COMPUTERS AND BIOMEDICAL RESEARCH 19, 63–79 (1986)

Representation and Semiautomatic Acquisition of Medical Knowledge in CADIAG-1 and CADIAG-2[1]

KLAUS-PETER ADLASSNIG*,† AND GERNOT KOLARZ‡

*Computer Science Division, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, California 94720; †Department of Medical Computer Sciences, University of Vienna, Garnisongasse 13, A-1090 Vienna, and ‡Ludwig Boltzmann Institute for Rheumatology and Focal Diseases, Kaiser-Franz-Ring 8, A-2500 Baden, Austria

Received April 18, 1985

CADIAG-1 and CADIAG-2 (Computer-Assisted DIAGnosis) are medical expert systems especially designed for ill-defined areas such as internal medicine. Both systems are being tested in the setting of a medical information system. With respect to their knowledge representation, CADIAG-1 has obvious advantages in totally ill-defined areas such as syndromes in internal medicine, whereas CADIAG-2 seems more suited for domains with basic laboratory programs, e.g., hepatology or gall bladder and bile duct diseases. The formalization of relationships between medical entities led to first-order predicate calculus formulas in the case of CADIAG-1 and to a model based on fuzzy set theory in the case of CADIAG-2. In both systems two kinds of relationships between medical entities are considered: (1) necessity of occurrence and (2) sufficiency of occurrence. Statistical interpretations using the 2 × 2 table paradigm yield a way to calculate these relationships automatically from samples of patient data. Results obtained by exploiting 3530 patient records from a rheumatological hospital are presented. The described application is a machine-learning program that allows inductive learning from examples under statistical uncertainty.

1. INTRODUCTION

CADIAG-1 and CADIAG-2 (Computer-Assisted DIAGnosis) have been developed for supporting diagnostic decisions in internal medicine. At present, each system contains about 300 diseases (the same diseases) from the fields of rheumatology and gastroenterology.

An early version of CADIAG-1 was developed by Spindelberger and Grabner (*1*) in 1968 and was applied to liver diseases by Gangl *et al.* (*2*) in 1969. The formal representation of relationships between medical entities is achieved by first-order predicate calculus formulas (*3*). Results obtained with CADIAG-1 are discussed in (*4*).

The development of CADIAG-2 is based on the experience gained with

[1] This work was supported by a Max Kade Fellowship for Scientific Research and supported in part by NSF-IST Grant 8320416 and NASA Grant NCC2-275.

0010-4809/86 $3.00

CADIAG-1. CADIAG-2 was designed and implemented by Adlassnig (*5*) in 1980. The formalization of medical entities is performed by applying the theory of fuzzy sets (Zadeh (*6*)). Fuzzy logic (*7*) serves as an appropriate inference mechanism. A detailed comparison of CADIAG-2 results with physicians' diagnoses can be found in (*4*, *8*).

Both expert systems are data-driven, rule-based systems. They contain single antecedent/consequent and compound antecedent/consequent rules. Single antecedents are formed either by symptoms S_i (signs, laboratory test results, findings) or by diseases D_j. Compound antecedents are constituted by logical combinations SC_l of symptoms and/or diseases. But only S_iD_j, SC_lD_j, S_iS_j, and D_iD_j rules are admitted. The association between the antecedent and the consequent in a rule is determined by the medical relationships between them. In CADIAG-1 as well as in CADIAG-2 two kinds of relationships are considered: the necessity of occurrence of an antecedent with the consequent and the sufficiency of occurrence of an antecedent for concluding the consequent.

The necessity relationship can be regarded as a backward implication from consequents to antecedents, i.e., the consequent implies to a certain degree the antecedent. It is applied in the sense of a modus tollens syllogism. This syllogism can be used to infer the absence of an antecedent from an excluded consequent.

In contrast, the sufficiency relationship is a forward implication. Consequents can be derived directly from given evidence in the antecedents, i.e., the antecedent implies to a certain degree the consequent. The appropriate syllogism is found in the modus ponens.

Knowledge that can be formalized by means of rules with compound antecedents presupposes a well-known inner structure of the knowledge domain. This structure indicates which pieces of evidence are suited to form compound antecedents and which are not. But this kind of knowledge is not always available. Very often, only single associations between surface appearances and underlying concepts such as diseases can be hypothesized from observations. Examples can easily be found in several domains of medicine, e.g., in the fields of syndromes in internal medicine or psychiatric diseases. Most of the knowledge that has been acquired about these diseases until now can only be expressed in single antecedent/consequent rules.

Furthermore, unstructured domains of knowledge are characterized by possessing more knowledge about the necessity than about the sufficiency of antecedents. In medicine, for instance, this is because the knowledge about the necessity of observations in order to diagnose diseases is provided by investigating sets of patients suffering from a certain disease and observing the occurrence of symptoms in these patients. Investigations concerning the sufficiency require sets of patients exhibiting these symptoms independently of having the particular disease or not. But such investigations are more difficult to carry out. They have to contain all patients of a catchment area of a medical institution or all patients that are included in a medical demographic study.

Table 1 illustrates the different levels of prestructuralization of knowledge and provides some examples.

TABLE 1

DIFFERENT LEVELS OF PRESTRUCTURALIZATION OF KNOWLEDGE DOMAINS DEMAND DIFFERENT TYPES OF RULES IN EXPERT SYSTEMS

Level	Typical rules	Domain examples and applications	Proportion of single : compound antecedents	Epistemological process
1	$(e_i \wedge \cdots \wedge e_n) \xrightarrow{\mu} h_j$ $\mu \in \{0,1\}$ or $\{-1,1\}$ Compound antecedents express prestructuralization of knowledge Certain inferences by applying propositional logic, predicate logic, or similar methods	Family relationships Taxonomy of animals or plants PROLOG (*9*) ONCOCIN (*10*): treatment management of cancer patients	small : large	↑
2	$(e_i \wedge \cdots \wedge e_n) \xrightarrow{\mu} h_j$ $\mu \in [0,1]$ or $[-1,1]$ Compound antecedents express prestructuralization of knowledge Uncertain inferences by applying probability theory, confirmation theory, or fuzzy set theory Additional uncertainty due to uncertain evidence in the antecedents	MYCIN (*11*): infectious diseases PROSPECTOR (*12*): mineral exploration EMYCIN/PUFF (*13*): pulmonary function AI/RHEUM (*14*): parts of rheumatology	small : large	↑
3	$e_i \xrightarrow{\mu} h_j$, $h_j \xrightarrow{\mu} e_i$ $\mu \in [0,1]$ or $[-1,1]$ Single antecedents express unstructured knowledge Uncertain inferences by applying probability theory, confirmation theory, or fuzzy set theory Additional uncertainty due to uncertain evidence in the antecedents	INTERNIST-1/CADUCEUS (*15, 16*): internal medicine CADIAG-2 (*4, 5, 8*): internal medicine ESDAT (*17*): primary medical care	large : small	↑
4	e_i A h_j Single antecedents express unstructured knowledge Prompting of hypotheses by applying pattern matching procedures or specificity measurements	CADIAG-1 (*1, 2, 4*): internal medicine RECONSIDER (*18*): general medicine	large : small	↑

Note. e—evidence; h—hypothesis, μ—degree of inference; A—association.

By considering expert systems as tools that extend the ability of human beings rather than substitute for them, expert systems seem especially helpful in unstructured domains of knowledge. The present proportion between rules with single and compound antecedents is about 23,000 : 80 in CADIAG-1 and about 23,000 : 110 in CADIAG-2.

After this introduction, the formal representation of medical knowledge in CADIAG-1 and CADIAG-2 is shown in Section 2. Section 3, then, describes

the statistical interpretations of the medical relationships applied in CADIAG-1 and CADIAG-2. The application of these statistical interpretations on 3530 patient records from a rheumatological hospital, including 103 rheumatological diseases and 905 symptoms, signs, test results, and findings, is discussed in Section 4. Together with a discussion of the achieved results, a brief description of related work published by other authors is given in Section 5.

2. Knowledge Representation in CADIAG-1 and CADIAG-2

In CADIAG-1, the two kinds of relationships between medical entities—necessity and sufficiency—are combined and yield the following easy-to-gather relationships between symptoms and diseases:

S_i OC D_j: obligatory occurrence and confirmation, i.e., the symptom has to be present in the patient in order to establish the diagnosis, and, if it is present, it confirms the diagnosis. If the symptom is definitely absent, the diagnosis is excluded.

$$\text{OC} \triangleq \textbf{if } S_i \textbf{ then } D_j \text{ or } \textbf{if not } S_i \textbf{ then not } D_j. \qquad [2.1]$$

Example 1: *if* endoprothesis of the knee in X-ray *then* arthroplasty of the knee or *if not* endoprothesis of the knee in X-ray *then not* arthroplasty of the knee.

S_i FC D_j: facultative occurrence and confirmation, i.e., the symptom does not have to be present in order to establish the diagnosis, but if the symptom occurs, the diagnosis is thus confirmed.

$$\text{FC} \triangleq \textbf{if } S_i \textbf{ then } D_j. \qquad [2.2]$$

Example 2: *if* intracellular uric acid crystals in joint effusion *then* gout.

S_i ON D_j: obligatory occurrence and nonconfirmation, i.e., the symptom has to be present in order to establish the diagnosis. Therefore, if the symptom is absent, the diagnosis is excluded.

$$\text{ON} \triangleq \textbf{if not } S_i \textbf{ then not } D_j. \qquad [2.3]$$

Example 3: *if not* onset of disease prior to 16th year of age *then not* juvenile rheumatoid arthritis.

S_i EX D_j: exclusion, i.e., if the symptom is present, the diagnosis is excluded.

$$\text{EX} \triangleq \textbf{if } S_i \textbf{ then not } D_j. \qquad [2.4]$$

Example 4: *if* Waaler Rose titer $\geq 1:64$ *then not* seronegative rheumatoid arthritis.

S_i FN D_j: facultative occurrence and nonconfirmation, i.e., the symptom does not have to be present in order to establish the diagnosis, and, if it occurs,

it does not confirm the diagnosis. The symptom certainly provides evidence for the diagnosis, but it only expresses the existence of an association between the exhibited symptom and the underlying disease.

$$\text{FN} \triangleq \textbf{if } S_i \textbf{ then may be } D_j. \qquad [2.5]$$

EXAMPLE 5: *if* elevated amylase in serum *then may be* acute pancreatitis.

In order to process the data of an individual patient, the representation of the considered relationships OC, FC, ON, EX, and FN as *if–then* statements is sufficient. More complex problems such as the verification of consistency of the CADIAG-1 knowledge base (see Barachini (*3*)) demanded an investigation of these medical relationships more deeply. This led to the following first-order predicate calculus formulas, where the predicate $S_i(\mathrm{p})$ means "patient p exhibits S_i" and the predicate $D_j(\mathrm{p})$ "patient p suffers from D_j." The other signs have the following interpretations: $\forall$ for every, $\exists$ there exists a, $\wedge$ and, $\rightarrow$ implies, $\neg$ not.

$$S_i \text{ OC } D_j \triangleq \forall \mathrm{p}(S_i(\mathrm{p}) \rightarrow D_j(\mathrm{p})) \wedge \forall \mathrm{p}(D_j(\mathrm{p}) \rightarrow S_i(\mathrm{p})) \wedge \exists \mathrm{p}(S_i(\mathrm{p}) \wedge D_j(\mathrm{p})) \qquad [2.6]$$

$$S_i \text{ FC } D_j \triangleq \forall \mathrm{p}(S_i(\mathrm{p}) \rightarrow D_j(\mathrm{p})) \wedge \neg \forall \mathrm{p}(D_j(\mathrm{p}) \rightarrow S_i(\mathrm{p})) \wedge \exists \mathrm{p}(S_i(\mathrm{p}) \wedge D_j(\mathrm{p})) \qquad [2.7]$$

$$S_i \text{ ON } D_j \triangleq \forall \mathrm{p}(D_j(\mathrm{p}) \rightarrow S_i(\mathrm{p})) \wedge \neg \forall \mathrm{p}(S_i(\mathrm{p}) \rightarrow D_j(\mathrm{p})) \wedge \exists \mathrm{p}(S_i(\mathrm{p}) \wedge D_j(\mathrm{p})) \qquad [2.8]$$

$$S_i \text{ EX } D_j \triangleq \forall \mathrm{p}(S_i(\mathrm{p}) \rightarrow \neg D_j(\mathrm{p})) \wedge \exists \mathrm{p}(S_i(\mathrm{p}) \wedge \neg D_j(\mathrm{p})) \wedge \exists \mathrm{p}(D_j(\mathrm{p}) \wedge \neg S_i(\mathrm{p})) \qquad [2.9]$$

$$S_i \text{ FN } D_j \triangleq \neg \forall \mathrm{p}(S_i(\mathrm{p}) \rightarrow D_j(\mathrm{p})) \wedge \neg \forall \mathrm{p}(D_i(\mathrm{p}) \rightarrow S_i(\mathrm{p})) \wedge \exists \mathrm{p}(S_j(\mathrm{p}) \wedge D_j(\mathrm{p})) \qquad [2.10]$$

SC_lD_j, S_iS_j, and D_iD_j relationships are defined in an analogous way.

Given a patient's symptoms, the rules for OC, FC, ON, and EX are applied to infer confirmed and excluded diagnoses. The FN relationships serve as a basis for generating meaningful diagnostic hypotheses, which are offered to the physician for subsequent confirmation or exclusion. But in order to avoid too many diagnostic hypotheses (in extended applications almost every diagnosis would be established as a diagnostic hypothesis by using the mere FN relationship), two selection procedures were developed to differentiate between more specific and less specific symptoms:

(1) Precomputation of unique symptom patterns in the medical knowledge base with the aim to match them with patient's symptoms during the diagnostic process and thus to establish diagnostic hypotheses.

(2) On-line selection of chief complaints during the diagnostic process that prompt possible diagnoses.

A more extended description of the diagnostic process and the precomputation of unique symptom patterns in CADIAG-1 can be found in (*4*).

In CADIAG-2, the necessity and sufficiency of occurrence are considered separately. They are termed frequency of occurrence (O) of symptom S_i with disease D_j and strength of confirmation (C) of symptom S_i for disease D_j. These relationships are interpreted as binary fuzzy relationships. They take their values $\mu_O(S_i,D_j)$ and $\mu_C(S_i,D_j)$ in the interval $[0,1] \cup \{v\}$ (v: no relationship). Linguistic terms such as *always*, *often*, *seldom*, *never*, *strong*, *weak*, etc., have been found semantically useful in order to characterize these relationships, although numerical values are stored in the knowledge base of CADIAG-2 (see (*19*)). The general form of the CADIAG-2 rules is

if antecedent **then** consequent **with** (O,C) [2.11]

where the relationship tuple (O,C) contains linguistic and numerical values λ_O and μ_O, and/or λ_C and μ_C.

EXAMPLE 6 (cf. Example 2): *if* intracellular uric acid crystals in joint effusion *then* gout *with* (λ_O = *seldom* [μ_O = 0.25], λ_C = *always* [μ_C = 1.00]).

EXAMPLE 7 (cf. Example 5): *if* elevated amylase in serum *then may be* acute pancreatitis *with* (λ_O = very often [μ_O = 0.90], λ_C = strong [μ_C = 0.70]).

SC_lD_j, S_iS_j, and D_iD_j relationships are treated in an analogous way. An example for an SC_lD_j rule is:

EXAMPLE 8: *if* low back pain, and limitation of motion of the lumbar spine, and diminished chest expansion, and the patient is male, and between 20 and 40 years of age *then may be* ankylosing spondylitis *with* (λ_O = very often [μ_O = 0.90], λ_C = strong [μ_C = 0.80]).

While in CADIAG-1 symptoms and diseases take their values in {1 = present, 0 = absent, v = not yet examined} and {1 = confirmed, $\frac{1}{2}$ = hypothesis, 0 = excluded, v = not yet generated}, respectively, in CADIAG-2 symptoms and diseases are fuzzy subsets of appropriate reference sets. Using this concept, uncertainty about the observation of pathological symptoms as well as the confirmation or exclusion of diseases can be expressed appropriately.

Symptom fuzzy sets S_i such as *normal body temperature*, *subfebrile body temperature*, or *high body temperature* are characterized by membership functions $\mu_{S_i}: X \rightarrow [0,1]$, which assign to every possible $x \in X$ a degree of membership of x in S_i. For example, a body temperature of 37.6°C has a degree of membership of 0.00 in the fuzzy subset normal body temperature, but the membership in subfebrile body temperature is 1.00 and in high body temperature is, say, 0.20. In formal terms, it is

$$\mu_{\text{normal body temperature}}(37.6°\text{C}) = 0.00$$

$$\mu_{\text{subfebrile body temperature}}(37.6°\text{C}) = 1.00$$

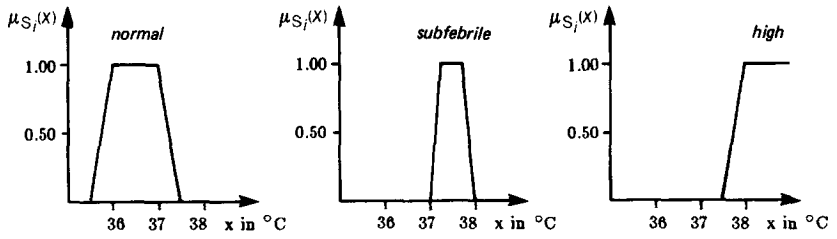

FIG. 1. Some fuzzy subsets for body temperature.

$$\mu_{\text{high body temperature}}(37.6°\text{C}) = 0.20.$$

The reference sets X contain all possible values x the symptoms may assume. In CADIAG-2, about 400 such membership functions are stored and applied in order to transfer patient data from the patient data base of the Vienna Medical Information System (*20*) to CADIAG-2. Similar functions are used for assigning numerical test results to symptoms in CADIAG-1, but here only 1.00/0.00 assignments are allowed. Furthermore, many of the functions are adjusted to sex and age of the individual patient to be diagnosed. Figure 1 shows the membership functions for normal, subfebrile, and high body temperature.

In regard to one patient, the degrees of membership $\mu_{S_i}(x)$ express binary fuzzy relationships $\mu_{PS}(P_k,S_i)$ between the patient and the symptoms exhibited. Diagnoses are also considered as fuzzy subsets. The binary fuzzy relationship $\mu_{PD}(P_k,D_j)$ expresses the degree to which this diagnosis can logically be concluded from the given medical evidence.

The inference process in CADIAG-2 is carried out using the fuzzy compositional rule of inference (*7*) that composes the binary fuzzy relationship values $\mu_{PS}(P_k,S_i)$ with $\mu_O(S_i,D_j)$ and with $\mu_C(S_i,D_j)$ in order to infer values $\mu_{PD}(P_k,D_j)$ (see Adlassnig *et al.* (*8*)).

Confirmed and excluded diagnoses in CADIAG-2 are inferred similarly to those in CADIAG-1, but diagnostic hypotheses are established differently. They are generated if the antecedent in a fully confirming rule is only partly present, i.e., the value of the symptom or the symptom combination lies in the range $\varepsilon < \mu_{PS}(P_k,S_i)$ or $\mu_{PSC}(P_k,SC_l) < 1$, or if the applicable rule shows a strength of confirmation of $\varepsilon \leq \mu_C(S_i,D_j)$ or $\mu_C(SC_l,D_j) < 1$, where ε is a threshold value to preclude diagnostic hypotheses with too little evidence. The ranking of the hypotheses is achieved by calculating support scores heuristically (an unnormalized function) and arranging the hypotheses according to their scores in descending order. A support score becomes higher the more symptoms explain the hypothesis under consideration and the higher the frequency of occurrence and the strength of confirmation values between these symptoms and the hypothesis are (see (*8*)). Table 2 shows the transition between CADIAG-1 and CADIAG-2 relationships.

TABLE 2

COMPARISON OF THE MEDICAL RELATIONSHIPS IN CADIAG-1 AND CADIAG-2

CADIAG-1 relationships	CADIAG-2	
	Frequency of occurrence	Strength of confirmation
OC	1.00	1.00
FC	v or $0.00 < \mu_O < 1.00$	1.00
ON	1.00	v or $0.00 < \mu_C < 1.00$
EX	0.00	0.00
FN	v^a or $0.00 < \mu_O < 1.00$	v^a or $0.00 < \mu_C < 1.00$

[a] At least one of the two relationships has to be $\neq v$ (v: no relationship).

3. STATISTICAL INTERPRETATION OF CADIAG-1 AND CADIAG-2 RELATIONSHIPS

Not only judgmental knowledge of an expert but also statistical knowledge, if available, can be collected in a knowledge base. A statistical interpretation of the relationships used in an expert system is then required.

In CADIAG-2, the frequency of occurrence (as the name already expresses) and the strength of confirmation are interpretable as statistical frequencies. Let us consider a 2 × 2 table (*21*) for analyzing the statistical relationships between a symptom S and a disease D (Table 3). The entries in the fields of Table 3 are absolute frequencies that can be calculated from samples of patient data with known symptoms and known diseases. They have the following meanings:

$F(S \cap D) = a$: number of patients with S and D
$F(S \cap \bar{D}) = b$: number of patients with S and $\bar{D}$
$F(\bar{S} \cap D) = c$: number of patients with $\bar{S}$ and D
$F(\bar{S} \cap \bar{D}) = d$: number of patients with $\bar{S}$ and $\bar{D}$

$F(S)$, $F(\bar{S})$, $F(D)$, and $F(\bar{D})$ are defined analogously, and N is the total number of patients considered for this calculation.

Now the statistical interpretation for the frequency of occurrence is

$$\mu_O(S,D) \triangleq \frac{F(S \cap D)}{F(D)} = \frac{a}{a + c} = F(S/D) \xrightarrow[\text{for}]{\text{estimation}} P(S/D). \quad [3.1]$$

and for the strength of confirmation

$$\mu_C(S,D) \triangleq \frac{F(S \cap D)}{F(S)} = \frac{\mathrm{a}}{\mathrm{a} + \mathrm{b}} = F(D/S) \xrightarrow[\text{for}]{\text{estimation}} P(D/S). \quad [3.2]$$

TABLE 3

2 × 2 TABLE FOR THE ANALYSIS OF STATISTICAL RELATIONSHIPS BETWEEN SYMPTOMS AND DISEASES

	D	$\bar{D}$	
S	$F(S \cap D)$	$F(S \cap \bar{D}$	$F(S)$
$\bar{S}$	$F(\bar{S} \cap D)$	$F(\bar{S} \cap \bar{D})$	$F(\bar{S})$
	$F(D)$	$F(\bar{D})$	N

or briefly

	D	$\bar{D}$	
S	a	b	$a + b$
$\bar{S}$	c	d	$c + d$
	$a + c$	$b + d$	$a + b + c + d$

Note. $\bar{S}$—not S; $\bar{D}$—not D.

Both relationships are based on $F(S \cap D)$, a set of patients who usually consult a physician because they show symptoms S of a disease D. The calculated frequencies allow an estimation for the conditional probabilities $P(S/D)$ and $P(D/S)$.

In general, a statistical analysis of the relationships between two entities (e.g., for adjusting normal and pathological ranges of tests in screening procedures to select risk patients in a population) differentiates between four statistical relationships (*22*):

Sensitivity: proportion of present symptoms (or positive test results) among the diseased

$$\frac{a}{a + c} = P(S/D). \qquad [3.3]$$

Specificity: proportion of absent symptoms (or negative test results) among the healthy

$$\frac{d}{b + d} = P(\bar{S}/\bar{D}). \qquad [3.4]$$

Positive predictive value: proportion of the diseased among the group with present symptoms (or positive test results)

$$\frac{a}{a+b} = P(D/S). \qquad [3.5]$$

Negative predictive value: proportion of healthy among the group with absent symptoms (or negative test results)

$$\frac{d}{c+d} = P(\bar{D}/\bar{S}). \qquad [3.6]$$

These four relationships are mutually independent, i.e., they cannot be calculated from each other. Only by additional usage of the prior probabilities $P(D)$ or $P(S)$, respectively, and applying the Bayes' theorem, such a calculation would be possible. Furthermore, one should notice that the following is valid:

$$P(\bar{S}/D) = 1 - P(S/D) \qquad [3.7]$$

$$P(S/\bar{D}) = 1 - P(\bar{S}/\bar{D}) \qquad [3.8]$$

$$P(\bar{D}/S) = 1 - P(D/S) \qquad [3.9]$$

$$P(D/\bar{S}) = 1 - P(\bar{D}/\bar{S}) \qquad [3.10]$$

As is easily recognizable, the frequency of occurrence corresponds to the sensitivity and the strength of confirmation to the positive predictive value.

At present, it seems to be most advantageous to consider those relationships for diagnostic inferences which are based on $F(S \cap D) = a$, a set of patients "known" to the physician and studied in medical investigations. The other two relationships are based on $F(\bar{S} \cap \bar{D}) = d$, a set of persons which can only be included in medical studies having an extended control group that contains persons who neither show symptom S nor have disease D. By using $P(S/D)$ and $P(D/S)$ only, it is guaranteed that the physician has at least the ability to estimate in which ranges the frequencies of occurrence and strengths of confirmation lie. Furthermore, he or she is thus able to check automatically calculated relationships for plausibility. In case of $P(\bar{S}/\bar{D})$ and $P(\bar{D}/\bar{S})$ that would be more difficult or even impossible. But to include calculated relationships in the knowledge base without physician's acceptance seems, at least at present, not advised (as explained in the next section).

At this juncture, it shall not remain unmentioned that the medical relationships employed in INTERNIST-1/CADUCEUS (*15, 16*)—the frequency weight and the evoking strength—correspond directly to the frequency of occurrence and the strength of confirmation in CADIAG-2. They are defined and utilized differently and obviously developed independently (see Spindelberger and Grabner (*1*) and Pople *et al.* (*23*)).

The statistical intepretations of the SC_iD_j, S_iS_j, and D_iD_j relationships are conducted analogously to that for S_iD_j relationships. Using Table 2, the counterpart relationships in CADIAG-1 can be interpreted similarly.

4. Semiautomatic Knowledge Acquisition in CADIAG-1 and CADIAG-2

Having a statistical interpretation of the medical relationships applied in a medical expert system, an automatic calculation of these relationships from patient samples can be carried out. In the following description only the calculation of S_iD_j relationships is considered.

The results of such a calculation can be used in different stages of the development of an expert system and for different purposes: as a means to acquire new knowledge by learning from examples, as a means to check already documented knowledge in order to verify it, and as a means to check stored patient records for documentation errors. In the cases of CADIAG-1 and CADIAG-2, the calculation was conducted for the rheumatological groups that contain at present 189 rheumatological diseases, among them:

Joint diseases.

Diseases of the spinal column.

Diseases of the soft tissue and connective tissue system.

Diseases of cartilage and bone.

Systemic diseases with facultative manifestations in the locomotor apparatus.

Regional pain syndromes.

In order to document these diseases, 905 symptoms, signs, laboratory test results, findings of causative agents, biopsy and histology findings, X-ray findings, ECG findings, and concomitant diseases were considered.

The state of documentation in both systems was different when the calculation of S_iD_j relationships was carried out:

The rheumatological diseases in CADIAG-1 were fully documented (about 20,000 S_iD_j relationships).

The rheumatological diseases in CADIAG-2 were partly documented (about 2000 S_iD_j relationships).

All intended SC_lD_j, S_iS_j, and D_iD_j relationships were fully documented in both systems.

As a test sample, 3530 extended patient records collected in a rheumatological hospital over a period of 2 years were available. The data were stored within the framework of the Vienna Medical Information System. But only 103 of the above-mentioned 189 diseases occurred in this sample. The frequency of these 103 diseases varied greatly. Some diseases such as tuberculous arthritis, pseudogout, juvenile rheumatoid arthritis, and Still's disease occurred only once or twice; others such as coxarthrosis, gonarthrosis, and spondylosis occurred several hundred times. This fact is caused, in general, by the different prevalences of these diseases in the entire population and, in particular, by the special type of hospital that cares for a certain group of patients.

In order to carry out the necessary calculations, the following steps were taken:

(1) Assignment of patient data from patient records, i.e., detailed observation data and numerical test results, to symptoms in the CADIAG systems; after that every symptom had a value from {1 = present, 0 = absent, v = not yet examined}; the assignment was done using the CADIAG-1 assignment functions mentioned briefly in Section 2.

(2) Assignment of clinicians' discharge diagnoses to diseases in the CADIAG systems.

(3) For each S_iD_j combination, a 2 × 2 table was established and filled in according to the number of cases with investigated symptoms; for every symptom/disease pair the following results were printed out:

Number of patients (actual sample size) N.
Frequency of the symptom $F(S)$.
Frequency of the disease $F(D)$.
Sensitivity $P(S/D)$.
Specificity $P(\bar{S}/\bar{D})$.
Positive predictive value $P(D/S)$.
Negative predictive value $P(\bar{D}/\bar{S})$.
Result of a significance test.
Correlation coefficient with sign test for positive or negative correlation.
Confidence interval test for $\mu_O(S,D)$ and $\mu_C(S,D)$ (for CADIAG-2 only).

In this way, 93,215 2 × 2 tables were calculated (see also Deschka (*24*)).

As a significance test, the χ^2 test was chosen, but according to the sample size, Yates continuity correction and Fisher's exact test were sometimes applied (*21*). As correlation test, the ϕ coefficient (*25*), whose range lies between 0 and 1, with a simple sign test for positive or negative correlation was performed.

The printout for every disease was divided into two parts. The first part served for the comparison of the physician's and the calculated relationships. The second part provided the calculated relationships for those symptoms that were not yet included in the documentation of the respective disease.

In case of CADIAG-1, the obligatory, confirming, and excluding relationships in the first part of the lists were checked for the respective frequencies, i.e., OC and ON demanded a $P(S/D) = 1.00$, OC and FC a $P(D/S) = 1.00$, and EX a $P(S/D) = 0.00$ and $P(D/S) = 0.00$. This was not always the case: some ONs had to be substituted by FNs, and several times patient data had to be corrected. After that, symptoms with high significance and/or high positive correlation were checked for FN relationships in the first and second parts of the lists. This led to several changes, but only 3–5% of all documented relationships had to be corrected (addition or deletion of an FN relationship). In general, the correlation coefficient was more expressive concerning the association between symptoms and diseases than the significance test.

After correcting the CADIAG-1 relationships, the first part of the CADIAG-2 lists was checked. The results were similar to those attained in CADIAG-1. But additionally, the CADIAG-1 lists formed the basis for the documentation of

diseases not yet documented in CADIAG-2, but already documented in CADIAG-1. Those which were documented in CADIAG-1 were also documented in CADIAG-2, but now with the respective sensitivity and positive predictive values as frequencies of occurrence and strengths of confirmation. In some first tests with patients suffering from osteoarthroses, the diagnostic results established by CADIAG-2 on the basis of the semiautomatic knowledge acquisition confirmed the applicability of this method. At present, fully automatic acquisition of knowledge is not possible for several reasons:

Insufficient sample sizes for some diseases (just for the most important ones in an expert system, the rare diseases).

Unrepresentative samples of patients (appropriate adjustment to the respective population is necessary).

Cooccurrence of diseases in the patients (*16, 26*) (sometimes essential biases in the frequencies are produced).

Effects of previous treatment (alteration and reduction of the symptom pattern of a disease).

5. Discussion

The application of two relationships in CADIAG-2—the frequency of occurrence alias sensitivity and the strength of confirmation alias positive predictive value—seems to be very useful. These relationships are very often available (sometimes, however, only after studying extended patient samples), and they are understandable to the physician. It is estimated that about 80% of the medical assertions about associations between symptoms and diseases in textbooks are related to the frequency of occurrence and only about 20% to the strength of confirmation. In a strong logical sense, a direct inference from symptoms to diseases is only possible using the strength of confirmation (modus ponens). The frequency of occurrence, however, can be applied to exclude diagnoses (modus tollens) or to support diagnostic hypotheses by means of heuristics such as established in CADIAG-2 (*8*) or in INTERNIST-1/CADUCEUS (*15, 16, 23*).

If a specific statement about the frequency of occurrence and the strength of confirmation is not possible at all, either because there are too few cases or the medical investigations are not advanced enough, the notation of the FN relationship in CADIAG-1 is a possibility to express associations between medical entities. A similar approach that uses these simple associations between symptoms and diagnoses is successfully made in RECONSIDER (*18*).

The additional application of the specificity $P(\overline{S}/\overline{D})$ and the negative predictive value $P(\overline{D}/\overline{S})$ could probably enrich the diagnostic inference process but concrete investigations have still to be conducted.

A disadvantage of using the pure positive predictive value is that it depends on the prior probability $P(D)$ of the disease in a certain population. This seems meaningful considering a statistical investigation of the population, but for diagnosing diseases in *one* patient it loses its meaning. On the contrary, for

establishing the *correct* diagnosis in one patient, the prior probability of a disease is irrelevant. The correct diagnosis is determined by (1) considering all *possible* diseases that can account for the observed symptoms, and (2) trying to find *further* evidence for or against the diseases in order to confirm or exclude them.

This whole issue is still being debated and a convincing solution is not yet in sight. At this point, however, we plead for a kind of positive predictive value that is prior probability free and with which one can infer numerically possible diseases in the patient rather than compute their probabilities. This seems especially useful in expert systems that are supposed to prompt rare diseases with the same strength as frequent ones, provided that they match the patient's symptom pattern.

But regardless of the problems discussed, the methods for generating medical knowledge described in Sections 3 and 4 are in practical usage very able to provide a good basis for a semiautomatic acquisition of medical relationships. The two-step documentation of relationships, i.e., first, CADIAG-1 documentation of relationships that express confirmation, exclusion, obligatory occurrence, and association and after that, CADIAG-2 documentation of quantified relationships turned out to be especially appropriate, also in cases in which no semiautomatic calculation formed a basis for the documentation (e.g., for CADIAG-2/PANCREAS (*27*)).

But let us now turn to some more general remarks on knowledge acquisition and to related work published by other authors. Knowledge acquisition is considered more and more to be the crucial point in constructing expert systems (*28*). According to Buchanan *et al.* (*28*), four possible ways to acquire knowledge can be distinguished:

From expert via knowledge engineer to knowledge base.
From expert via editor program to knowledge base.
From data via induction program to knowledge base.
From textbook via language processing program to knowledge base.

The most common process of knowledge acquisition is that of extracting knowledge from a human expert, usually done by a knowledge engineer as mediator. Afterward, one tries to formalize the provided knowledge, or at least parts of it, and to store it in a knowledge base. This is a hard, time-consuming, and sometimes tedious task. Some limited help in this process can be expected from an editing program that makes the knowledge engineer no longer necessary. An example for an editing program, which also advises the user on potential modifications to the knowledge base, is the TEIRESIAS program developed by Davis (*29*). Other examples are the knowledge acquisition parts of CADIAG-1 and CADIAG-2, which provide interfaces that allow the expert physician to interact directly with the knowledge bases. The main characteristic is their ability to accept medical terms in natural language, to process them under consideration of synonymous terms, orthographic variants, and different suffixes, and to trigger whole sections of related symptoms if the entered medical term cannot be matched uniquely (*30*).

The availability of cases with known outcome suggests the development of programs that are able to learn from these cases. From the epistemological point of view, this type of knowledge formation is known as induction, which is a form of inference of laws from accumulated simple facts (see Harre (*31*) also for a critique on inductivism).

In Carbonell *et al.* (32), a general account of machine learning is presented. One of the first inductive machine learning programs was META-DENDRAL, described in Buchanan and Mitchell (*33*). Other work is reported in Hayes-Roth and McDermott (*34*) and in Michie (*35*). These systems accomplish inductive inference tasks on examples that contain no uncertainty, neither in their data nor in their outcome. Programs that cope with inductive machine learning under uncertainty are described in Blum (*36, 37*) and Drastal and Kulikowski (*38*).

Programs that employ sample cases to refine stored knowledge rather than generate it are described in Politakis and Weiss (*39*), Weiss and Kulikowski (*40*), Kulikowski (*41*), and Anderson (*42*).

A fourth source of knowledge and probably the most productive one can be found in textbooks. A successful attempt to extract knowledge from a highly structured medical dictionary of diseases resulted in the already mentioned RECONSIDER system (*18*). This system shows a very good performance in prompting possible diagnoses.

The semiautomatic knowledge acquisition systems of CADIAG-1 and CADIAG-2 accomplish inductive machine learning in large steps. Until now, the programs have been applied twice, after 1250 and 3530 cases, respectively. The increase of the case numbers yielded an improvement of the results, but there are limitations. If several hundred cases of a disease are evaluated, the next hundred no longer improve the results. But, on the other hand, there are at present only cases for 103 diseases from a total of 189, so a further collection of these rare diseases will improve the overall outcome.

The automatic generation of rules with compound antecedents, which could enhance the diagnostic process, with the described method, is limited by the combinatorial explosion that would occur. For example, the systematic checking of all symptom pairs for their statistical relationships to all diseases would make the calculation of $\binom{905}{2} \times 103 = 42{,}133{,}180$ 2×2 tables necessary. A reduction of the computational effort could be achieved by taking the mutual exclusiveness of many symptoms into account. A further improvement could be made by taking only those symptoms with a high specificity into consideration. These are those either with a high strength of confirmation or which were documented only rarely.

In the end, some implementation details can be given. The semiautomatic knowledge acquisition systems for CADIAG-1 and CADIAG-2 consist of several batch programs, which run on an IBM 4341 Model 2. The main program that calculates the respective frequencies needs about 2 Mbyte of processing storage. The total calculation time for 3530 patients, which includes the transfer of patient data from the patient data base to an intermediate file, the subsequent frequency calculations, and the 2×2 table calculations, is about 15 hr.

REFERENCES

1. SPINDELBERGER, W., AND GRABNER, G. Ein Computerverfahren zur diagnostischen Hilfestellung. *In* "Computer in der Medizin—Probleme, Erfahrungen, Projekte" (K. Fellinger, Hrsg.), pp. 189–221. Verlag Brüder Hollinek, Wien, 1968.
2. GANGL, A., GRABNER, G., AND BAUER, P. Erste Erfahrungen mit einem Computerprogramm zur Differentialdiagnose der Lebererkrankungen. *Wien. Z. Inn. Med. & Ihre Grenzgeb.* **50,** 553 (1969).
3. BARACHINI, F. "Konsistenzprüfung von Wissensbasen medizinischer Expertensysteme." Ph.D. Diss., Technical University of Vienna, Vienna, 1984.
4. ADLASSNIG, K.-P., KOLARZ, G., SCHEITHAUER, W., EFFENBERGER, H., AND GRABNER, G. CADIAG: Approaches to computer-assisted medical diagnosis. *Comput. Biol. Med.* **15,** 315 (1985).
5. ADLASSNIG, K.-P. A fuzzy logical model of computer-assisted medical diagnosis. *Methods Inf. Med.* **19,** 141 (1980).
6. ZADEH, L. A. Fuzzy sets. *Inf. Control* **8,** 338 (1965).
7. ZADEH, L. A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Trans. Syst. Man. Cybernet.* **SMC-3,** 28 (1973).
8. ADLASSNIG, K.-P., KOLARZ, G., AND SCHEITHAUER, W. Present state of the medical expert system CADIAG-2. *Methods Inf. Med.* **24,** 13 (1985).
9. CLOCKSIN, W. F., AND MELLISH, C. S. "Programming in PROLOG." Springer–Verlag, Berlin/Heidelberg/New York, 1981.
10. SHORTLIFFE, E. H., SCOTT, A. C., BISCHOFF, M. B., CAMPBELL, A. B., VAN MELLE, W., AND JACOBS, C. D. ONCOCIN: An expert system for oncology protocol management. "Proceedings, International Joint Conference on Artificial Intelligence," pp. 876–881, 1981.
11. SHORTLIFFE, E. H. "Computer-Based Medical Consultation, MYCIN." Elsevier, New York/Oxford/Amsterdam, 1976.
12. DUDA, R. O., GASCHNIG, Y. G., AND HART, P. Model design in the PROSPECTOR consultant system for mineral exploration. *In* "Expert Systems in the Micro-Electronic Age" (D. Michie, Ed.), pp. 153–167. Edinburgh Univ. Press, Edinburgh, 1979.
13. AIKINS, Y. S., KUNZ, Y. C., AND SHORTLIFFE, E. H. PUFF: An expert system for interpretation of pulmonary function data. *Comput. Biomed. Res.* **16,** 199 (1983).
14. KINGSLAND III, L. C., LINDBERG, D. A. B., AND SHARP, D. C. AI/RHEUM: A consultant system for rheumatology. *J. Med. Syst.* **7,** 221 (1983).
15. MILLER, R. A., POPLE, H. E., AND MYERS, J. D. INTERNIST-I, an experimental computer-based diagnostic consultant for general internal medicine. *N. Engl. J. Med.* **307,** 468 (1982).
16. MILLER, R. A. INTERNIST-1/CADUCEUS: Problems facing expert consultant programs. *Methods Inf. Med.* **23,** 9 (1984).
17. HORN, W. ESDAT—Decision support for primary care. "Proceedings, MEDINFO 83," pp. 484–487. North–Holland, Amsterdam/New York/Oxford, 1983.
18. TUTTLE, M. S., SMERERTZ, D. D., BLOIS, M. S., AND NELSON, N. "Expertness" from structured text? RECONSIDER: A diagnostic prompting program. "Proceedings, Conference on Applied Natural Language Processing, Santa Monica, California," pp. 124–131, 1983.
19. ADLASSNIG, K.-P., AND KOLARZ, G. CADIAG-2: Computer-assisted medical diagnosis using fuzzy subsets. *In* "Approximate Reasoning in Decision Analysis" (M. M. Gupta and E. Sanchez, Eds.), pp. 219–247. North–Holland, Amsterdam/New York/Oxford, 1982.
20. GRABNER, G., AND GRABNER, H. The Vienna general medical information system. "Proceedings, MEDIS 1975, Tokyo," pp. 156–163. Kansai Institute of Information Systems, 1975.
21. EVERITT, B. S. "The Analysis of Contingency Tables." Chapman & Hall, London, 1977.
22. BENNETT, B. M. On indices for diagnostic screening. *Biom. Z.* **17,** 303 (1975).
23. POPLE, H. E., MYERS, J. D., AND MILLER, R. A. DIALOG: A model of diagnostic logic for internal medicine. "Proceedings, International Joint Conference on Artificial Intelligence," pp. 848–855, 1975.
24. DESCHKA, W. "Symptom- und Krankheitshäufigkeiten in der computerunterstützten Medizinischen Diagnostik." Master thesis, University of Vienna, Vienna, 1983.

25. CLAUSS, G., AND EBNER, H. "Grundlagen der Statistik—Für Psychologen, Pädagogen und Soziologen." Volkseigener Verlag Volk und Wissen, Berlin, 1978.
26. ADLASSNIG, K.-P., AND GRABNER, G. Approaches to computer-assisted diagnosis in gastroenterology. *EDV in Med. Biol.* **11,** 74 (1980).
27. ADLASSNIG, K.-P., SCHEITHAUER, W., AND GRABNER, G. CADIAG-2/PANCREAS: An artificial intelligence system based on fuzzy set theory to diagnose pancreatic diseases. *In* "Proceedings, Third International Conference on System Science in Health Care," pp. 396–399. Springer–Verlag, Berlin/Heidelberg/New York/Tokyo, 1984.
28. BUCHANAN, B. G., BAARSTOW, D., BECHTEL, R., BENNETT, J., CLANCEY, W., KULIKOWSKI, C., MITCHELL, T., AND WATERMAN, D. A. Constructing an expert system. *In* "Building Expert Systems" (F. Hayes-Roth, D. A. Waterman, and D. B. Lenat, Eds.), pp. 127–167. Addison–Wesley, London, 1983.
29. DAVIS, R. Interactive transfer of expertise: Acquisition of new inference rules. *Artificial Intelligence* **12,** 121 (1979).
30. ADLASSNIG, K.-P. "Ein computerunterstütztes medizinisches Diagnosesystem unter Verwendung von fuzzy Teilmengen." Ph.D. Diss., Technical University of Vienna, Vienna, 1983.
31. HARRE, R. "The Philosophies of Science: An Introductory Survey." Oxford University Press, Oxford, 1972.
32. CARBONELL, J. G., MICHALSKI, R. S., AND MITCHELL, T. M. Machine learning: A historical and methodological analysis. *AI Mag.* Fall 83, **69** (1983).
33. BUCHANAN, B. G., AND MITCHELL, T. M. Model-directed learning of production rules. *In* "Pattern-Directed Inference Systems" (D. A. Waterman and F. Hayes-Roth, Eds.), pp. 297–312. Academic Press, New York, 1978.
34. HAYES-ROTH, F., AND MCDERMOTT, J. Knowledge acquisition from structural descriptions. "Proceedings, International Joint Conference on Artificial Intelligence," pp. 356–362, 1977.
35. MICHIE, D. Inductive rule generation in the context of the fifth generation. "Proceedings, International Machine Learning Workshop, Monticello, Illinois," pp. 65–70, 1983.
36. BLUM, R. L. Automating the study of clinical hypotheses on a time-oriented database: The RX-project. "Proceedings, MEDINFO 80," pp. 450–460. North–Holland, Amsterdam/New York/Oxford, 1980.
37. BLUM, R. L. Representation of empirically derived causal relationships. "Proceedings, International Joint Conference on Artificial Intelligence," pp. 268–271, 1983.
38. DRASTAL, G. A., AND KULIKOWSKI, C. A. Knowledge-based acquisition of rules for medical diagnosis. *J. Med. Syst.* **6,** 433 (1982).
39. POLITAKIS, P., AND WEISS, S. M. Using empirical analysis to refine expert system knowledge bases. *Artificial Intelligence* **22,** 23 (1984).
40. WEISS, S. M., AND KULIKOWSKI, C. A. "A Practical Guide to Designing Expert Systems." Rowman and Allanheld Publishers, New Jersey, 1984.
41. KULIKOWSKI, C. A. Knowledge acquisition and learning in EXPERT. "Proceedings, International Machine Learning Workshop, Monticello, Illinois," pp. 71–73, 1983.
42. ANDERSON, J. R. Knowledge compilation: The general learning mechanism. "Proceedings, International Machine Learning Workshop, Monticello, Illinois," pp. 203–212, 1983.