

CADIAG-2: COMPUTER-ASSISTED MEDICAL DIAGNOSIS  
USING FUZZY SUBSETS

KLAUS-PETER ADLASSNIG, GERNOT KOLARZ

Department of Medical Computer Sciences  
(Director: Prof.Dr. G. Grabner)  
University of Vienna

A-1090 Vienna, Garnisongasse 13, Austria  
and

Ludwig Boltzmann Institute for Rheumatology and Focal Diseases  
(Directors: Prof.Dr. F. Endler and Prof.Dr. N. Thumb)

A-2500 Baden, Kaiser-Franz-Ring 8, Austria

A model of a computer-assisted medical diagnostic system using fuzzy subsets has been developed and implemented. Two aspects of symptom-disease relationships - occurrence and confirmability - are documented linguistically by medical experts or on the basis of medical database evaluations. Symptom combinations and their relationships to diseases are part of the stored medical knowledge.

Given patients' symptom patterns, the diagnostic process provides confirmed and excluded diagnoses as well as diagnostic hypotheses. Detailed explanations for the diagnostic results are given and proposals for further examinations on the patients are made.

An extended application in rheumatology with 1.000 cases including about 2.000 symptoms, signs and test results and 186 diseases is at present being tested.

Keywords: Computer-assisted diagnosis, medical consultation, fuzzy subsets, fuzzy logic, fuzzy relationships, acquisition of medical knowledge, confirmed diagnoses, excluded diagnoses, diagnostic hypotheses, rheumatology, hepatology, coagulation defects, pancreatic diseases.

## 1. INTRODUCTION

In medical science it is seldom possible to work with exact definitions, descriptions or assertions. In medical diagnosis there is very rarely a sharp boundary between diseases and the appearance of more than one disease in

the patient at the same time destroys the expected symptom patterns of disease hypotheses, which makes the diagnostic and therapeutic decision more difficult. The assignment of laboratory test results to the ranges normal or pathological is arbitrary in

borderline cases, the intensity of pain can only be described verbally and depends on the subjective estimation of the patient, and precise relationships between symptoms, signs, test results, findings, i.e. any observation on the patient (short: symptoms) and diagnoses can very seldom be found in descriptions of diseases.

Thus you might read in /1/ the following statements about acute pancreatitis:

- Acute pancreatitis is almost always connected with sickness and vomiting (p. 678).
- Typically, acute pancreatitis begins with sudden aches in the abdomen (p. 677).
- The case history frequently reports about ulcerus ventriculi and duodeni (p. 676).

In /2/ one can find similar medical descriptions, e.g. for the general differential diagnosis of icterus:

- Bilirubinurie excludes the hemolytic icterus but bilirubin is detectable with hepatocellular or cholestatic icterus (p. 561).

In the field of rheumatology, ZÖLLNER /3/ writes about the attack of gout:

- The typical attack of gout occurs almost always in men after the 25th year of life (p. 677).
- It is characterized by sudden start out of full health, by restriction to one joint, by strong aches, by the intensive inflammation with rubor, swelling and heat and by the almost always good therapeutic success with colchicin (p.677).

- But the attack of gout does not always proceed in the above-mentioned course (p. 677).

In /4,5/ one can read similarly weak assertions and descriptions of diseases.

Firm relationships between symptoms or combinations of them and diagnoses sometimes appear and it is necessary to use them, but they are rare. So, KÖLARZ gives in /6/, partly according to the American Rheumatism Association, clearly defined criteria for the rheumatology diseases rheumatic fever, rheumatoid arthritis, Sjögren's disease, psoriatic arthritis, Reiter's disease, Behcet's disease, Bechterew's disease, systematic lupus erythematosus, sclerodermia and gout. Certain combinations of symptoms prove these diagnoses if excluding symptoms are missing.

In the computer-assisted diagnostic system CADIAG-1 based on tri-valued logic /7-16/ the firm relationships proving, excluding and obligatory (as well as the weak relationships facultative and not proving) between symptoms and diseases are used in the diagnostic process.

The theory of fuzzy subsets developed by L.A.ZADEH /17/ (general literature see /18-28/, fuzzy bibliographies see /29-31/) is an attempt at a mathematical theory of vagueness and imprecision. This theory seems very useful in the development of computer-assisted medical diagnostic models because of its power to formalize soft and complex concepts.

Such a model is able to handle weak and firm medical assertions like

- symptom  $S_i$  is often present and seldom proving for diagnosis  $D_j$  or

- strongly increased symptom  $S_i$  is almost always present and always proving for diagnosis  $D_j$ .

On this basis, the computer-assisted diagnostic system CADIAG-2 establishes fuzzy logical relationships between the patient's symptom pattern and diagnoses in order to deduce medical diagnoses logically and to reason them. In the case of an insufficient diagnostic output, proposals for further examinations on the patient are established by the system. Co-occurrence of diseases which is very common in internal medicine is treated by restarting the diagnostic procedure at the end of a diagnostic iteration with unexplained patient's symptoms.

A survey on medical diagnosis and fuzzy sets is given in a companion paper in the same edition /32/.

## 2. INTENDED AIMS

The following requirements are imposed on the computer-aided medical consultation system CADIAG-2:

- 1) Medical diagnostic knowledge shall be stored in form of logical relationships between symptoms and diseases, between symptoms, between diseases and between symptom combinations and diseases. The system shall be capable of gathering and formalizing empirical and judgemental as well as statistical knowledge.

Relationships between single symptoms and single diseases shall express the degree of certainty to infer from one symptom to a specific diagnosis. Relationships between symptoms and between diseases shall be introduced in order to obtain hierarchical dependencies, i.e. super-

and sub-relations among symptoms and among diseases. Relationships between symptom combinations and diseases shall incorporate the kind of medical knowledge that considers clinically found combinations of symptoms having a high evidence for diagnoses, or combinations of symptoms proposed as internal guideline principles by medical associations for the confirmation or exclusion of diagnoses.

- 2) The logical relationships may be fuzzy. They do not necessarily correspond to Boolean or tri-valued logic /7-16/.

- 3) Frequent as well as rare diagnoses shall be offered after analyzing the patient's medical data. The only criterion for offering diagnoses shall be that they match patient's symptom pattern.

- 4) The diagnostic process shall be performed iteratively. In this way it will be possible to confirm or exclude diagnostic hypotheses step by step. Furthermore, it shall be possible to put in confirmed diagnoses at the beginning of an iteration. The diagnostic procedure shall then erase symptoms explained by confirmed diagnoses and continue with the unexplained symptoms. Such a procedure shall help to solve the problem of co-occurrence of diseases in a patient that is very common in internal medicine /8,13,33,34/.

- 5) Explanations for every diagnostic output during the consultation session of CADIAG-2 shall be given by the system. This central requirement is imposed to insure that

- CADIAG-2 is accepted by the medical community and that
- CADIAG-2 is used for educational pur-

poses in medical students' and professional training.

6) Meaningful proposals for patient's further examination shall be provided, with consideration of the financial costs of the examination and its risk for the patient.

7) Retrospective applications for scientific studies as well as prospective use in clinical practice shall be possible.

### 3. CADIAG-2 SYSTEM: THEORETICAL ASPECTS <sup>1)</sup>

#### 3.1. General considerations

Let  $\Sigma = \{S_1, S_2, \dots, S_m\}$  be the set of symptoms,  $\Delta = \{D_1, D_2, \dots, D_n\}$  the set of diseases or diagnoses <sup>2)</sup> resp. and  $\Pi = \{P_1, P_2, \dots, P_q\}$  the set of patients under consideration. The cardinality of  $\Sigma, \Delta$  and  $\Pi$  is given by  $m, n$  and  $q$ .  $\Sigma, \Delta$  and  $\Pi$  are non-fuzzy sets.

REMARK 1:

Every  $S_i \in \Sigma, 1 \leq i \leq m$ , is a fuzzy subset of a reference set <sup>3)</sup>  $\Omega = \{x_1, x_2, \dots\}$  characterized by a membership function  $\mu_{S_i}(x)$ . Set  $\Omega$  contains all possible values  $x$  which can be assumed by  $S_i$ . The membership function  $\mu_{S_i}(x)$  defines the strength of affiliation of  $x \in \Omega$  in  $S_i$ .

Example:

Let  $S_i$  be potassium, increased in mmol/liter then  $\Omega = \{1.0 \dots 15.0\}$  (more than 15.0 mmol and less than 1.0 mmol liter are outside the possible range) defines the reference set. Assuming a normal range of measured potassium from about 3.5 mmol/liter to about 5.2 mmol/liter yields

$$\begin{aligned} \mu_{\text{potassium, increased}}(5.2) &= 0.5 & \text{or} \\ \mu_{\text{potassium, increased}}(6) &= 1. \end{aligned}$$

REMARK 2:

Every  $D_i \in \Delta, 1 \leq i \leq n$ , is a fuzzy subset of a reference set  $\Pi = \{P_1, P_2, \dots, P_q\}$  characterized by a membership function  $\mu_{D_i}(p)$ . The function  $\mu_{D_i}(p)$  assigns to every patient  $P_j \in \Pi$  a degree of membership of  $P_j \in \Pi$  in  $D_i$  (Figure 1).

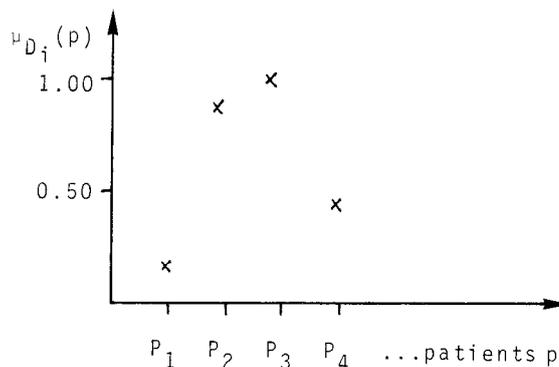


Fig. 1: Diagnosis  $D_i$  as a fuzzy subset of the reference set  $\Pi$  that contains all patients  $p$  under consideration.

The following definition is equivalent to the above-mentioned:

Every  $P_j \in \Pi, 1 \leq j \leq q$ , is a fuzzy subset of a reference set  $\Delta = \{D_1, D_2, \dots, D_n\}$  characterized by a membership function  $\mu_{P_j}(d)$ . The function  $\mu_{P_j}(d)$  assigns to every diagnosis  $D_i \in \Delta$  a degree of membership of  $D_i \in \Delta$  in  $P_j$  (Figure 2).

It will be noticed that for  $d=D_i$  and  $p=P_j$

$$\mu_{D_i}(P_j) = \mu_{P_j}(D_i). \tag{1}$$

Example:

Given three patients  $P_1, P_2$  and  $P_3$  and let  $D_1$  be gout, then, for example,  $\mu_{\text{gout}}(P_1) = 1$ , i.e. patient  $P_1$  suffers from

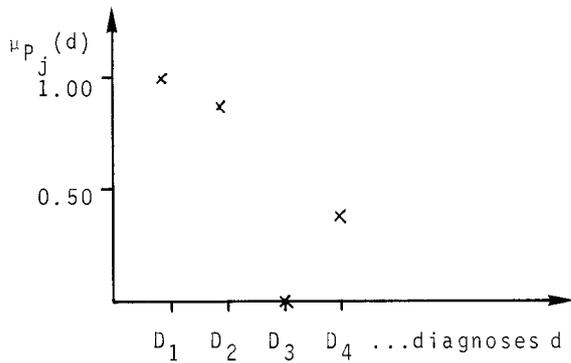


Fig. 2: Patient  $P_j$  as a fuzzy subset of the reference set  $\Delta$  that contains all diagnoses  $d$  under consideration.

gout,  $\mu_{\text{gout}}(P_2)=0.7$ , i.e. patient  $P_2$  shows signs of gout, and  $\mu_{\text{gout}}(P_3)=0$ , i.e. patient  $P_3$  does not suffer from gout. It is obvious that  $\mu_{\text{gout}}(P_1)=\mu_{P_1}(\text{gout})$ , etc.

REMARK 3:

From a geometrical point of view (see ALBIN /38/), patients  $p$  are considered as points in a  $m$ -dimensional space which is set by symptoms  $S_i \in \Sigma$ ,  $1 \leq i \leq m$ . Every patient  $p$  exhibiting symptom values  $x_i \in \Omega_i$  is associated to degrees of membership  $\mu_{D_j}(p) \in [0,1]$ , where  $D_j \in \Delta$ ,  $1 \leq j \leq n$ . The membership values  $\mu_{D_j}(p)$  show the affiliation of patient  $p$  to diagnoses  $D_j$ . In other notation, patient  $p$  is always represented by a symptom value pattern, thus

$$\mu_{D_j}(p) = \mu_{D_j}(x_1, x_2, \dots, x_m) \tag{2}$$

where  $D_j \in \Delta$ ,  $1 \leq j \leq n$ , and  $x_1 \in \Omega_1$ ,  $x_2 \in \Omega_2, \dots, x_m \in \Omega_m$ . Patients suffering from the same disease show a similar <sup>4)</sup> symptom value pattern. They form clusters in the space. Every cluster represents a disease but because  $\mu_{D_j}(p)$  is not necessarily 0 or 1 the disease clusters do not show sharp boundaries. To illustrate this see Figure 3.

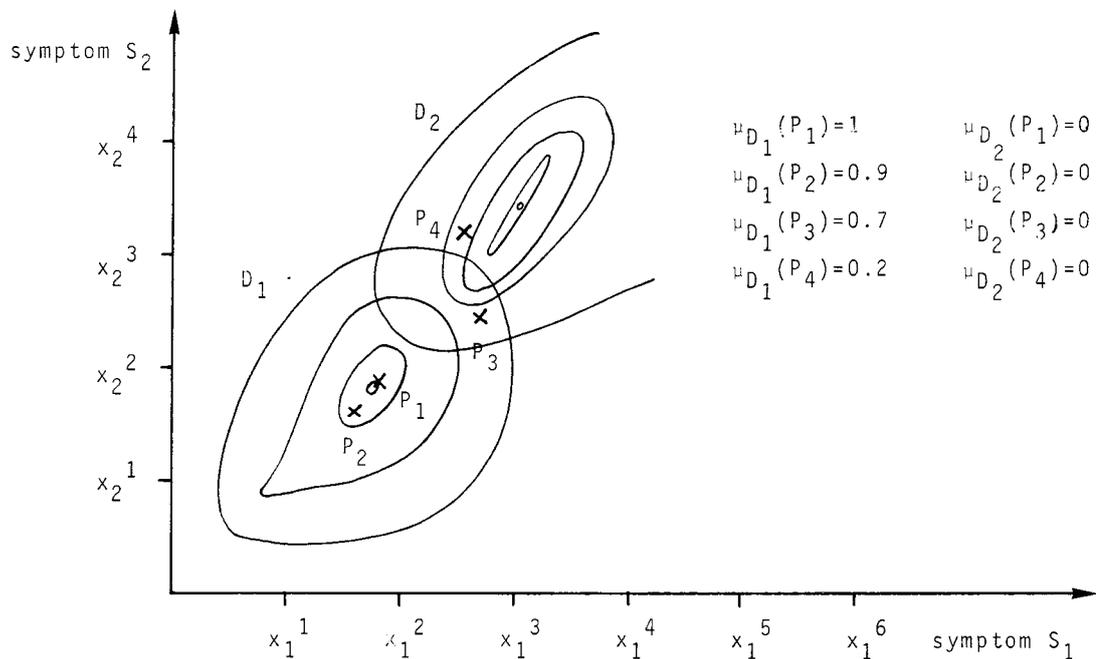


Fig. 3: Patients  $P_i$  suffering from the same disease  $D_j$  form clusters without sharp boundaries.

3.2. Representation and acquisition of medical knowledge

3.2.1. Symptom-disease fuzzy relationships

Two aspects of a symptom  $S_i$  are of essential value in order to find out its relation to a disease  $D_j$ :

- 1) occurrence of  $S_i$  in case of  $D_j$  and
- 2) confirmability of  $S_i$  for  $D_j$ .

Occurrence provides knowledge about the presence of symptom  $S_i$  with a patient  $p$  when he suffers from disease  $D_j$ . From a statistical point of view occurrence is the frequency of  $S_j$  with  $D_j$ .

Confirmability describes the power of symptom  $S_i$  to confirm a certain diagnosis  $D_j$  if it occurs in patient  $p$ . Example:

Symptom dyspnoe often occurs in patients suffering from scleroderma but dyspnoe has such a small discriminating value in differential diagnosis of rheumatic diseases that its confirmability for scleroderma is low. In contrast, symptom chorea minor occurs very seldom at patients suffering from rheumatic fever but if it occurs it confirms rheumatic fever.

In terms of fuzzy set theory occurrence of  $S_i$  at  $D_j$  is defined as a fuzzy subset  $O$  of the reference set  $\Phi_0(x) = \{0, 1, 2, \dots, 100\}$ , where  $x$  means the frequency of  $S_i$  in  $x$  of one hundred cases of  $D_j$ . Similarly, confirmability of  $S_i$  for  $D_j$  is determined as a fuzzy subset  $C$  of the reference set  $\Phi_C(x) = \{0, 1, 2, \dots, 100\}$ , i.e.  $S_i$  has been confirmed  $D_j$  in  $x$  of one hundred cases.

It must be stated that the above-defined reference sets  $\Phi_0$  and  $\Phi_C$  are based on

relative values of frequencies. They have been defined relative to 100. Also, the frequencies do not have to be statistically determined. They can be directly defined, estimated subjectively or based on judgemental knowledge (see paragraph 3.2.2.)

The membership functions of  $O$  and  $C$  are very simple. They are:

$$\mu_O(x) = f_1(x; 1, 50, 99) \quad , x \in \Phi_0 \quad (3)$$

$$\mu_C(x) = f_1(x; 1, 50, 99) \quad , x \in \Phi_C \quad (4)$$

where  $f_1(x; \alpha, \beta, \gamma)$  is the following standardized function (see ZADEH /39/):

$$f_1(x; \alpha, \beta, \gamma) = \begin{cases} 0 & \text{for } x \leq \alpha \\ 2 \left( \frac{x-\alpha}{\gamma-\alpha} \right)^2 & \text{for } \alpha < x \leq \beta \\ 1 - 2 \left( \frac{x-\gamma}{\gamma-\alpha} \right)^2 & \text{for } \beta < x < \gamma \\ 1 & \text{for } x > \gamma. \end{cases} \quad (5)$$

The membership functions  $\mu_O(x)$  and  $\mu_C(x)$  are shown in Figure 4.

It is also possible to use another monotonously increasing function to define  $O$  and  $C$ , e.g.  $\mu_O(x) = \mu_C(x) = \frac{x}{100}$ .

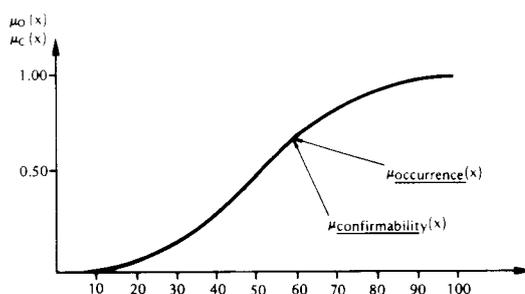


Fig. 4: Membership functions  $\mu_O(x)$  for occurrence of  $S_i$  with  $D_j$  and  $\mu_C(x)$  for confirmability of  $S_i$  for  $D_j$ .

3.2.2. Acquisition of symptom-disease fuzzy relationships

There are two different ways (and a combination of the two) to determine the occurrence and confirmability relationships between symptoms and diseases:

- 1) Linguistic documentation by medical experts and
- 2) Medical database evaluation by statistical means.

3.2.2.1. Linguistic documentation

Symptom-disease relationships can be documented by predefined terms that express the interconnections between the considered medical entities. A medical expert or, in order to advance the acceptance of the documented knowledge by the medical community, a team of diagnostic experts document  $S_i D_j$  occurrence relationships as well as  $S_i D_j$  confirmability relationships by answering the following questions:

- 1) Occurrence: "How often does  $S_i$  occur with  $D_j$ ?"
- 2) Confirmability: "How strongly does  $S_i$  confirm  $D_j$ ?"

In this way judgemental medical knowledge founded either directly on medical textbooks (examples have been given in paragraph 1.) or on a clinician's experience can be incorporated. No statistical or computational assumptions (availability of patients data to estimate probabilities of symptoms or diseases in a population or to estimate conditional probabilities of symptoms that appear in case of a disease, availability of sample cases with or without fixed assignments to diseases in order to adapt discriminant equations or to com-

pute disease clusters, etc.) have to be fulfilled to determine  $S_i D_j$  relationships.

The possible set of physician's answers to specify occurrence and confirmability may be the fuzzy subsets  $O_i$  and  $C_i$ ,  $1 \leq i \leq 9$ , and the answer "unknown" (Table 1).

Table 1: Fuzzy subsets for occurrence and confirmability for physician's linguistic documentation of symptom-disease relationships.

i	<u>occurrence</u> $O_i$	<u>confirmability</u> $C_i$
1	<u>always</u> 5)	<u>always</u> 6)
2	<u>almost always</u>	<u>almost always</u>
3	<u>very often</u>	<u>very often</u>
4	<u>often</u>	<u>often</u>
5	<u>unspecific</u>	<u>unspecific</u>
6	<u>seldom</u>	<u>seldom</u>
7	<u>very seldom</u>	<u>very seldom</u>
8	<u>almost never</u>	<u>almost never</u>
9	<u>never</u> 7)	<u>never</u>
	unknown	unknown

The fuzzy subsets  $O_i$  and  $C_i$  have the same reference sets  $\Phi_{O_i}$  and  $\Phi_{C_i}$  as the reference sets of  $O_i$  and  $C_i$ , i.e.

$$\Phi_{O_i} = \Phi_{O_i} \quad \text{and} \quad \Phi_{C_i} = \Phi_{C_i}, \quad 1 \leq i \leq 9. \quad (6)$$

The fuzzy subsets  $O_1, O_2, O_4, O_5, O_6, O_8$  and  $O_9$  are defined by the following membership functions 8) :

$$\begin{aligned}
 \mu_{\text{always}}(x) &= f_1(x; 97, 98, 99) \\
 \mu_{\text{almost always}}(x) &= f_1(x; 80, 85, 90) \\
 \mu_{\text{often}}(x) &= f_1(x; 40, 60, 80) \\
 \mu_{\text{unspecific}}(x) &= f_2(x; 20, 50) \\
 \mu_{\text{seldom}}(x) &= 1 - f_1(x; 20, 40, 60)
 \end{aligned} \quad (7)$$

$$\begin{aligned} \mu_{\text{almost never}}(x) &= 1 - f_1(x; 10, 15, 20) \\ \mu_{\text{never}}(x) &= 1 - f_1(x; 1, 2, 3) \end{aligned}$$

where  $f_2(x; \alpha, \beta)$  is also a standardized function (see /39/)

$$f_2(x; \alpha, \beta) = \begin{cases} f_1(x; \beta - \alpha, \beta - \frac{\alpha}{2}, \beta) & \text{for } x \leq \beta \\ 1 - f_1(x; \beta, \beta + \frac{\alpha}{2}, \beta + \alpha) & \text{for } x > \beta. \end{cases} \quad (8)$$

By using the modified form of the following operation on fuzzy subsets (see /25/), also called linguistic hedge /19,40/ or linguistic modifier /39/,

$$\text{concentration: } \mu_{\text{very } A}(x) = [\mu_A(x)]^2 \quad (9)$$

the determination of  $O_3$  and  $O_7$  leads to

$$\begin{aligned} \mu_{\text{very often}}(x) &= [\mu_{\text{often}}(x)]^4 \\ \mu_{\text{very seldom}}(x) &= [\mu_{\text{seldom}}(x)]^4. \end{aligned} \quad (10)$$

The definition of  $C_1, C_2, \dots, C_9$  will be performed in analogy to the operation for  $O_1, O_2, \dots, O_9$ . The membership functions for  $O_i$  and  $C_i$  are shown in Figure 5.

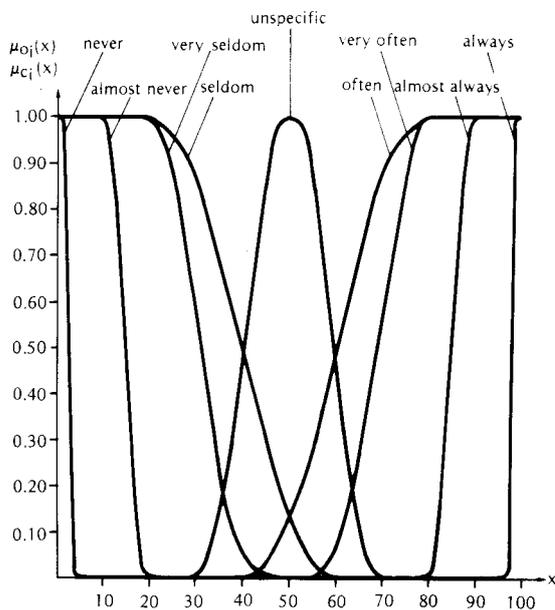


Fig. 5: Fuzzy membership functions of  $O_i$  and  $C_i$ .

Having defined the fuzzy subsets  $O_i$  and  $C_i$ , the following question arises:

What numerical degree of affiliation of symptom  $S_i$  to disease  $D_j$  shall be selected when the medical expert who documents the  $S_i D_j$  relationships chooses a linguistic representation  $O_i$  or  $C_i$  for characterizing the symptom-disease relationships?

The answer is given in the following manner:

It seems understandable that, at the moment of documenting a certain  $S_i D_j$  relation, the physician imagines "fuzzy" ranges for describing occurrence and confirmability. If, for example, the physician claims that  $S_i$  very often appears with  $D_j$ , then he selects a "fuzzy" range from about 70 to 80 occurrences of  $S_i$  with  $D_j$  in his mind (relative to 100) and terms this "fuzzy" range very often. The same is valid for confirmability. If, for example, the physician chooses always in order to describe confirmability of  $S_i$  for  $D_j$  then he considers a "fuzzy" range of about 99 to 100, i.e. medical studies have produced the result that in 99 or 100 cases  $S_i$  confirms  $D_j$ .

In order to get suitable ranges for all linguistic terms  $O_i$  and  $C_i$  that are labels of fuzzy subsets it seems rational to consider those ranges of the reference sets  $\Phi_0$  and  $\Phi_C$  where the membership functions  $\mu_{O_i}(x) \geq 0.5$  and  $\mu_{C_i}(x) \geq 0.5$  resp. Because of

$$\begin{aligned} O_1 \subset O_2 \subset O_3 \subset O_4 & \quad \text{and} \\ O_9 \subset O_8 \subset O_7 \subset O_6, & \end{aligned} \quad (11)$$

with  $\subset$  as the inclusion <sup>10)</sup> of fuzzy subsets which is defined (see /17/) as

$$A \subset B \text{ if } \mu_A(x) \leq \mu_B(x), x \in X,$$

where A and B are fuzzy subsets of the reference set X, the ranges are overlapping. Assuming ranges  $r_{0_i}$  given by  $0_i$  we exclude

$r_{0_1}$  from  $r_{0_2}$ ,  $r_{0_2}$  from  $r_{0_3}$ ,  $r_{0_3}$  from  $r_{0_4}$   
and  
 $r_{0_9}$  from  $r_{0_8}$ ,  $r_{0_8}$  from  $r_{0_7}$ ,  $r_{0_7}$  from  $r_{0_6}$

to overcome this problem.

To give an example, consider  $0_2 = \text{almost always}$  that defines  $r_{0_2} = 85-100$ . But because  $0_1 = \text{always}$  occupies the range 98-100 we calculate  $r_{0_2} = 85-98$ . The analogous way is valid for the computation of ranges  $r_{C_i}$  for  $C_i$ .

Subsequently, we calculate the means  $x_{0_i}$  and  $x_{C_i}$  from the determined  $r_{0_i}$  and  $r_{C_i}$ . We now consider the means to be the numerical representatives of the linguistic terms  $0_i$  and  $C_i$  (11). By obtaining the means  $x_{0_i}$  and  $x_{C_i}$  which are elements of the reference sets  $\mu_{0_i}(x)$  and  $\mu_{C_i}(x)$  we are able to determine  $\mu_{0_i}(x_{0_i})$  and  $\mu_{C_i}(x_{C_i})$  for every  $i$ ,  $1 \leq i \leq 9$ .

Figure 6 shows the considered ranges  $r_{0_i}$  and  $r_{C_i}$  and the calculated means  $x_{0_i}$  and  $x_{C_i}$ .

Table 2 presents a survey of the determined  $\mu_{\text{occurrence}}(x)$  and  $\mu_{\text{confirmability}}(x)$  based on the linguistic possibilities to document  $0_i$  and  $C_i$ .

3.2.2.2. Medical database evaluation

In Hospital Information Systems /14,15, 41,42/ medical databases are applied to store both patients personal and medical data. On the one hand patients medical data are the basis for daily routine in.hospital (test result tables, drug tables, statistics of available beds, etc.).

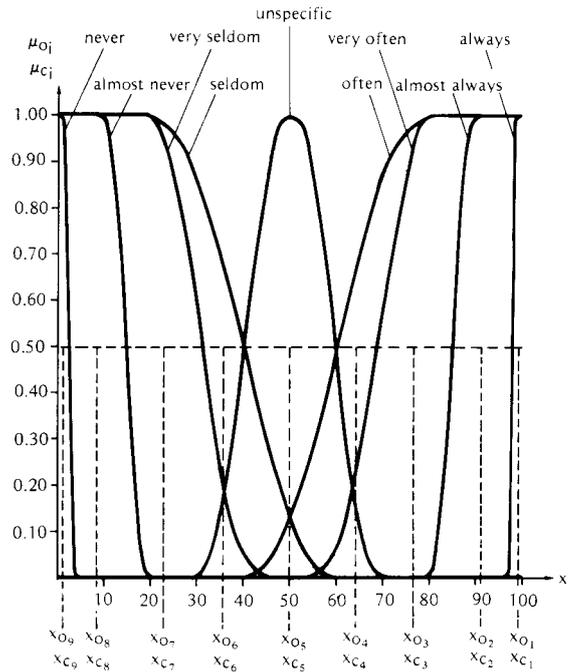


Fig. 6: Ranges  $r_{0_i}$  and  $r_{C_i}$  and the calculated means  $x_{0_i}$  and  $x_{C_i}$  for the fuzzy subsets  $0_i$  and  $C_i$ .

On the other hand patients' medical data are utilized to carry out medical research (analysis of symptom patterns that characterize diseases, course of patient recovery from diseases after administration of certain drugs, etc.).

If a medical database contains a big amount of medical records of patients including all relevant symptoms as well as final diagnoses established by clinicians at discharge of patients from hospital or after the treatment period of outpatients, it can be used to calculate  $S_i D_j$  relationships. With respect to the two aspects of a  $S_i D_j$  relationship

Table 2: Determination of  $\mu_{\text{occurrence}}(x)$  and  $\mu_{\text{confirmability}}(x)$  from the linguistic possibilities to document  $O_i$  and  $C_i$ .

Fuzzy Subsets $O_i$ or $C_i$	Ranges $r_{O_i}$ or $r_{C_i}$	Means $x_{O_i}$ or $x_{C_i}$	Degrees of Membership $\mu_{O_i}(x_{O_i})$ or $\mu_{C_i}(x_{C_i})$
<u>always</u>	98 to 100	99	1.00
<u>almost always</u>	85 to 98	91.5	$\sim 0.99$
<u>very often</u>	$\sim 69$ to 85	$\sim 77$	$\sim 0.90$
<u>often</u>	60 to $\sim 69$	$\sim 64.5$	$\sim 0.75$
<u>unspecific</u>	40 to 60	50	0.50
<u>seldom</u>	$\sim 31$ to 40	$\sim 35.5$	$\sim 0.25$
<u>very seldom</u>	15 to $\sim 31$	$\sim 23$	$\sim 0.10$
<u>almost never</u>	2 to 15	8.5	$\sim 0.01$
<u>never</u>	0 to 2	1	0.00

the following considerations are made:

#### Consideration 1:

Occurrence of a symptom  $S_i$  in case of a disease  $D_j$  is defined by the relative frequency of occurrence of  $S_i$  with  $D_j$ .

Assuming  $x_2$  cases of  $D_j$  and counting a frequency of  $x_1$  for  $S_i$ , we calculate a relative frequency  $x_0$  by

$$x_0 = \frac{x_1}{x_2} \cdot 100. \quad (12)$$

We are now able to compute a degree of membership of  $x_0$  in the fuzzy subset occurrence by applying (3)

$$\mu_{O_i}(x_0) = f_1(x_0; 1, 50, 99). \quad (13)$$

Example:

Let us assume 67 patients suffering from gout. In 13 cases we found tophi at the patients. We calculate

$$x_0 = \frac{13}{67} \cdot 100 = 19.4.$$

This yields

$$\mu_{O_i}(x_0) = f_1(19.4; 1, 50, 99) \approx 0.07 \quad (12).$$

#### Consideration 2:

Confirmability of a symptom  $S_i$  for a disease  $D_j$  is defined by the relative frequency of occurrence of  $D_j$  with  $S_i$ . Assuming  $x_4$  cases of  $S_i$  and counting a frequency of  $x_3$  for  $D_j$  we calculate a relative frequency  $x_C$  by

$$x_C = \frac{x_3}{x_4} \cdot 100. \quad (14)$$

In analogy to the above we are able to compute a degree of membership of  $x_C$  in the fuzzy subset confirmability by applying (4)

$$\mu_{C_i}(x_C) = f_1(x_C; 1, 50, 99). \quad (15)$$

Example:

Let us assume 13 cases of tophi. We found in all these cases the diagnosis gout. We calculate

$$x_C = \frac{13}{13} \cdot 100 = 100$$

and obtain

$$P_C(x_C) = f_1(100; 1, 50, 99) = 1.00^{13}.$$

### 3.2.2.3. Problems and conclusions

A lot of problems appear in practical attempts at determining symptom-disease relationships:

#### 1) Linguistic documentation

##### Problem 1:

$S_i D_j$  relationships can be documented linguistically but only after the symptom-disease relationships have been analysed in medical studies. But because not all  $S_i D_j$  relationships will have been investigated,  $P_0(x)$  or  $P_C(x)$  are often impossible to define<sup>14)</sup>.

##### Problem 2:

Medical studies have often yielded different results considering the same symptom-disease relationship where the percentage of occurrence covers two or more linguistic terms. In /1/ (p. 676), for example, we read that "...C...could find abusus of alcohol in 70% and K... and H...in 40% of the cases of acute pancreatitis."

#### 2) Medical database evaluation

##### Problem 1:

Patients go to a hospital several times suffering from several diseases. The first problem to get statistical material is to consider patients' data only from one stay

in hospital or from one series of visits in an outpatient department. In patient's records such a period is characterized by the dates of admission, discharge and the final diagnoses established by the responsible medical person /14,15/.

##### Problem 2:

A medical history, a general status, X-ray, ECG, EEG, histological, ultrasonic, and rheumatological findings, etc. are made during one stay in hospital. Additionally, several laboratory tests are performed. For the most part, examinations are two, three and more times in the course of one illness period. Now the question arises which examination results should be considered with respect to the final diagnoses. In CADIAG-2, this problem is solved by taking the first values of these examinations that are performed twice or more times during one stay in hospital, to avoid effects of therapy.

Due to these considerations the results are well-established patients' symptom patterns connected to final diagnoses (Table 3).

##### Problem 3:

In order to count frequencies  $S_i$  with  $D_j$  and  $D_j$  with  $S_i$ , symptom patterns with only one final diagnosis were necessary. In general, such data are not available in internal medicine. The process of studying and solving this problem has not been completed yet. So far, the frequencies of  $S_i$  with  $D_j$  are calculated by taking every symptom  $S_i$  in patients' symptom patterns into account considering a certain disease  $D_j$  (vice versa for calculating the frequencies of  $D_j$  with  $S_i$ ).

Table 3: Patients' pattern in a medical database which are the starting point for statistical evaluations.

patients	patients with final diagnoses	symptom pattern
Patient 1	D <sub>1</sub> D <sub>2</sub> D <sub>3</sub> D <sub>4</sub>	S <sub>1</sub> S <sub>2</sub> S <sub>3</sub> S <sub>7</sub> ...
Patient 1 <sup>15)</sup>	D <sub>3</sub> D <sub>5</sub> D <sub>6</sub>	S <sub>2</sub> S <sub>10</sub> S <sub>20</sub> S <sub>30</sub> ...
Patient 2	D <sub>7</sub>	S <sub>100</sub> S <sub>101</sub> S <sub>102</sub> ...
Patient 3	D <sub>4</sub> D <sub>6</sub> D <sub>7</sub> D <sub>8</sub> D <sub>9</sub> D <sub>10</sub>	S <sub>3</sub> S <sub>10</sub>
Patient 4	D <sub>11</sub> D <sub>12</sub>	S <sub>100</sub> S <sub>101</sub> S <sub>102</sub> ...
:	:	
:	:	

Example (from Table 3):

Occurrence:

S <sub>1</sub> with D <sub>1</sub> : 1 of 1 → x <sub>0</sub> =100	μ <sub>0</sub> (x <sub>0</sub> )=1
S <sub>2</sub> with D <sub>1</sub> : 1 of 1 → x <sub>0</sub> =100	μ <sub>0</sub> (x <sub>0</sub> )=1
:	:
S <sub>1</sub> with D <sub>3</sub> : 1 of 2 → x <sub>0</sub> =50	μ <sub>0</sub> (x <sub>0</sub> )=0.5
:	:
S <sub>7</sub> with D <sub>6</sub> : 0 of 2 → x <sub>0</sub> =0	μ <sub>0</sub> (x <sub>0</sub> )=0

Confirmability:

D <sub>1</sub> with S <sub>1</sub> : 1 of 1 → x <sub>C</sub> =100	μ <sub>C</sub> (x <sub>C</sub> )=1
D <sub>1</sub> with S <sub>2</sub> : 1 of 2 → x <sub>C</sub> = 50	μ <sub>C</sub> (x <sub>C</sub> )= 0.5
:	:
D <sub>3</sub> with S <sub>1</sub> : 1 of 1 → x <sub>C</sub> =100	μ <sub>C</sub> (x <sub>C</sub> )= 1
:	:
D <sub>6</sub> with S <sub>7</sub> : 0 of 1 → x <sub>C</sub> =0	μ <sub>C</sub> (x <sub>C</sub> )=0

But it is obvious that with the use of this method, errors can and will appear.

Example:

Hypertension is a very common disease not directly related to any rheumatic disease. Elevated blood pressure is the obligatory symptom in hypertension, retinal disorders may occur. In a given patient gonarthrosis, a degenerative knee joint disease, and hypertension due to constriction of blood vessels may be found. If the patient's symptom pattern is used for the described evaluation, high blood pressure and eye disturbances are linked not only to hypertension but also to gonarthrosis, a disease which certainly has nothing to do with eye troubles.

3.2.3. Symptom-symptom fuzzy relationships

Apart from the knowledge about symptom-disease relationships, medical knowledge also includes judgement about relations among symptoms. Symptoms very often exclude the presence of other symptoms. On the other hand, symptoms often imply the presence of symptoms or have to be necessarily present with other symptoms. CADIAG-2 takes this into account by means of always and never occurrence and always and never confirmability relationships among symptoms.

Always occurring and always confirming relationships express hierarchical dependencies between symptoms. The sub-term confirms the super-term. It follows that the super-term is obligatory, i.e. it is always occurring in respect of each of its sub-terms. If the super-symptom is set to be "not present" then every sub-symptom has to be set to "not present".

Table 4: Symptom-symptom and disease-disease relationships considered in CADIAG-2 with their degrees of membership in occurrence and confirmability.

S <sub>i</sub> S <sub>j</sub> - and D <sub>i</sub> D <sub>j</sub> -relationships			
<u>occurrence</u>	<u>confirmability</u>	$\mu_0(x)$	$\mu_C(x)$
<u>always</u> ( <u>obligatory</u> and <u>proving</u> )	<u>always</u>	1	1
<u>unknown</u> ( <u>facultative</u> and <u>proving</u> )	<u>always</u>	-	1
<u>always</u> ( <u>obligatory</u> and <u>not proving</u> )	<u>unknown</u>	1	-
<u>never</u> ( <u>excluding</u> )	<u>never</u>	0	0
<u>unknown</u>	<u>unknown</u>	-	-

Examples:

Symptom GOT, normal excludes GOT, increased, GOT, highly increased, GOT, decreased. Further, diffuse swelling, right hand as well as diffuse swelling, left hand confirm diffuse swelling, hand or recurrent high fever confirms high fever. But if the diffuse swelling, hand is denied, it can be concluded that neither a diffuse swelling of the right hand nor of the left hand is present.

CADIAG-2 only considers strong symptom-symptom relationships. The relation-

ships taken into account and the attached degrees of membership in the fuzzy subsets occurrence of S<sub>i</sub> for S<sub>j</sub> and confirmability of S<sub>i</sub> for S<sub>j</sub> are shown in Table 4. The fuzzy subsets occurrence and confirmability are defined in the same manner as described in paragraph 3.2.1.

During the diagnostic consultation, CADIAG-2 uses S<sub>i</sub>S<sub>j</sub> relationships in a symptom-symptom inference module. The input of the module consists of patients' symptom patterns established from the patients' medical database /14,15/ or from physician's input. The output that is the starting point for the inner diagnostic process <sup>16)</sup> consists of an enlarged symptom pattern including the original and the inferred symptoms attached with their cause.

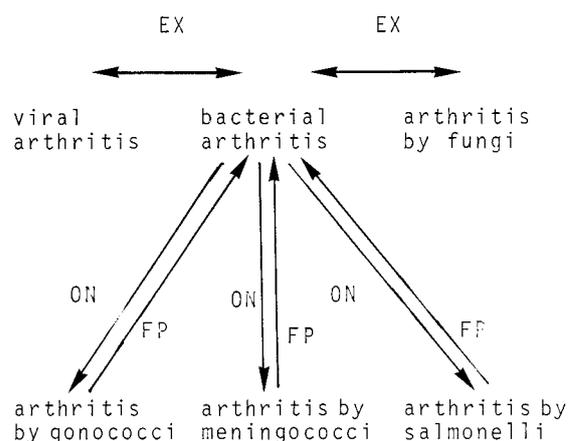
#### 3.2.4. Disease-disease fuzzy relationships

Medical knowledge about diseases to a great extent includes hierarchical concepts. In order to incorporate hierarchical disease dependencies into CADIAG-2, disease-disease relationships have been established.

Examples:

A bacterial arthritis can be caused by gonococci, meningococci, salmonelli or others. Therefore, sub-terms of bacterial arthritis are bacterial arthritis by gonococci, - by meningococci, - by salmonelli etc. On the other hand, an acute arthritis by fungi or viruses excludes a bacterial arthritis (see Figure 7).

In analogy to the definition of relationships among symptoms, only strong disease-disease relationships are taken into



EX: excluding ( $\mu_0(x)=0; \mu_C(x)=0$ )

FP: facultative and proving ( $\mu_C(x)=1$ )

ON: obligatory and not proving ( $\mu_0(x)=1$ )

Fig. 7: Section of disease-disease relationships in CADIAG-2.

account <sup>17)</sup> (see Table 4).

Disease-disease relationships are used in CADIAG-2, in a similar way as symptom-symptom relationships, in a disease-disease inference module. The module is activated in two different parts of CADIAG-2. Firstly, the input of the module consists of a patient's known diseases (patient's past history or result of a previous iteration of CADIAG-2) or secondly, of the diagnostic output of the inner diagnostic process. In the first case, the patients known disease pattern will be enlarged. In the second case it will be ascertained if confirmed diagnoses imply other diagnoses or if excluded diagnoses are compulsory for diagnoses not yet excluded.

### 3.2.5. Symptom combination-disease fuzzy relationships

A physician often attributes a low value to a single symptom but if it appears together with other symptoms, this combination can be a strong indication for a specific diagnosis. Computer-assisted diagnostic systems have to incorporate this kind of medical knowledge.

Most clinically found symptom combinations give only a slight evidence for diseases (see MYCIN /43/). Furthermore, established clinical symptom combinations are often not accepted in the entire medical community to be a definite criterion to prove or exclude a disease. Different medical schools insist on their own medical concepts.

Therefore the possibility to establish "soft" or "fuzzy" relationships between symptom combinations and diseases should be given. Strong, proving combinations of symptoms for diseases with consideration of excluding criteria have been established and were published as international guiding principles (see, for example, the diagnostic criteria of the American Rheumatism Association for rheumatoid arthritis /44/ or of the American Heart Association for rheumatic fever /45/), but they are rare.

In CADIAG-2, symptom combinations are considered to be fuzzy logical combinations (LC) <sup>18)</sup> with symptoms and diseases as fuzzy variables connected by the fuzzy logical operations "conjunction", "disjunction", "negation", "minimal" and "maximal" (see below). Moreover, combinations have fuzzy relationships to diseases.

## 1. Example:

The symptom combination low back pain, limitation of motion of the lumbar spine, and diminished chest expansion in a man between 20 and 40 years is highly characteristic for ankylosing spondylitis. But in various rare diseases, e.g. psoriatic spondylitis, Reiter's syndrome, enteropathic arthritis, the same symptom combination may occur.

## 2. Example (from American Heart Association publication /45/):

In order to diagnose rheumatic fever the following criteria have to be considered:

Major criteria:

1. Carditis (SCP1)
2. Polyarthritis (SCP2)
3. Synderham's chorea (SCP3)
4. Erythema marginatum (SCP4)
5. Subcutaneous nodules (SCP5)

Minor criteria:

6. Fever (SCP6)
7. Arthralgia (SCP7)
8. History of previous rheumatic fever or rheumatic cardiopathia (SCP8)
9. Elevated ESR (SCP9)
10. Positive C-reactive protein or leucocytosis (SCP10)
11. Prolonged PR-interval (SCP11)

The diagnosis rheumatic fever is established if two major criteria or one major criterion and two minor criteria are fulfilled and if, additionally, a preceding streptococcal infection (SCP12) (elevated antistreptolysin titer or streptococci in throat swab or precedent scarlatina) can be detected. The absence of the latter

should make the diagnosis questionable, except in situations where rheumatic fever is only discovered after a long latent period from the antecedent infection (e.g., Synderham's chorea or low grade-carditis).

Because of the fact that symptoms considered to be diagnostic criteria are often "higher level" symptoms that express a pathophysiological state or similar concepts<sup>19)</sup> a two level definition of symptom combinations has been realized in CADIAG-2.

In the first step the "higher level" symptoms are established from directly observed symptoms  $S_i \in \Sigma$  and given diseases  $D_j \in \Delta$  of the patient. "Higher level" symptoms are called symptom combinations primary (SCP) in CADIAG-2. In this way the establishing of the "higher level" symptom

"joint aches in at least one joint when subjected to passive motion or pressure" yields "pain on passive motion or tenderness in one MCP v in one PIP v in one DIP v in one wrist v in one elbow v in one shoulder v in one acromioclavicular v in one sternoclavicular v in one jaw v in one hip v in one knee v in one ankle v in one toe joint."

This has been defined as a major criterion for rheumatoid arthritis by the American Rheumatism Association /44/. In practical use the sets  $\Sigma$  and  $\Delta$  are given by using hierarchical thesauri that form the basis for the documentation of symptom-disease relationships. Beside this, every "higher level" symptom is named and included in a thesaurus of SCP. The decision as to whether or not a medical term is considered to be a symptom or a SCP is made by the documenting physician. So, for

Table 5: Rules of the contextfree language that defines the syntactical structure of symptom combination primary and secondary.

rules		
<symptom combination secondary>	::=<identifier secondary>=<expression secondary>;	
<identifier secondary>	::=SCS<integer>	← mapping in a thesaurus of symptom combinations secondary
<expression secondary>	::=<factor secondary> <factor secondary> <factor secondary> *	
<factor secondary>	::= <variable secondary>	
<variable secondary>	::=<identifier primary> <expression secondary> <mima-term secondary>	
<mima-term secondary>	::= 1 1 * <integer>/<integer>:<variable list secondary>	← "left" integers="right" integer=number of expressions secondary in variable list secondary
<variable list secondary>	::=<expression secondary> <variable list secondary> *	
<symptom combination primary>	::=<identifier primary>=<expression primary>;	
<identifier primary>	::=SCP<integer>	← mapping in a thesaurus of symptom combinations primary
<expression primary>	::=<factor primary> <factor primary> <factor primary> *	
<factor primary>	::= <variable primary>	
<variable primary>	::=<symptom> <disease> (<expression primary>) <mima-term primary>	
<symptom>	::=S<integer>	← mapping in a symptom thesaurus
<disease>	::=D<integer>	← mapping in a diagnosis thesaurus
<mima-term primary>	::= 1 1 * <integer>/<integer>:<variable list primary>	← "left" integers≤"right" integer=number of expressions primary in variable list primary
<variable list primary>	::=<expression primary> <variable list primary> *	
<integer>	::=<digit> <integer>	
<digit>	::=0 1 2 3 4 5 6 7 8 9	

instance, we included carditis in the thesaurus of SCP because it is determined by "pancarditis v endocarditis v pericarditis" which are all defined as symptoms in Z.

Table 5 presents the rules of a context-free language that defines the syntactical structure of SCP. The definition of the language has been carried out in extended BACKUS-NAUR form (see /49/). The syntax can be analyzed by a recursive descent parser /49-51/.

Examples:

SCP1=S1vS2vS3;  
(carditis=pancarditis v endocarditis v pericarditis)

SCP2=D96vD97;  
(clinical pattern of an acute infectious arthritis=acute arthritis, pyogenic microorganism v acute arthritis, non-pyogenic microorganism)

Symptom and diagnoses are considered to be fuzzy variables taking their values in [0,1]. Let  $x_1, x_2, \dots$  be fuzzy variables, in general the following operations are defined:

$$\begin{aligned} \text{conjunction}(v): x_1 \wedge x_2 &= \text{MIN}(x_1, x_2) \\ \text{disjunction}(v): x_1 \vee x_2 &= \text{MAX}(x_1, x_2) \\ \text{negation } (\neg): \neg x_1 &= 1 - x_1 \end{aligned} \quad (16)$$

and

$$\begin{aligned} \text{minimal } (\uparrow): & \uparrow N_1/N_2: x_1; x_2; \dots; x_{N_2}; \\ \text{maximal } (\downarrow): & \downarrow N_1/N_2: x_1; x_2; \dots; x_{N_2}; \end{aligned} \quad (17)$$

where  $N_1$  and  $N_2$  are natural numbers with  $N_1 \leq N_2$ .

Connectives (16) are well-known (see DUBOIS and PRADE /19/), the "minimal" and "maximal" connectives (17) (20) have to be explained:

In Boolean logical sense  $\uparrow N_1/N_2: x_1; \dots; x_{N_2};$  means that at least  $N_1$  of the  $N_2$  logical variables  $x_1; x_2; \dots; x_{N_2}$  have to be true to make the result true /9/. Another notation would be (example):

$$\uparrow 2/3: x_1; x_2; x_3; \equiv (x_1 \wedge x_2) \vee (x_1 \wedge x_3) \vee (x_2 \wedge x_3).$$

In fuzzy logic, this latter notation is rational and yields

$$\uparrow N_1/N_2: x_1; x_2; \dots; x_{N_2}; \equiv \text{MIN}(N_1 \text{ greatest values of } x_1; x_2; \dots; x_{N_2}); \quad (18)$$

As a pendant the "maximal" operation was introduced. In Boolean logic

$\downarrow N_1/N_2: x_1; x_2; \dots; x_{N_2};$  means that at the most  $N_1$  of the  $N_2$  logical variables are or can yet be true to make the result true /9/. In equivalent notation (example):

$$\downarrow 2/4: x_1; x_2; x_3; x_4; \equiv \sim((x_1 \wedge x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge x_4) \vee (x_1 \wedge x_3 \wedge x_4) \vee (x_2 \wedge x_3 \wedge x_4));$$

This makes obvious that the maximal operator is expressable in fuzzy logic by

$$\downarrow N_1/N_2: x_1; x_2; \dots; x_{N_2}; \equiv 1 - \text{MIN}(N_1 + 1 \text{ greatest values of } x_1; x_2; \dots; x_{N_2}); \quad (19)$$

Let us summarize the first step in establishing symptom combinations. Symptom  $S_i$  and diseases  $D_j$  are connected

by the logical operations "conjunction", "disjunction", "negation", "minimal" and "maximal" to symptom combinations primary  $SCP_k \in \Lambda_1 = \{SCP_1, SCP_2, \dots, SCP_s\}$ , where  $\Lambda_1$  is the non-fuzzy set of SCP taken into account and  $s$  its cardinality. Given fuzzy values  $x_i = \mu_{S_i}(x)$  for  $S_i$  and  $x_j = \mu_{D_j}(x)$  for  $D_j$  a degree of affiliation  $\mu_{SCP_k}(x_1, x_2, \dots, x_n)$

is calculated with  $x_i, x_j \in x_1, x_2, \dots, x_n$  and  $n$  as the number of  $S_i$  and  $D_j$  taken into account to establish  $SCP_k$ . The SCP rules are of the form:

$$LC(S_i, D_j) \xrightarrow{\mu_{SCP_k}(x_i)} SCP_k \quad (20)$$

In the second step of the two level definition of symptom combinations, fuzzy logical combinations of "higher level" symptoms are established that have fuzzy relationships to diseases. These combinations are called symptom combinations secondary (SCS) (21).

$SCS_1 \in \Lambda_2 = \{SCS_1, SCS_2, \dots, SCS_t\}$  where  $\Lambda_2$  defines the non-fuzzy set of the SCS taken into account, and  $t$  its cardinality. Table 5 gives the rules to define the syntactical structure of SCS. The syntax can be analyzed by a recursive descent parser.

Example (see 2<sup>nd</sup> example stated above): Considering the major and minor criteria of rheumatic fever as SCP, the SCS is given by

$$\begin{aligned} SCS1 = & ((\uparrow 2/5: SCP1; SCP2; SCP3; SCP4; SCP5; \vee \\ & (\uparrow 1/5 : SCP1; SCP2; SCP3; SCP4; SCP5; \wedge \\ & \uparrow 2/6: SCP6; SCP7; SCP8; SCP9; SCP10; SCP11;)) \\ & \wedge SCP12) \vee (SCP3 \wedge SCP8); \end{aligned}$$

Let us summarize the second step in establishing symptom combinations.  $SCP_k$

are connected by the logical operations "conjunction", "disjunction", "negation", "minimal" and "maximal" to  $SCS_1 \in \Lambda_2$ . Given fuzzy values  $x_i = \mu_{SCP_k}(x)$  for  $SCP_k$ , a degree of affiliation  $\mu_{SCS_1}(x_1, x_2, \dots, x_n)$  is calculated with  $x_i \in x_1, x_2, \dots, x_n$  and  $n$  as the number of  $SCP_k$  taken into account to establish  $SCS_1$ . The SCS rules are of the form:

$$LC(SCP_k) \xrightarrow{\mu_{SCP_k}(x_i)} SCS_1. \quad (21)$$

Now, in analogy to symptom-disease fuzzy relationships, two aspects are of high value to find out relationships between SCS and diseases:

- 1) Occurrence of  $SCS_1$  in case of  $D_j$  and
- 2) Confirmability of  $SCS_1$  for  $D_j$ .

These two aspects can be defined in the same way as for symptom-disease relationships. The membership functions for occurrence and confirmability are analogous to Figure 4. The fuzzy relationships between  $SCS_1$  and  $D_j$  are defined linguistically by medical experts. The possible set is analogous to Table 1.

Examples (see 1<sup>st</sup> and 2<sup>nd</sup> examples above):

1. Example:

The SCS for ankylosing spondylitis often occurs in ankylosing spondylitis, but as mentioned above this symptom combination does not confirm the diagnosis ankylosing spondylitis.

2. Example:

The SCS for rheumatic fever often occurs with rheumatic fever, but if it occurs, it always confirms rheumatic fever.

### 3.3. Symptoms as fuzzy subsets

All observable properties of a patient are used in medical diagnosis. These can be symptoms, signs, laboratory test results and complaints related to the patient (also information about his family and his living circumstances) at any observation time. Let us term every property as a symptom.

As stated above, every symptom  $S_i \in \Sigma$ ,  $1 \leq i \leq m$ , is a fuzzy subset of a reference set  $\Omega = \{x_1, x_2, \dots\}$  characterized by a membership function  $\mu_{S_i}(x)$ .

Let us now distinguish two different kinds of reference sets  $\Omega$ .

1.  $\Omega = \{x_1, x_2, \dots\}$  contains continuous values, e.g. as  $\Omega$  for  $S =$  increased alkaline phosphatase.
2.  $\Omega = \{0, 1\}$ , i.e. the symptom  $S_i$  is a binary symptom, which can be absent (0) or present (1) (see also /52, 53/), e.g. as  $\Omega$  for  $S =$  drug history, medicine against rheumatism (see Figure 8).

Most of the laboratory test results and some of the signs observed during physical examination (e.g. SCHÖBER distance, normal  $\geq 4$  cm, pathological  $< 4$  cm) provide continuous values as a result of the investigation, and most of symptoms of patient's history, complaints, evidence or infection, X-ray findings, etc. are binary symptoms. Some of the binary symptoms are real ones, e.g. drug history, medicine against rheumatism, others are not, e.g. headache. But we do not want to establish fuzzy subsets to define the severity level of symptoms, because so far it does not seem practical to calculate numerical values for, e.g. severe headache.

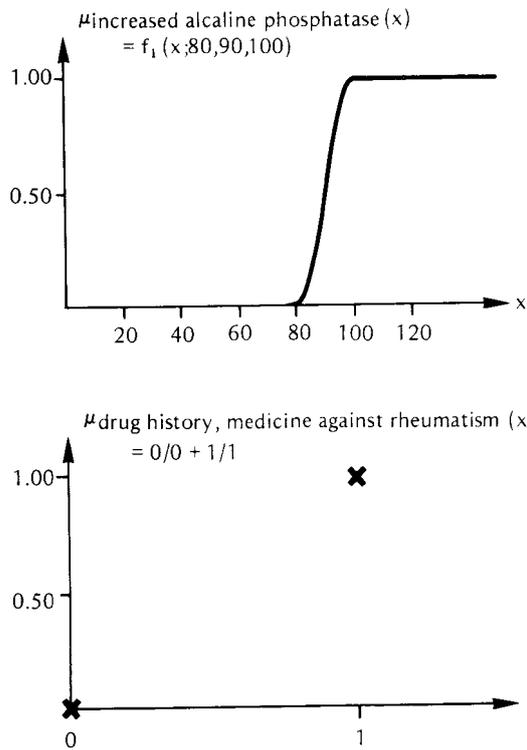


Fig. 8: Symptom membership functions.

All symptoms taken into account may now be defined to have one of the two kinds of reference sets. Uninvestigated symptoms are not considered.

3.4. Diagnostic fuzzy indications

Given a patient's symptom pattern both the

- symptom-disease relationships and the
- symptom combination-disease relationships and the
- diseases-disease relationships

yield fuzzy diagnostic indications that are the basis for establishing confirmed and excluded diagnosis as well as diagnostic hypotheses.

3.4.1. Indications from symptom-disease relationships

Let us introduce a binary fuzzy relation  $R_0$  which is a fuzzy subset of  $\Sigma \times \Delta$  and characterized by the two parameter membership function  $\mu_{R_0}(S_i, D_j)$  in the interval  $[0, 1]$ . The  $\mu_{R_0}(S_i, D_j)$  are identical to  $\mu_0(x)$  which has been documented for  $S_i$  with regard to  $D_j$  as occurrence relationship.

The confirmability relationships form the elements of the binary fuzzy relation  $R_C \subseteq \Sigma \times \Delta$ . Thus,  $R_C$  is defined by  $\mu_{R_C}(S_i, D_j)$  which is equal to  $\mu_C(x)$  considering the symptom  $S_i$  and diagnosis  $D_j$ .

Furthermore, we introduce the binary fuzzy relation  $R_S \subseteq \Sigma \times \Sigma$ . The sets  $\Pi$  and  $Z$  are given when the symptom patterns of the patients under consideration are available. Symptoms which have not been investigated but are considered in the documentation work of the physician have to be dropped from the relation matrices  $R_P$  and  $R_C$ . Now, the functions  $\mu_{R_S}(p, S_i)$  which characterize the fuzzy relation  $R_S$  are identical with  $\mu_{S_i}(x)$  calculated for the patient  $p$  and corresponding to the defined fuzzy subset  $S_i$ .

Finally, four different fuzzy indications are calculated by means of fuzzy relation compositions:

1.  $S_i D_j$  occurrence indication  $R_1 = R_S \circ R_0$   
 $\mu_{R_1}(p, D_j) = \max_{S_i} \min \{ \mu_{R_S}(p, S_i); \mu_{R_0}(S_i, D_j) \}$  (22)
2.  $S_i D_j$  confirmability indication  $R_2 = R_S \circ R_C$   
 $\mu_{R_2}(p, D_j) = \max_{S_i} \min \{ \mu_{R_S}(p, S_i); \mu_{R_C}(S_i, D_j) \}$  (23)
3.  $S_i D_j$  non-occurrence indication  $R_3 = R_S \circ (1 - R_0)$   
 $\mu_{R_3}(p, D_j) = \max_{S_i} \min \{ \mu_{R_S}(p, S_i); 1 - \mu_{R_0}(S_i, D_j) \}$  (24)

4.  $S_i D_j$  non-symptom indication  $R_4 = (1 - R_S) \circ R_0$   
 $\mu_{R_4}(p, D_j) = \max_{S_i} \min \{1 - \mu_{R_S}(p, S_i); \mu_{R_0}(S_i, D_j)\}$  (25)

where  $p \in \Pi, S_i \in \Sigma, 1 \leq i \leq m$  and  $D_j \in \Delta, 1 \leq j \leq n$ .

3.4.2. Indications from symptom combination-disease relationships

The binary fuzzy relation  $R_0$  which is a fuzzy subset of  $\Lambda_2 \times \Delta$  is characterized by the membership function  $\mu_{R_0}(SCS_1, D_j)$ . The  $\mu_{R_0}(SCS_1, D_j)$  express the  $SCS_1 D_j$  occurrence relationships documented by a medical expert.

A further  $R_C \subset \Lambda_2 \times \Delta$  is introduced. The characteristic function  $\mu_{R_C}(SCS_1, D_j)$  expresses the  $SCS_1 D_j$  confirmability relationship for  $SCS_1$  and a disease  $D_j$ . Moreover, it is necessary to introduce a binary fuzzy relation  $R_{SCS} \subset \Pi \times \Lambda_2$ . Given patients  $p$  with their symptoms  $S_i$  and known diseases  $D_j$ , the  $\mu_{R_{SCS}}(p, SCS_1)$  can be calculated as illustrated in Figure 9.

In analogy to (22), (23), (24), and (25), relations between patients  $p$  and disease  $D_j$  can be determined by

1.  $SCS_1 D_j$  occurrence indication  $R_5 = R_{SCS} \circ R_0$   
 $\mu_{R_5}(p, D_j) = \max_{SCS_1} \min \{ \mu_{R_{SCS}}(p, SCS_1); \mu_{R_0}(SCS_1, D_j) \}$  (26)

2.  $SCS_1 D_j$  confirmability indication  $R_6 = R_{SCS} \circ R_C$   
 $\mu_{R_6}(p, D_j) = \max_{SCS_1} \min \{ \mu_{R_{SCS}}(p, SCS_1); \mu_{R_C}(SCS_1, D_j) \}$  (27)

3.  $SCS_1 D_j$  non-occurrence indication  $R_7 = R_{SCS} \circ (1 - R_0)$   
 $\mu_{R_7}(p, D_j) = \max_{SCS_1} \min \{ \mu_{R_{SCS}}(p, SCS_1); 1 - \mu_{R_0}(SCS_1, D_j) \}$  (28)

4.  $SCS_1 D_j$  non-SCS indication  $R_8 = (1 - R_{SCS}) \circ R_0$   
 $\mu_{R_8}(p, D_j) = \max_{SCS_1} \min \{ 1 - \mu_{R_{SCS}}(p, SCS_1); \mu_{R_0}(SCS_1, D_j) \}$  (29)

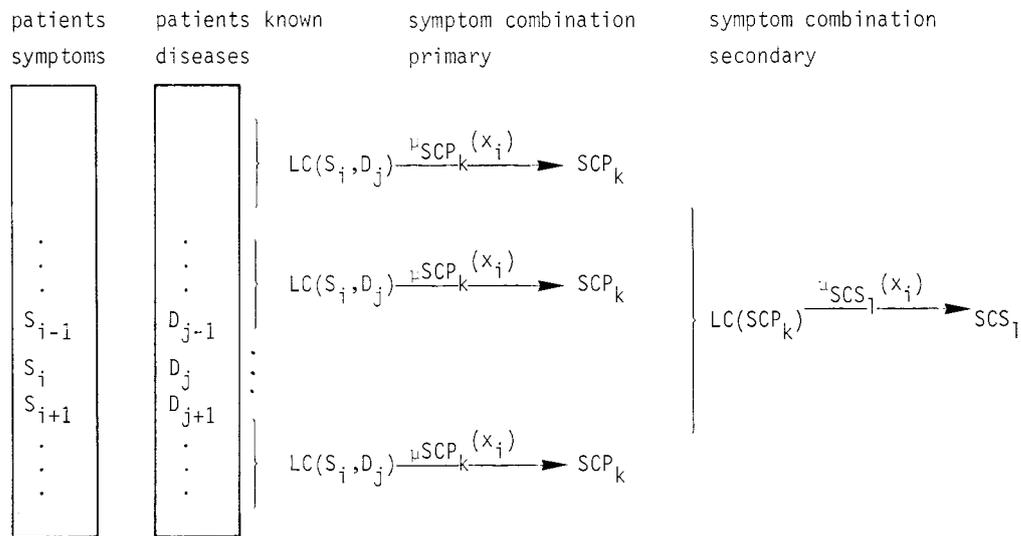


Fig. 9: Process of calculating the membership functions  $\mu_{SCS_1}(x)$  for  $SCS_1$ .

where  $p \in \Pi$ ,  $SCS_i \in \Lambda_2$ ,  $1 \leq i \leq s$  and  $D_j \in \Delta$ ,  $1 \leq j \leq n$ .

3.4.3. Indications from disease-disease relationships

As stated in paragraph 3.2.4, disease-disease fuzzy relationships are established in CADIAG-2 and can be used to infer diagnoses from confirmed or excluded diagnoses provided by the last diagnostic iteration. The representation of this consideration by means of compositions of fuzzy relations yields the following:

1.  $D_i D_j$  occurrence indication  $R_9 = R_D \circ R_0^*$   

$$\mu_{R_9}(p, D_j) = \max_{D_i} \min\{\mu_{R_D}(p, D_i); \mu_{R_0^*}(D_i, D_j)\} \quad (30)$$

2.  $D_i D_j$  confirmability indication  $R_{10} = R_D \circ R_c^*$   

$$\mu_{R_{10}}(p, D_j) = \max_{D_i} \min\{\mu_{R_D}(p, D_i); \mu_{R_c^*}(D_i, D_j)\} \quad (31)$$

3.  $D_i D_j$  non-occurrence indication  $R_{11} = R_D \circ (1 - R_0^*)$   

$$\mu_{R_{11}}(p, D_j) = \max_{D_i} \min\{\mu_{R_D}(p, D_i); 1 - \mu_{R_0^*}(D_i, D_j)\} \quad (32)$$

4.  $D_i D_j$  non-disease indication  $R_{12} = (1 - R_D) \circ R_0^*$   

$$\mu_{R_{12}}(p, D_j) = \max_{D_i} \min\{1 - \mu_{R_D}(p, D_i); \mu_{R_0^*}(D_i, D_j)\} \quad (33)$$

3.5. Diagnostic results

By using the above-defined relations  $R_1, R_2, \dots, R_{12}$ , only three categories of diagnostic results are distinguished:

1) Confirmed diagnoses

All  $D_j$  are put out as confirmed diagnoses if either

$$\mu_{R_2}(p, D_j) = 1.00 \quad (34)$$

or

$$\mu_{R_6}(p, D_j) = 1.00 \quad \text{or} \quad (35)$$

by considering  $\mu_{R_2}(p, D_j)$  or  $\mu_{R_6}(p, D_j)$

as  $\mu_{R_D}(p, D_i)$  in (31),

$$\mu_{R_{10}}(p, D_j) = 1.00. \quad (36)$$

2) Excluded diagnoses

All  $D_j$  are displayed as excluded diagnoses if either

$$\mu_{R_3}(p, D_j) = 1.00 \quad (37)$$

$$\mu_{R_4}(p, D_j) = 1.00 \quad \text{or} \quad (38)$$

$$\mu_{R_7}(p, D_j) = 1.00 \quad (39)$$

$$\mu_{R_8}(p, D_j) = 1.00. \quad (40)$$

or, by considering  $\mu_{R_2}(p, D_j)$ ,  $\mu_{R_6}(p, D_j)$  or  $\mu_{R_{10}}(p, D_j)$  as  $\mu_{R_D}(p, D_i)$  in (32),

$$\mu_{R_{11}}(p, D_j) = 1.00 \quad (41)$$

or, by considering  $\mu_{R_3}(p, D_j)$ ,  $\mu_{R_4}(p, D_j)$ ,  $\mu_{R_7}(p, D_j)$ ,  $\mu_{R_8}(p, D_j)$  and  $\mu_{R_{11}}(p, D_j)$  as  $\mu_{R_D}(p, D_i)$  in (33),

$$\mu_{R_{12}}(p, D_j) = 1.00. \quad (42)$$

3) Diagnostic hypotheses

All diagnoses  $D_j$  where a degree of membership  $\mu_H(p, D_j)$  of either

$$0.50 < \mu_H(p, D_j) = \max\{\mu_{R_1}(p, D_j); \mu_{R_2}(p, D_j)\} \quad (43)$$

or

$$0.50 < \mu_H(p, D_j) = \max\{\mu_{R_5}(p, D_j); \mu_{R_6}(p, D_j)\} \quad (44)$$

or

$$0.50 < \mu_H(p, D_j) = \max\{\mu_{R_9}(p, D_j); \mu_{R_{10}}(p, D_j)\}. \quad (45)$$

is calculated are considered to be diagnostic hypotheses.

3.6. Reasonings for diagnostic results

The acceptance of a computer-assisted medical consultation system is highly dependent on the ability of the system

Table 6: Reasonings for confirmed diagnoses.

a) $D_j$ proved by a symptom $S_i(R_2)$				
Name of $S_i$	Numerical value incl. units of $S_i$	$\mu_{R_S}(p, S_i)$ and the function that calculated the degree of membership	$\mu_{R_0}(S_i, D_j)$ and its linguistic correspondence	$\mu_{R_C}(S_i, D_j)$ and its linguistic correspondence
b) $D_j$ proved by a symptom combination secondary $SCS_1(R_6)$				
Name of $SCS_1$		$\mu_{R_{SCS}}(p, SCS_1)$ and the logical combination of $SCS_1$	$\mu_{R_0}(SCS_1, D_j)$ and its linguistic correspondence	$\mu_{R_C}(SCS_1, D_j)$ and its linguistic correspondence
Names of $SCP_k$		$\mu_{R_{SCP}}(p, SCP_k)$ and the logical combinations of $SCP_k$		
Names of $S_i$	Numerical value incl. units of $S_i$	$\mu_{R_S}(p, S_i)$ and the function that calculated the degree of membership		
Names of $D_j$		$\mu_{R_D}(p, D_j)$ and its cause		
c) $D_j$ proved by a diagnosis $D_i(R_{10})$				
Name of $D_i$		$\mu_{R_D}(p, D_i)$	$\mu_{R_0}(D_i, D_j)$ and its linguistic correspondence	$\mu_{R_C}(D_i, D_j)$ and its linguistic correspondence

to explain its diagnostic output. In CADIAG-2, reasonings for diagnoses displayed as confirmed or excluded diagnoses or as diagnostic hypotheses are given on request.

The following kinds of reasonings for confirmed diagnoses are possible (see Table 6).

The reasoning process for excluded

diagnoses and diagnostic hypotheses are performed similarly.

### 3.7. Proposals for patient's further examination

To give the physician proposals for further examinations of the patient after the first diagnostic consultation of CADIAG-2 is one of the main objectives in establishing computer-

aided medical decision aids. In rheumatology, for example, the first diagnostic consultation starts with symptoms and signs from patient's rheumatological history. If no confirmed diagnoses can be found or if some symptoms in the patient's symptom pattern remain unexplained by the established confirmed and excluded diagnoses and diagnostic hypotheses, CADIAG-2 proposes further examinations. In the first step CADIAG-2 proposes class-1 symptoms, e.g. symptoms that can be routinely checked. In the second step it proposes class-2 symptoms, e.g. simple lab tests, and class-3 symptoms, e.g. tests that are expensive to carry out or invasive procedures. Thus the tests to be performed can be optimized with respect to their financial costs and their risk for the patient.

3.8. Explanation of patient's symptoms by diagnostic results

The aim of the diagnostic iterations is to obtain diagnostic results which together explain patient's symptom pattern, i.e. no symptoms, signs, or test results that the patient shows must remain unexplained. If symptoms remain unexplained then the physician consulting CADIAG-2 has to decide if an unimportant symptom is in question or not. For example, if headache remains unexplained then the physician has to decide if it originates in, e.g. "too much liquor last night", or if a brain tumor is possible and the diagnostic process has to be continued. Table 7 shows the different kinds of explaining patient's symptoms.

Table 7: Explanation of patient's symptoms by diagnostic results.

diagnostic results	explain
confirmed diagnoses and diagnostic hypotheses $D_j$	present symptoms $S_i$ with $S_i D_j$ relationships $\mu_0(S_i, D_j) > 0.50$ or $\mu_C(S_i, D_j) > 0.50$ .
excluded diagnoses	present symptoms $S_i$ with $S_i D_j$ relationships $\mu_0(S_i, D_j) < 0.50$ .
—	absent symptoms have not to be explained.

4. ILLUSTRATIVE EXAMPLE

Let us consider a patient with the following symptom pattern:

Table 8: Section of patients' symptom pattern.

symptoms	measured or present	fuzzy value $\mu(p, S_i)$
$S_1$	} 37.5°C	1.00
$S_2$		0.60
$S_3$	Yes	1.00
$S_4$	Yes	1.00
$S_5$	Yes	1.00
$S_6$	Yes	1.00
$S_7$	Yes	1.00
$S_8$	?	?
$S_9$	Yes	1.00

Table 9: Section of symptom-disease relationships.

Disease $D_j$		$D_1$ <u>rheumatic fever</u>		$D_2$ <u>infectious arthritis by staphylococci</u>	
		$\mu_0(S_i, D_1)$	$\mu_C(S_i, D_1)$	$\mu_0(S_i, D_2)$	$\mu_C(S_i, D_2)$
$S_1$	<u>fever</u>	<u>often</u>		<u>very often</u>	
$S_2$	<u>elevated fever</u>	<u>often</u>		<u>very often</u>	
$S_3$	<u>hydrops of the knee</u>	<u>seldom</u>		<u>seldom</u>	
$S_4$	<u>carditis</u>	<u>very often</u>	<u>often</u>	<u>almost never</u>	
$S_5$	<u>joint pain</u>	<u>almost always</u>		<u>almost always</u>	
$S_6$	<u>erythema</u>	<u>often</u>		<u>very seldom</u>	
$S_7$	<u>previous tonsillitis</u>	<u>very often</u>		<u>very seldom</u>	
$S_8$	<u>synovial fluid examination staphylococci</u>	<u>never</u>	<u>never</u>	<u>always</u>	<u>always</u>
$S_9$	<u>AST elevated</u>	<u>almost always</u>		<u>seldom</u>	

We now imagine the following documented symptoms of rheumatic fever and infectious arthritis by staphylococci (see Table 9).

In the first diagnostic iteration the following diagnostic results are calculated:

Table 10: Results of the medical diagnostic process.

Diagnostic results	Reasonings for diagnostic results	Proposals for further examinations
Diagnostic hypotheses:		
$D_1$	$S_5=1.00$ $O^0=\underline{\text{almost always}}$ $S_4=1.00$ $C=\underline{\text{often}}$	$S_8$ $O^0=\underline{\text{never}}$ $C=\underline{\text{never}}$
$D_2$	$S_5=1.00$ $O^0=\underline{\text{almost always}}$	$S_8$ $O^0=\underline{\text{always}}$ $C=\underline{\text{always}}$

Performing the synovial fluid examination on patient P and finding  $S_8=\text{staphylococci}$  would prove  $D_2=\text{infectious arthritis by staphylococci}$  as the correct diagnosis of P.

### 5. DISCUSSION AND CRITISM

So far the CADIAG-2 system has been designed and implemented but not yet tested with a representative number of hospital patients. Design considerations and the presentation of the control structure of CADIAG-2 are not included in this paper.

Nevertheless, some main points that give an impression of the scope of CADIAG-2 and of the way of establishing it in clinical practice shall be mentioned.

The frame and rule based medical know-

ledge system of CADIAG-2 consists of about 2.500 symptoms and 308 diseases from the fields of rheumatology (186), hepatology (74), coagulation defects (34) and pancreatic diseases (10). The documentation of strong relationships between symptoms and diseases (occurrence: always or never and/or confirmability: always or never) as well as among symptoms and among diseases has been finished. SCS that can confirm rheumatic fever, rheumatoid arthritis, Sjögren's disease, psoriatic arthritis, Reiter's disease, Behcet's disease, Bechterew's disease, systematic lupus erythematosus, sclerodermia or gout have been implemented.

Patients' records from a rheumatological hospital are gathered and stored by means of a hospital information system /14,15/. Patients' data consists of

- 1) Physical examination  
(513 binary positions in documentary system  
⇓  
282 binary symptoms in CADIAG-2)

- 2) Rheumatological history  
(372 binary positions in documentary system  
⇓  
158 binary symptoms in CADIAG-2)
- 3) Rheumatological examination  
(1227 binary positions in documentary system  
⇓  
453 binary symptoms in CADIAG-2)
- 4) Laboratory tests  
(80 lab tests  
⇓  
305 binary symptoms in CADIAG-2)
- 5) X-ray findings  
(78 binary symptoms in CADIAG-2)
- 6) Histological findings  
(75 binary symptoms in CADIAG-2)

and the final rheumatological diagnoses confirmed by the rheumatologist. By means of the final diagnosis, the diagnostic results provided by CADIAG-2 can be com-

Table 11: Status of the medical knowledge system of CADIAG-2.

	relationships occurrence	confirmability	rheumatology	hepatology	coagulation defects	pancreatic diseases
S <sub>i</sub> D <sub>j</sub>	<u>always</u>	<u>always</u>	18	1	2	0
	<u>always</u>	-	189	113	67	0
	-	<u>always</u>	34	13	4	2
	<u>never</u>	<u>never</u>	536	54	558	0
S <sub>i</sub> S <sub>j</sub>	<u>always</u>	<u>always</u>	0	0	0	0
	<u>always</u>	-	419	4	0	0
	-	<u>always</u>	419	4	0	0
	<u>never</u>	<u>never</u>	2.395	143	786	34
D <sub>i</sub> D <sub>j</sub>	<u>always</u>	<u>always</u>	0	-	0	-
	<u>always</u>	-	54	-	6	-
	-	<u>always</u>	54	-	6	-
	<u>never</u>	<u>never</u>	1.370	-	0	-
SCS <sub>i</sub> D <sub>j</sub>	-	<u>always</u>	10	-	-	-
	<u>always</u>	-	1	-	-	-

pared and the accuracy of the method can be calculated. The connection of CADIAG-2 to the medical database of the hospital information system has been realized. The CADIAG-2 database interface includes the calculation of degrees of membership in fuzzy symptom subsets given numerical lab test results (about 300 appropriate membership functions were defined) and the evaluation of logical combinations (similar to  $SCP_k$  and  $SCS_j$ ) for the simplest documentary data.

CADIAG-2 is planned to be used

- 1) in retrospective mode to evaluate and improve the medical knowledge system (about 1.000 patients' records are available) and
- 2) in prospective mode to optimize the number of necessary lab tests, X-ray, puncture findings, etc.

Moreover, CADIAG-2 provides a deep insight in a physician's mental diagnostic process because of its explanation and reasoning capabilities and is highly suitable in medical diagnostic training.

#### FOOTNOTES

- 1) A first theoretical approach to CADIAG-2 was made by ADLASSNIG /35/.
- 2) In the context of this paper the only difference between the terms "disease" and "diagnosis" will be the following (see /36,37/:
  - "Diseases" are described by nosograms that have been established in the field of nosology. For example: "Acute pancreatitis is characterized by sudden aches of the abdomen, sickness, vomiting, etc.".

"Syndrome" is considered to be a synonym to "disease".

- The term "diagnosis" always appears in conjunction with patients. "Diagnosis" is an idiogram about an individuuum. For example: "Smith suffers from acute pancreatitis".
- 3) The term "reference set" is used in KAUFMANN /21-24/. ZADEH /17,28/ calls it "universe of discourse". DUBOIS and PRADE /19/ name it "field of reference" or "universe" for short.
  - 4) On the basis of a defined metric and distance in a space, similarity between  $P_i$  and  $P_j$  means a small distance  $d(P_i, P_j)$  between  $P_i$  and  $P_j$  (see DUBOIS and PRADE /19/, KAUFMANN /21/).
  - 5) In terms of CADIAG-1 /7-16/, "Symptom  $S_i$  always occurs in case of disease  $D_j$ ." means that the presence of  $S_i$  is obligatory if  $D_j$  shall be confirmed.
  - 6) "Symptom  $S_i$  always confirms disease  $D_j$ ." means if  $S_i$  occurs then  $S_i$  proves the presence of  $D_j$ .
  - 7) "Symptom  $S_i$  never occurs with  $D_j$ ." means if  $S_i$  occurs then  $D_j$  has to be excluded.
  - 8) The parameters of the membership functions have not been proved in clinical practice yet. If appropriate the parameters can be changed.
  - 9)  $[\mu_A(x)]^4$  for  $\mu_{\text{very } A}(x)$  seems to fit physicians' experience better.
  - 10) ZADEH /17/, DUBOIS and PRADE /19/ and KAUFMANN /21/ call it "inclusion". KICKERT /25/ terms it "containment".
  - 11) To consider the means to be the numerical representatives of the linguistic terms is arbitrary. But the  $x_{0_i}$  and  $x_{c_i}$  are suitable values with the linguistically expressed strength of  $S_i, D_j$  relationships in an ordered manner. They are helpful for numerical computation and never appear on screen during the medical diagnostic consultation.
  - 12) Comparing with the linguistic approach  $x_0 \approx 19.40$  is in the range of very seldom.

- 13) In linguistic terms, tophi are always confirming gout (tophi are proving gout). /2/ AMMANN, R.: Ikterus. In SIEGENTHALER, W. (Hrsg.): Differentialdiagnose innerer Krankheiten. Georg Thieme Verlag Stuttgart, 1975.
- 14) In the similar tri-valued logical approach of CADIAG-1 /10/, about 40% of the symptom-disease relationships have been found unknown. /3/ ZÖLLNER, N.: Gicht. In GROSS, R., P.SCHÖLMERICH (Hrsg.): Lehrbuch der Inneren Medizin. F.K.Schattauer Verlag Stuttgart-New York, 1977.
- 15) Patient 1 stayed in hospital two different times. /4/ OVERZIER, C.: Systematik der Inneren Medizin: Daten, Fakten, Übersichten. Georg Thieme Verlag Stuttgart, 1975.
- 16) Inner diagnostic process terms the part of the consultation system that indicates confirmed and excluded diagnoses as well as diagnostic hypotheses given a certain symptom pattern. /5/ SCHALDACH, H. (Hrsg.): Wörterbuch der Medizin. VEB Verlag Volk und Gesundheit Berlin, 1974.
- 17) Certainly, soft relationships like almost always, often, seldom, etc. (see paragraph 3.2.1.) can also be used to express symptom-symptom and disease-disease relationships. In that case well-formed and extensive statistical material about relations among symptom and among diseases have to be available. /6/ KOLARZ, G.: Diagnosekriterien entzündlich rheumatischer Erkrankungen. In SINGER, F. (Hrsg.): Entzündlich-rheumatische Erkrankungen - ihre Bedeutung für den praktizierenden Arzt. Pensionsversicherungsanstalt der Arbeiter und Fa.Merck Sharp & Dohme Wien, 1977, 39-42.
- 18) KAUFMANN /21/ terms such combinations "functions of fuzzy variables". DUBOIS and PRADE /19/ name it "fuzzy expressions". /7/ 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.
- 19) WEISS, et al. /46,47/ call it "pathophysiological state" and the relationships between observations (history, symptoms, signs and laboratory tests) and states are associational. LINDBERG, et al. /48/ denote these combinations "intermediate hypotheses". /8/ ADLASSNIG, K.-P., G.GRABNER: Approaches to Computer-Assisted Diagnosis in Gastroenterology. EDV in Medizin und Biologie 11 (1980) 74-80.
- 20) In a first consideration,  $\uparrow$  and  $\downarrow$  were introduced to handle major and minor criteria in rheumatology. But these simple operations have been found to be a useful and practical aid for the definition of symptom combinations in general. /9/ ADLASSNIG, K.-P., G.GRABNER: The Viennese Computer-Assisted Diagnostic System. Its Principles and Values. AUTOMEDICA 3 (1980) 141-150.
- 21) In WEISS, et al. /46,47/ a similar concept is used. "Classification table rules" are established that are Boolean combinations of confirmed or denied states. /10/ 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.
- REFERENCES /11/ BAUER, P., A.GANGL, G.GRABNER: Ein Computer-Verfahren zur Zuordnung eines Krankheitsbildes zu einer Diagnose-Gruppe. Wien.Z.inn.Med. 51 (1970) 497-509.
- /1/ KASPER, H., H.SOMMER. Klinik der akuten Pankreatitis. In SCHWIEGK, H. (Hrsg.): Handbuch der Inneren Medizin, 3.Band: Verdauungsorgane, Teil 6. Springer-Verlag Berlin-Heidelberg-New York, 1976.

- /12/ BAUER, P., A.GANGL, G.GRABNER, O.JAHN: Ein Computer-Verfahren zur Unterstützung des Arztes bei der Erstellung von Differentialdiagnosen. *Impuls* 10 (1968) 705-712.
- /13/ GANGL, A., G.GRABNER, P.BAUER: Erste Erfahrungen mit einem Computerprogramm zur Differentialdiagnose der Lebererkrankungen. *Wien.Z.inn.Med.* 50 (1969) 553-586.
- /14/ GRABNER, G., H.GRABNER: The Viennese General Medical Information System. *Proceedings of MEDIS'75 TOKYO*, Kansai Institute of Information Systems, 1975, 156-163.
- /15/ GRABNER, H., G.GRABNER: Aims and Structure of the Vienna General Medical Information System WAMIS. In ANDERSON, J., J.M.FORSYTHE (Eds.): *MEDINFO 74*. North-Holland Publishing Company Amsterdam-New York-Oxford, 1974, 375-379.
- /16/ 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.
- /17/ ZADEH, L.A.: Fuzzy Sets. *Information and Control* 8 (1965) 338-353.
- /18/ 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.
- /19/ DUBOIS, D., H.PRADE: *Fuzzy Sets and Systems: Theory and Applications*. Academic Press New York-London, 1980.
- /20/ 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.
- /21/ KAUFMANN, A.: *Introduction to the Theory of Fuzzy Subsets. Volume I. Fundamental Theoretical Elements*. Academic Press New York-San Francisco-London, 1975.
- /22/ KAUFMANN, A.: *Introduction à la Théorie des Sous-Ensembles Flous. Tome II. Application à la Linguistique et à la Sémantique*. Masson Paris, 1975.
- /23/ 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.
- /24/ KAUFMANN, A.: *Introduction à la Théorie des Sous-Ensembles Flous. Tome IV. Compléments et Nouvelles Applications*. Masson Paris, 1977.
- /25/ KICKERT, W.J.M.: *Fuzzy Theories on Decision-Making*. Martinus Nijhoff Social Sciences Division Leiden-Boston-London, 1978.
- /26/ NEGOITA, C.V., D.A.RALESCU: *Applications of Fuzzy Sets to System Analysis*. Birkhäuser Verlag Basel-Stuttgart, 1975.
- /27/ WECHLER, W.: *The Concept of Fuzziness in Automata and Language Theory*. Akademie-Verlag Berlin, 1978.
- /28/ 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.
- /29/ GAINES, B.R., KOHOUT, L.J.: *The Fuzzy Decade: A Bibliography of Fuzzy Systems and Closely Related Topics*. *Int.J.Man-Machine Studies* 9 (1977) 1-68.
- /30/ KANDEL, A., R.R.YAGER: *A 1979 Bibliography on Fuzzy Sets, Their Applications, and Related Topics*. 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, 621-744.
- /31/ TOTTH, H.: *Fuzzy Bibliography (Stand vom 20.02.1982). Versuch einer Weiterführung von KANDEL/YAGER: A 1979 Bibliography on Fuzzy Sets, Their Application and Related Topics*. Personal Communication.
- /32/ ADLASSNIG, K.-P.: *A Survey on Medical Diagnosis and Fuzzy Subsets*. In GUPTA, M.M., E.SANCHEZ (Eds.): *Fuzzy Information and Decision Processes*. North-Holland Publishing

- Company Amsterdam-New York-Oxford.  
In press.
- /33/ 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 Intelli-  
gence Tbilisi, 1975, 848-855.
- /34/ POPLE, H.E.: The Formation of Com-  
posite Hypotheses in Diagnostic  
Problem Solving - An Exercise in  
Synthetic Reasoning. Proceedings of  
the 5th International Joint Con-  
ference on Artificial Intelligence  
Cambridge, 1977, 1030-1037.
- /35/ ADLASSNIG, K.-P.: A Fuzzy Logical  
Model of Computer-Assisted Medical  
Diagnosis. Meth.Inform.Med. 19  
(1980), 141-148.
- /36/ SADEGH-ZADEH, K.: Krankheitsbegriffe  
und nosologische Systeme. METAMED  
1 (1977), 4-41.
- /37/ SADEGH-ZADEH, K.: Grundlagenprobleme  
einer Theorie der klinischen Praxis,  
Teil 1: Explikation des medizini-  
schen Diagnosebegriffs. METAMED 1  
(1977), 76-102.
- /38/ ALBIN, M.A.: Fuzzy Sets and Their  
Application to Medical Diagnosis  
and Pattern Recognition. PhD diss.  
University of California, Berkeley,  
1975.
- /39/ ZADEH, L.A.: A Fuzzy-Algorithm  
Approach to the Definition of  
Complex or Imprecise Concepts. In  
BOSSEL, H., S.KLACZKO, N.MÜLLER  
(Eds.): System Theory in the Social  
Sciences. Birkhäuser Verlag Basel-  
Stuttgart, 1976, 202-282.
- /40/ ZADEH, L.A.: A Fuzzy Set-Theoretic  
Interpretation of Linguistic  
Hedges. J.Cybern. 2 (1972), 4-34.
- /41/ GRAMS, R.R.: Medical Information  
Systems: The Laboratory Module.  
The Humana Press Clifton-New Jersey,  
1979.
- /42/ SHANNON, R.H.(Ed.): Hospital Infor-  
mation Systems. North Holland  
Publishing Company Amsterdam-New  
York-Oxford, 1979.
- /43/ SHORTLIFFE, E.H.: Computer-Based  
Medical Consultations: MYCIN.  
Elsevier New York-Oxford-Amsterdam,  
1976.
- /44/ ROPES, M.W., G.A.BENNETT, S.COBB,  
R.JACOX, R.A.JESSAR: 1958 Revision  
of Diagnostic Criteria for Rheuma-  
toid Arthritis. Arthr.Rheum. 2 (1959)  
16-20.
- /45/ American Heart Association,  
Rheumatic Fever Committee of the  
Council on Rheumatic Fever and  
Congenital Heart Disease. Circu-  
lation 43 (1971), 983-988.
- /46/ WEISS, S., C.A.KULIKOWSKI, A.SAFIR:  
Glaucoma Consultation by Computer.  
Comput.Biol.Med. 8 (1978) 25-40.
- /47/ WEISS, S.M., C.A.KULIKOWSKI,  
S.AMAREL, A.SAFIR: A Model-Based  
Method for Computer-Aided Medical  
Decision-Making. Artificial Intelli-  
gence 11 (1978) 145-172.
- /48/ LINDBERG, J.A.B., G.C.SHARP,  
L.C.KINGSLAND, III, S.M.WEISS,  
S.P.HAYES, H.UENO, S.E.HAZELWOOD:  
Computer Based Rheumatology Con-  
sultant. In LINDBERG, D.A.B.,  
S.KAIHARA (Eds.): MEDINFO 80.  
North-Holland Publishing Company  
Amsterdam-New York-Oxford, 1980,  
1311-1315.
- /49/ WIRTH, N.: Compilerbau. B.G.  
Teubner Stuttgart, 1977.
- /50/ AHO, A.V., J.D.ULLMAN: Principles  
of Compiler Design. Addison-  
Wesley Publishing Company, 1978.
- /51/ GRIES, D.: Computer Construction  
for Digital Computers. John Wiley  
and Sons, Inc., 1971.
- /52/ 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.
- /53/ TAUTU, P., G.WAGNER: The Process  
of Medical Diagnosis: Routes and  
Mathematical Investigations. Meth.  
Inform.Med. 1 (1978), 1-10.