlas composed of digitized histologic and ultrastructural images of tumors.

**DIFFERENTIAL DIAGNOSIS** - a module providing information about commonalities or differences of pairs of tumor types selected by the user.

**REFERENCES** - a module providing bibliographic information about the sources of data stored in the system.

**QUIZ** - a computer-based quiz module including multiple-choice questions and simulations of the diagnostic process based on the analysis of histopathological and ultrastructural slides.

In this paper we present our approach to designing a consulting system (expert system in AI terminology) and adapting it to a tutoring environment.

## 2 The problem area

1 to 2% of all malignant tumors originate in the central nervous system (CNS) (Colmant and Noltenius, 1988). In childhood, brain tumors are the second most common group of tumors, surpassed in frequency only by leukemia (Percy et al., 1972). Despite major advances in neuroimaging (CT and MRI), tissue diagnosis of brain tumors is the basis for predicting outcome and determining treatment (Becker and Halliday, 1987). The modern WHO classification of brain tumors is based on histogenetic classification of tumors by Bailey and Cushing (1926) (Zulch, 1981). CNS tumors are classified according to histology, presuming that differing histology indicates different pathologic, biologic, kinetic, and metabolic growth characteristics. A grading system of brain tumors was suggested by Kernohan and Sayre (Kernohan and Sarye, 1952). Grade I is considered to be most benign and grades II, III, and IV indicate increasing degrees of malignancy. Although tumors have a heterogeneous

composition, it is usually limited, allowing a histological diagnosis to be confirmed on the basis of major cell type (Zulch, 1979). The histological examination of brain tumors provides information about architectural pattern and morphological characteristics of tumor cell populations, and biological behavior of tumors. Electron Microscopy makes a valuable contribution in diagnosing Schwannoma, demonstrating cilia, cell junctions, and intracytoplasmic filaments in ependymoma, differential diagnosis of small-cell CNS tumors, poorly differentiated gliomas and meningeal tumors (Becker and Halliday, 1987; Scheithauer and Bruner, 1987). Immunocytochemistry finds applications in the demonstration of cytoplasmic and cell surface determinants that serve as glial, neuronal or other markers of differentiation (Bonin and Rubinstein, 1984). The most frequently used of the immunohistochemical markers is GFAP (glial fibrillary acid protein) which enables differential diagnosis between glial and non-glial tumors, especially when the morphological features are equivocal and when traditional staining methods have been incapable of firmly establishing the nature of a neoplasm (Russell and Rubinstein, 1989). The expression of neurofilament protein subunits in tumor cells helps to establish the neuronal origin or the neuronal differentiation of cells (Russell and Rubinstein, 1989). Some brain tumors have a tendency to appear at a certain site. Some brain tumors occur only in certain age groups. To a certain degree there are associations between the site of the tumor, the age of the patient, and the histological subtype within one tumor type. In children, most gliomas are of low grade (I–II) whereas in adults high grade lesions (III–IV) are more common. Four brain tumors make up 88% of CNS tumors in childhood: astrocytoma, ependymoma, craniopharyngioma, and PNET (primitive neuroepithelial tumors including neuroblastoma, medulloblastoma, ependymblastoma, pineoblastoma) (Becker and Halliday, 1987). In adults the most common tumors are gliomas (astrocytoma, anaplastic astrocytoma, glioblastoma multiforme, and oligodendroglioma) and meningioma (Codd and Kurland, 1985). The anatomical areas most frequently involved in childhood CNS tumors lie within the posterior fossa (cerebellum, brain stem) whereas the majority of gliomas in adults arise within the cerebral hemispheres, a site rather more unusual in children (Tobias and Hayward, 1990). Low-grade tumors are distinguished histologically from higher grade tumors by the absence of certain features, especially cellular pleomorphism, high cell density, mitotic activity, necrosis and endothelial proliferation, but these features are not absolute indicators of malignancy (Becker and Halliday, 1987). Prominent cellular pleomorphism does not always indicate a histologically malignant neoplasm; e.g., pleomorphic xanthoastrocytoma is usually much more benign in behavior than might be expected from its pleomorphic appearance (Burger and Fuller, 1991). Endothelial proliferation, whose presence is an ominous sign in astrocytoma, is an expected pathologic finding in pilocytic astrocytoma, benign, slowly growing tumor (Rekate and Rakfal, 1991). The presence of mitoses in oligodendroglioma does not correspond to higher grade as it does in other gliomas; e.g., astrocytoma or ependymoma. Accuracy in diagnosing brain tumors is important for the purposes of prognostication and therapy. The lack of precisely defined entities in the classification of brain tumors makes differential diagnosis difficult and can be the source of overdiagnosis of malignancy, especially in pediatric neoplasms of CNS (Burger and Fuller, 1991). Tumors of central neuroepithelial origin constitute the most important and numerous (40–45 %) group of intracranial tumors. Their complexity is due to the considerable number of cell types involved in neoplasia, and in many cases to interpretative difficulties (Russell

and Rubinstein, 1989). The main group of tumors considered by our system arise from cells derived from the neuroepithelium. The second group is composed of meningioma, a tumor which forms 13.4 to 19.2 % of all primary brain tumors and is the second most common brain tumor of adults (Russell and Rubinstein, 1989). The third group consists of tumors which are frequently the source of misdiagnosis and should be distinguished from neuroepithelial tumors and meningioma through differential diagnosis. Malformative tumors, lymphomas, tumors of the pituitary and metastatic tumors are not considered by the prototype version of the system.

### 3 Design considerations

The maxim: ‘a picture can speak a thousand words’ is very pertinent for pathology. Memory capacity for visual information is much greater than for verbal information (Anderson, 1990). The knowledge of an experienced pathologist is composed of thousands of images and the cognitive basis is *holistic* (Dreyfus and Dreyfus, 1986) and *contextual* (Anderson, 1990) image recognition. An expert pathologist has difficulties in verbalizing the way he/she arrives at conclusions. He/she does not remember the exact visual details or spatial relations in a picture but rather remembers an abstract representation that captures the picture’s meaning and recognizes thousands of special cases he/she has analyzed before. A training pathologist seems to form diagnostic hypothesis and draw conclusions from sets of relevant features defining disease categories. After enough experience with pathological analysis with a particular condition, the image pattern associated with that condition is stored in memory, and when a similar pattern is seen, the image memory is triggered and the diagnosis comes to mind. There is no decomposition of the image pattern into features and no need for rules associating conditions with features. Configurations of elements through repeated exposure comes to be recognized as single units or *chunks* (Miller, 1956). These perceptual patterns directly index the expert’s knowledge store (Slatter, 1987). High-level general knowledge determines the interpretation of the low-level perceptual units (Anderson, 1990). Pathologists call diagnosis making *pattern-seeing*.

If symbolic computation is to be used to model human problem solving skills we were forced to decompose knowledge (images) into features and relationships. Asking the expert for rules we forced him to regress to the level of a beginner who makes inferences using facts and rules. The proceduralization of knowledge and the automation of cognitive skills that accompany the development of expertise, serve to make expert thinking less accessible to introspection as well as to exploration by the knowledge engineer. Since experts seem to use *holistic similarity recognition* (Dreyfus and Dreyfus, 1986), imitating them would mean duplicating that pattern recognition process rather than using the typical expert system approach. Expert systems should not be expected to perform as well as human experts, nor should they be seen as simulations of human expert thinking.

Being aware of the limitations of expert systems we decided to exploit another expert sys-

tems capabilities. We planned a tutor which would teach knowledge that was prerequisite to learning diagnostic skills in pathology and would facilitate the users' mastery of basic skills before they would learn advanced skills themselves. To find concepts for designing a consulting module which would make the transition from the tutoring environment to learning on one's own smoother and easier we analyzed the diagnostic process in pathology taking into consideration the different levels of experience of pathologists.

We presented to a trainee pathologist, a general pathologist and an expert neuropathologist sample cases of brain tumors and instructed them to think out loud while performing a diagnostic task. They reported sets of features they observed and their diagnostic implications. We asked pathologists their opinions about the importance of a particular feature to the diagnosis. Each of the cases was a source of three different sets of features. We asked the expert to give his diagnostic suggestion for a certain combination of features. Sets of possible diagnoses for a particular set of features were established by the expert using a percentage scale. In this way a diagnostic minimum and maximum (the smallest and the biggest set of features defining a disease category) was provided by the users for each diagnostic category. We analyzed diagnostic models formed by the trainee pathologist, the general pathologist and the expert. The simplified model of the trainee pathologist was characterized by the smallest sets of features and the biggest sets of diagnostic hypotheses. She recognized: tumor cellularity, nuclear morphology, and nuclear/cytoplasmic ratio, but she lacked some concepts for diagnostic categories and demonstrated some difficulties with cell structure recognition and tumor cells origination. The intermediate model of general pathologists was characterized by larger sets observed features and smaller sets of diagnostic hypotheses than was the case for the novice. They recognized tumor cells origination, and tumor grade, but they lacked concepts for tumor subtypes for some tumor types. The expert's diagnostic model was characterized by the biggest sets of features and the smallest sets of diagnostic hypotheses (usually a single diagnosis was suggested). We presumed that the set of features observed by the expert formed a disease category. The diagnosis making technique of the expert was different from the techniques the expert used to train the students, as had been reported in other studies (Elstein et al., 1978). The trainee pathologist collected data first (the way she was taught during her studies) and then formulated diagnostic hypothesis, whereas the experienced pathologist recognized critical data and formulated diagnostic hypothesis and then collected additional data to confirm or reject hypothesis. Experts collect the most critical data early (Wortman, 1972; Elstein et al., 1972, Eistein et al., 1978). Early hypothesis formation is important in directing problem-solving activities, so that data gathering is directed (not inductive) (Balla, 1985). During diagnosis making the expert asked questions directed at testing specific hypotheses. The expert modified his questions to fit the diagnostic problems while the novice tended to ask a firm set of questions in constant order in a manner she was taught during her studies. A similar observation was reported by de Donbal (de Donbal, 1978). The heuristic (goal oriented) search is a technique employed by the expert when dealing with familiar problems. To use it one must understand the structure of the problem and know the likely outcomes. Information gathering is oriented in such a way that the data can be used to reach the chosen goal. The expert has a clear understanding of the end goal from the earliest stages (Balla, 1985). The novice requires



sub-goals to direct his/her short-term memory through the problem-solving process. He/she uses the recognition method (recognition by the one or two critical cues) which involves the reduction of a problem to something that is known and recognized. The expert knew what was important and so he was able to recognize and appreciate the significance of critical cues. The expert's information processing depends on collection and translation of present data into terms referable to past experience. The novice was able to make references to theoretical knowledge and to past experience of other pathologists. The experienced pathologist was able to absorb more information on looking at the image than his inexperienced colleague. It is likely that with experience people will be able to remember more associations and find it easier to deal with more relevant factors at the same time. The experienced pathologist had no difficulty in remembering details of a complicated history presented by the patient, whereas his inexperienced colleague needed to make constant reference to external sources of knowledge and to external memory (notes) to supplement short term (working) memory. These observations seemed to confirm that only a limited number of problems can be tackled at the same time, and if the "system" receives too much information, it can be easily overloaded because of the limitations of short term memory and channel capacity (Miller, 1956). What information meant to doctors depended on their understanding of the significance of each individual piece of information, as well as special meaning they gave to seeing all these features in combination. A piece of information was seen by the expert in the context of the situation (other features) which were being assessed. The actual importance of a particular piece of information depended on the circumstances (other features). Certain bits of information had more significance than others. One piece of information was composed of a set of features, not of a single piece of evidence. It was not enough to obtain the data; it was also necessary to judge it (put the data on a scale before attempting to draw inferences from it) in the context in which it was used. The novice had more difficulty in attaching weights to cues than did the expert. The weight attached to a cue referred to the subjective assessment of the significance of the data obtained by the doctor. It is impossible to make rational decisions without correct weighing of cues (Coles et al., 1980). The expert seemed to estimate likelihoods subjectively and intuitively to combine them to arrive at his degree of certainty about the presence of a particular disease. The expert knew prior probabilities, and without realizing, intuitively was doing "statistics". To express probability he used qualitative expressions: certain, almost certain, probable, improbable, likely, unlikely, common, frequent, rare, never, possible. We observed, as did Kong et al. (Kong et al., 1986) that even when the pathologist knew the numbers, he used qualitative expressions and he felt more comfortable with them than with exact values. As far as diagnostic certainty was concerned, the expert never achieved a 100% level of confidence in his judgement; he always left room for doubt.

Our conclusion derived from an analysis of the diagnostic process is that to design a consulting (tutoring) module as an expert system the main problems to be solved are:

1. identify a set of consensus features and diseases;
2. design of hierarchical structure of information and develop a consensus about the structure of the knowledge base (diseases, features, and their values);

3. acquire probabilistic dependencies among features and diseases and achieve a grasp of the probabilities and of the practical significance of some diagnoses;
4. gather data in ways that recognize critical cues (that will useful data) for specific choices with strong distinguishing values, assess the significance of the data, and attach appropriate weight to a cue;
5. form early hypothesis about alternative choices;
6. assess the reliability of data and eliminate observer's errors;
7. test hypotheses in manner that recognizes the influence of new information on prior probability;
8. design strategies for selecting questions that reduce the uncertainty in the differential diagnosis quickly;
9. construct a final hypothesis and estimate the level of certainty.

## 4 Knowledge acquisition

At the knowledge acquisition stage we wanted to identify the different classes of users likely to use a system, and to analyze their requirements and the types of knowledge they bring to bear on a problem.

### 1. Decision tables:

As the first stage of knowledge acquisition we decided to use the decision tables. A decision table is a tabular representation of factors to consider in making a decision (conditions), steps to be taken when a certain combination of conditions exists (actions), and specific combinations of conditions and the actions to be taken under those conditions (rules) (Fergus, 1969). We used a primary decision table in the form of encoded possible variation of disease categories and secondary decision tables in the form of sets of features recognized by the trainee pathologist or general pathologist formulating and evaluating diagnostic hypotheses under question. The secondary decision tables were matched to the primary decision table to find possible diagnoses.

#### *Primary decision table:*

We defined series of features divided into 5 sets (clinical, macroscopic, histologic, ultra-structural, immunohistochemical data) and disease categories based on textbooks, patients records and consultation with two experts. The features are each structured into sets of 2 to 11 lists of values. Most of these values are mutually exclusive and exhaustive (e.g., the feature *cellular density* can take on any one of the values *cells back-to-back*, *cells closely packed*,

*diffuse/loose pattern*), but some features can have multiple values at one time (e.g., feature *shape of cells* can take on any combination of the values *polygonal/stellate*, *elongated/spindle*, *bizarre/monstrual*, *round*, *uniform cells*, *polymorphism*). We asked the experts to establish independently sets of features for each diagnosis taking into consideration different varieties of tumor within one tumor type. The experts weighted features for each diagnostic category using scale: mandatory/common, rare or absent. We marked feature combinations in the table.

The primary decision table was used as a simplified matrix of the experts' memory organization about five levels (clinical, macroscopic, histologic, ultrastructural, and immunohistochemical) of disease entities (diagnostic units). See Table 1 through Table 6 that follow:

code	diagnosis
D1	Astrocytoma fibrillary
D2	Astrocytoma pilocytic
D3	Astrocytoma gemistocytic
D4	Subependymal giant cell astrocytoma
D5	Astroblastoma
D6	Pleomorphic xanthoastrocytoma
D7	Anaplastic astrocytoma
D8	Oligodendroglioma
D9	Mixed oligoastrocytoma
D10	Anaplastic oligodendroglioma
D11	Ependymoma
D12	Subependymoma
D13	Ependymoblastoma
D14	Anaplastic ependymoma
D15	Choroid plexus papilloma
D16	Anaplastic choroid plexus papilloma
D17	Pineocytoma
D18	Pineoblastoma
D19	Ganglioneuroma
D20	Ganglioglioma
D21	Neuroblastoma
D22	Glioblastoma multiforme
D23	Giant cell glioblastoma
D24	Gliosarcoma
D25	Medulloblastoma
D26	Germinoma
D27	Capillary hemangioblastoma
D28	Meningioma
D29	Anaplastic meningioma
D30	Schwannoma
D31	Neurofibroma

Table 1: Set of diagnoses.

code	clinical features
f1	Age:
f1.1	<15
f1.2	15–30
f1.3	>30
f2	Tumor location:
f2.1	Meninges
f2.2	Cerebral hemisphere
f2.3	Suprasellar area
f2.4	Pineal region
f2.5	Lateral ventricles
f2.6	III ventricle
f2.7	IV ventricle
f2.8	Cerebellar hemisphere
f2.9	Midline cerebellum
f2.10	Brain stem
f2.11	Cerebellar pontine angle
f3	Phakomatoses

Table 2: Set of clinical features.

code	macroscopic features
f4	Borders:
f4.1	Well delineated
f4.2	Ill defined
f5	Consistency:
f5.1	Soft
f5.2	Firm
f6	Cut surface:
f6.1	Homogenous
f6.2	Nodular/papillary
f6.3	Gelatinous
f6.4	Whitish
f6.5	Yellowish
f6.6	Reddish
f6.7	Grayish
f6.8	Multicolored
f7	Cysts
f8	Necroses
f9	Calcifications
f10	Hemorrhages

Table 3: Set of macroscopic features.

*Secondary decision tables:*

code	histological features
f11	Cellular density:
f11.1	Cells back-to-back
f11.2	Cells closely packed
f11.3	Diffuse, loose pattern
f12	Pattern:
f12.1	Fascicular distribution of cells
f12.2	Perivascular pseudorosettes
f12.3	Rosettes with or without lumen
f12.4	Pseudopalisades around necroses
f12.5	Palisading
f12.6	Tubular pattern
f12.7	Whorl formations
f12.8	Cellular nests
f12.9	Channels of capillary structure
f12.10	Disorganized
f13	Shape of cells:
f13.1	Round
f13.2	Polygonal/stellate
f13.3	Elongated/spindle
f13.4	Bizarre/monstrual
f13.5	Uniform cells
f13.6	Polymorphism
f14	Nuclear morphology:
f14.1	Dense nuclei/scanty cytoplasm
f14.2	Round nuclei/clear cytoplasm/distinct cell borders
f14.3	Vesicular nuclei/prominent nucleoli
f14.4	Oval nuclei/fine sparse chromatin/indistinct nucleoli
f14.5	Oval nuclei/coarse clumped chromatin
f14.6	Polymorphism/multiple nuclei/irregular nuclearmembrane
f14.7	Nuclei with pseudoinclusions
f14.8	Elongated nuclei
f14.9	Nuclei peripherally located (>50% of cells)
f15	Mitoses:
f15.1	Absent
f15.2	Occasionally
f15.3	Numerous/abnormal
f16	Necroses
f17	Vascularization:
f17.1	Thin wall vessels
f17.2	Hyalinized vessels' walls
f17.3	Vascular endothelial proliferation
f18	Microcystic degenerative changes
f19	Calcification/psammoma bodies
f20	Proliferation of collagen and reticulin fibers
f21	Lipid droplets/foamy cytoplasm
f22	PAS - positive material (mucin)
f23	Rosenthal fibres
f24	Foreign cells

Table 4: Set of histological features.

code	ultrastructural features
f25	Glial filaments
f26	Cilia
f27	Junctional complexes
f28	Dense-core vesicles
f29	Microtubules
f30	Synaptic structures
f31	Basal membrane
f32	Blepharoplasts
f33	Axons regular with neurotubules
f34	Axons surrounded by membrane stacks
f35	Axons abnormal, irregular
f36	Luse bodies
f37	Weibel-Palade bodies
f38	Concentric myelin-like bodies
f39	Concentric RES around lipid droplets
f40	Glycogen
f41	Enlargement of mitochondria with atypical crests

Table 5: Set of ultrastructural features.

code	immunohistochemical data
f42	GFAP:
f42.1	Most of cells GFAP-positive
f42.2	Cells GFAP-positive occasionally
f42.3	Areas of cells GFAP-positive and GFAP-negative
f42.4	Negative
f43	Neurofilaments:
f43.1	Positive
f43.2	Negative

Table 6: Set of immunohistochemical data.

Features combinations observed by the trainee pathologist, the general pathologist, and the expert during analysis of different tumor types were organized in the form of secondary decision tables which were then matched to the primary decision table to find possible diagnostic hypotheses for a particular set of features. For sets of features observed by the trainee pathologist and the general pathologist, sets of possible diagnoses were derived from the primary decision table and the order of diagnoses within particular set of diagnoses was established by the expert using a percentage scale. Sets of features observed by the expert were used to verify diagnostic entities.

Secondary decision tables were used as users' models of the perceptual units (features combinations) under consideration. See Table 7 and Table 8.

Decision tables provided:

- features to consider in diagnosing brain tumors,
- diagnostic entries,
- importance of features for particular diagnoses,
- critical cues for particular sets of diagnostic hypotheses,
- correlations between the certain sets of features and each of the members of the sets of diagnostic hypotheses in the form of degrees of certainty (in the percentage scale),
- rules with AND relationship among applicable conditions (set of features), that is that is, for a rule to be satisfied the first applicable condition AND the second AND the third must exist; like wise all of the applicable actions (diagnoses) must be taken,
- rules with OR relationship among conditions (features) within a rule; therefore, whenever one of the conditions is met the rule applies,
- NOT conditions, for the negative condition (absence of a feature).

2. In the next step we wanted to find procedural rules for diagnosis making in the domain of primary brain tumors. We asked the expert to provide examples which were submitted to XiRule (Expertech, Ltd., 1986) to induce chains of rules. We compared discriminative features and chains of rules induced by the computer program from the expert's examples with rules derived from interviews with the expert and analysis of the literature. The order of the features taken into consideration by the computer program was different from the order suggested by the expert. The rationale for questions selected by the computer program often was not acceptable to the expert. Although these questions were undoubtedly the most discriminating for the diseases, they were not natural questions to ask according to the problem solving protocol followed by the human expert. We discussed these differences with the expert and noticed that the human as a decision maker seeks more information than is necessary. We determined the discriminatory powers of the questions (critical cues)



Macroscopic features						
Feature	F1	x		x		x
	F2		x		x	o
	F3			x		
	F4	x	x	x	x	
	F5	x			x	x
Diagnosis	D1	D				
	D2		D			
	D3			D		
	D4				D	
	D5					D

D = presence of a diagnosis, x = common feature, o = rare feature

Table 7: Primary decision table - sets of features per diagnosis.

Case nr.1		
Macroscopic level		
Feature	F1	
	F2	y
	F3	
	F4	y
	F5	y
Diagnosis	D1	10%
	D2	10%
	D3	5%
	D4	70%
	D5	5%

y = presence of a feature

% = degree of certainty of a particular diagnosis for a set of features present in case nr.1

Table 8: Secondary decision table - set of features present in case nr.1.

and reordered the rules according to the expert's suggestions. The first critical features concerned the location of a brain lesion and the patients' age. We asked the expert to determine the order of common outcomes (tumor types) for particular combinations of tumor sites and patient's age based on prior probabilities derived from his experience. We did not performed statistical analysis of data to compare with expert's prior probability because of lack of a big enough database. Research has shown that using the expert as the sole source of knowledge can lead to sub-optimum results (Shortliffe and Buchanan, 1984; Michalski and Chilausky, 1980).

3. The first-pass knowledge base consisted of factual and procedural rules derived from an examination of the decision tables and the expert's examples. It was far easier for the expert and other users to criticize a working version of a program than to describe how it should behave. The users could comment on what they needed at each step of diagnosis making. We used fast prototyping as a method for eliciting the expert's knowledge and the user's requirements. Using the initial prototype in later knowledge acquisition enabled early detection of mistakes and easier correction of the primary decision table. Rapid prototyping to elicit feedback was a successful method for acquiring knowledge and debugging the knowledge base. We used the expert system shell XiPlus (Expertech, Ltd., 1986) not only as a knowledge acquisition tool but also as a tool for implementation of the prototype version since this offered a productive development environment for easy rule construction and modification, supported forward and backward chaining, and had built "how" and "why" functions into the inference engine. It also enabled us to check for circular arguments, to check for redundancy or duplication and to trace the system as it ran. We tested the system gradually in four steps because the number of possible routes through the final program could be extremely large and difficult to test exhaustively in complete system. A set of hypothetical problems (test cases) with common and unusual features was run through the system and the system's diagnostic suggestions were compared with the expert's conclusions. The diagnostic accuracy for the first-pass knowledge base was 17/31 (54.8%). We decided to add more structures to the interview process in an attempt to improve the knowledge base.

4. We used event recall interviews to induce case-based reasoning, and limited information tasks and diagnosis making under constraint (time and feature limitations) to learn about the importance of particular chunks of information and to acquire rules of thumb used by the expert in uncertain circumstances.

5. The second-pass knowledge base was used to check a modified model of reasoning. Sample cases were used for debugging and to evaluate system performance. Testing revealed diagnostic accuracy in 34/50 (68%) of the sample cases studied.

6. The third-pass knowledge base was tested on 30 cases used to verify the first and the second-pass knowledge bases which were misdiagnosed, and on 20 cases described in the literature. The diagnostic accuracy achieved was 41/50 (82%).

7. In the next stage, testing of the knowledge base was performed by the general pathologist

and the pathologist in training on 110 cases. 79 cases from the Department of Oncology, Medical University of Lodz (Poland) were diagnosed based on glass slides and electron micrographs which provided data for histologic and ultrastructural diagnosis. Macroscopic and immunohistochemic data were not tested for these cases. Additional 31 sample cases (one for each diagnostic category) provided data for all levels of diagnosis. The results were that the general pathologist with the assistance of the system achieved 91.1% (72/79) of diagnostic accuracy for clinical cases and 91.8% (101/110) for all (clinical and sample) cases tested. The pathologist in training achieved 69.6% (55/79) of diagnostic accuracy for clinical cases and 71.8% (79/110) of diagnostic accuracy for all cases.

## 5 Knowledge organization

A central problem-solving task within brain tumors pathology is the classification of sets of features into disease categories. The brain tumors classification proposed by WHO offers diagnostic entities based on cell type (histological features). See Table 9.

One of the most important factors is the prior probability of a disease being present with different frequencies depending on the location and patient's age, the first discriminating features in diagnosing brain tumors. The significance of the next diagnostic cues varies depending on the initial set of values. To design structural hierarchies of information in the reasoning system we classified tumors considered by the system using location of the tumor lesion and the patient's age as the initial critical cues (clinical features). See Table 10.

## 6 Knowledge representation

In designing the consulting module, we decided not to imitate the expert's diagnostic process (using some estimates of how likely or how uncertain various conclusions might be) and not to present the expert's diagnostic process directly to the trainee. Instead, the system incorporates the expert's memory organization in simplified form as a store of possible feature combinations defining disease categories (sets of diagnostic units). We matched the sets of perceptual units (features observed by the users), provided to the system as answers to questions, with sets of diagnostic units. A rule-based approach seemed appropriate for this goal. Since the expert describing tumor types used sentences similar to rules ("If ... then ..."), and real cases with established diagnoses used to verify the diagnostic definitions in the decision table also corresponded to rules, we presumed that brain tumors could be defined and categorized by rules. Rule-based formalism was also used for encoding knowledge in GUIDON (Clansey, 1987). Choosing this type of knowledge representation, we realized that we would have to predict all diagnostic situations that can occur. Usually a set of features make up a unit of information rather than a single feature. Frequently each of two or more symptoms independently has only a very low correlation with a disease state, but the group

Table 9: WHO brain tumors classification  
(limited to tumors considered by the system)

---

Tumors of neuroepithelial tissue
Glial tumors
Astrocytic tumors
Astrocytoma fibrillar
Astrocytoma pilocytic
Astrocytoma gemistocytic
Subependymal giant cell astrocytoma
Astroblastoma
Pleomorphic xanthoastrocytoma
Anaplastic astrocytoma
Oligodendroglial tumors
Oligodendroglioma
Mixed oligoastrocytoma
Anaplastic oligodendroglioma
Ependymal tumors
Ependymoma
Subependymoma
Ependymoblastoma
Anaplastic ependymoma
Choroid plexus tumors
choroid plexus papilloma
Anaplastic choroid plexus papilloma
Pineal tumors
Pineocytoma
Pineoblastoma
Neuronal tumors
Ganglion cell tumors
Ganglioneuroma
Ganglioglioma
Neuroblastoma
Poorly differentiated tumors
Glioblastoma
Glioblastoma multiforme
Giant cell glioblastoma
Gliosarcoma
Medulloblastoma
Germ cell tumors
Germinoma
Tumors of meningeal tissues
Meningioma
Anaplastic meningioma
Tumors of nerve sheath cells
Schwannoma
Neurofibroma
Tumors of blood vessel origin
Capillary hemangioblastoma

---

Table 10: Tumor type according to location and age (Part 1)  
 (the most common tumors in particularly sites ordered by prior probability)

---

Meninges	
	Meningioma
	Anaplastic meningioma
Cerebral hemispheres	
Adults	
	Glioblastoma multiforme
	Anaplastic astrocytoma
	Astrocytoma fibrillary
	Astrocytoma gemistocytic
	Oligodendroglioma
	Mixed oligoastrocytoma
	Anaplastic oligodendroglioma
	Astroblastoma
	Gliosarcoma
	Ganglioneuroma
	Ganglioglioma
Adolescents	
	Astrocytoma fibrillary
	Mixed oligoastrocytoma
	Ependymoma
	Anaplastic ependymoma
	Pleomorphic xanthoastrocytoma
	Giant cell glioblastoma
	Oligodendroglioma
	Gangliocytoma
	Ganglioglioma
Children	
	Astrocytoma fibrillary
	Ependymoma
	Mixed oligoastrocytoma
	Anaplastic astrocytoma
	Pleomorphic xanthoastrocytoma
	Giant cell glioblastoma
	Neuroblastoma
	Ependymoblastoma
	Astroblastoma
	Oligodendroglioma
	Gangliocytoma
	Ganglioglioma

---

Table 10: (Part 2)

---

Suprasellar area (hypothalamus/optic chiasm)	
Adults	<ul style="list-style-type: none"> <li>Meningioma</li> <li>Anaplastic meningioma</li> <li>Astrocytoma fibrillary</li> <li>Glioblastoma multiforme</li> <li>Schwannoma</li> <li>Neurofibroma</li> </ul>
Adolescents	<ul style="list-style-type: none"> <li>Astrocytoma pilocytic</li> <li>Germinoma</li> </ul>
Children	<ul style="list-style-type: none"> <li>Astrocytoma pilocytic</li> <li>Germinoma</li> </ul>
Pineal region	
Adults	<ul style="list-style-type: none"> <li>Pineocytoma</li> <li>Astrocytoma fibrillary</li> <li>Subependymoma</li> <li>Germinoma</li> </ul>
Adolescents	<ul style="list-style-type: none"> <li>Germinoma</li> <li>Pineocytoma</li> <li>Pineoblastoma</li> <li>Astrocytoma fibrillary</li> </ul>
Children	<ul style="list-style-type: none"> <li>Germinoma</li> <li>Pineoblastoma</li> <li>Pineocytoma</li> <li>Astrocytoma fibrillary</li> </ul>
Lateral ventricles	
Adults	<ul style="list-style-type: none"> <li>Ependymoma</li> <li>Subependymal giant cell astrocytoma</li> <li>Subependymoma</li> <li>Oligodendroglioma</li> </ul>
Adolescents	<ul style="list-style-type: none"> <li>Ependymoma</li> <li>Subependymal giant cell astrocytoma</li> <li>Choroid plexus papilloma</li> </ul>
Children	<ul style="list-style-type: none"> <li>Ependymoma</li> <li>Choroid plexus papilloma</li> <li>Anaplastic choroid plexus papilloma</li> <li>Subependymal giant cell astrocytoma</li> </ul>

---

Table 10: (Part 3)

---

III ventricle	
Adults	Gangliocytoma Ganglioglioma Subependymoma Germinoma Meningioma
Adolescents	Choroid plexus papilloma Gangliocytoma Ganglioglioma
Children	Choroid plexus papilloma Ependymoma Astrocytoma fibrillary Germinoma
IV ventricle	
Adults	Choroid plexus papilloma Subependymoma Capillary hemangioblastoma
Adolescents	Ependymoma Medulloblastoma
Children	Ependymoma Medulloblastoma Ependymoblastoma
Cerebellar hemisphere	
Adults	Capillary hemangioblastoma Astrocytoma fibrillary Anaplastic astrocytoma Glioblastoma multiforme Medulloblastoma
Adolescents	Astrocytoma pilocytic Astrocytoma fibrillary Medulloblastoma Capillary hemangioblastoma
Children	Astrocytoma pilocytic Medulloblastoma Capillary hemangioblastoma

---

Table 10: (Part 4)

---

Midline cerebellum	
Adults	<ul style="list-style-type: none"> <li>Meningioma</li> <li>Capillary hemangioblastoma</li> <li>Astrocytoma fibrillary</li> </ul>
Adolescents	<ul style="list-style-type: none"> <li>Medulloblastoma</li> <li>Ependymoma</li> <li>Astrocytoma pilocytic</li> <li>Astrocytoma fibrillary</li> <li>Capillary hema</li> </ul>
angioblastoma	
Children	<ul style="list-style-type: none"> <li>Medulloblastoma</li> <li>Ependymoma</li> <li>Astrocytoma pilocytic</li> <li>Astrocytoma fibrillary</li> <li>Capillary hemangioblastoma</li> </ul>
Brain stem	
Adults	<ul style="list-style-type: none"> <li>Oligodendroglioma</li> <li>Astrocytoma fibrillary</li> </ul>
Adolescents	<ul style="list-style-type: none"> <li>Astrocytoma pilocytic</li> <li>Astrocytoma fibrillary</li> <li>Anaplastic astrocytoma</li> <li>Glioblastoma multiforme</li> <li>Oligodendroglioma</li> <li>Ependymoblastoma</li> </ul>
Children	<ul style="list-style-type: none"> <li>Astrocytoma pilocytic</li> <li>Astrocytoma fibrillary</li> <li>Anaplastic astrocytoma</li> <li>Glioblastoma multiforme</li> <li>Oligodendroglioma</li> <li>Ependymoblastoma</li> </ul>
Cerebellar pontine angle	<ul style="list-style-type: none"> <li>Schwannoma</li> <li>Meningioma</li> <li>Neurofibroma</li> <li>Ependymoma</li> </ul>

---



of symptoms taken together has a high correlation and hence great discriminatory power. It would be more powerful expressively if we considered the probability of various combinations of symptoms given the disease, but with  $n$  symptoms we have  $2^n$  possible combinations of symptoms. With  $n$  symptoms and  $m$  possible diseases, we have  $2^n \times m$  possible pairs of symptom sets with diseases. It is impossible to divine all possible feature combinations. But we were interested only in combinations of features that were possible from the pathological point of view. The selection of a limited number of possible combinations of features was justified by analysis of clinical cases (not ideal cases from textbooks but real cases encountered in practice) which were diagnosed using the same feature sets implemented in the system. The diversity of the real cases encoded in the secondary decision tables provided many unusual combinations of circumstances.

The inferencing mechanism in a knowledge-based computer system operates like an automated theorem prover. A theorem consists of a set of hypotheses and a conclusion that follows from the hypotheses. In a knowledge-based system, the data input by the user corresponds to the hypotheses of the theorem. The inferencing mechanism then uses this data together with the knowledge incorporated into the system (corresponding to the definitions and postulates used in proving a theorem) to derive a diagnosis or other recommendation as the conclusion of the theorem. Balla (Balla, 1985) indicates that doctors do not use this kind of logical procedure in making a diagnosis. Results of studies of how doctors reason clearly showed that few of the subjects had any difficulty drawing positive conclusions from positive information, but thinking in negative terms or using negative information was difficult for them even when the required steps were logically valid. In designing the knowledge base to support this kind of theorem-proving model of reasoning, we were not trying to model the expert's diagnosis making process. In fact, we did not try to model the human decision making process at all. Instead, we used the logic-based approach to describe diagnostic categories in much the same way that the experts did when they described clinical cases using the logical connectors AND, OR, and NOT which were considered in decision tables. The logic for using the decision tables can be stated: If conditions Then consequences. Decision logic tables are just clear and concise and are capable of showing meaningful relationships in a similar manner. They are, in fact, based on the If... Then concept. The "if" area is made up of all conditions or tests that are required to determine the conclusions or actions (the "then" area) (McDaniel, 1968). In this way, we represented more than 2000 combinations of features which were relevant to the diagnosis of primary brain tumors. These statements representing different combinations of features became the antecedent conditions for our rules. Except that there is no measure of certainty or uncertainty represented in these rules, they capture relevant combinations of features identified by the experts.

Rules in the knowledge base are in conditional form:

IF antecedent condition THEN consequent.

The antecedent condition of a rule can be any simple or complex statement formulated using the logical connectives AND, OR, and NOT, that represent some possible combination of

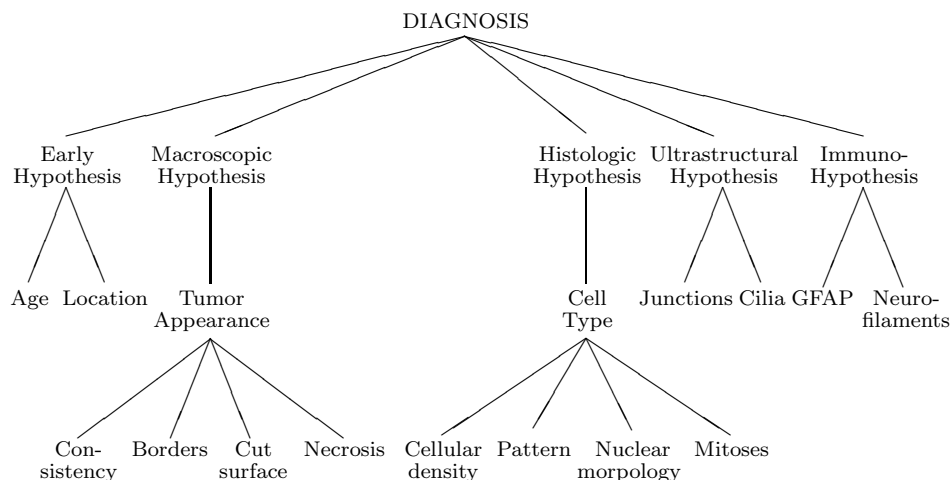


Figure 1: Diagnostic graph

features or symptoms. These may contain variables for tumor location, age of patient, etc., and conditions on these parameters. The consequent of a rule a new statement representing one or more facts that can be added to working memory whenever the antecedent condition is established. Symbolically, we represent a rule with antecedent  $A$  and consequent  $C$  as  $A \Rightarrow C$ .

What strategy does the system use to arrive at a diagnosis? There are several possible strategies, but all of them involve construction of a proof. A proof in this case is a graph with nodes labeled with statements and connected by arcs. There is one top node in the graph which is labeled by a diagnosis. Below this top node are all nodes labeled by the antecedent conditions of some rule that links these conditions with the diagnosis associated with the top node or “parent” node of these nodes. The statement associated with each of these “child” nodes can in their turn be the consequents of other rules in the system, and the antecedents of these rules become the labels for the “children” of these nodes (the “grandchildren” of the top node). Additional branches in the graph may be generated by other rules. The graph or proof is complete when all the bottom nodes are labeled by facts that have been established by information provided by the user. So the top node represents a possible diagnosis, the bottom nodes represent the initial facts about the case, and the middle nodes represent intermediate conclusions that the system draws during its search for a diagnosis. See Figure 1.

How does the system construct a proof and establish a possible diagnosis? There are two basic methods called *backward chaining* and *forward chaining*. The rules have the general form

$$\text{IF } A_1 \text{ AND } A_2 \text{ AND } \dots \text{ AND } A_n \text{ THEN } C$$

where each  $A_i$  represents an antecedent condition and C represents a consequent. An  $A_i$  may represent some basic feature of the case or it may represent an intermediate conclusion; C may represent an intermediate conclusion or a possible diagnosis. The rules have this same form in both backward and forward chaining systems, but the rules are used differently in the search for a diagnosis.

In backward chaining, the system selects a rule with a possible diagnosis as consequent and looks for facts that establish the antecedents or other rules that support the antecedents. If other rules are used, then the system looks for facts or rules that support their antecedents, and so on until the system finds facts to establish all the antecedents of all the rules or it hits an antecedent for which there is no support. Because the system starts with a possible diagnosis and works backward to try to confirm the diagnosis, this kind of reasoning is also called *goal driven*.

In forward chaining, the system starts at the other end of the rule. It finds any rule all of whose antecedents are supported by facts and adds the consequent of that rule to working memory. It then finds another rule all of whose antecedents are supported and “fires” it. As each rule is fired, the set of facts in working memory that can be used to satisfy other rules grows larger. The system continues in this manner until a possible diagnosis is reached or until no more rules are satisfied. Because this kind of reasoning starts with the initial data and draws all the conclusions it can that are supported by the data, this kind of reasoning is also called *data driven*.

Our knowledge base is composed of 480 deduction rules and the inference mechanism including both, goal driven and data driven operations applies these to a set of facts (clinical, macroscopic, histologic, ultrastructural, and immunohistochemical features) which are input by the user. In the rules, we used the operator IS to represent cases involving exclusive and exhaustive values of features and the operator INCLUDES for features that can have multiple values. Different partial knowledge bases were designed for each of the diagnostic levels.

A preliminary hypothesis-identification step is based on the most obvious clinical features present: tumor location and age of patient.

At the next stage, the user can choose the level to be used to differentiate among early diagnoses. This enables the user to use macroscopic, histologic, ultrastructural, or immunohistochemical data to further differentiate between tumor types. If the user wants to investigate the early diagnostic hypotheses at the macroscopic level, the system will ask appropriate questions about macroscopic findings and, based on the user’s answers to these questions, one of the suggested hypotheses is promoted. Questions are selected to narrow the number of diseases under consideration and the relevant rules are fired based on the answers to these questions. See Figure 2.

By observing how the system behaves, the user can determine what kinds of cues have the most discriminative value in a particular diagnostic situation. At each level the user is given

Figure 2: Sample MacLib rules in English and Xiplus form.

---

IF location is midline cerebellum  
AND age is < 15  
THEN early hypothesis is  
medulloblastoma/ependymoma/astrocytoma pilocytic/astrocytoma fibrillary/  
capillary hemangioblastoma  
AND ask level of diagnosis

WHEN the level of diagnosis is macroscopic  
AND early hypothesis is  
medulloblastoma/ependymoma/astrocytoma pilocytic/astrocytoma fibrillary/  
capillary hemangioblastoma  
THEN command load Macroscopic Diagnosis  
AND ask borders  
AND ask cut surface  
AND ask consistency  
AND ask necrosis

IF borders is ill delineated  
AND cut surface includes reddish OR grayish  
AND consistency is soft  
AND necrosis is yes  
THEN tumor is probable poorly differentiated

IF early hypothesis is  
medulloblastoma/ependymoma/astrocytoma pilocytic/astrocytoma fibrillary/  
capillary hemangioblastoma  
AND tumor is probable poorly differentiated  
THEN macroscopic hypothesis is probable medulloblastoma/ependymoma

WHEN the level of diagnosis is histologic  
AND early hypothesis is  
medulloblastoma/ependymoma/astrocytoma pilocytic/astrocytoma fibrillary/  
capillary hemangioblastoma  
THEN command load Histologic Diagnosis  
AND ask cellular density  
AND ask pattern  
AND ask nuclear morphology  
AND ask mitoses

IF cellular density is cells back-to-back  
AND pattern includes rosettes with/without lumen OR disorganized  
AND nuclear morphology includes dense nuclei/scanty cytoplasm  
OR polymorphism/multiple nuclei/irregular nuclear membrane  
AND mitoses is numerous/abnormal  
THEN cell type is poorly differentiated

---

---

```

IF early hypothesis is
medulloblastoma/ependymoma/astrocytoma pilocytic/astrocytoma fibrillary/
capillary hemangioblastoma
AND cell type is poorly differentiated
THEN histologic hypothesis is medulloblastoma

IF histologic hypothesis is medulloblastoma
AND junctional complexes is yes
AND cilia is no
THEN ultrastructural hypothesis is medulloblastoma

IF early hypothesis is medulloblastoma/ependymoma/astrocytoma pilocytic/astrocytoma fibrillary/
capillary hemangioblastoma
AND gfap is cells gfap-positive occasionally OR negative
AND neurofilaments is positive
THEN immuno-hypothesis is medulloblastoma

```

---

Figure 2: (Part 2)

diagnostic suggestions and the final diagnosis depends on his/her judgment. The user can ask why a certain question is being asked and what is the importance of the question to the final diagnosis. Digitized pathological images are linked to the reasoning system as help files. These files present templates for each feature that a pathologist can reproducibly identify. By a close comparison of the stored images with the histological material being analyzed, the user can improve the reliability of his/her structure identification.

## 7 Summary

The main goal of the project was to develop a program to assist in diagnosis making by teaching the diagnostic procedure to residents and by offering pathologists external sources of reference knowledge to refresh their personal knowledge and to support the most frequently occurring diagnostic tasks: differentiation among diagnoses, confirmation of diagnoses, and the search for possible diagnoses. The diagnostic accuracy achieved by a trainee pathologist and a general pathologist using the system has shown that the assistance of the system can enhance their performance in the domain of brain tumors and make it in the case of a general pathologist almost comparable to that of a neuropathologist.

Since there is no well defined monolithic diagnostic process that machines can handle with complete confidence, our aim was not to model the diagnostic process of the expert but to modify the way the novice is taught how to solve diagnostic problems. A beginner makes inferences using facts and rules (similar to a computer program) but with talent and a great deal of experience the beginner develops into an expert, who immediately sees what to do without consciously applying rules. In the early stages of learning problem-solving skills, the

learner has to discover a procedure for arriving at the solution to a problem. At this stage, for different diagnostic situations the computer program can offer in an interactive manner critical cues with the highest discriminatory power which enable differentiation among diagnostic hypotheses. At the next stage the user jumps intuitively to the solution in situations that were similar to the ones with which he/she has become already familiar. Acquisition of problem solving skills moves from abstract rules to particular cases. Textbooks give lists of features when describing diseases. Such lists are unhelpful unless accompanied by factual information in the context of differential diagnosis. The weighing of data is usually poorly described in standard texts and less experienced physicians have difficulties with correct data assessment. There is a need for goal orientation and early hypothesis formation. The choice of the first hypothesis depends on an evaluation of prior probabilities and of the utility of diagnostic categories. Rather than learn lists of features, the novice should learn the correct weights for cues, especially for those that are of high discriminatory value, and those that are highly specific for certain diseases (Balla, 1985). The information contain in reference or text books is not organized in this way. The computer-assisted management of visual and verbal information and the stepwise presentation of this information in the context of the diagnostic situation enables the trainee pathologist to figure out what to do in a particular situation. The computer program provides the user with procedural and diagnostic suggestions for decision making in an interactive manner. A computer dialogue lets the user explore a domain of knowledge in a dynamic fashion, consulting expertise more crisply than is possible with the self-interpretation of a textbook.

The system was developed by physicians interested in the domain of primary brain tumors as well as applications of artificial intelligence in medicine, without a computer scientist acting as an intermediary. In our opinion, direct participation by domain specialists in the development of knowledge bases for diagnostic support systems promises a partial solution both to the problem of the knowledge acquisition bottleneck and to the problem of the lack of acceptance of such systems by the medical community.

## **Acknowledgment**

The authors wish to thank Drs.: Pawel Liberski, Janusz Alwasiak, Wielislaw Papierz, Anna Szadowska, and Wojciech Biernat from the Medical University of Lodz, Poland for providing their knowledge, assistance and criticism in designing the knowledge base, and for their help in evaluation and testing of the system.

The part of this work was supported by The Kosciuszko Foundation, New York.

## References

- [1] Anderson, J. R.: Cognitive Psychology and Its Implications. W. H. Freeman and Company, New York (1990).
- [2] Baak, J. P.: The Principles and Advances in Quantitative Pathology. *Anal Quant Cytol*, 9, pp. 89–95 (1987).
- [3] Balla, J.I.: The diagnostic process. A model for clinical teachers. Cambridge University Press, Cambridge (1985).
- [4] Becker, L.E. and Halliday, W.C.: Central Nervous System Tumors of Childhood. *Perspect Pediatr Pathol*, vol. 10, pp. 86–134 (1987).
- [5] Bonnin, J.M. and Rubinstein, L.J.: Immunohistochemistry of central nervous system tumors. Its contribution to neurosurgical diagnosis. *J Neurosurg*, 60: 1121 (1984).
- [6] Burger, P.C. and Fuller, G.N.: Pathology-Trends and Pitfalls in Histologic Diagnosis, Immunopathology, and Applications of Oncogene Research. *Neurol Clin*, vol. 9, no. 2, pp. 249–272 (1991).
- [7] Chang, E.; McNeeley, M. and Gamble, K.: Strategies for choosing the next test in an expert system. In: Proceedings of the American Association of Medical Systems and Informatics Congress. American Association of Medical Systems and Informatics, Bethesda, pp. 198–202 (1984).
- [8] Clancey, W.J.: Knowledge-Based Tutoring. The GUIDON Program. The MIT Press, Cambridge (1987).
- [9] Codd, M.B. and Kurland, L.T.: Descriptive Epidemiology of Primary Intracranial Neoplasms. *Prog Exp Tumor Res*, vol. 29, pp. 1–11 (1985).
- [10] Coles, L. S.; Brown, B. W.; Engelhard, C.; Halpern, J. and Fries, J. R.: Determining the most valuable clinical variables: a stepwise multiple logistic regression program. *Methods Inf Med*, 19, pp. 42–49 (1980).
- [11] Colmant, H. J. and Noltenius, H.: Tumors of the central nervous system. In H. Noltenius (Eds.) *Human Oncology. Pathology and Clinical Characteristics*. Urban & Schwarzenberg, Baltimore-Munich, pp. 751–840 (1988).
- [12] de Dombal, F. T.: Medical diagnosis from a clinical point of view. *Methods Inf Med*, 17, pp. 28–35 (1978).
- [13] Dreyfus, H. L. and Dreyfus, S. E.: *Mind over Machine*. Free Press, New York (1986).
- [14] Elstein, A.S.; Shulman, I. S.; Jason, H. and Loupe, M. J.: Methods and theory in the study of medical inquiry. *J Med Educ*, 47, pp. 85–92 (1972).

- [15] Fergus, R. M.: Decision Tables - What, Why, and How. In D. Couger and R. Knapp (Eds): Systems Analysis Techniques. John Wiley & Sons, Inc., New York, pp. 162–179 (1974).
- [16] Griffiths, M. and Palissier, C.: Algorithmic Methods for Artificial Intelligence. Chapman and Hall, New York (1987).
- [17] Gyde, O.H.: Computer assisted bone marrow aspirate. In Proceedings of The Second International Conference on Artificial Intelligence Systems (Expert Systems) as Diagnostic Consultants for the Cytologic and Histologic Diagnosis of Cancer, March 13–15, Chicago (1988).
- [18] Heckerman, D.E.; Horvitz, E.J. and Nathwani, B.N.: Update on the Pathfinder Project. In Proceedings of the Thirteenth Symposium on Computer Applications in Medical Care, Washington, IEEE Computer Society Press, pp. 203–207 (1989).
- [19] Horvitz, E.J.; Heckerman, D.E.; Nathwani, B.N. and Fagan, L.M.: Diagnostic Strategies in the Hypothesis-Directed PATHFINDER System. In Proceedings of the Eighth Symposium on Computer Applications in Medical Care, Washington, IEEE Computer Society Press, pp. 630–636 (1984).
- [20] Horvitz, E.J.; Heckerman, D.E.; Ng, K.C. and Nathwani, B.N: Heuristic Abstraction in the Decision-Theoretic Pathfinder System. In Proceedings of the Thirteenth Symposium on Computer Applications in Medical Care, Washington, IEEE Computer Society Press, pp. 178–182 (1989).
- [21] Hughes, M.L.; Shank, R.M. and Stein, E.S.: Decision Tables. MDI Publications, Wayne, (1968).
- [22] Kernohan, J.W. and Sayre, G.P.: Tumors of the central nervous system. Atlas of tumor pathology, fasc. 35. Armed Forces Institute of Pathology, Washington (1952).
- [23] Kong, A.; Barnett, G. O.; Mosteller, F. and Youtz, C.: How medical professionals evaluate expressions of probability. N Engl J Med 4, vol. 315, no. 12, pp. 740–744 (1986).
- [24] McDaniel, H.: Introduction to Decision Logic Tables. John Wiley & Sons, Inc., New York (1968).
- [25] Michalski, R.S. and Chilausky, R.L.: Learning by being told and learning from examples. Int J Pol Anal Info Sys, 4, pp. 125–160 (1980).
- [26] Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol Rev, 63, pp. 81–97 (1956).
- [27] Miller, R.A.; Pople, E.P. and Myers, J.D.: Internist-1: An experimental computer-based diagnostic consultant for general internal medicine. N Engl J Med, 307, pp. 476–486 (1982).



- [28] Nathwani, B.N.; Heckerman, D.E.; Horvitz, E.J. and Lincoln, T.L.: Integrated Expert Systems and Videodisc in Surgical Pathology: An Overview. *Hum Pathol* 21:11–27, (1990).
- [29] Percy, A.K.; Elveback, L.R.; Okazaki, H. and Kurland, L.T.: Neoplasms of the central nervous system. Epidemiologic considerations. *Neurology*, 22:40 (1972).
- [30] ReKate, H.L. and Rakfal, S.M.: Low-grade Astrocytomas of Childhood. *Neurol Clin*, vol. 9, no. 2, pp. 423–440 (1991).
- [31] Russell, D.S. and Rubinstein, L.J.: *Pathology of Tumours of the Nervous System*. Williams & Wilkins, Baltimore, V-th Edition (1989).
- [32] Schank, R.C.: Where's the AI. *AI Magazine*, Winter, pp. 38–49 (1991).
- [33] Scheithauer, B.W. and Bruner, J.M.: Central Nervous System Tumors. *Clin Lab Med*, vol. 7, no. 1, pp. 157–179 (1987).
- [34] Shortliffe, E.H.: *Computer-based medical consultations: MYCIN*. Elsevier, New York (1976).
- [35] Shortliffe, E.H. and Buchanan, B.G.: *Rule-based expert systems*. Addison-Wesley Pub., Reading, Mass (1984).
- [36] Tobias, J.S. and Hayward, R.D.: Brain and spinal cord tumours in children. In D.G.T. Thomas (Eds.): *Neuro-oncology. Primary malignant brain tumours*. The John Hopkins University Press, Baltimore, pp. 164–192 (1990).
- [37] Wortman, P.M.: Medical diagnosis: an information processing approach. *Comput Biomed Res*, 5, pp. 315–328 (1972).
- [38] Zulch, K.J.: *Histological typing of tumours of the central nervous system*. World Health Organization, Geneva (1979).
- [39] Zulch, K.J.: Historical development of the classification of brain tumors and new proposal of the World Health Organization (WHO). *Neurosurg Rev*, 4, pp. 123–127 (1981).