

A SURVEY ON MEDICAL DIAGNOSIS AND FUZZY SUBSETS

KLAUS-PETER ADLASSNIG

Department of Medical Computer Sciences
(Director: Prof.Dr. G. Grabner)
University of Vienna
A-1090 Vienna, Garnisongasse 13, Austria

Fuzzy set theory with its capability of defining inexact medical entities as fuzzy subsets, with its linguistic approach providing an excellent approximation to medical texts as well as its power of approximate reasoning seems to be perfectly appropriate for designing and developing computer-assisted diagnostic, prognostic and treatment recommendation systems. The detailed survey of fuzzy approaches to medical decision making shows the wide and deep interest of researchers all over the world to find a solid fuzzy theoretical basis for the representation of medical knowledge as well as for the diagnostic process and to use it in practical applications.

Keywords: fuzzy set theory, medical diagnosis, survey.

1. INTRODUCTION

Since the very beginning of developing computer-assisted diagnostic methods and medical decision aids in the years 1958 /1/ and 1959 /2/, a great number of medical diagnostic models applied in different medical fields with varying results (bibliographies in computer-assisted diagnosis can be found in /3-5/) have been proposed and a certain number of them has been tested.

The first logical approach to medical decision making (as opposed to numerical models ¹⁾ based on Bayes' Theorem, likelihood ratio, linear and nonlinear discriminant analysis, clustering methods, factor and principal component analysis, decision analysis, adaptive threshold logic unit, perceptron algorithm, etc.) has been suggested by LEDLEY and LUSTED in 1959 /2/ (see also

LEDLEY /6/).

Since that year, several systems using Boolean logic to express functional and etiological relationships between symptoms and diseases have been proposed /7/ and established /8-16/. In 1969 GANGL, et al. /11/ proved the usefulness of such approaches. The described batch system considered 82 liver diseases and 323 symptoms (history, physical examination, laboratory tests, X-ray findings, special tests). In each of the 20 test cases the clinically confirmed diagnosis was put out as diagnostic hypothesis (among 3-6 others). These results were very encouraging.

Later then, as a logical consequence, the symptoms and diagnoses were not only considered to be present or absent as in two-valued systems but the truth-values "present", "absent", or "not

examined" were assigned to symptoms and "present" (confirmed), "absent" (excluded), and "possible" (hypothesis) to diagnoses. Thus tri-valued logic was introduced as a means for the representation of medical knowledge /17-20/.

In the report about the computer-assisted medical diagnostic consultation system CADIAG-1 /19/, the tri-valued logical system of KLEENE /21/ seems to be adequate to formalize medical judgement in an abstract and computable way, understandable by and acceptable to the medical community. This system has now been established for 308 diseases from internal medicine (rheumatology, hepatology, coagulation defects and pancreatic diseases) and takes about 2.500 symptoms into account.

The possibility of expressing medical knowledge in terms of predicate logic was also mentioned in /19/.

In the last ten years, some very powerful computer-based medical decision and consultation systems have been developed by researchers working in the field of artificial intelligence. SHORTLIFFE, et al. give in /22/, among others, an excellent introduction to these symbolic reasoning approaches.

The MYCIN system (see /22,23/) has been designed to assist physicians in the detection of the cause of infectious diseases and to recommend therapies. The medical knowledge about the diseases considered is represented in IF-THEN production rules associated with certainty factors.

CASNET /22,24,25/ offers associational links between Boolean combinations of tests and single pathophysiological

causal states that carry confidence factors. Furthermore, state network rules associated with a given likelihood for the effect S_j to follow from the occurrence of the cause S_i are used in CASNET. Finally, classification rules that are Boolean combinations of confirmed or denied states implying diagnostic statements are incorporated in order to represent medical knowledge. CASNET was created for glaucoma consultation.

The EXPERT system /26-28/ permits the definition of rules with confidence levels and hierarchical concepts to convert medical knowledge into a computable form. It has been applied in rheumatology, endocrinology and ophthalmology /29/.

A further system applicable in the field of internal medicine is called INTERNIST /22,30,31/. Symptom-disease (frequency of occurrence) as well as disease-symptom relationships (evoking strength) are stored on the basis of an ad hoc scoring method.

MYCIN, CASNET and EXPERT are logical in the sense that relationships between all kinds of observations on the patient form the basis for IF-THEN inferences (modus ponens in classical logic, implication in terms of logical operations) but the inferences do not necessarily have a definite character. They are associated with certainty factors, confidence factors, likelihoods, or confidence levels that express the fuzziness, the non-definiteness of the relationships. The "evoking strength" of the INTERNIST system also belongs into this category.

Other authors who use IF-THEN statements

but without any elements of uncertainty are BUTTON /32/ and GIESZL /33/.

Fuzzy set theory with its capability of defining inexact medical entities as fuzzy subsets, with its linguistic approach providing an excellent approximation to medical texts as well as its power of approximate reasoning, seems to be perfectly appropriate for designing and developing computer-assisted diagnostic, prognostic and treatment recommendation systems. A detailed review of fuzzy approaches to medical diagnosis published until now follows in the next paragraphs.

In order to conclude this section it should be mentioned that disease descriptions using the theory of formal languages /34/ or automata /35,36/ as well as methods using decision table techniques /37/ also belong into the category of logical diagnostic systems.

2. REVIEW OF PUBLISHED METHODS

2.1. Symptoms, signs, and test results and fuzzy subsets

ALBIN /38/, PEREZ-OJEDA /39/, MOON, et al. /40/, and ESGBUE and ELDER /41/ propose the application of fuzzy subsets to determine normal or pathological ranges as well as the boundaries for low, normal, high or normal, slightly decreased, decreased, etc. for clinical or diagnostic tests. The membership functions of these fuzzy subsets define the affiliation strength of a numerical test result in the fuzzy subsets under consideration. In this way, ALBIN /38/ presents appropriate membership functions for normal and long IVY bleeding time, for decreased, slightly decreased, and normal platelet

count, for normal, slightly long, and long thrombin time, etc. In /41/, for example, a suitable linear function for the fuzzy subset α_t of abnormal cholesterol levels, expressed in milligrams per 100 milliliters of serum, is given by

$$\mu_{\alpha_t}(r_t) = \begin{cases} 0 & \text{for } r_t < 260 \\ \frac{r_t}{340} - \frac{26}{34} & \text{for } 260 \leq r_t \leq 600 \\ 1 & \text{for } r_t > 600. \end{cases} \quad (1)$$

where $r_t \in \alpha_t$ is any possible numerical result of the cholesterol test t and α_t expresses the set including all possible test results. Thus, the degree of abnormality of the cholesterol test result is reflected by the membership function $\mu_{\alpha_t}(r_t) \in [0,1]$. Performing the cholesterol^t test on a patient and getting a result of $r_t < 260$ mg/100 ml means that the patient shows a normal cholesterol level. When a test result $r_t > 600$ mg/100 ml is obtained, the patient under consideration reveals an abnormal cholesterol level. Furthermore, a test result between 260 and 430 mg/100 ml ($r_t = 430$ mg/100 ml is the value where $\mu_{\alpha_t}(r_t)$ becomes 0.5) seems to be "more^t normal" than a test result between 430 and 600 mg/100 ml.

ESGBUE and ELDER /41/ propose the use of fuzzy subsets to characterize the severity levels of nonbinary symptoms such as headache or cyanosis. The membership functions would reflect the painfulness or blueness of these symptoms. But compared to the laboratory test descriptions using fuzzy subsets, the definition of the reference sets for symptom fuzzy sub-

sets is less natural and clear. A subjective (patient or physician) assignment of the observed symptom value (patient's pain, or blue discoloration) in the fuzzy subset is needed.

PEREZ-OJEDA /39/ and MOON, et al. /40/ use linguistic modifiers (also termed hedges /42-44/) to calculate the degree of membership of a test result in the fuzzy subset S_2 out of the degree of membership in S_1 . Linguistic modifiers are for example very, more or less, not, etc. By using them, MOON, et al. /40/ compute the degree of membership of the result x of the urine sodium concentration test in the fuzzy subset very high urine sodium concentration $\mu_{S_2}(x)$ out of high urine sodium concentration $\mu_{S_1}(x)$ by performing

$$\mu_{S_2}(x) = \mu_{\text{very } S_1}(x) = [\mu_{S_1}(x)]^2. \quad (2)$$

In /41/, ESGBUE and ELDER introduce fuzzy subsets to determine diseases or disorders in the patient's past history which might be missed or remain undiagnosed. The presence or absence of such diseases might then be decided on from past undiagnosed symptoms, but the physician only considers the "prominent" symptoms of past diseases. The fuzzy subset Σ_j , containing the prominent symptoms of past disease j is defined as

$$\Sigma_j = \{(x_{ij}, \mu_{\Sigma_j}(x_{ij})) / x_{ij} \in \beta_j, \mu_{\Sigma_j}(x_{ij}) \geq \alpha\} \quad (3)$$

where $\mu_{\Sigma_j}(x_{ij}) \in [0, 1]$ and $0 \leq \alpha \leq 1$.

The set of possible symptoms for past undiagnosed disease j is β_j , where $1 \leq j \leq p$ and p is the number of undiagnosed diseases in the patient's past history. Wherever the physician designates the membership function $\mu_{\Sigma_j}(x_{ij})$ of symptom x_{ij} for disease j over a specified level α , the symptom of β_j becomes a "promi-

nent" symptom.

2.2. Diseases/diagnoses and fuzzy subsets

SMETS, et al. /45/ mention the fuzziness of diagnostic terms like arterio-sclerosis or angina pectoris. Mostly, diseases are not clearly defined entities. It is often impossible to determine precisely the symptoms related to a disease. Diagnoses are defined to be fuzzy subsets with symptoms as elements. The symptoms are combined with a degree of membership which characterizes the intensity of belonging to the fuzzy subset that represents the disease under consideration.

ALBIN /38/ considers patients suffering from a certain disease as points in the n -dimensional symptom space S , where n is the number of symptoms taken into account. Now, every given patient with disease D_j exhibiting symptoms $(S_1, \dots, S_n) \in S$ is combined with a degree of membership $\mu_{D_j}(S_1, \dots, S_n) \in [0, 1]$ that shows the affiliation of the patient in D_j . Because $\mu_{D_j}(S_1, \dots, S_n)$ is not necessarily 1 for $i=j$ and 0 for $i \neq j$, patients suffering from the same disease do not form clusters with sharp boundaries.

2.3. Fuzzy relations between symptoms and diseases

In /39/, PEREZ-OJEDA proposes a network presentation of medical knowledge (associative memory, semantic network) in which the facts, symptoms, diseases, etc., are linked by relations. Six types of nodes are considered: disease-complex nodes, disease nodes, symptom-complex nodes, symptom nodes, statement nodes (relations) and test nodes. PEREZ-OJEDA

translates medical statements like

"Acute pyelonephritis usually presents bladder irritation and infection." and "Acute pyelonephritis presents occasionally fever, or chills, and malaise." into

OR (usually (acute pyelonephritis, AND (bladder irritation, infection)), occasionally (acute pyelonephritis, AND (OR (fever, chills), malaise))).

The relations used in this information structure (usually, occasionally, always, almost always, etc.) are frequency or probabilistic modifiers. They are represented as fuzzy subsets, e.g.

always=0.1/0.7+0.3/0.8+0.6/0.9+1.0/1.0,

of the universe of discourse

U= 0+0.1+0.2+0.3+0.4+0.5+0.6+0.7+0.8+0.9+1.0.

A network in the fashion described above has been built which consists of 23 disease-complex and disease nodes, 101 symptom nodes and 227 statements.

SANCHEZ /46,47/ starts by considering the introduction of a fuzzy relation $R \subset S \times D$ between symptoms S and diseases D . SANCHEZ calls this "medical knowledge" expressing associations between symptoms and diseases. Let, further, A be a fuzzy subset of S related to the patient, then the computation of the max-min²⁾ composition $B = A \circ R$ ³⁾ is assumed to describe the state of the patient in terms of diagnoses as a fuzzy subset B of D , characterized by its membership function

$$\mu_B(d) = \max_{s \in S} \min\{\mu_A(s); \mu_R(s,d)\}, \quad d \in D. \quad (4)$$

If we consider several patients belonging

to a set P and define a relation $Q \subset P \times S$, equation (4) becomes

$$\mu_T(p,d) = \max_{s \in S} \min\{\mu_Q(p,s); \mu_R(s,d)\}, \quad (5)$$

where $p \in P$ and T is a fuzzy relation on $P \times D$. The composition is written as

$$T = Q \circ R. \quad (6)$$

In fuzzy logic, (6) is called a compositional rule of inference (ZADEH /44/, BELLMANN and ZADEH /42/ or a fuzzy meta-implication (KAUFMANN /48/).

Moreover, the composition $B = A \circ R$ of a fuzzy subset A with a fuzzy relation R corresponds to the concept of the fuzzy conditional statement: IF A THEN B BY R (see /42,44,52,53/).

Applied in the field of medical diagnosis, the conditional statement allows to infer medical diagnoses B out of patient's symptoms A by means of the "medical knowledge"⁴⁾ which is presented in a relation matrix R established by medical specialists.

If the relationships between symptoms and diseases are considered to be either 0 (no relationships) or 1 (strong relationship), e.g. $\mu_R(s,d) \in \{0,1\}$, and patient's symptoms can be evaluated as 0 (absent) or 1 (present), e.g. $\mu_A(s) \in \{0,1\}$, then the max-min composition (4) or (5), respectively, is reduced to a Boolean sum-product composition (see also /46-48/).

JOLY, et al. /54/ report an application of the max-min composition in the field of cardiopathies. Cardiologists have determined a fuzzy relation R in order to reflect the degrees of association between five cardiological symptoms and

the diagnoses normal⁵⁾, left ventricular hypertrophy and valvular cardiopathy. Five symptoms have been partitioned into classes (7 binary and 16 quantitative symptoms). The results obtained by the diagnostic procedure carried out for 21 patients with disparate cardiopathies were used "to confirm the classical concepts about cardiac function."

Another problem treated by SANCHEZ is related to equation (6). Knowing R and Q in (6), it is easy to find T. However, it is often as interesting to find R from given T and Q. In medicine, the problem frequently arises to determine degrees of association between symptoms and diseases from real patient data in Q and known diagnoses in T that have been established by medical specialists.

SANCHEZ presents in /46,47/ a solution of the problem by obtaining the greatest R for $T=Q \circ R$ (see also DUBOIS and PRADE /43/). But because of getting the greatest solution for R from the solution set, which also can be void, R may contain degrees of association between symptoms and diseases that are stronger than in reality. Thus it depends on the control structure of the computer-based diagnostic system that uses the computed R in which way results are interpreted: "diagnostic hypotheses or confirmed diagnoses".

The problem to find Q from given T and R can be considered in analogy to the problem mentioned above. This task presents itself when studying different fuzzy subsets of S, i.e. patient's symptom evaluations contained in Q that keep the composition $T=Q \circ R$ without changing T and R. The corresponding

medical question would be: "How far can patients' data vary until other diagnoses have to be made?".

Obtaining the greatest Q of possible Q's is demonstrated by SANCHEZ in /46, 47/. In /46,47,55/ the converse problems of (6) treated in the last two sections are also investigated by SANCHEZ in terms of Brouwerian logic.

An application of a converse problem of (6) to the differential diagnosis of icterus is given by SOULA, et al./56/. In this example four diseases and ten symptoms are taken into account.

Furthermore, in /57/, SANCHEZ has investigated the applicability of inverses of fuzzy relations in medical diagnosis.

SANCHEZ and SAMBUC /53/ and SAMBUC, et al. /58/ propose an interesting application of ϕ -fuzzy sets⁶⁾ to the medical diagnostic problem. ϕ -fuzzy sets take their values in the set of intervals of the form $[a_1, a_2]$ that are included in the real interval $[0,1]$. To illustrate how symptoms are represented in this way, we will quote an example from /58/ for the sign "pain on palpation" in patient p

$\neg(p)=[0,0]$ absence of pain on palpation
 $\mu(p)=[0.8,0.8]$ marked pain
 $\nu(p)=[0,1]$ totally indeterminate value
 $\lambda(p)=[0.2,0.6]$ there is hesitation between mild and relatively marked pain.

After defining suitable connectives for ϕ -fuzzy variables, ϕ -fuzzy functions of the form hyperthyreoidism = (tachycardia OR weight loss) AND (thermophobia OR increased thirst) AND (in-

creased seric T4 values OR increased seric T3 values) are evaluated. The following possible results are noted in /58/:

hyperthyroidism=[0,1]: Impossible to determine from the available signs whether or not the patient is hyperthyroid.

hyperthyroidism=[0.7,0.8]: The patient suffers from frank hyperthyroidism.

hyperthyroidism=[0.35,0.45]: The patient presents a tableau of border-line hyperthyroidism.

ADLASSNIG proposes in /59/ an approach to computer-assisted medical diagnosis using fuzzy subsets that is based on two different relationships between symptoms S_i and diseases D_j :

- 1) $S_i D_j$ -occurrence relationships and
- 2) $S_i D_j$ -confirmability relationships.

Occurrence and confirmability are considered to be linguistic variables with always, almost always, very very often, very often, rather often, often, more or less often, unknown, more or less seldom, seldom, rather seldom, very seldom, very very seldom, almost never, never as elements (see BELLMAN and ZADEH /42/). Single $S_i D_j$ -indications can be deduced from the $S_i D_j$ -occurrence and $S_i D_j$ -confirmability relationships. Examples for fuzzy subsets used to determine single $S_i D_j$ -indications are strong and weak. Finally, one obtains total indications of the patient's symptom pattern to disease D_j by consolidating the single $S_i D_j$ -indications.

TUSCH /60,61/ uses this model /59/ in a slightly modified form for the cranial computer tomography. The application

considers five tumor diagnoses: malignomata, semimalignomata, metastases, malformation tumors and benignomata. 25 symptoms gathered by seven different examinations describe every case: number of foci, structure of foci (native), edemata, localisation of edemata, form and position of ventricles, sulci and cisterns. The symptoms are dichotomous, with "symptom present" and "symptom absent/not investigated" as the two distinct values. TUSCH examines different algorithms in order to calculate total indications of the patient's symptom pattern to diseases D_j . The efficiency of the procedures lies between 55 and 76% compared to physicians' diagnoses. 802 tumor cases were used to perform this calculation. The $S_i D_j$ -occurrence and $S_i D_j$ -confirmability relationships have been documented linguistically by a neuroradiologist.

2.4. Symptom combinations and fuzzy subsets

ALBIN /38/ proposes a fuzzy linguistic approach to medical diagnosis. In /38/, it is said that physicians normally try to fit patients to certain prototypes of diseases. As a first approximation to such an approach, disease prototypes can be defined as shown in Table 1.

The definition of fuzzy membership functions for long S_1 , decreased S_2 , normal S_2 , etc., then allows the calculation of degrees of membership in the fuzzy subset long S_1 , decreased S_2 , normal S_2 , etc. given symptom values measured on the patient. Fuzzy disease membership values of patients can now be computed by using the MAX definition

Table 1: Disease prototype definitions for two hemorrhagic disorders (from /38/, p. 128).

Disorder	S ₁	S ₂	S ₃	S ₄	S ₅
	I.V. Bleeding Time	Platelet Count	QUICK Time	Partial Thromboplastin Time	Thrombin Time
D ₁ <u>Thrombocytopenia</u>	<u>Long</u>	<u>Decreased</u>	<u>Normal</u>	<u>Normal</u>	<u>Normal</u>
D ₂ <u>Von Willebrand's Disease</u>	<u>Long</u>	<u>Normal</u>	<u>Normal</u>	<u>Normal</u> or <u>Long</u>	<u>Normal</u>

for fuzzy set union and MIN definition for fuzzy set intersection, for example:

$$\begin{aligned} \mu_{D_1}(p) &= \mu_{\text{long}} S_1(12 \text{ min.}) \\ &\wedge \mu_{\text{decreased}} S_2(60\,000/\text{mm}^3) \\ &\wedge \mu_{\text{normal}} S_3(12 \text{ sec.}) \\ &\wedge \mu_{\text{normal}} S_4(37 \text{ sec.}) \\ &\wedge \mu_{\text{normal}} S_5(18 \text{ sec.}) \\ &= 0.95 \end{aligned}$$

Patient p, in our example, fits the prototype definition of D₁ almost exactly and if additionally computed fuzzy disease membership values $\mu_{D_i}(p)$, where $i \neq 1$, are much smaller than $\mu_{D_1}(p)$, diagnosis D₁=thrombocytopenia is presumably the correct diagnosis for patient p.

In /62/, SANCHEZ, et al. propose a similar linguistic approach to the classification of hyperlipoproteinemias. But in the diagnostic procedure described, predetermined bandwidths of the symptoms depending on the measuring accuracy of the laboratory performing the tests are additionally taken into account. The application involves five lipoprotein tests determined by electro-

phoretic methods. The degrees of membership in three types of dyslipoproteinemia were calculated for five test cases.

MOON, et al. /40/ present symptom combinations as AND and OR functions. The calculation of the symptom combinations is performed using the fuzzy intersection for AND and the algebraic sum for OR. A different way of presenting the OR function is mentioned: MOON, et al. use the convex combination

$$\mu_D(S_1, \dots, S_k) = \frac{\sum_{i=1}^k w_i \mu_{S_i}}{\sum_{i=1}^k w_i} \quad (7)$$

where μ_{S_i} is the ceiling function of μ_{S_i} and the weights w_i are set equal to the posterior probabilities $p(D/S_i)$ calculated for the single symptom S_i .

CERUTTI and PIERI /63/ also use linguistic variables as an alternative to numerical variables in order to characterize complex relationships between symptoms and diseases. Primary linguistic terms (small, medium, large)

that express intervals on the numerical symptom scale compounded with linguistic hedges (very, slightly, more or less) are connected by AND and OR to compose fuzzy logical combinations. By evaluating the fuzzy logical combination, a membership function is obtained that calculates non-fuzzy values (for details see /63/).

2.5. Medical diagnosis using procedural fuzzy concepts

In /64/, WECHSLER proposes a procedural knowledge base system using fuzzy concepts. Instead of using fuzzy membership degrees, classes of functions f_i^p are established whose task is to define implicitly fuzzy sets. A class is related to a concept and it contains one or more functions f_i^p where p stands for the property to be checked and i relates to some context, e.g. to the observer i . For example:

$$f_i^{\text{age}}(\text{teenage}) = (\text{age } 12, \text{age } 18)$$

$$f_i^{\text{age}}(\text{young}) = (\text{age } 17, \text{age } 30)$$

Subsequently, two kinds of order-decisions that determine whether or not a given object has the property in question are presented:

- 1) Middle order: The closer the value of the property of a given object is to the middle of the interval defined by f_i^p the better it fits the property to be checked. For example:
Age 18 fits "teenage" better than "young".
- 2) LB order: The first member of the interval defined in f_i^p is the best.

Then the medical knowledge is embedded in procedures and a possible program

checking for ulcers may look like ⁷⁾:

```
(LAMBDA(+D ULCERS)
  (PROGRAM ()
    (EXISTS Patient pain-location slightly
      below the top of
      xiphoid process  $f_u^{p1}$ )
    (EXISTS Patient time of pain breakfast  $f_u^{p2}$ ))).
```

2.6. Fuzzy cluster analysis in medical diagnosis

In /38/, ALBIN has developed a fuzzy clustering approach using an elliptic metric to define distance in the symptom space $S_1 \times \dots \times S_n = S = R^n$. The proposed fuzzy clustering method is compared with a Bayesian approach on the basis of some test patients.

FORDON /65/, and FORDON and BEZDEK /66/ present a very interesting application of fuzzy clustering techniques to the diagnosis of hypertensive patients. Different distance measurements are used. In one example, a fuzzy k-means algorithm diagnosed 218 essential and renovascular hypertensive cases described by four continuous parameters. The overall efficiency of the algorithm was 80.7%.

In /67/, BEZDEK applied the fuzzy ISODATA clustering algorithm to feature selection of medical data. In the numeric example, 11 binary symptoms of 107 stomach disease patients who had either hiatal hernia or gallstones were considered.

A further paper proposing the use of fuzzy clustering methods in medical diagnosis is the one by ESOGUE and ELDER /68/. That paper gives an instructive example of applying a criterion function that is a modified form of the minimum MINKOWSKI metric to

rheumatic valvular heart diseases.

2.7. Fuzzy decision analysis in medical diagnosis

Imprecision in decision analysis is modeled using fuzzy set theory in NOMURA, et al. /69/ and MISHIMA, et al. /70/. These reports propose an interesting model that represents probability or utility by a membership function. The model selects an optimal treatment by comparing among degrees of truth of some propositions on expected utilities. This method is applied to managing patients with pancreatic diseases centered around primary cancer. The decision process starts from a point at which a patient is very likely to be suffering from any one of four pancreatic diseases taken into account. In clinical practice the procedure helps to decide whether or not a radical operative invasion against the supposed disease should be performed.

2.8. Bayesian approach and fuzzy subsets

In /38/, ALBIN argued that it is entirely reasonable to assign

$$\mu_{D_i}(S_1, \dots, S_n) = P(D_i/S_1, \dots, S_n) \quad (8)$$

under the condition that good Bayesian values can be obtained for the probability indices $P(D_i/S_1, \dots, S_n)$.

"Calling $P(D_i/S_1, \dots, S_n)$ the value of the membership function μ_{D_i} at (S_1, \dots, S_n) in R^n is much more reasonable than calling it a probability, since the manner of estimating $P(D_i/S_1, \dots, S_n)$ is very inexact. Physicians' estimates based on their own medical knowledge have also been used in computing $P(D_i/S_1, \dots, S_n)$

from $P(S_1, \dots, S_n/D_i)$ and $P(D_i)$, so these estimates should be used with equal validity for $\mu_{D_i}(S_1, \dots, S_n)$."

quoted from /38/, p. 84

2.9. Miscellaneous

CHENG and McINNIS describe in /71/ an algorithm for multiple attribute (symptoms), multiple alternative (diagnoses) decision problems (see also KICKERT /72/). The relative importance of a symptom j denoted by μ_{W_j} , and the rating $\mu_{r_{ij}}$ that indicates the importance of the symptom j with respect to disease i have been established for 12 blood chemistry tests considering 78 diseases. On this basis the computer program calculates possible diseases which are ranked and arranged in descending order.

In FOX /73/, some critical remarks concerning WECHSLER's paper /64/ are put forward. FOX emphasizes that clinical and professional acceptability of computer-based medical consultation systems is of primary importance. Technology should be planned with this in mind.

OHSATO, et al. /74/ established a fuzzy relation in a preoperative nursing care system. The relation between preoperative patient's state and postoperative state is determined by using patient's attributes before and after the operation estimated by a nurse, e.g. degree of seriousness of preoperative and postoperative patient's states.

3. FINAL REMARKS

The detailed survey of fuzzy approaches to medical decision making shows the wide

and deep interest of researchers all over the world to find a solid fuzzy theoretical basis for the representation of medical knowledge as well as for the diagnostic process and to use it in practical applications.

The survey was made with the intention of studying fuzzy approaches in medicine in order to use the knowledge for developing the general fuzzy computer-assisted diagnostic system CADIAG-2. CADIAG-2 is described in a companion paper in the same edition /75/. A first theoretical approach to CADIAG-2 was made in /76/.

FOOTNOTES

- 1) Such categorization is a little bit arbitrary. It shall only emphasize the difference between methods mainly based on numerical calculations in order to find out diseases shown by the patient and methods using logical inference for the most part to obtain patient's diagnoses.
- 2) Instead of the max-min composition the more general term sup-min composition is often used (see for example DUBOIS and PRADE /43/).
- 3) SANCHEZ /46,47/ and KAUFMANN /48-51/ write $B=ROA$, which is equivalent to the notation used above.
- 4) At this juncture it becomes necessary to mention that the "medical knowledge" matrix R reflects only a very small part of the entire (diagnostic) medical knowledge.

By establishing the matrix R a physician would ask:
 "What does the degree of association between a symptom S_i and a disease D_j really mean? Is it the frequency of occurrence of S_i at disease D_j ? Is it a certainty factor of "confirmed by"? Is it an evoking strength /30,31/ of S_i for D_j or what else?"

Furthermore a symptom, sign, test

result, or finding often gains in importance with respect to a certain disease if it occurs with other symptoms or signs. For example, the symptom sacroileitis may occur in several diseases, as ankylosing spondylitis, juvenile arthritis, psoriatic arthropathy, Reiter's syndrome, Crohn's disease, and others. In combination with other nonspecific symptoms, e.g. concomitant mouth ulcer and conjunctivitis the diagnosis "Reiter's syndrome" can be established.

Next, the "degree of association" expressed as real number on $[0,1]$ does not remain constant with respect to time, place and individual. It changes with sex, age, season. For instance, in spring and autumn influenza is more probable as well as in times of a flu epidemic.

Finally, concepts used in medicine, such as hierarchy of symptoms and diseases, co-occurrence of diseases at the same patient and at the same time, evidence of symptoms depending on their time of duration, handling of unexplained symptoms at the end of the diagnostic process, etc. cannot be taken into account by only one "medical knowledge" matrix.

- 5) In classification theory, the normal, healthy, typical, unchanged pattern or class representative from which the abnormal, ill, untypical pattern is to be separated is always regarded as one class.
- 6) DUBOIS and PRADE /43/ term them interval-valued fuzzy sets.
- 7) The example is expressed in notation of QA4 (see /64/), a question-answering language implemented in LISP. Furthermore, it shall be mentioned that in order to diagnose ulcer the procedure is incomplete with regard to the medical point of view.

REFERENCES

- /1/ LIPKIN, M., J.D.HARDY: Mechanical Correlation of Data in Differential Diagnosis of Hematological Diseases. J.Am.Med.Ass. 166 (1958) 113-125.

- /2/ LEDLEY, R.S., L.B.LUSTED: Reasoning Foundations of Medical Diagnosis. Science 130 (1959) 9-21.
- /3/ ROGERS, W., B.RYACK, G.MOELLER: Computer-Aided Medical Diagnosis: Literature review. Int.J. Bio-Medical Computing 10 (1979) 267-289.
- /4/ WAGNER, G., P.TAUTU, U.WOLBER: Problems of Medical Diagnosis: A Bibliography. Meth.Inform.Med. 17 (1978) 55-74.
- /5/ WARDLE, A., L.WARDLE: Computer-Aided Diagnosis: A Review of Research. Meth.Inform.Med. 17 (1978) 15-28.
- /6/ LEDLEY, R.S.: Use of Computers in Biology and Medicine. McGraw-Hill Book Company New York, 1965.
- /7/ GREMY, F., M.GOLDBERG: Decision Making Methods in Medicine. Springer Verlag Berlin-Heidelberg-New York, 1977, 419-459.
- /8/ ADLASSNIG, K.-P., T.GERGELY, H.GRABNER, G.GRABNER: A Computer-Assisted System for Diagnostic Decision Making-Online Usage of the Database of the Medical Information System WAMIS. In SHIRES, D.B., H.WOLF (Eds.): MEDINFO 77. North-Holland Publishing Company Amsterdam-New York-Oxford, 1977, 213-218.
- /9/ BAUER, P., A.GANGL, G.GRABNER: Ein Computer-Verfahren zur Zuordnung eines Krankheitsbildes zu einer Diagnoses-Gruppe. Wien Z. Inn. Med. 11 (1970) 407-509.
- /10/ BAUER, P., A.GANGL, G.GRABNER, O.JAHN: Ein Computer-Verfahren zur Unterstützung des Arztes bei der Erstellung von Differential-Diagnosen. Impuls 10 (1968) 705-712.
- /11/ GANGL, A., G.GRABNER, P.BAUER: Erste Erfahrungen mit einem Computerprogramm zur Differentialdiagnose der Lebererkrankungen. Wien.Z.inn.Med. 50 (1969) 553-586.
- /12/ SPINDELBERGER, W., G.GRABNER: Ein Computerverfahren zur diagnostischen Hilfestellung. In FELLINGER, K. (Hrsg.): Computer in der Medizin - Probleme, Erfahrungen, Projekte. Verlag Brüder Hollinek Wien, 1968, 189-221.
- /13/ FIESCHI, M., M.JOUBERT, D.FIESCHI, M.ROUX: Medical Diagnostic Aid: Application of the Method of Refutation of DAVIS and PUTNAM. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 764-768.
- /14/ PIRTKIEN, R.: Computereinsatz in der Medizin. Georg Thieme Verlag Stuttgart, 1971.
- /15/ SCHILD, W., B.GAVISH, B.LUNENFELD: Computer-Assisted Medical Diagnosis with an Application to Endocrinology. IBM Israel Scientific Center, Technical Report 014, 1974.
- /16/ SPECHTMEYER, H., H.E.WICHMANN, R.GROSS: Kliniknahe computerunterstützte Diagnostik am Beispiel der Hämatologie. Dt.med.Wschr. 103 (1978) 545-549.
- /17/ ADLASSNIG, K.-P., G.GRABNER: Approaches to Computer-Assisted Diagnosis in Gastroenterology. EDV in Medizin und Biologie 11 (1980) 74-80.
- /18/ ADLASSNIG, K.-P., G.GRABNER: The Viennese Computer-Assisted Diagnostic System. Its Principles and Values. AUTOMEDICA 3 (1980) 141-150.
- /19/ ADLASSNIG, K.-P., G.KOLARZ, F.LIPOMERSKY, I.GRÖGER, G.GRABNER: CADIAG-1: A Computer-Assisted Diagnostic System on the Basis of Symbolic Logic and its Application in Internal Medicine. In O'MOORE, R.R., B.BARBER, P.L.REICHERTZ, F.ROGER (Eds.): Medical Informatics Europe 82. Springer-Verlag Berlin-Heidelberg-New York, 1982, 495-505.
- /20/ SALAMON, R., M.BERNADET, F.GREMY: Ternary Algebra: Theoretical Aspects and Possible Applications. In DOMBAL, F.T. de, F.GREMY (Eds.): Decision Making and Medical Care: Can Information Science Help? North-Holland Publishing Company Amsterdam-New York-Oxford, 1976, 505-513.
- /21/ RESCHER, N.: Many-valued logic. McGraw-Hill Book Company New York, 1969.
- /22/ SHORTLIFFE, E.H., B.G.BUCHANAN, E.A.FEIGENBAUM: Knowledge Engineering for Medical Decision Making: A Review of Computer-Based Clinical Decision Aids. Proceedings of the IEEE 67 (1979) 1207-1224.

- /23/ SHORTLIFFE, E.H.: Computer-Based Medical Consultations: MYCIN. Elsevier New York-Oxford-Amsterdam, 1976.
- /24/ WEISS, S., C.A.KULIKOWSKI, A.SAFIR: Glaucoma Consultation by Computer. *Comput.Biol.Med.* 8 (1978) 25-40.
- /25/ WEISS, S.M., C.A.KULIKOWSKI, S.AMAREL, A.SAFIR: A Model-Based Method for Computer-Aided Medical Decision-Making. *Artificial Intelligence* 11 (1978) 145-172.
- /26/ LINDBERG, D.A.B., G.C.SHARP, L.C.KINGSLAND, III, S.M.WEISS, S.P.HAYES, H.UENO, S.E.HAZELWOOD: Computer Based Rheumatology Consultant. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 1311-1315.
- /27/ UENO, H., D.A.B.LINDBERG, L.C.KINGSLAND, III, J.H.ÉSTHER: Design of a Criteria-Based Rheumatology Consulting System for a Microcomputer. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 1316-1320.
- /28/ WEISS, S.M., C.A.KULIKOWSKI: EXPERT: A System for Developing Consultation Models. *Proceedings of the 6th International Joint Conference on Artificial Intelligence Tokyo, 1979, 942-947.*
- /29/ WEISS, S.M., C.A.KULIKOWSKI: Data Base Analysis Aids for Designing Consultation Systems. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 794-797.
- /30/ POPLE, H.E., J.D.MYERA, R.A.MILLER: DIALOG: A Model of Diagnostic Logic for Internal Medicine. *Proceedings of the 4th International Joint Conference on Artificial Intelligence Tbilisi, 1975, 848-855.*
- /31/ POPLE, H.E.: The Formation of Composite Hypotheses in Diagnostic Problem Solving - An Exercise in Synthetic Reasoning. *Proceedings of the 5th International Joint Conference on Artificial Intelligence Cambridge, 1977, 1030-1037.*
- /32/ BUTTON, K.F.: Laboratory Diagnosis by Computer. *Comput.Biol.Med.* 3 (1973) 131-136.
- /33/ GIESZL, L.R.: An Interactive Intelligence System for Medical Decision Aids. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 809-814.
- /34/ LEDLEY, R.S.: Syntax-Directed Concept Analysis in the Reasoning Foundations of Medical Diagnosis. In JACQUEZ, J.A. (Ed.): *Computer Diagnosis and Diagnostic Methods*. Charles C.Thomas Springfield, 1972, 152-168.
- /35/ BARTHEL, D.: Automatische Diagnosestellung. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig* 17 (1968) 515-528.
- /36/ BARTHEL, D.: *Prinzipien der automatischen Diagnosestellung*. VEB Georg Thieme Leipzig, 1970.
- /37/ KAMSTIESS, R., H.G.DINKLOH, M.MÜLLER: Entscheidungstabellentechnik als Mittel der Verständigung zwischen Medizinern und Systemanalytikern. *Informatik-Spektrum* 2 (1980) 79-85.
- /38/ ALBIN, M.A.: *Fuzzy Sets and Their Application to Medical Diagnosis and Pattern Recognition*. PhD diss. University of California, Berkeley, 1975.
- /39/ PEREZ-OJEDA, A.: *Medical Knowledge Network: A Database for Computer-Aided Diagnosis*. Thesis. University of Toronto, Toronto, 1976.
- /40/ MOON, R.E., S.JORDANOV, A.PEREZ, I.B.TURKSEN: Medical Diagnostic System with Human-Like Reasoning Capability. In SHIRES, D.B., H.WOLF (Eds.): MEDINFO 77. North-Holland Publishing Company Amsterdam-New York-Oxford, 1977, 115-119.
- /41/ ESOGBUE, A.O., R.C.ELDER: Fuzzy Sets and the Modelling of Physician Decision Processes, Part I: The Initial Interview-Information Gathering Session. *Fuzzy Sets and Systems* 2 (1979) 279-291.
- /42/ BELLMAN, R.E., L.A.ZADEH: Local and Fuzzy Logics: Memorandum NO.ERL-M584, Electronics Research Laboratory, College of Engineering, University of California, Berkeley 94720, 11 May 1976.
- /43/ DUBOIS, D., H.PRADE: *Fuzzy Sets and Systems: Theory and Applications*. Academic Press New York-London, 1980.

- /44/ ZADEH, L.A.: Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. IEEE Transactions on Systems, Man, and Cybernetics, Vol.SMC-3, No.1, January 1973, 28-44.
- /45/ SMETS, PH., H.VAINSEL, R.BERNARD, F.KORNREICH: Bayesian Probability of Fuzzy Diagnosis. In SHIRES, D.B., H.WOLF (Eds.): MEDINFO 77. North-Holland Publishing Company Amsterdam-New York-Oxford, 1977, 213-218.
- /46/ SANCHEZ, E.: Medical Diagnosis and Composite Fuzzy Relations. In GUPTA, M.M., R.K.RAGADE, R.R.YAGER (Eds.): Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Company Amsterdam-New York-Oxford, 1979, 437-444.
- /47/ SANCHEZ, E.: Compositions of Fuzzy Relations. In GUPTA, M.M., R.K.RAGADE, R.R.YAGER (Eds.): Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Company Amsterdam-New York-Oxford, 1979, 421-433.
- /48/ KAUFMANN, A.: Introduction to the Theory of Fuzzy Subsets. Volume I. Fundamental Theoretical Elements. Academic Press New York-San Francisco-London, 1975.
- /49/ KAUFMANN, A.: Introduction à la Théorie des Sous-Ensembles Flous. Tome II. Application à la Linguistique et à la Sémantique. Masson Paris, 1975.
- /50/ KAUFMANN, A.: Introduction à la Théorie des Sous-Ensembles Flous. Tome III. Applications à la Classification et à la Reconnaissance des Formes aux Automates et aux Systèmes, au Choix des Critères. Masson Paris, 1975.
- /51/ KAUFMANN, A.: Introduction à la Théorie des Sous-Ensembles Flous. Tome IV. Compléments et Nouvelles Applications. Masson Paris, 1977.
- /52/ SANCHEZ, E.: Resolution of Eigen Fuzzy Sets Equations. Fuzzy Sets and Systems 1 (1978) 69-74.
- /53/ SANCHEZ, E., R.SAMBUC: Relations floues. Fonction ϕ -floues. Application à l'aide au diagnostic en pathologie thyroïdienne. In LAUDET, M., J.ANDERSON, F.BEGON (Eds.): Medical Data Processing. Taylor and Francis Ltd. London, 1976, 513-524.
- /54/ JOLY, H., E.SANCHEZ, J.GOUVERNET, J.VALT: Applications of Fuzzy Set Theory to the Evaluation of Cardiac Function. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 91-95.
- /55/ SANCHEZ, E.: Solutions in Composite Fuzzy Relation Equations: Application to Medical Diagnosis in Brouwerian Logic. In GUPTA, M.M., G.N.SARIDIS, B.R.GAINES (Eds.): Fuzzy Automata and Decision Processes. Elsevier North-Holland New York, 1977, 221-234.
- /56/ SOULA, G., J.GOUVERNET, A.BARRE, J.L.SAN MARCO: Application of Fuzzy Relations to Medical Decision Making. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 844-848.
- /57/ SANCHEZ, E.: Inverses of Fuzzy Relations. Application to Possibility Distributions and Medical Diagnosis. Fuzzy Sets and Systems 2 (1979) 75-86.
- /58/ SAMBUC, R., E.SANCHEZ, J.GOUVERNET, J.L.SAN MARCO: ϕ -Fuzzy Functions. Application to Medical Diagnosis. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 784-788.
- /59/ ADLASSNIG, K.-P.: Ein einfaches Modell zur medizinischen Diagnostik mit fuzzy Teilmengen. EDV in Medizin und Biologie 13 (1982) in press.
- /60/ TUSCH, G.: Fuzzy Mengen und Anwendungen. Diplomarbeit. Universität Hannover, Hannover, 1980.
- /61/ TUSCH, G.: Ein fuzzy Algorithmus zur diagnostischen Klassifizierung in der cranialen Computertomographie (CCT). In BRAUER, W. (Ed.): GI-11. Jahrestagung. Springer-Verlag Berlin-Heidelberg-New York, 1981, 598-605.
- /62/ SANCHEZ, E., J.GOUVERNET, R.BARTOLIN, L.VOVAN: Linguistic Approach in Fuzzy Logic of the WHO Classification of Dyslipoproteinemias. International Congress on Applied Systems Research and Cybernetics, Acapulco, 1980.
- /63/ CERUTTI, S., C.T.PIERI: A Method for the Quantification of the Decision-Making Process in a Computer-Oriented

- Medical Record. Int.J.Bio-Medical Computing 12 (1981) 29-57.
- /64/ WECHSLER, H.: A Fuzzy Approach to Medical Diagnosis. Int.J.Bio-Medical Computing 7 (1976) 191-203.
- /65/ FORDON, W.A.: Computer-Aided Differential Diagnosis of Hypertension. PhD diss. Purdue University, Purdue, 1976.
- /66/ FORDON, W.A., J.C.BEZDEK: The Application of Fuzzy Set Theory to Medical Diagnosis: In GUPTA, M.M., R.K. RAGADE, R.R.YAGER (Eds.): Advances in Fuzzy Set Theory and Applications. North-Holland Publishing Company Amsterdam-New York-Oxford, 1979, 445-461.
- /67/ BEZDEK, J.C.: Feature Selection for Binary Data-Medical Diagnosis with Fuzzy Sets. National Computer Conference, AFIPS Press, 1976, 1057-1068.
- /68/ ESOGBUE, A.O., R.C.ELDER: Fuzzy Sets and the Modelling of Physician Decision Processes. Part II: Fuzzy Diagnosis Decision Models. Fuzzy Sets and Systems 3 (1980) 1-9.
- /69/ NOMURA, Y., M.NAKAMURA, H.MISHIMA, N.KATO: An Application of Fuzzy Decision Analysis to the Management of Pancreatic Cancer. Automedica 3 (1980) 131-139.
- /70/ MISHIMA, H., N.KATOH, T.HATTORI, Y. NOMURA, M.NAKAMURA: A Fuzzy Decision Analysis for the Management of Pancreatic Cancer. In LINDBERG, D. A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 830-834.
- /71/ CHENG, Y.-Y.M., B.McINNIS: An Algorithm for Multiple Attribute, Multiple Alternative Decision Problems Based on Fuzzy Sets with Application to Medical Diagnosis. IEEE Transactions on Systems, Man, and Cybernetics, SMC-10 (1980) 645-650.
- /72/ KICKERT, W.J.M.: Fuzzy Theories on Decision-Making. Martinus Nijhoff Social Sciences Division Leiden-Boston-London, 1978.
- /73/ FOX, J.: Some Observations on Fuzzy Diagnosis and Medical Computing. Int.J.Bio-Medical Computing 8 (1977) 269-275.
- /74/ OHSATO, A., T.SEKIGUCHI, T.NAKAI, K.YAMAZAKI: Information Processing for Nursing by Fuzzy Relation. In LINDBERG, D.A.B., S.KAIHARA (Eds.): MEDINFO 80. North-Holland Publishing Company Amsterdam-New York-Oxford, 1980, 62-66.
- /75/ ADLASSNIG, K.-P., G.KOLARZ: CADIAG-2: Computer-Assisted Medical Diagnosis Using Fuzzy Subsets. In GUPTA, M.M., E.SANCHEZ (Eds.): Fuzzy Information and Decision Processes. North-Holland Publishing Company Amsterdam-New York-Oxford. In press.
- /76/ ADLASSNIG, K.-P.: A Fuzzy Logical Model of Computer-Assisted Medical Diagnosis. Meth.Inform.Med. 19 (1980), 141-148.