Tatra Mt. Math. Publ. **16** (1999), 255–272

TATRA MOUNTAINS
Mathematical Publications

THEORETICAL COMPARISON OF INFERENCE IN CADIAC AND MYCIN-LIKE SYSTEMS

Milan Daniel

ABSTRACT. CADIAG and MYCIN-like systems are expert systems with numerous applications. CADIAG–2 is a fuzzy expert system based on max-min inference, while inference in CADIAG–IV is an iterative data driven process, using fuzzy numbers. MYCIN-like systems use combining functions to calculate the global weights (degrees) of suggested diagnoses. This paper brings a comparison of CADIAG–2, CADIAG–IV and MYCIN-like systems. It studies relations between them and shows how CADIAG–IV can be embedded into MYCIN-like systems and vice-versa. Some remarks regarding their common generalization are given as well.

1. Introduction

CADIAG–2 is a medical diagnostic rule-based expert system based on methods of fuzzy set theory, composition of fuzzy relations [1, 2]. It is presented as an example of a fuzzy expert system in basic monographs as [21], [32]. CADIAG–IV [27] is a quite new generation of the system with its inference engine based on different ideas [3, 4].

Undoubtedly, one of the first medical diagnostic expert systems was MYCIN [29, 5], the so called "Grandfather of expert systems" (see [26]). The essential version of MYCIN called EMYCIN [30] has become extremely influential as a pioneering rule-based expert system shell. But it has turned out that the processing of uncertainty in a compositional (truth-functional) way makes it impossible to consider the results of MYCIN-like systems as conditional beliefs (conditional probabilities, see [19], [20]).

A systematic theoretical analysis of MYCIN-like systems (as developed in [10, 11, 12, 16, 17]) was presented in full detail in Chap. VI – IX of the monograph [15] and it was summarized in [18]. In particular, a method of guarded use of

2000 Mathematics Subject Classification: 04A72, 68T35.
Keywords: expert systems, rule based systems, fuzzy relations, fuzzy inference, MYCIN-like systems, combining functions, vagueness, uncertainty.

Partial support by the grant No. 1030601 of the GAASCR (GA AV ČR) is acknowledged.

MYCIN-like systems, that offers partial control over the probabilistic soundness of the results of MYCIN-like systems, has been suggested.

Our aim is to present a theoretical study of some relations between CADIAG (and similar fuzzy approaches) and MYCIN-like systems showing close relations. The results show that systems like CADIAG–2 or CADIAG–IV are subjected to a very similar criticism as that of MYCIN-like systems, but, on the other hand, they can benefit from the (partially) positive results of a cautious use of MYCIN-like systems. The principal ideas and results of the comparison of CADIAG–2 and MYCIN-like systems and of their common generalization are presented in [7]. This paper brings a brief review of results on CADIAG–2, it brings relations and a comparison of CADIAG–2 and CADIAG–IV, and relations and a comparison of CADIAG–IV and MYCIN-like systems from the point of view of their inference and uncertainty processing. The paper is a free continuation of [7], but the original paper is not supposed to be necessarily read to understand the present one.

The paper is organized as follows: Section 2 brings a brief description of CADIAG–2, CADIAG–IV, MYCIN-like systems and their problems and advantages. In section 3, a generalization of max-min composition to conorm-min composition, generalized Conorm-CADIAG–2, and ideas of including negative knowledge into CADIAG–2 are recalled. Main results on embedding of CADIAG–2 into MYCIN-like systems are presented as well. A generalized N-MYCIN-like system is introduced in Section 4. Section 4, describes how and under which conditions CADIAG–IV can be theoretically embedded into MYCIN-like systems. An embedding of MYCIN-like systems into CADIAG–IV is also studied. It is shown under which conditions CADIAG–IV gives the same results as corresponding MYCIN-like system and vice-versa. The last Section 5 suggests ideas of a notion of a generalized, compositional rule-based expert system. In the end, some final remarks and conclusions are stated.

2. What are CADIAG and MYCIN-like systems?

2.1 Preliminaries.

We suppose that the reader is familiar with basic notions of fuzzy sets as *fuzzy set* and its *membership function*, *fuzzy relation* and composition of fuzzy relations, specially *max-min composition*, *triangular norm* or *t-norm*, and *t-conorm* or *s-norm*, etc. For more details and examples, see [7, 32].

An *ordered Abelian group* (OAG) $G = (G, \oplus, \ominus, o, \leq)$, is a set G with an associative and commutative binary operation $\oplus$, with a neutral element o ($x \oplus o = x$) with a unary operation of inverse $\ominus$ ($x \oplus (\ominus x) = o$) and with a linear ordering such that monotonicity holds ($x \leq y \rightarrow x \oplus z \leq y \oplus z$). For the reason of uncertainty processing we extend an OAG with absorbing *extremal elements* $\bot$ and $\top$, where $\bot \oplus \top$ is not defined.

A usual example of OAG is the additive group of reals $\mathbb{R} = (\mathbb{R}, +, -, 0, \leq)$, which can be extended to $\mathbb{R}^+$ with ∞ and $-\infty$.

As another example of OAG$^+$ we present $\mathrm{MC} = \big([-1,1], \oplus_{\mathrm{MC}}, -, 0, \leq\big)$, i.e., the ordered interval $[-1,1]$ of reals, with 0 as a neutral element, with the usual $-$ for inverse elements, with extremal elements -1, 1 and with the following binary operation:

$$
\begin{aligned}
x \oplus_{\mathrm{MC}} y &= x + y - x \cdot y && \text{for } x, y \geq 0\,,\\
x \oplus_{\mathrm{MC}} y &= x + y + x \cdot y && \text{for } x, y \leq 0\,,\\
x \oplus_{\mathrm{MC}} y &= \frac{x + y}{1 - \min\big\{|x|,\, |y|\big\}} && \text{for } x \cdot y \leq 0\,.
\end{aligned}
$$

This is the combining function of EMYCIN (see [30]). Note that $\mathbb{R}^+$ and MC are isomorphic.

It is a well-known fact that a binary operation $\oplus$ (addition) of any OAG$^+$ on the ordered interval $[-1,1]$ restricted onto $[0,1]$ defines a t-conorm. For more information on OAGs see [9] or [15].

2.2 Medical expert system CADIAG.

CADIAG — Computer Assisted DIAGnosis — is a consulting expert system designed for internal medicine. It is intended to be an active assistant to physician in diagnostic situations.

CADIAG considers two classes of medical entities:

- symptoms, signs, test results, and findings (S_i),
- diseases, and diagnoses (D_j).

Besides them, a combination of symptoms and a combination of symptoms and diseases are considered as well.

In this paper we will focus on two quite different generations of the system: on CADIAG–2 [1, 2] and CADIAG–IV [4, 27]. Both of them have several different medical applications.

2.2.1 CADIAG–2.

Inference of CADIAG–2 is based on the fuzzy set theory, specially on the max-min composition of fuzzy relations.

Symptoms S_i take values μ_{S_i} in $[0,1] \cup \nu$, which indicate the degree to which a patient exhibits symptom S_i (value ν implies that symptom S_i has not yet been examined). Similarly, diseases D_j take values μ_{D_j}, fuzzy values $0 < \mu_{D_j} < 1$ represent possible diagnoses, while values $\mu_{D_j} = 1$ and $\mu_{D_j} = 0$

correspond to confirmed and excluded diagnoses, respectively. A set of symptoms is denoted by $\Sigma = \{S_1, \dots S_m\}$; $\Delta = \{D_1, \dots, D_n\}$ is a set of diseases and $\Pi = \{P_1, \dots P_r\}$ is a set of patients.

To represent patient–symptom relationships we can consider either a fuzzy relation $R_{\mathrm{PS}} \underset{\sim}{\subsetneq} \Pi \times \Sigma$, where $\mu_R(P_i, S_j)$ is defined as the degree in which it is true that the patient P_i has the symptom S_j, or a *symptom combination*, which is a simultaneous presence/absence of simultaneous presence of symptoms, i.e., a conjunction or negation of a conjunction of symptoms.

Two binary fuzzy relations $R^c_{\mathrm{SD}}, R^o_{\mathrm{SD}} \underset{\sim}{\subsetneq} \Sigma \times \Delta$ were introduced to express symptom–disease relationships, *confirmation relation* R^c_{SD} represents the so called *strength of confirmation* (c), while *occurrence relation* R^o_{SD} represents the so called *frequency of occurrence* (o).

The relationships between medical entities are given in the form of rules

$$\text{IF (antecedent) THEN (consequent) WITH } (o, c)\,.$$

Parameters o and c contain numeric and/or linguistic values which correspond to frequency of occurrence μ^o and strength of confirmation μ^c. E.g., IF (ultrasonic of pancreas is pathological) THEN (pancreatic cancer) WITH (0.75 $\sim$ often, 0.25 $\sim$ weak).

It is stated [1] that CADIAG–2's relationships can be interpreted statistically, so that parameters are conditioned frequencies, i.e., $\mu^o_{\mathrm{SD}}(S_i, D_j) = F(S_i|D_j)$ and $\mu^c_{\mathrm{SD}}(S_i, D_j) = F(D_j|S_i)$. This interpretation is often used in the case of numeric values.

As it is presented in [7], this statistical interpretation of membership degrees of fuzzy relations together with the compositional inference (see bellow) is *the main questionable point of* CADIAG–2 *and similar systems.*

- **Fuzzy logical inference in CADIAG–2**

The compositional inference rule proposed by Zadeh ([31] 1973) is adopted as an inference mechanism. In fact, it means an application of the max-min composition of fuzzy relations, $R_{\mathrm{PD}} = R_{\mathrm{PS}} \circ R_{\mathrm{SD}}$:

$$\mu_{\mathrm{PD}}(P_q, D_j) = \max_{S_i \in \Sigma} \min\left\{\mu_{\mathrm{PS}}(P_q, S_i),\ \mu^c_{\mathrm{SD}}(S_i, D_j)\right\}. \tag{1}$$

Because of the features of physicians' thinking and of distinguishing confirmation and occurrence relations, there are three types of inference rules/compositions:

1. confirmation (by present symptoms) $R^1_{\mathrm{PD}} = R_{\mathrm{PS}} \circ R^c_{\mathrm{SD}}$,
2. exclusion (by present symptoms) $R^2_{\mathrm{PD}} = R_{\mathrm{PS}} \circ \left(1 - R^c_{\mathrm{SD}}\right)$,
3. exclusion (by absent symptoms) $R^3_{\mathrm{PD}} = \left(1 - R_{\mathrm{PS}}\right) \circ R^o_{\mathrm{SD}}$.

All three inference rules/compositions are max-min compositions.

Analogous rules are defined also for another relations, but only symptom–disease and symptom combination–disease relations are used in practical applications. In this text we can consider a single symptom as a special case of symptom combination (as one-element elementary conjunction of symptoms).[1]

The value $\mu_{\mathrm{PS}}(P_q, E)$, for a symptom combination E, is computed from the values $\mu_{\mathrm{PS}}(P_q, S_i)$ for S_i occurring in E, using truth functions of fuzzy logic.

- **Results of a diagnostic process and their interpretation**

There are three different inference rules for every IF – THEN rule. They are handled separately. Thus, inference process produces, for a fixed patient, three different resulting values for every diagnosis.

Diagnosis D_j is identified as a *confirmed diagnosis*, i.e., $\mu_{D_j} = 1$ ($\mu_{\mathrm{PD}}(P_q, D_j) = 1$ for fixed patient P_q), if $\mu^1_{\mathrm{PD}}(P_q, D_j) = 1$.

Diagnosis D_j is determined as an *excluded diagnosis*, i.e., $\mu_{D_j} = 0$, if $\mu^2_{\mathrm{PD}}(P_q, D_j) = 1$ or $\mu^3_{\mathrm{PD}}(P_q, D_j) = 1$.

In other cases we speak of *possible diagnoses*. Moreover, if a diagnosis D_j can be simultaneously confirmed and excluded, we shall call D_j a diagnostic contradiction.

Possible diagnoses are compared according to values μ_{D_j}. There is also another implemented method of comparison according to the so called number of points, for details see [1] and [7].

For overview of applications, see, e.g., [25].

CADIAG–2 was subsequently improved/modified and implemented as CADIAG–2/E, and CADIAG–3 in Vienna, but we will skip them in this text. Conorm-CADIAG–2 — a theoretical generalization of CADIAG–2 — was suggested in [7], see section 3.1. An idea of including negative knowledge into CADIAG–2 is presented in [7] as well.

2.2.2 CADIAG–IV.

A new generation of CADIAG — CADIAG–IV [4] — which is a part of the MedFrame project [27] — is under development at the Department of Medical Computer Science in Vienna. It tries to avoid problems of previous generations of CADIAG.

CADIAG–IV assigns two values to every medical term T: a strength of evidence T^+, and a strength of counterevidence T^-, from interval $[0, 1]$; in full generality, fuzzy numbers can be used. In correspondence with it, there are two types of rules in CADIAG–IV: rules that give evidence for, and rules that give counterevidence against the right-hand side. The occurrence value that was used

[1] In the rest of this paper, $E \in \Sigma$, for a symptom combination E, denotes a combination of symptoms from Σ.

in CADIAG–2 to deduce exclusions is no longer used for the inference process in CADIAG–IV.

Thus, evidence is inferred separately from counterevidence, similarly as in the case of *confirmation relation* R^{c}_{PD} and *exclusion relation* R^{e}_{PD} in Conorm-CADIAG–2 [7], but any total value corresponding to value $\mu^{\mathrm{tot}}_{\mathrm{PD}} = \mu^{c}_{\mathrm{PD}} \oplus -\mu^{e}_{\mathrm{PD}}$ is not computed in CADIAG–IV. Hence, there are two resulting values for every diagnosis D_i, i.e., D_i^+, D_i^-.

There are no fuzzy relations R_{SD}, R_{PS}, R_{SD} considered in CADIAG–IV. The IF – THEN rules are given in the following form

$$r = (l_r, b_r, \sigma_r, n_r)\,,$$

where $l_r \in \mathcal{L}$ is a left-hand side operator applied on the set of medical entities from knowledge base, b_r is an occurrence value (a weight of rule), $\sigma_r \in \{+,-\}$ expresses a type of rule, and n_r is succedent (right-hand side of the rule) — a medical entity from knowledge base. A set of operators $\mathcal{L}$ is the smallest set of logic operators containing all projections π_i (selection of ith medical entity from knowledge base) which is closed under $\neg$, $\wedge$. Disjunction $\vee$ and the operator atleast^n_k are considered in $\mathcal{L}$ as well. Disjunction is defined by $a \vee b = \neg(\neg a \wedge \neg b)$, and the operator atleast^n_k (at least k arguments from n ones are true[2]) is expressible by $\vee$ and $\wedge$, i.e., by $\neg$ and $\wedge$ only, as well. A fuzzification of operators from $\mathcal{L}$ is used for evidence of weights of rules which are represented by fuzzy numbers, see [3].

There are other implicit natural limitations for knowledge bases as, e.g., symptoms (patient data) not allowed in right-hand side of rules, etc.

A rule base of CADIAG–IV as it is described in [3, 4] is more general than that one of CADIAG–2, in the sense that it admits chains of rules and even cycles.

• An Inference in CADIAG–IV

An inference in CADIAG–IV [3, 4] is an iterative data-driven process. Contributions of rules are summarized by operation $\oplus$, $x \oplus y = x + y - xy$ (i.e., by probabilistic sum, or by positive cone of MYCIN's group operation, see further). In CADIAG–IV, negation $\neg$ is defined as $\neg(x^+, x^-) = (x^-, x^+)$, conjunction $\wedge$ is defined as $x \wedge y = (x^+y^+, x^- \oplus y^-)$, A contribution of a rule is the weight of its left-hand side multiplied by σ^+ (resp. by σ^-).

For *low knowledge bases* (where operators from $\mathcal{L}$ are applied only on symptoms, and n_r represents a diagnosis, there are no chains of rules) the inference is

[2]In fact, operator atleast^n_k which is used in CADIAG–IV is a simplification (an abbreviation) of it due to more efficient computing, hence atleast^n_k used in CADIAG–IV is not expressible by $\neg$ and $\wedge$, but we shall omit it in this text. In CADIAG–IV there holds atleast^n_k = multiple of k maximal x_i^+ (multiple of $n-k+1$ maximal x_j^-), where $i,j \in \{1,\ldots,n\}$.

similar to that one from CADIAG–2, but different operations are used: $\oplus$ instead of $\max$, and multiplication instead of $\min$.

For general knowledge bases, the inference is, in fact, a forward chaining of rules. We may obtain a cyclic inference for a cyclic knowledge base. The iterative process continues as long as values of medical entities are changing.

Fact 2.1. *From the theoretical point of view, the iterative inference of* CADIAG–IV *as it is described in* [3, 4] *allows an endless cyclic process.*

Note:

- From the theoretical point of view, a cyclic inference terminates after a finite number of inference steps if and only if value 1 is propagated from outside at least in one entity from every inference cycle (i.e., not to cycles of rules, where corresponding components of values of all medical entities are equal to 0).
- From the implementational point of view, all cyclic inferences of CADIAG–IV finish, because of finite representation of reals in a computer. Thus, weights of diagnoses $(D_i^\sigma)^{(r)}$ after r iterations are not distinguishable from $(D_i^\sigma)^{(r+1)} > (D_i^\sigma)^{(r)}$ due to precision of computing (for corresponding σ and for a sufficiently large finite r).
- It is a question whether $(D_i^\sigma)^{(r)}$, which is not distinguishable from $(D_i^\sigma)^{(r+1)} > (D_i^\sigma)^{(r)}$, is distinguishable from $\lim\limits_{t\to\infty}(D_i^\sigma)^{(t)}$ or not.

Fact 2.2. *There is not known any reasonable interpretation for a cyclic inference process. It seems that there is no reasonable interpretation of it.*

On the other hand, an idea of iterative inference should be usefully used in acyclic iterations over cycles in a knowledge base from a certain class of cyclic knowledge bases.

Hence, inference with a cyclic knowledge base must be improved, in future, to be reasonably usable. Moreover, no cyclic applications of CADIAG–IV have been implemented till now. Thus, further in this text, we can consider for CADIAG–IV only acyclic knowledge bases.

2.3 MYCIN-like systems.

Intuitively, MYCIN-like systems are rule-based knowledge systems with weighted rules of the form $E \Rightarrow H(w)$ (with the meaning "if the antecedent E holds then the succedent H holds with the weight w") with an inference engine similar to that of (E)MYCIN. (The expert system PROSPECTOR is a famous example)[3].

[3] In our context, we can mention also an investigation on MYCIN-like systems with fuzzy connectives, see [23, 24].

These systems deal with *belief* via combining functions, i.e., in a compositional way, which is a questionable matter in the same way as inference in CADIAG.

To analyse these systems formally, a definition of MYCIN-like systems was suggested (in [16, 17]) which formalizes — *mutatis mutandis* — the main common features of these systems. Since we will not need their definition in its full generality (i.e., allowing *acyclic chains of rules*) we present here a particular case of this notion, sufficient for our purposes.

One has *symptoms* $S_1, \ldots, S_n$ and *diagnoses* $D_1, \ldots, D_m$. Rules have the form $E \Rightarrow D(w)$, where E is an *elementary conjunction* of some symptoms, i.e., the conjunction of some symptoms and of some other negated symptoms, e.g., $S_1 \,\&\, \neg S_3 \,\&\, S_5$ means "symptom S_1 is present, and S_3 absent, and S_5 present". Mostly we will work with the interval $[-1,1]$ as the set of weights: 1 meaning "certainly yes", -1 meaning "certainly no", and 0 meaning "no preference". (This can be isomorphically shifted to $[0,1]$, with 0 meaning "certainly no" and $\frac{1}{2}$ "no preference".)

A *rule base* is a set Θ of rules. A user assigns weights to symptoms (default value being "no preference"); thus formally he/she defines a mapping q assigning to each symptom S its weight $q(S)$. From this a system computes the weights of each used elementary conjunction of symptoms, using truth functions (here called *combining functions*) of fuzzy logic over $[-1,1]$: $\mathrm{NEG}(x) = -x$ for negation, $\mathrm{CONJ}(x,y) = \min(x,y)$ for conjunction. We will write $W(E|q)$ for the weight of E given by the questionnaire q.

Each rule $R = \big(E \Rightarrow D(w)\big)$ determines the *contribution* $C(R|q)$ of R given q, defined from the weight $a = W(E|q)$ of the antecedent given by q, and from the weight w of the rule, where $\mathrm{CTR}(a,w) = 0$ if $a \le 0$, $\mathrm{CTR}(a,w) = \min\{a,w\}$ if $a,w \ge 0$, $\mathrm{CTR}(a,w) = -\min\{a,-w\}$ if $a > 0,\ w < 0$.

The main point comes now: A *global weight* of a diagnosis D given q, $W(D|q)$ is computed using an operation $\oplus$ as the $\oplus$-sum of contributions of all rules whose succedent is D (rules *leading* to D):

$$W(D|q) = C(R|q) \oplus \cdots \oplus C(R'|q)\,,$$

where $R, \ldots, R'$ are all rules from Θ leading to D. *Basic assumption:* the ordered set of weights $[-1,1]$ (or $[0,1]$ if desired) together with $\oplus$ is an extended ordered Abelian group.

This completes the definition of a MYCIN-like system and related notions. Summarizing, *a particular* MYCIN-*like system is fully determined by a set* Θ *of rules and by a group operation* $\oplus$.[4]

[4] Note that this formalism makes rules of the form $E \Rightarrow \neg D(w)$ superfluous; due to our interpretation of negation, such a rule is equivalently represented as $E \Rightarrow D(-w)$. The inference mechanism just combines both positive and negative contributions to the global weight of D.

Classical examples:

(1) EMYCIN*:* $\oplus_{\mathrm{MC}}$ is the operation described above.

(2) PROSPECTOR*:* presented with weights $[0,1]$ and $\oplus$ using $x \oplus_{PP_0} y = \frac{xy}{xy+(1-x)(1-y)}$.

Note again that $\oplus_{\mathrm{MC}}$ and $\oplus_{PP_0}$ are isomorphic, see [10], [15].

2.4 Problems and disadvantages of investigated systems.

The evident common advantage of both MYCIN-like systems and CADIAG is that they are simple and (often) give reasonable (acceptable) results.

Their common weakness is the compositional (truth-functional) processing of belief: they *propagate* the data on a patient through the rules, *compose* the effects (contributions) of the rules and *infer* a number attached to a diagnosis: *What is this number?* One would like to get something like (an approximation of) the belief of the presence of D, given the knowledge Θ and data q. But do we get this? *Not necessarily!*

The simplest test case follows: the presence of symptom S_1 gives weight x to D, the presence of S_2 gives weight y to the same diagnosis D; what about the presence of both S_1 and S_2? If the weight is a relative frequency (relative probability), then observe that $P(D\,|\,S_1\&S_2)$ *is obviously not a function* of $P(D|S_1)$, $P(D|S_2)$; CADIAG–2 just takes $\max\{P(D|S_1),\ P(D|S_2)\}$; a MYCIN-like system (with probabilities linearly transformed to $[-1,1]$) takes $P(D|S_1) \oplus P(D|S_2)$ which is larger than both summands if they are both positive, smaller if they are both negative, and otherwise it is between them. (Some authors call such an operation *compensatory*, see, e.g., [22]).

The question what is better has no answer: both can be wrong. On the one hand, in MYCIN-like systems and in CADIAG–IV, one can have very similar rules whose contributions sum up to unwontedly large weight; on the other hand, CADIAG–2 does not pay attention to several *independent* rules confirming the same diagnosis with an equal weight, it just gives you the common weight as a result. The problem is partially overcame by using of a disjunction in antecedent of such rules in CADIAG–IV.

From the positive side, both systems share the property that they can never get certainty (value 1) from uncertain contributions.

In MYCIN-like systems, a rule with a negative weight may diminish an effect of a rule with a positive weight — this feature is absent in the max-min inference of CADIAG–2 (which is one of its weaknesses), it is absent also in CADIAG–IV which does not combine D_j^+ with D_j^- to make a final value $D = D_j^+ \oplus D_j^-$.

Let us illustrate this on examples concerning CADIAG–2. They will be discussed again later as a motivation for the generalization of CADIAG.

Example 1. Let us have a knowledge base including the following three rules:

R_1: IF S_1 THEN D_1 WITH $(c_1 = 0.7)$,

R_2: IF S_2 THEN D_2 WITH $(c_2 = 0.7)$,

R_3: IF S_3 THEN D_2 WITH $(c_3 = 0.7)$.

Let the patient P_q have symptoms S_1, S_2, S_3 with the same degree, e.g., 1, i.e., $\mu_{R_{\rm PS}}(P_q, S_1) = \mu_{R_{\rm PS}}(P_q, S_2) = \mu_{R_{\rm PS}}(P_q, S_3) = 1$. In this case, CADIAG–2 generates the following hypotheses:

– from R_1: we get $\mu_{D_1} = 0.7$, that means disease D_1 with degree 0.7,

– from R_2, R_3: we get $\mu_{D_2} = 0.7$, that means disease D_2 with degree 0.7.

CADIAG–2 concludes both diagnoses D_1 and D_2 with the same degree 0.7, because $\max\{\mu_{R_{\rm SD}}(S_2, D_2),\ \mu_{R_{\rm SD}}(S_3, D_2)\} = \max\{0.7, 0.7\} = 0.7$ (the operation max is idempotent). Due to nonidempotency of group operation, this is a problem neither of CADIAG–IV nor of MYCIN-like systems.

Example 2. Let us consider a disease D supported by either of the symptoms S_1 and S_2 separately, but not by their conjunction. For a part of patient database, see [7].

Let us have rules:

R_1: IF S_1 THEN D WITH $(o_1 = 0.4166,\ c_1 = 0.7)$,

R_2: IF S_2 THEN D WITH $(o_2 = 0.4166,\ c_2 = 0.7)$,

R_3: IF $S_1 \& S_2$ THEN D WITH $(o_3 = 0.0059,\ c_3 = 0.0333)$.

Let us further suppose that a patient P has both of the symptoms S_1 and S_2. In this case we get the same value $\mu(P, D) = 0.7$ as in the case in which only one of the symptoms S_1 or S_2 is present. A contribution of rules R_1 and R_2 cannot be decreased by any other rule (with exception of a rule with $o = 1$). But from the intuitive point of view it is obvious that $\mu(P, D)$ should have a far less value; to keep statistical interpretation it should be even $\mu(P, D) = 0.0333$. There is the same problem in CADIAG–IV. We get $D^+ = 0.7 \oplus 0.7 \oplus 0.0333$, and there is no comparison of $(0.7,\ D^-)$ and of $(0.7 \oplus 0.7,\ D^-)$ etc.

To analyze and understand the processing of beliefs in fuzzy expert systems similar to CADIAG–2 and CADIAG–IV, we suggest *first* to analyze their structures, relating it to the structure of MYCIN-like systems and to propose a possible generalization of both of them (this is the main purpose of this paper); and, *second* to apply and generalize the method of (partial) improvement of probabilistic properties of MYCIN-like systems developed by Hájek and Valdés (Möbius transform, the "method of guarded use", see [15]) to CADIAG–2 and to CADIAG–IV.

3. Embedding of CADIAG–2 into MYCIN-like systems

3.1 Conorm-CADIAG–2.

Let us consider once again a problem of distinguishing simple or multiple contributions, see Example 1. Observe that the operation max (which composes contributions in max-min inference) is idempotent, i.e., $\max(x,x) = x$ for all x. We know that the operation max is a t-conorm. We can generalize max-min inference to *conorm-min inference* in the following way:

$$\mu_{\mathrm{PD}}(P_q, D_j) = \bigvee_{E_i \in \Sigma} \min\left\{\mu_{\mathrm{PS}}(P_q, E_i)\,,\ \mu^c_{\mathrm{SD}}(E_i, D_j)\right\}, \qquad (2)$$

where $\bigvee$ stands for a t-conorm.

If we take a non-idempotent t-conorm, e.g., the operation $\oplus_{\mathrm{MC}}$ from subsection 2.1 which is restricted to interval $[0,1]$, then we obtain

$$\mu_{\mathrm{PD}}(P_q, D_2) = \min\left\{1, \mu^c_{\mathrm{SD}}(S_2, D_2)\right\} \vee \min\left\{1, \mu^c_{\mathrm{SD}}(S_3, D_2)\right\} = 0.7 \vee 0.7 > 0.7$$

(using $\oplus_{\mathrm{MC}}$ we get $0.7 \oplus_{\mathrm{MC}} 0.7 = 0.91$). This enables us to distinguish the cases of diagnoses D_1 and D_2 from Example 1. (But again: this is sometimes desirable, sometimes not.)

From now on the generalized CADIAG–2 with conorm-min inference (2) instead of max-min inference (1) will be called *Conorm-CADIAG–2*.

3.2 Embedding of Conorm-CADIAG–2 into MYCIN-like systems.

Let us assume Conorm-CADIAG–2 with a t-conorm given by a group operation $\oplus$ of any OAG^+, thus we have the following conorm-min inference:

$$\mu_{\mathrm{PD}}(P_k, D_j) = \bigoplus_{E_i} \min\left\{\mu_{\mathrm{PS}}(P_k, E_i)\,,\ \mu^c_{\mathrm{SD}}(E_i, D_j)\right\}, \qquad (3)$$

where E_i is a symptom or symptom combination.

In [7], it is shown that we can construct a MYCIN-like system which gives exactly the same results as our Conorm-CADIAG–2. To accomplish this, we need some auxiliary notions.

A *rule base* Θ *given by the confirmation relation* R^c_{SD} *and by the occurrence relation* R^o_{SD} is a rule base Θ of MYCIN-like system with rules of the form

$$\begin{aligned}
&E_i \Rightarrow D_j\big(\mu^c_{\mathrm{SD}}(E_i, D_j)\big)\,, && \\
&E_i \Rightarrow \bar{D}_j(1) && \text{if} \quad \mu^c_{\mathrm{SD}}(E_i, D_j) = 0\,,\\
&\bar{E}_i \Rightarrow \bar{D}_j(1) && \text{if} \quad \mu^o_{\mathrm{SD}}(E_i, D_j) = 1\,,
\end{aligned}$$

where E_i is a symptom or symptom conjunction, $\bar{E}_i$ is an absent symptom or combination of absent symptoms, D_j is a diagnosis, and $\bar{D}_j$ is a new auxiliary diagnosis[5] corresponding to D_j. (Note: Here, all the weights are nonnegative numbers.)

Furthermore, a *questionnaire q given by patient–symptom relation* R_{PS} for a patient P_k is a mapping q assigning to each single symptom S_i its weight $q(S_i) = \mu_{\mathrm{PS}}(P_k, S_i)$.

Results inferred with rule base Θ given by R^c_{SD} and R^o_{SD} are interpreted as follows:

1) diagnosis D_j is confirmed if $W_\Theta(D_j|q) = 1$,

2) diagnosis D_j is excluded if $W_\Theta(\bar{D}_j|q) = 1$,

3) other diagnoses are possible with weight $W_\Theta(D_j|q)$.

Thus we have the following embedding theorem:

THEOREM 3.1. *Let R^c_{SD} be a confirmation relation, R^o_{SD} be an occurrence relation, and R_{PS} be a patient–symptom relation, let $\bigvee$ be a t-conorm given by an Abelian group operation $\oplus$. Further, let Θ be a rule base given by R^c_{SD} and R^o_{SD}, and let q be a questionnaire given by R_{PS} for a patient P_k.*

*Then the Conorm-*CADIAG–2 *with the conorm $\bigvee$ gives the same results as a corresponding* MYCIN*-like system (i.e.,* MYCIN*-like system using the operation $\oplus$, the rule base Θ, and the questionnaire q).*

3.3 Including negative knowledge into CADIAG–2.

We have seen in Subsection 2.4 that CADIAG–2 cannot decrease contribution of one rule or of a group of rules by another rule(s). We extend CADIAG–2 with a new symptom (combination)–disease relation: an *exclusion relation* $R^e_{\mathrm{SD}} \subsetneq \Sigma \times \Delta$, represented by μ^e_{SD} in $[0, 1]$. The value $\mu^e_{\mathrm{SD}}(E_i, D_j)$ indicates the degree in which the present symptom combination E_i excludes (or disconfirms) the diagnosis D_j.

We suppose that it is impossible that symptom combination E_i both confirms and excludes diagnosis D_j. Thus we make the following consistency assumption:

$$\mu^c_{\mathrm{SD}}(E_i, D_j) = 0 \qquad \text{or} \qquad \mu^e_{\mathrm{SD}}(E_i, D_j) = 0\,.$$

That means there are two different fuzzy degrees for a couple of present symptom combinations, but at least one of them is equal to zero.

We define exclusion of disease D_j for patient P_q from observed symptom combination E_i: $R^e_{\mathrm{PD}} = R_{\mathrm{PS}} \circ R^e_{\mathrm{SD}}$. Using conorm-min composition, we get values $\mu^c_{\mathrm{PD}}(P_q, D_j)$ and $\mu^e_{\mathrm{PD}}(P_q, D_j)$.

[5]When first considered, it looks rather easier and sufficient to include rules $S_i \Rightarrow D_j(-1)$ instead of using of auxiliary diagnosis. But, such rules always bring their effect to diagnosis, while we want use them only for totally sure exclusion, i.e., only in situations of sure S_i (sure $\neg S_i$ respectively), i.e., only if $\mu_{\mathrm{PS}}(P, S_i) = 1$ (or $\mu_{\mathrm{PS}}(P, S_i) = 0$ respectively).

Using it, we define the relation of a total degree (for confirmation and exclusion) of diagnosis D_j by patient P_q from the observed symptoms as follows:

$$\mu^{\mathrm{tot}}_{\mathrm{PD}}(P_q, D_j) = \mu^{c}_{\mathrm{PD}}(P_q, D_j) \oplus -\mu^{e}_{\mathrm{PD}}(P_q, D_j)\,. \tag{4}$$

This completes our definition of Conorm-CADIAG–2 with negative knowledge.

Analogously as for Conorm-CADIAG–2 in previous subsection, we can formulate and prove that the embedding theorem for Conorm-CADIAG–2, which is extended by the handling of negative knowledge, takes the form described in [7].

4. Embedding of CADIAG–IV into MYCIN-like systems and vice-versa

Due to operation $\oplus$, CADIAG–IV has some advantages/disadvantages of Conorm-CADIAG–2. It would be useful to include negative knowledge in CADIAG–IV in the same way as it is introduced in [7]. Negative knowledge is really similarly cumulated: positive one for confirmation separately from negative one for exclusive one (analogy of μ^c and μ^e), but it is not combinated in total value for/against a given medical entity (there is no analogy of μ^{tot}) in CADIAG–IV.

On a first glance, a difference between CADIAG–IV and MYCIN-like systems is caused by more complicated antecedents allowed in CADIAG–4 and by the way of representing of conjunction and of rule contribution: multiplication in CADIAG–IV while minimum in MYCIN-like system.

We want to make a generalization analogical to CADIAG–2 and Conorm-CADIAG–2, but this time in the family of MYCIN-like systems. We want to replace min by a general t-norm and to get ‘Norm-MYCIN-like system’. It is not as simple as in the case of Conorm-CADIAG–2. Minimum is not used as a t-norm in MYCIN-like systems, because it is applied on interval of weights $[-1, 1]$. After isomorphical transformation of weights onto interval $[0, 1]$ we get another operation for conjunction. But, the operation of minimum which represents conjunction on MYCIN-like systems is expressible by the t-norm min and its dual conorm max as follows:

$$\begin{aligned}
x \,\&\, y &= \min(x, y) = t(x, y) && \text{for } x, y \geq 0\,,\\
x \,\&\, y &= \min(x, y) = -\max(-x, -y) = -s(-x, -y) && \text{for } x, y \leq 0\,,\\
x \,\&\, y &= \min(x, y) = -\min(-x, y) = -t(-x, y) && \text{for } x \leq 0 \leq y\,,\ |x| \leq |y|\,,\\
x \,\&\, y &= \min(x, y) = -\max(-x, y) = -s(-x, y) && \text{for } x \leq 0 \leq y\,,\ |x| \geq |y|\,.
\end{aligned}$$

A contribution of a rule is

$$\begin{array}{ll} \mathrm{CTR}(a,w) = \min(a,w) = t(a,w) & \text{for } a \geq 0\,,\ w \geq 0\,, \\ \mathrm{CTR}(a,w) = -\min(-a,w) = -t(-a,w) & \text{for } a \geq 0\,,\ w \leq 0\,, \\ \mathrm{CTR}(a,w) = 0 & \text{for } a \leq 0\,. \end{array}$$

We can use the same procedure for any t-norm t and its dual conorm s. Hence we get a generalization of MYCIN-like systems. From now on, let us call it *N*-MYCIN-*like system* given by its group operation and by a t-norm t.

For multiplication and $\oplus$ we obtain:

$$\begin{array}{ll} x \,\&\, y = xy & \text{for } x,y \geq 0\,, \\ x \,\&\, y = -(-x \oplus -y) = x \oplus y & \text{for } x,y \leq 0\,, \\ x \,\&\, y = -(-xy) = xy & \text{for } x \leq 0 \leq y\,,\ |x| \leq |y|\,, \\ x \,\&\, y = -(-x \oplus y) = x \oplus -y & \text{for } x \leq 0 \leq y\,,\ |x| \geq |y|\,. \end{array}$$

A contribution of a rule is

$$\begin{array}{ll} \mathrm{CTR}(a,w) = aw & \text{for } a \geq 0\,,\ w \geq 0\,, \\ \mathrm{CTR}(a,w) = -(-aw) = aw & \text{for } a \geq 0\,,\ w \leq 0\,, \\ \mathrm{CTR}(a,w) = 0 & \text{for } a \leq 0\,. \end{array}$$

Theorem 4.1. *Let us have a knowledge base θ for* CADIAG–IV *with left-hand sides of rules in a form of elementary conjunctions. Then there exists a knowledge base for a N*-MYCIN-*like system given by group operation $\oplus$ and by multiplication on $[0,1]$ as a t-norm operation such that the N*-MYCIN-*like system gives the same results as* CADIAG–IV.

Idea of proof: We will introduce new auxiliary medical entities $\overline{X}_i$ for every X_i from θ and put it in a new knowledge base Θ. For every rule $r_i = (K_1\&\ldots\&K_k, w_j, +, \pi_m)$ from θ we create and put a rule R_i: $L_1\&\ldots\&L_k \Rightarrow D_m(w_j)$ into Θ, and for every $r_j = (K_1\&\ldots\&K_k, w_i, -, \pi_m)$ we put there R_j: $L_1\&\ldots\&L_k \Rightarrow \overline{D}_m(-w_j)$, where $L_l = X_k$ if $K_l = \pi_k$ and $L_l = \neg X_k$ if $K_l = \neg\pi_k$.

If there is an evidence (S_i^+, S_i^-) for symptom S_i, then we will answer S_i^+ for S_i while S_i^- for $\overline{S}_i$ within a process of consultation with N-MYCIN-like system given by group operation $\oplus$, by t-norm min and by the knowledge base Θ.

THEOREM 4.2. *Let us have a low*[6] *knowledge base* Θ *for a* N-MYCIN-*like system given by group operation* $\oplus$ *and by multiplication on* $[0,1]$ *as a t-norm operation. Then there exists a knowledge base* θ *for* CADIAG–IV *such that, after* $\oplus$*-summarization* $D = D^+ \oplus -D^-$, CADIAG–IV *gives the same results as the considerated* N-MYCIN-*like system.*

Idea of proof: For every rule $R_j: L_1 \& \ldots \& L_k \Rightarrow D_m(w_j)$, where L_i is either A_k or $\neg A_k$ for some proposition A_k from Θ, we create and put into θ either a rule $r_j = (K_1 \& \ldots \& K_k, w_j, +, \pi_m)$ if $w_j > 0$ or $r_j = (K_1 \& \ldots \& K_k, -w_j, -, \pi_m)$ if $w_j < 0$, where $K_i = \pi_k$ if $L_i = A_k$ or $K_i = \neg\pi_k$ if $L_i = \neg A_k$.

If there is $q(S_i) \geq 0$, then we put $S_i^+ = q(S_i)$, $S_i^- = 0$; otherwise, we put $S_i^+ = 0$, $S_i^- = -q(S_i)$.

Note: In order that the last theorem were valid for general knowledge bases with chain of rules, it would be necessary to compute $X = X^+ \oplus -X^-$ for all intermediate mathematical entities, what means in fact full acceptation of processing of negative knowledge described in [7].

5. A generalization of compositional rule based systems

The results from previous chapter and from [7] tend to a definition of a general system which would be a generalization of CADIAG–2, CADIAG–IV (resp. their modifications which were introduced), and MYCIN-like systems together.

But there is still a lot of open problems here. E.g., a formulation and validity of embedding theorems for knowledge bases with general formula using conjunction, disjunction and negation in antecedents of rules. Among steps to such a generalization there is also a generalized Möbius transformation, see [6], a useful tool for construction of knowledge bases. Another problem is the question of general knowledge bases with chains of rules.

6. Conclusion

In this paper we have studied relations of CADIAG–2 and of CADIAG–IV with MYCIN-like systems.

It was explained that both approaches have both advantages and weak points, so that neither CADIAG nor MYCIN-like systems are better or worse in a general sense. To repeat once again: the common weak point is the *compositionality*, which is difficult to defend if weights of rules are relative frequencies.

[6] Just as in our definition, see Section 2.3, no chains of rules are admitted.

Even if the inference in CADIAG–2 and CADIAG–IV are based on different ideas and motivations which are not similar to those of MYCIN-like systems, it was shown that from the theoretical point of view, their inference MYCIN-like systems are as good as both CADIAG–2 and CADIAG–IV and, moreover, CADIAG–IV is as good as MYCIN-like systems as well.

Advantages and weak points of different types of investigated systems which is possible to improve were studied. Obviously, the investigated systems are different in a lot of details. Nevertheless, the nature of their inference is almost the same. First of all, the only principal difference, i.e., inference above cyclic knowledge bases in CADIAG–IV, should be more elaborated and improved.

It is possible to use usefully all types of investigated systems in practise, but it is necessary to keep in mind their weak points, a common nature of their inference.

Finally, remarks to a common generalization of CADIAG–2, CADIAG–IV and MYCIN-like systems were presented. This opens the way to an analysis of CADIAG-like systems similar to that of MYCIN-like systems.

Acknowledgement. I am grateful to Laurenz Brein for a discussion and explanation concerning CADIAG–IV.

REFERENCES

[1] ADLASSNIG, K. P.: *Fuzzy set theory in medical diagnosis*, IEEE Trans. Systems, Man Cybernet. **SMC-16** (1986), 260–265.

[2] ADLASSNIG, K. P.: *Update on* CADIAG–2*: a fuzzy medical expert system for general internal medicine*, in: Progress in Fuzzy Sets and Systems (W. H. Janko, M. Roubens, and H.-J. Zimmermann, eds.), Kluwer Academic Publishers, 1990, pp. 1–6.

[3] BREIN, L.: *Das Medizinische Expertensystem* CADIAG–IV*: Inferenzproceß, Erklärungen und Untersuchungsvorschläge*, Master thesis, Technische Universität Wien, 1997.

[4] BREIN, L.—ADLASSNIG, K. P.—KOLOUSEK, G.: *Rule Base and Inference process of the Medical Expert System* CADIAG–IV., Draft version, 1998.

[5] BUCHANAN, B.—SHORTLIFF, E.: MYCIN *Experiments of the Stanford Heuristic Programming Project*, Addition-Wesley, Reading, MA, 1984.

[6] DANIEL, M.: *Generalized Möbius transformation of knowledge bases*, Kybernetika **34** (1998), 91–104.

[7] DANIEL, M.—HÁJEK, P.—NGUYEN, P. H.: CADIAG–2 *and* MYCIN-*like systems*, Artificial Intelligence in Medicine **9** (1997), 241–259.

[8] DUBOIS, D.—PRADE, H.: *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, New York, 1980.

[9] FUCHS, L.: *Partially ordered algebraic structures*, Pergamon Press, Oxford, 1963.

[10] HÁJEK, P.: *Combining functions for certainty degrees in consulting systems*, Int. J. Man-Machine Studies **22** (1985), 59–76.

[11] HÁJEK, P.: *Towards a probabilistic analysis of* MYCIN-*like expert systems*, in: COMPSTAT '88, Physica-Verlag, Copenhagen, Heidelberg, 1988.

[12] HÁJEK, P.: *Towards a probabilistic analysis of* MYCIN-*like expert systems II*, in: Artificial Intelligence and Information-Control Systems of Robots (I. Plander, ed.), Amsterdam, North-Holland, 1989, pp. 45–54.

[13] HÁJEK, P.—HÁJKOVÁ, M.: *The expert system shell EQUANT-PC: Philosophy, structure and implementation*, Computational Statistic Quarterly **4** (1990), 261–267.

[14] HÁJEK, P.—HÁJKOVÁ, M.—HAVRÁNEK, T.—DANIEL, M.: *The expert system shell EQUANT-PC: Brief information*, Kybernetika **25** (1989), (1–3):suppl. 4–9.

[15] HÁJEK, P.—HAVRÁNEK, T.—JIROUŠEK, R.: *Uncertain Information Processing in Expert Systems*, CRC Press, Boca Raton, 1992.

[16] HÁJEK, P.—VALDÉS, J. J.: *Algebraic foundations of uncertainty processing in rule-based expert systems (group-theoretical approach).*, Computers and Artificial Intelligence **9** (1990), 325–334.

[17] HÁJEK, P.—VALDÉS, J. J.: *Generalized algebraic foundations of uncertainty processing in rule-based expert systems (dempsteroids)*, Computers and Artificial Intelligence **10** (1991), 29–42.

[18] HÁJEK, P.—VALDÉS, J. J.: *An analysis of* MYCIN-*like expert systems*, Mathware **1** (1994), 45–68.

[19] HECKERMAN, D. F.: *Probabilistic interpretations for mycin's certainty factor*, in: Uncertainty in Artificial Intelligence Pages, Amsterdam, North Holland, 1986, pp. 167–196.

[20] JOHNSON, R.: *Independence and bayesian updating methods*, in: Uncertainty and Artificial Intelligence (L. Kanal and J. Lemmer, eds.), Amsterdam, North-Holland, 1986, pp. 197–200.

[21] KLIR, G. J.—FOLGER, T. A.: *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall International, 1988.

[22] MESIAR, R.: *Compensatory operators based on triangular norms and conorms*, in: Proceedings of EUFIT '95, Aachen, August 1995, pp. 131–135.

[23] MESIAR, R.—PÍŠ, P.: *Fuzzy model of inexact reasoning in medical consulting systems*, in: Proceedings of AICSR'89 (I. Plander, ed.), Elsevier, Štrbské Pleso, 1989, pp. 155–158.

[24] MESIAR, R.—PÍŠ, P.: *Fuzzy model of inexact reasoning in medicine*, Comp. Meth. and Progr. in Biomedicine **30** (1989), 1–8.

[25] NGUYEN, P. H.: *Some contributions to fuzzy expert systems with emphasis on medicine.*, Technical Report Ph.D. Dissertation 9327636, Technische Universität Wien, 1995.

[26] PUPPE, F.: *Systematic Introduction to Expert Systems*, Springer-Verlag, Berlin, 1993.

[27] SAGEDER, B.—BOEGL, K.—ADLASSNIG, K. P.—KOLOUSEK, G.—TRUMMER, B.: *The knowledge model of MedFrame/*CADIAG–IV, in: Medical Informatics Europe '97C (Pappas, N. Maglaveras, and J.-R. Scherrer, eds.), IOS Press, 1997, pp. 629–633.

[28] SANCHEZ, E.: *Medical diagnosis and composite fuzzy relations*, in: Advances in Fuzzy Set Theory and Applications (M. M. Gupta, R. K. Ragade, and R. R. Yager, eds.), North-Holland, New York, 1979, pp. 437–444.

[29] SHORTLIFFE, E.—BUCHANAN, B.: *A model of inexact reasoning in medicine.*, Math. Biosci. **23** (1975), 351.

[30] Van MELLE, W.: *A domain of independent system aids in constructing knowledge-based consultation programs*, Technical Report, Ph.D. Thesis, Stanford University, CA, 1980.

[31] ZADEH, L. A.: *Outline of a new approach to the analysis of complex systems and decision processes*, IEEE Transaction on Systems, Man, and Cybernetics **3** (1973), 28–44.

[32] ZIMMERMANN, H. J.: *Fuzzy Set Theory — and Its Applications*, Kluwer Academic Publisher, Dordrecht, 1991.

Received March 31, 1998

Institute of Computer Science
Academy of Sciences of the Czech Republic
Pod vodárenskou věží 2
CZ–182 07 Praha 8
CZECH REPUBLIC
E-mail: milan@uivt.cas.cz