Kybernetika

VOLUME *42* (2006), NUMBER 6

The Journal of the Czech Society for Cybernetics and Information Sciences

Published by:
Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic

Editorial Office:
Pod Vodárenskou věží 4, 182 08 Praha 8

Kybernetika is a bi-monthly international journal dedicated for rapid publication of high-quality, peer-reviewed research articles in fields covered by its title.

Kybernetika traditionally publishes research results in the fields of Control Sciences, Information Sciences, System Sciences, Statistical Decision Making, Applied Probability Theory, Random Processes, Fuzziness and Uncertainty Theories, Operations Research and Theoretical Computer Science, as well as in the topics closely related to the above fields.

The Journal has been monitored in the Science Citation Index since 1977 and it is abstracted/indexed in databases of Mathematical Reviews, Current Mathematical Publications, Current Contents ISI Engineering and Computing Technology.

Kybernetika. Volume *42* (2006) ISSN 0023-5954, MK ČR E 4902.

Published bi-monthly by the Institute of Information Theory and Automation of the Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. — Address of the Editor: P. O. Box 18, 182 08 Prague 8, e-mail: kybernetika@utia.cas.cz. — Printed by PV Press, Pod vrstevnicí 5, 140 00 Prague 4. — Orders and subscriptions should be placed with: MYRIS TRADE Ltd., P. O. Box 2, V Štíhlách 1311, 142 01 Prague 4, Czech Republic, e-mail: myris@myris.cz. — Sole agent for all "western" countries: Kubon & Sagner, P. O. Box 34 01 08, D-8 000 München 34, F.R.G.

Published in December 2006.

KYBERNETIKA — VOLUME *42* (2006), NUMBER 6, PAGES 723 – 748

A THEORETICAL COMPARISON OF DISCO AND CADIAG–II–LIKE SYSTEMS FOR MEDICAL DIAGNOSES

Tatiana Kiseliova

In this paper a fuzzy relation-based framework is shown to be suitable to describe not only knowledge-based medical systems, explicitly using fuzzy approaches, but other ways of knowledge representation and processing. A particular example, the practically tested medical expert system Disco, is investigated from this point of view. The system is described in the fuzzy relation-based framework and compared with CADIAG-II-like systems that are a "pattern" for computer-assisted diagnosis systems based on a fuzzy technology. Similarities and discrepancies in – representation of knowledge, patient's information, inference mechanism and interpretation of results (diagnoses) – of the systems are established.

This work can be considered as another step towards a general framework for computer-assisted medical diagnosis.

Keywords: fuzzy relations, medical diagnoses

AMS Subject Classification: 03B52, 03E72, 62F15, 93E12, 92C50

1. INTRODUCTION

Following a classification done in [12], medical knowledge-based systems (KBS) comprise medical consultation systems and medical expert systems, that are aimed to support physicians in decision making. To deal with uncertainty of medical knowledge, the fuzzy sets mechanism has been successfully applied in different KBS. Sanchez [29] was the first, who introduced a fuzzy relation-based framework to medical diagnosis. The author separated the problem of medical diagnosis into three stages: determination of symptoms, medical knowledge, diagnosis. These parts can be considered as components of medical KBS – knowledge acquisition, knowledge representation, knowledge processing – in the narrow sense. The representations of Sanchez are formalized by fuzzy relations and a composition of fuzzy relations is used as an inference mechanism. Since the end of 1970s the fuzzy relations area together with fuzzy control were successfully applied at the University of Vienna Medical School: several generations of medical consultation CADIAG-systems were developed (see, for example, [1, 3, 4, 19, 31]). They are the representative examples of fuzzy technology applications in medical expert systems.

In this paper we develop an idea, that a fuzzy relation based framework is suitable to describe not only medical systems, explicitly using fuzzy approaches, but other ways of knowledge representation and processing.

Several works in this direction have been done recently. The school of Adlassnig continues to develop a new generation of expert system MedFrame/CADIAG-IV, where fuzzy relations are extended with type-2 fuzzy relations and negative associations [3, 4, 19]. Kolousek [18] considered various kinds of knowledge representation schemes, such as graph-based, rule-based, tables, time information in the fuzzy relation-based framework. Hájek et al. [5, 6, 10, 11] have done deep investigations to understand MYCIN and CADIAG-like systems. Among them are E(MYCIN), PROSPECTOR [10, 20, 32] and CADIAG, CADIAG-II, CADIAG-IV [4, 31], both types defined as compositional rule based systems. Conditions to embed CADIAG-II, CADIAG-IV into MYCIN-like systems and vice versa were proved. Proposals for generalization of the compositional rule based systems were discussed. It was mentioned that this process is far from its final stage. Open problems such as formulation and validity of embedding theorems for knowledge bases with general formula using conjunction, disjunction and negation in antecedents of rules are among others to be further developed [5].

In this paper we compare the system Disco based on discrimination analysis (adapted from [2] in [13]) with CADIAG-II and with generalization of both CADIAG-II and MYCIN-like systems [6]. Our intention is to show that Disco can be described in the fuzzy relation-based terminology and to establish similarities and discrepancies of systems in representation of *medical knowledge, patient's information, inference mechanism, diagnosis.* A formal representation of these components, denoted as

$$\langle R_{SD}, R_{PS}, R_{PS} \circ R_{SD}, R_{PD} \rangle \tag{1}$$

correspondingly, will be described in the next sections. It serves as a point of departure for a comparison of the above mentioned systems and their generalization. The results of a comparison can be considered as a contribution to the possible generalization of the different fuzzy relation-based systems. Although their "facades" often differ on the first glance. Under "the possible generalization" or "general framework" we understand the possible complete description of the different variants, representations, interpretations of the above listed components in (1).

Disco, based on the discrimination analysis, solves similar to CADIAG-II-like systems problems, but in a different way. Disco handles incomplete information, partial inconsistency, fuzzy description of relations in a natural way and avoids many problems associated with probabilistic approaches. Thus it is interesting also to compare how the questionable point of CADIAG-II – interpretation of membership degrees of fuzzy relations together with the compositional inference – is solved in Disco. The properties of the systems are different. But it will be shown that they can be described in the same language that allows to find similarities and differences between these systems. This "unification" in representation does not relate to the quality of results. But the description of Disco in the fuzzy relation-based framework will show that the discrimination-based system is not a separate approach valid only for a particular application, but a part of a unique approach. From one side this view allows to understand better the described processes and from another imparts a nice flexibility and variety of representations under the unique theoretical background. It should be mentioned, that Disco was tested for venal diseases at the Tbilisi Medical University and gave satisfactory results [13].

In the next section we give a description of CADIAG-II-like systems and their generalization – Conorm-CADIAG-II – including negative knowledge. Three main aspects of the systems, namely, determination of symptoms, medical knowledge and diagnosis are interpreted in terms of fuzzy sets theory in details. This clarification is examined, because not everything, described linguistically, can be considered as fuzzy sets; sometime it leads to a misunderstanding. We begin the next section with basic notions and definitions of the fuzzy domain. System Disco is described in Section 3 as was originally proposed in [2] and further specified in [13]. Translation of Disco to the general fuzzy relation-based framework is described in Section 4. Sections are summarized by main characteristics of systems needed for the following comparison in Section 5. Finally, Section 6 concludes the discussion.

2. CADIAG–II LIKE SYSTEMS

CADIAG-II [1] and other similar systems [23] are computer assisted medical diagnosis systems for different applications. All of them are based on fuzzy set theory and corresponding fuzzy logic reasoning mechanisms. In particular, they are based on fuzzy rules and an inference procedure – a composition of fuzzy relations – is applied.

In the next section we introduce main denotations and definitions of the fuzzy approach needed for the following discussions. An important remark is, that defining a fuzzy set F as a mapping from a universe of discourse to the unit interval, we identify F and its membership function μ_F. The same concerns fuzzy relations, since fuzzy relations are a special case of fuzzy sets.

2.1. Basic notions for the fuzzy relation–based framework

2.1.1. Fuzzy sets

Let U be a collection of objects (universe of discourse), for example, the set of real numbers, or the set of real numbers between 0 and 10, etc.

Definition 1. Assume U is a universe. A fuzzy set F [1] is defined as a mapping:

$$\mu_F : U \to [0, 1].$$

In the framework of fuzzy set theory μ_F is called also the *membership function*. This terminology stresses the idea that for each $x \in U$, $\mu_F(x)$ indicates the corresponding membership value [7, 15, 17, 33, 38, 41]. We identify F and μ_F to simplify the notations.

[1] or a fuzzy subset F of U.

2.1.2. Power sets

Definition 2. Let U be a universe. The set of all fuzzy sets (the fuzzy power set) on U is defined as

$$\mathfrak{F}(U) =_{\text{def}} \{F|F : U \to [0,1]\}.$$

In the case of crisp sets this definition corresponds to the definition of a power set:

$$\mathfrak{P}(U) =_{\text{def}} \{A|A \subset U\}.$$

2.1.3. Type-2 fuzzy sets

A type-2 fuzzy set $\tilde{F}$ of a set U is a fuzzy set whose degrees of membership are themselves fuzzy sets.

Definition 3. A type-2 fuzzy set $\tilde{F}$ is a mapping:

$$\tilde{F} : U \to \mathfrak{F}(V)$$

where $\mathfrak{F}(V)$ is the fuzzy power set of a universe V (Definition 2).

2.1.4. Basic set-theoretical operations on fuzzy sets

Let F, G be fuzzy sets, defined on the same universe U.

Definition 4. The *intersection* $F \cap G$ is point-wise defined as

$$(F \cap G)(x) =_{\text{def}} \min(F(x), G(x)),\ \ x \in U.$$

Definition 5. The *union* $F \cup G$ is point-wise defined as

$$(F \cup G)(x) =_{\text{def}} \max(F(x), G(x)),\ \ x \in U.$$

Definition 6. The *complement* of fuzzy set F is defined as

$$\overline{F}(x) =_{\text{def}} 1 - F(x),\ \ x \in U.$$

2.1.5. t-norms and t-conorms

The concept of triangular norms (t-norms) comes from the ideas of probabilistic metric spaces originally proposed in [22, 30] and actively pursued by many authors, especially [14].

In fuzzy sets, triangular norms and conorms play the key role by providing generic models for intersection and union operations on fuzzy sets.

Definition 7. Let $\tau : [0,1] \times [0,1] \to [0,1]$ be a binary operation [14, 33], τ is called *t-norm* iff the following properties $T_1,\ T_2,\ T_3,\ T_4$ hold:

T_1: (boundary condition) $\forall\, x \in [0,1]$:

$$\tau(x,1) = x$$

T_2: τ is monotone, i. e., $\forall\, x, x', y, y' \in [0,1]$:

$$\text{if } x \leq x' \text{ and } y \leq y' \text{ then } \tau(x,y) \leq \tau(x',y')$$

T_3: τ is commutative, i. e., $\forall\, x, y \in [0,1]$:

$$\tau(x,y) = \tau(y,x)$$

T_4: τ is associative, i. e., $\forall\, x, y, z \in [0,1]$:

$$\tau(x, \tau(y,z)) = \tau(\tau(x,y), z).$$

For each t-norm holds:

T_0: $\tau(0,1) = \tau(1,0) = \tau(0,0) = 0$ and $\tau(1,1) = 1$, i. e., τ is an extension of the boolean conjunction. T_0 is an immediate consequence of the commutativity, monotonicity and the boundary condition of t-norms.

Examples of t-norms:

$$\tau_m(x,y) =_{\text{def}} \min(x,y)$$
$$\tau_b(x,y) =_{\text{def}} \max(0, x+y-1)$$
$$\tau_a(x,y) =_{\text{def}} x \cdot y.$$

Sometimes $\tau_b(x,y)$ *(bounded difference)* is called *Łukasiewicz* t-norm and $\tau_a(x,y)$ *(algebraic product)* is called *product* t-norm as well [14].

Definition 8. Let $\sigma : [0,1] \times [0,1] \to [0,1]$, σ is called *t-conorm* iff the following properties $S_1,\ S_2,\ S_3,\ S_4$ hold:

S_1: (boundary condition) $\forall\, x \in [0,1]$:

$$\sigma(x,0) = x$$

S_2: σ is monotone, i. e., $\forall\, x, x', y, y' \in [0,1]$:

$$\text{if } x \leq x' \text{ and } y \leq y' \text{ then } \sigma(x,y) \leq \sigma(x',y')$$

S_3: σ is commutative, i. e., $\forall\, x, y \in [0,1]$:

$$\sigma(x,y) = \sigma(y,x)$$

S_4: σ is associative, i. e., $\forall\, x, y, z \in [0,1]$:

$$\sigma(x, \sigma(y,z)) = \sigma(\sigma(x,y), z).$$

For each t-conorm is:

$S_0 : \sigma(0,0) = 0$ and $\sigma(0,1) = \sigma(1,0) = \sigma(1,1) = 1$, i. e., σ is an extension of the boolean disjunction.

Examples of t-conorms:

$$\sigma_m(x,y) =_{\text{def}} \max(x,y)$$
$$\sigma_b(x,y) =_{\text{def}} \min(1, x+y)$$
$$\sigma_a(x,y) =_{\text{def}} x+y-x\cdot y.$$

Sometimes $\sigma_b(x,y)$ (*bounded sum*) is called *Łukasiewicz* t-conorm and $\sigma_a(x,y)$ *(algebraic sum)* is called *probabilistic sum* [14] as well.

2.1.6. Aggregation operators

As was mentioned in [14] aggregation (fusion) of several input values into a single output value is an indispensable tool not only of mathematics or physics, but of majority of other sciences. If it is assumed that the number of input values is fixed, say $n \in \mathbb{N}$, an aggregation operator is a real function of n variables, and both inputs and outputs are from the unit interval $\mathrm{I} = [0,\ 1]$, then n-ary aggregation operator h is always a mapping: $\mathrm{h} : [0,1]^n \to [0,1]$.

2.1.7. Averaging operators

Averaging operators are special aggregation operators and defined as follows.

Definition 9. An averaging operator A is a mapping $\mathrm{A} : [0,1]^n \to [0,1]$ that is characterized by the following set of axioms [16]:

(A1) idempotency – for all $a \in [0,1]$,

$$\mathrm{A}(\underbrace{a, a, a, \dots, a}_{n \text{ times}}) = a$$

(A2) monotonicity – for any pair of n-tuples $\langle a_1, a_2, \dots, a_n\rangle \in [0,1]^n$ and $\langle b_1, b_2, \dots, b_n\rangle \in [0,1]^n$, if $a_k \le b_k$ for all $k = 1, \dots, n$, then

$$\mathrm{A}(a_1, a_2, \dots, a_n) \le \mathrm{A}(b_1, b_2, \dots, b_n).$$

It is significant that any function A, that satisfies these axioms, gives values that, for any n-tuple $\langle a_1, a_2, \dots, a_n\rangle \in [0,1]^n$, lie in the closed interval defined by the inequalities

$$\min(a_1, a_2, \dots, a_n) \le \mathrm{A}(a_1, a_2, \dots, a_n) \le \max(a_1, a_2, \dots, a_n). \qquad (2)$$

Examples of averaging operators are *generalized means, ordered weighted averaging operators (OWA), λ-averages* [8, 9, 17, 37].

Definition 10. (Yager [36]) An ordered weighted averaging (OWA) operator of dimension n is a mapping $\mathrm{OWA} : \mathbb{R}^n \to \mathbb{R}$ characterized by an n-dimensional vector W, called the weighting vector, such that its components $w_j,\ j = 1, \dots, n$, lie in the unit interval and sum to one. The OWA operator is defined as

$$\mathrm{OWA}(a_1, \dots, a_n) = \sum_{j=1}^{n} w_j b_j$$

where b_j is the jth largest of the a_i.

The arithmetic mean M, the harmonic mean H, the quadratic mean Q, and $\mathrm{M_p}$ for $p \in]0, \infty[$ [2], p-mean, and the geometric mean G are given by, respectively,

$$\mathrm{M}(x_1, \dots, x_n) = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{3}$$

$$\mathrm{H}(x_1, \dots, x_n) = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} \qquad \mathrm{Q}(x_1, \dots, x_n) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}}$$

$$\mathrm{M_p}(x_1, \dots, x_n) = \left(\frac{1}{n} \sum_{i=1}^{n} x_i^p \right)^{\frac{1}{p}} \qquad \mathrm{G}(x_1, \dots, x_n) = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}.$$

2.1.8. Fuzzy relations

Definition 11. Let $U_1, \dots, U_n$ be non-empty universa, $U_1 \times \cdots \times U_n$ be a product space. A fuzzy relation R on $U_1 \times \cdots \times U_n$ is defined as

$$R : U_1 \times \cdots \times U_n \to [0, 1].$$

Fuzzy relations are only a special case of fuzzy sets (with U expanded to $U_1 \times \cdots \times U_n$), therefore the standard set operations for fuzzy sets are valid for fuzzy relations, too.

2.1.9. Type-2 fuzzy relations

By analogy to the definition of type-2 fuzzy sets (Definition 3) a type-2 fuzzy relation is defined as follows:

Definition 12. A type-2 fuzzy relation $\tilde{R}$ is a mapping:

$$\tilde{R} : U_1 \times U_2 \to \mathfrak{F}(U)$$

where $\mathfrak{F}(U)$ is the fuzzy power set of a universe U.

[2] $]0, \infty[$ denotes the open interval between 0 and ∞.

2.2. Predefined knowledge base in CADIAG–II–like systems

Let Θ_C denote a knowledge base of CADIAG-II-like systems. This knowledge base consists of fuzzy *IF–THEN* rules that describe the relationships between

- symptoms, signs, test results and findings – for all these medical entities we use a term "symptom" or its abbreviation "S";

and

- diseases, diagnoses – with denotation "D".

To build these fuzzy relations the crisp sets of patients $\Pi = \{p_1, \ldots, p_r\}$, symptoms $\Sigma = \{s_1, \ldots, s_m\}$ and diseases $\Delta = \{d_1, \ldots, d_n\}$ under consideration are used. For example, Δ can denote rheumatic diseases, Σ are symptoms of these diseases, and Π are investigated patients for rheumatic diseases.

Fuzzy relations between symptoms and diseases are defined as $R_{SD} : \Sigma \times \Delta \to [0, 1]$. CADIAG-II-like systems use two types of symptom-disease relations: confirmation R^c_{SD} and occurrence R^o_{SD} relations. These fuzzy relations estimate each symptom-diagnose connection from two perspectives, that are strength of confirmation and frequency of occurrence.

The rules in CADIAG-II-like systems have the form:

$$\textit{IF } (\text{antecedent}) \ \textit{THEN } (\text{consequent}) \ \textit{WITH } (o, \ c). \tag{4}$$

The rules are used in the following way: *IF* symptom(s), *THEN* disease(s), with a degree "o" of occurrence and a degree "c" of confirmation. CADIAG-II-like systems use not only relations between a symptom and a disease, but between a symptoms combination and disease(s) as well. A symptom combination is a simultaneous presence (conjunction) of symptoms or its negation.

A single symptom can be considered as a special case of a symptom combination (as one-element elementary conjunction of symptoms) [5]. We will use notation s_i for a single symptom and a combination of symptoms in the text of the paper until the explicite distinguishing will be needed.

Parameters o, c contain numeric and/or linguistic information, for example, crisp numbers, fuzzy numbers. In our notation they coincide with R^c_{SD} and R^o_{SD}.

These relations can be interpreted statistically and linguistically.

If a statistical way is taken, $R^o_{SD}(s_i, d_j)$ and $R^c_{SD}(s_i, d_j)$ are derived from relative frequencies, i. e., numerically [6]:

$$R^o_{SD}(s_i, d_j) = f(s_i|d_j)$$

and

$$R^c_{SD}(s_i, d_j) = f(d_j|s_i)$$

where $f(s_i|d_j) = \frac{f(d_j \cap s_i)}{f(d_j)}$, $f(d_j|s_i) = \frac{f(d_j \cap s_i)}{f(s_i)}$. $f(s_i|d_j)$ is a conditional frequency of s_i given d_j, $f(d_j|s_i)$ is a conditional frequency of d_j given s_i, $f(d_j \cap s_i)$ is the absolute frequency of joint occurrence of d_j and s_i, $f(d_j)$ and $f(s_i)$ are absolute frequencies of d_j and s_i, correspondingly.

A linguistic way opens a possibility to estimate relations R^o_{SD} and R^c_{SD} using fuzzy sets representing values of a linguistic variable, e. g., *almost always, often,* *medium, seldom, very seldom, almost never* for linguistic variables *occurrence* and *confirmation*.

These fuzzy sets are defined as mappings from $[0,1]$ to $[0,1]$.

The numerical and linguistic definitions of symptom-disease relations seem to belong to different types of fuzzy sets. If numerical values represent the membership degrees of fuzzy sets R^o_{SD} or R^c_{SD} in points (s_i, d_j), linguistic values are close to the fuzzy sets type-2, where degrees of membership of a fuzzy relation are fuzzy sets themselves [21]. To unify the representation of symptom-disease relations, the statistical way can also be represented by fuzzy relations type-2, where numerical values $f(s_i|d_j)$ and $f(d_j|s_i)$ are fuzzy singletons. In this way the numerical and linguistic representations can be uniquely described.

The possibility to define fuzzy relations as fuzzy intervals was discussed in [19]. It is based on the assumption, that different patient settings influence all-purpose consultant systems. In particular, this approach was realized in CADIAG-IV [4].

In the rest of our paper we will consider the numerical representation of symptom-disease relations as "ordinary" fuzzy relations.

2.3. The inference mechanism

To use the inference mechanism of a fuzzy approach described below, information about a patient (or a group of patients) to whom a diagnosis will be established, has to be available.

Information about patients' symptoms in CADIAG-II-like systems is presented in the form of a fuzzy relation $R_{PS} : \Pi \times \Sigma \to [0,1]$. Clearly, if one patient is observed, this relation can be written as a fuzzy set $S_p : \Sigma \to [0,1]$, where each element of the fuzzy set S_p shows to which degree it is true, that a patient p has symptom s_i.

A composition of fuzzy relations introduced by L. Zadeh [39] and interpreted by him as a $\max - \min$ composition, later were adopted by Sanchez for medical diagnoses [29] with the same interpretation. This approach is used in CADIAG-II as follows:

$$R_{PD} =_{\text{def}} R_{PS} \circ R_{SD} \tag{5}$$

(5) represents a composition of fuzzy relations and $\forall d_j \in \Delta,\ \forall p_q \in \Pi$:

$$R_{PD}(p_q, d_j) =_{\text{def}} \max_{s_i \in \Sigma} \min\{R_{PS}(p_q, s_i); R_{SD}(s_i, d_j)\} \tag{6}$$

where $R_{PD} : \Pi \times \Delta \to [0,1]$ are inferred possible diagnoses for the patient(s).

For one patient schema (5) takes the following form:

$$D_p =_{\text{def}} S_p \circ R_{SD} \tag{7}$$

and

$$D_p(d_j) = \max_{s_i \in \Sigma} \min\{S_p(s_i); R_{SD}(s_i, d_j)\} \tag{8}$$

corresponds to the $\max - \min$ composition (6), where $D_p : \Delta \to [0,1]$ is a fuzzy set of possible diagnoses for a patient.

Adlassnig [1] has found two fuzzy relationships, namely, occurrence and confirmability for the scheme (6).

As was mentioned in [6], "because of features of physicians' thinking" there are three types of inference rules/compositions for a final diagnosis:

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

The similar rules, but with symptom combinations, are used in CADIAG-II-like systems. They are denoted as R^4_{PD}, R^5_{PD}, R^6_{PD}.

2.4. Some remarks concerning the inference mechanism in CADIAG-II-like systems

As was already told in the previous section, the knowledge base, i. e., symptom-disease relations, can be defined numerically and linguistically. The numerical representation fits well for the $\max - \min$ composition of fuzzy relations (6): both information about patient symptoms S_p (in the case of (8)) and symptom-disease relations are ordinary fuzzy sets.

A linguistic representation often needs additional clarifications. Not everything, described linguistically, can be considered as fuzzy sets. Sometimes words like *always, often, medium, seldom, never* are "coded" by numbers, e. g., $1, .75, .5, .25, 0$, that take part in the calculations together with S_p. The values of S_p are numbers from $[0, 1]$.

Although information about patient's symptoms in CADIAG-II-like systems is given by a fuzzy set, i. e., assuming all values between 0 and 1, a physician (not a patient himself!) often uses three values to estimate patient's symptoms practically. Thus, investigated patient symptoms are considered as *present, not present, non-applicable*, that may be "coded" by, e. g., 1, 0, $\frac{1}{2}$, correspondingly. *Non-applicable* often has several meanings: for example, "a symptom has not been examined", "my experience tells me nothing about this symptom", etc.

Assigning only unique numbers to words seems to be a rather naive way of medical knowledge qualification. More realistic is to consider intervals that represent the verbal physician opinion. If a physician says that a symptom s_i often meets d_j, he may mean, that approximately in 95% s_i meets a disease d_j. This medical knowledge can be represented as interval $[0.92, 0.98]$ or a fuzzy number.

How to build these intervals or fuzzy sets is a question of the knowledge acquisition. The problem is not easy at all and has consistently been described as a bottleneck in the development of computer-consultant systems. Some models of knowledge acquisition were, in particular, discussed in [4, 19].

Assume now that the problem of knowledge acquisition is solved and a linguistic representation of symptom-disease relations is, for example, fuzzy sets *almost always, often, medium, seldom, very seldom, almost never*. As was already told in the previous section, a symptom-disease relation is considered to be a fuzzy set type-2. To visualize this, a symptom-disease relation can be represented in a tabular form (e. g., Table, Section 3), where each element of the table is a fuzzy set, for example, *almost always, often, medium, seldom, very seldom*, or *almost never*.

In the next steps we return to the ordinary representation of fuzzy sets.

2.5. Interpretation of inference results

The results of a multiple applications of scheme (5) for different types of a symptom-disease relation have to be interpreted for obtaining the patient(s) diagnosis.

In CADIAG-II-like systems the possibility to classify all diagnoses in three classes − confirmed, excluded, possible − is predefined. A diagnosis d_i is confirmed (i. e., $D_p(d_i) = 1$) iff there exists a fully present symptom s_j ($S_p(s_j) = 1$) which has full/maximal contribution to the diagnosis, i. e., $R^c_{SD}(s_j, d_i) = 1$. A diagnosis d_i is excluded (i. e., $D_p(d_i) = 0$) by a present symptom (by R^2_{PD}) iff there exists a fully present symptom s_j ($S_p(s_j) = 1$) which has 0 (i. e. negative) contribution to the diagnosis, i. e., $R^c_{SD}(s_j, d_i) = 0$. A diagnosis d_i is excluded (i. e., $D_p(d_i) = 0$) by an absent symptom (by R^3_{PD}) iff there exists a fully absent symptom s_j ($S_p(s_j) = 0$) which has full/maximal occurrence for the diagnosis, i. e., $R^o_{SD}(s_j, d_i) = 1$. Analogically, a diagnosis is confirmed if $R^4_{PD} = 1$ and excluded, if $R^3_{PD} = 1$ or $R^5_{PD} = 1$ or $R^6_{PD} = 1$.

On other cases the diagnoses are considered as possible. In general, a diagnosis d_i is denoted as a generated diagnostic hypothesis, if $\epsilon \leq D_p(d_i) \leq 1$, where $D_p(d_i) = \max\{R^1_{PD}(p, d_i), R^4_{PD}(p, d_i)\}$, i. e., negative parts of inference R^2_{PD}, R^3_{PD}, R^5_{PD} and R^6_{PD} are fully ignored. Threshold value ϵ is usually taken to be equal to 0.01. There are two other categories for a diagnosis: not generated diagnosis, if $0 < D_p(d_i) < \epsilon$, and diagnostic contradictions if a diagnosis should be confirmed and excluded in the same time. See also [6, 28].

Let us mention that although R_{SD} can be statistically interpreted (see Section 2.2.), there is no statistical interpretation of results of the inference: the resulting values are just weights, not frequencies or probabilities. The diagnoses can be compared by these weights.

2.6. Conorm-CADIAG-II with negative rules

Conorm-CADIAG-II generalizes CADIAG-II twofold [6]:

1. max-min composition is substituted by t-conorm-min composition: $\forall d_j \in \Delta$, $\forall p_q \in \Pi$:

$$R_{PD}(p_q, d_j) =_{\text{def}} \bigvee_{s_i \in \Sigma} \min\{R_{PS}(p_q, s_i); R_{SD}(s_i, d_j)\} \quad (9)$$

   where $\bigvee$ is a t-conorm.

2. Negative knowledge is introduced in CADIAG-II as an *exclusion* relation $R^e_{SD} : \Sigma \times \Delta \to [0, 1]$. The value $R^e_{SD}(s_i, d_j)$ indicates the degree in which the present symptom (combination) excludes (or disconfirms) the disease d_j.

A proposal, that a symptom s_i (or a given group of symptoms) cannot both confirm and exclude a given diagnosis d_j, leads to the following assumption: $R^e_{SD}(s_i, d_j) = 0$ or $R^c_{SD}(s_i, d_j) = 0$ at a given time.

In accordance with previous discussions (Section 2.2.), $R^c_{SD}(s_i, d_j)$, $R^e_{SD}(s_i, d_j)$ represent an already defined parameter c and a new parameter e for the rule (4), correspondingly. Considered symptoms exclude/confirm the diseases for a patient due to the scheme (5), i. e.,

$$R^e_{PD} =_{\text{def}} R_{PS} \circ R^e_{SD} \tag{10}$$

$$R^c_{PD} =_{\text{def}} R_{PS} \circ R^c_{SD} \tag{11}$$

and both are interpreted as t-conorm$-\min$ composition (9). The extension of Conorm-CADIAG-II with negative knowledge encompasses the main features of CADIAG-II without necessity of special types of inference rules R^2_{PD} and R^3_{PD} for exclusion of diagnoses.

The generalization of CADIAG-II-like systems was at first introduced to compare CADIAG-II and MYCIN-like systems and to escape some their shortcomings and disadvantages. For example, CADIAG-II cannot decrease contribution of one rule by another rule. For combination of positive and negative contributions extended Abelian group operation $\oplus$ defined on $[-1, 1]$ is used in [6].

Definition 13. (M. Daniel, P. Hájek, and H. Nguyen [6]) An ordered Abelian group $G = (G, \oplus, \ominus, 0, \leq)$, is a set G with an associative and commutative binary operation $\oplus$, with a neutral element 0 ($x \oplus 0 = x$), with a unary operation of inverse $\ominus$ ($x \oplus (\ominus x) = 0$) and with a linear ordering such as monotonicity holds ($x \leq y \rightarrow x \oplus z \leq y \oplus z$).

An extended ordered Abelian group is an extension of an ordered Abelian group with extremal elements, the greatest one $\top$ and the least one $\bot$, such that $\top \oplus x = \top$, $\top = \ominus\bot$, where $\top \oplus \bot$ is not defined. An example of the extended ordered Abelian group is $MC = ([-1, 1], \oplus_{MC}, -, 0, \leq)$ where

$$x \oplus_{MC} y =_{\text{def}} \begin{cases} x + y - x \cdot y, & \text{if } y \text{ and } x \text{ are positive;} \\ \frac{x+y}{1-\min\{|x|,|y|\}}, & \text{if } y \cdot x \in]-1, 0]; \\ x + y + x \cdot y, & \text{if } y \text{ and } x \text{ are negative.} \end{cases} \tag{12}$$

In Conorm-CADIAG-II positive contributions are aggregated separately by a t-conorm and by the same t-conorm the negative contributions are aggregated. After these performances the group operation $\oplus$ is used. If t-conorm, denoted as $\oplus'$, is a positive part (restricted onto $[0, 1]$) of an (extended) Abelian group operation, there can be two group operations used: $\oplus'$ which serves as t-conorm and $\oplus$ which plays role of group operation for combination of positive and negative results together.

If $\oplus' = \oplus$, there is the only one operation (it must be an (extended) Abelian group operation) in such a special case of Conorm-CADIAG-II-like system. In this special case it is not necessary to combine positive and negative contribution separately as the operation is commutative and associative. This special case of Conorm-CADIAG-II is very close to MYCIN-like systems [6], it is a special case of MYCIN-like system "translated" to the language of CADIAG-II.

A *total* degree can be calculated as follows:

$$R^{tot}_{PD}(p, d_j) =_{\text{def}} R^c_{PD}(p, d_j) \oplus -R^e_{PD}(p, d_j) \tag{13}$$

where $\oplus$ is an extended ordered Abelian group operation on $[-1, 1]$.

Thus, for every diagnosis its confirmation is decreased according to its exclusion, represented as negative confirmation.

Let us emphasize again: diagnoses in CADIAG-II-like systems can be confirmed and excluded in the same time and this total diagnostic contradiction is an important feature of these systems.

In [6, 10] it was discussed the possibility to transform the interval $[-1, 1]$ to $[0, 1]$, sending 0 to 0.5. But then a transformed operation would not be a t-conorm, but an ordered Abelian group operation on $(0, 1)$ that is a compensatory operation [14]. To preserve t-conorm-min inference, the set of values $[-1, 1]$ is taken in Conorm-CADIAG-II.

2.7. Some important items of CADIAG-II-like systems

Let us summarize information about CADIAG-II-like systems that will serve as a point of a comparison for the system Disco.

Predefined information in CADIAG-II-like systems in the form of *IF–THEN* rules is formalized by a fuzzy relation R_{SD} that connects symptoms and diseases. This relation is described from two perspectives, as frequency of occurrence and strength of confirmation. These two characteristics can be represented statistically and linguistically. To unify them the relations R^c_{SD} and R^o_{SD} can be defined as fuzzy sets type-2. The representation of R_{SD} is not limited by these two characteristics: R_{SD} can be an exclusion relation as well that is applied in Conorm-CADIAG-II. Max-min composition of fuzzy relations is applied as an inference mechanism in CADIAG-II-like systems. Conorm-CADIAG-II generalizes CADIAG-II by substitution of max-min composition with t-conorm-min composition and introduction of negative knowledge.

Due to at least two above mentioned characteristics of R_{SD} (R^c_{SD} and R^o_{SD}) and the presence or absence of patient's symptoms, the inference scheme (5) in CADIAG-II-like systems deduces the predefined characteristics of the patient's diagnosis: confirmed, excluded or possible. In Conorm-CADIAG-II the inference rule R^e_{PD} is applied for exclusion of diagnoses instead of R^2_{PD} and R^3_{PD}. The results of negative and positive contributions are combined by a group operation in Conorm-CADIAG-II.

3. DISCRIMINATION IN MEDICAL DIAGNOSES

For convenience let us call the following considered system, based on the discrimination analysis, Disco. The system was elaborated for application/utilization in [13], where the mechanism from [2] was further developed, i. e., better specified and more precise described the less class of possible computations.

Disco is a computer assisted medical diagnosis system which was practically implemented. The system was tested for venal diseases at the Tbilisi Medical University and gave satisfactory results [13]. This means that a decision done by the computer program, based on this method, in most cases coincided with the opinion of an experience physician that established a diagnosis for the investigated patient. This was checked for several real cases. However, due to the known economic situation on post-soviet countries, practical investigations were not further continued.

As was already mentioned in the Introduction, Disco solves the same problems as CADIAG-II-like systems. Following the main goal of this paper – to investigate a general framework for representation of CADIAG-II-like systems and Disco, based on fuzzy relations, i. e., the general language for both type of systems – we will concentrate on some theoretical aspects of Disco approach necessary for this underline purpose.

First, let us shortly describe the mechanism of Disco proposed in [2] (Sections 3.1, 3.2, 3.3) and then, in the following sections, some specifications done in [13], needed to represent this system in the same language as CADIAG-systems, will be presented.

3.1. Predefined knowledge base in Disco

The knowledge base Θ_D of the Disco system is presented in a tabular form (Table)

Table. An initial table
for a symptom-disease connection.

	d_1	d_2	$\dots$	d_n
s_1	f_{11}	f_{12}	$\dots$	f_{1n}
s_2	f_{21}	f_{22}	$\dots$	f_{2n}
$\vdots$	$\vdots$	$\vdots$	f_{ij}	$\vdots$
s_m	f_{m1}	f_{m2}	$\dots$	f_{mn}

where relations between symptoms and diagnoses are defined, i. e., the data are taken from patient records – historical cases. Diagnoses are proven in these historical cases: they are confirmed by experts-physicians for each case. These relations are described by relative frequencies $\{f_{ij},\ i = 1, \dots, m,\ j = 1, \dots, n\}$ where each f_{ij} denotes the proportion of patients with disease d_j and symptom s_i in the entire sample of patients with disease d_j. For example, $f_{23} = 0.7$ denotes 70% of those patients who suffered d_3 presented symptom s_2.

3.2. Positive and negative discrimination tables

The next step is to transform the initial table into two ones, so called the positive and negative discrimination tables:

$$p_{ij} =_{\text{def}} \left[\sum_{k=1, k \neq j}^{n} \text{LargeRatio} \left(\frac{f_{ij}}{f_{ik}} \right) \right] / (n-1) \tag{14}$$

$$n_{ij} =_{\text{def}} \left[\sum_{k=1, k \neq j}^{n} \text{LargeRatio} \left(\frac{f_{ik}}{f_{ij}} \right) \right] / (n-1) \tag{15}$$

for $i = 1, \ldots, m,\ j = 1, \ldots, n$. As presented in [2], LargeRatio is a fuzzy set, LargeRatio : $\mathbb{R} \to [0,1]$. Clearly, $p_{ij}, n_{ij} \in [0,1]$. A heuristic explanation of the positive and negative discrimination measures is as follows. p_{ij} represents the accumulated belief that symptom i is more indicative of *disease* j than of any remaining disease, while n_{ij} represents the accumulated belief that symptom i is more indicative of *not disease* j.

3.3. Inference with positive and negative discrimination values

Information about a patient in Disco is a set of symptoms from Σ of "yes–no" type, i. e., a patient data S_p in this case consists of values 1 and 0; 1 means the patient has a correspondent symptom, whereas 0 points out that a symptom is absent. All symptoms used in the system are always examined for any patient.

An inference mechanism described in [2] selects from tables $\{p_{ij}\}$ and $\{n_{ij}\}$ those rows corresponding to the symptoms of a considered patient and based on the getting new tables – $\{p'_{ij}\}$ and $\{n'_{ij}\}$ $i = 1, \ldots, q,\ q \le m,\ j = 1, \ldots, n$ – a diagnosis is established.

It means that rows for $S_p(s_i) = 0$ are ignored and just those for $S_p(s_i) = 1$ are used.

A diagnosis can be defined as follows:

$$D_p(d_j) =_{\text{def}} \frac{1}{2}\left(\text{Large}(\pi_j) + \text{Small}(\nu_j)\right), \quad j = 1, \ldots, n, \tag{16}$$

where:

$$\pi_j =_{\text{def}} \frac{\sum_{i=1}^{q} p'_{ij}}{|S_p|} \tag{17}$$

$$\nu_j =_{\text{def}} \frac{\sum_{i=1}^{q} n'_{ij}}{|S_p|}. \tag{18}$$

As presented in [2], Large : $[0,1] \to [0,1]$ is a fuzzy set monotonously increasing, Small : $[0,1] \to [0,1]$ is a fuzzy set monotonously decreasing, $|S_p|$ is the cardinality of a patient symptoms set that consists of present/absent symptoms, i. e., $S_p(s_i) = 1$ or $S_p(s_i) = 0$; each $D_p(d_i)$ is a number from $[0,1]$ that points out the belief of the disease d_i for a patient.

4. TRANSLATION OF DISCO TO THE GENERAL FRAMEWORK (1)

As was already mentioned in the Introduction, Disco can be described in the fuzzy relation-based framework (1) and this way similarities and differences with CADIAG-like systems can be established. Below we will discuss them. Notice that in [13] a solution of some special particular problems were considered, for example, construction of a membership function, estimation of information (entropy), etc. Some of solved problems will be presented in the next section to help to translate Disco to the general framework (1).

4.1. Medical knowledge, R_{SD}

It can be seen, a statistical interpretation of a symptom-disease occurrence relation in Disco is defined in the same way as in CADIAG-II-like systems. Therefore f_{ij} from Table can be denoted as $R^o_{SD}(s_i, d_j)$. It is necessary to underline that there is only one source of data in Disco (one table of frequencies $\{f_{ij}, i = 1\ldots, m, j = 1\ldots, n\}$) whereas there are two sources both in CADIAG-II-like systems (R^c and R^o) and Conorm-CADIAG-II (R^c and R^e).

One can mention, that expressions $\frac{f_{ij}}{f_{ik}}$ and $\frac{f_{ik}}{f_{ij}}$ in positive and negative discrimination tables (see Section 3.2.) are the likelihood ratios [40].

Ratio $\frac{f_{ij}}{f_{ik}} = 1$ says that the symptom s_i has the same frequency among the patients with diagnosis d_j and d_k, that the symptom s_i is equally likely in both the diagnoses d_j and d_k.

$\frac{f_{ij}}{f_{ik}} > 1$ says that the symptom s_i is more likely for the diagnosis d_j than d_k.

$\frac{f_{ij}}{f_{ik}} < 1$ says that the symptom s_i is less likely for the diagnosis d_j than d_k.

Notice, that f_{ik} refers frequency of patients with disease k, regardless of disease j, i. e., some of those patients can have both disease j and k.

If all ratios $\frac{f_{ij}}{f_{ik}}, k = 1, \ldots, n,\ k \neq j$ are summed up and then divided by $n - 1$, the resulting value shows, how likely is symptom among patients with disease than with different disease(s). The higher this mean value is, the likelier is the symptom among patients with the disease relative to patients with different disease(s). The role of the fuzzy set LargeRatio is to emphasize the large ratios with the help of membership degrees, because these large ratios are important when indicativeness of the symptom for the disease is established.

Absence of a disease means that a person has a different disease from Δ than that one in question. It is assumed that any person in question has at least one disease from Δ, i. e., a patient can have two (or even more diseases), i. e., there are no data from reference people without any disease.

In general, the fuzzy set LargeRatio reflexes an opinion of an expert what he understands under large ratios for the data from the Table. In [13] LargeRatio was build based on Yager's method [35], whereas in [2] no particular properties were established for LargeRatio and no particular method was supposed to construct this function. Also in [13] it was shown that LargeRatio can be considered to be monotonic increasing, can be piecewise linear or an s-type function. If f_{ik} in (14) or f_{ij} in (15) are equal 0, LargeRatio takes the value 1. LargeRatio($\frac{0}{0}$) is defined to be 0. In some sense it is not a very good choice, because it says the same as $\frac{0}{f_{ik}}$ says for very big f_{ik}: s_i is definitely much more likely for d_k than for d_j. But it is not true as $\frac{0}{0}$ says only that symptom s_i is relevant neither to diagnosis d_j nor to d_k. A neutral element would be more appropriated in this case.

No special assumptions were done for LargeRatio(1). The ratio $\frac{f_{ij}}{f_{ij}}$ is excluded from (14) and (15) because a comparison of symptom s_i with itself does not add any information to the indicativeness of this symptom for a disease d_j.

The above explanation of the positive and negative discrimination values leads to the assumption that, in general, p_{ij} and n_{ij} can be considered as answers to the questions "how strongly does symptom s_i confirm disease d_j" and "how strongly does symptom s_i confirm non disease d_j (disconfirm disease d_j)", correspondingly. Thus, using the terminology of Conorm-CADIAG-II-like systems (Section 2.6) p_{ij} and n_{ij} can be denoted as $R^c_{SD}(s_i, d_j)$ and $R^e_{SD}(s_i, d_j)$ to represent knowledge of both systems in the same language. Thus, a different knowledge represented in the same way as it is used in CADIAG-II-like systems: relations R^c_{SD} and R^e_{SD} in Disco contain transformed (by LargeRatio) arithmetical means of likelihood ratios of frequencies. Whereas CADIAG-II-like relations R^c_{SD} and R^e_{SD} can contain relative frequencies.

4.2. Patient's information S_p and inference mechanism $D_p = S_p \circ R_{SD}$

All symptoms used in the system are always examined for any patient and $S_p(s_i)$ consists of 0s and/or 1s.

To formalize the selection of corresponding to patient's symptoms rows from tables $\{p_{ij}\}$ and $\{n_{ij}\}$ the following equations can be used:

$$R^{c'}_{SD} =_{\text{def}} \Omega_s \times R^c_{SD} \tag{19}$$

$$R^{e'}_{SD} =_{\text{def}} \Omega_s \times R^e_{SD} \tag{20}$$

where Ω_s is a diagonal matrix of dimension $(m \times m)$. Diagonal entries ω_{ij} are elements of S_p, i.e., $\omega_{ii} = S_p(s_i), i = 1, \dots, m$. Fuzzy relations are processed as matrices in (19) and (20) and $\times$ is a usual multiplication of matrices. Matrices R^c_{SD} and $R^{c'}_{SD}$ (R^e_{SD} and $R^{c'}_{SD}$) have the same size, but some of rows are equal to 0s. None of the rows are really selected or physically removed.

π_j and ν_j can be described as $D^1_p(d_j)$ and $D^2_p(d_j)$, $j = 1, \dots, n$ and

$$D^1_p =_{\text{def}} \frac{S_p \times R^{c'}_{SD}}{|S_p|}, \tag{21}$$

$$D^2_p =_{\text{def}} \frac{S_p \times R^{e'}_{SD}}{|S_p|}. \tag{22}$$

Notice that $|S_p| = m$ and S_p is in a form of a row vector here. All values including 0s are in row S_p.

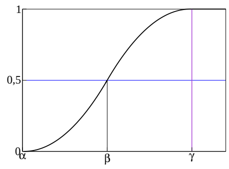
The final description of D_p is as follows: for $j = 1, \dots, n$

$$D_p(d_j) =_{\text{def}} \left(\text{Large}(D^1_p(d_j)) + \text{Small}(D^2_p(d_j))\right)/2. \tag{23}$$

The heuristic explanation of the fuzzy set D^1_p is as follows. A mean value of each column in p'_{ij} shows the mean strength of symptoms, presented by a patient for a disease. This can be taken as a strength of the disease itself. Values of fuzzy set Large indicate this strength for each disease from Δ. Meanwhile fuzzy set Small allows to assign small weights to values, that have large characteristics for non disease.

In general, there are no relations between Large and Small, between Large and LargeRatio in [2]. They can be constructed autonomously. In [13] an s-type membership function was taken to build Large and a z-type membership function was taken for Small. It is another difference between [2] and [13]. For example, an s-type membership function in Disco is defined as follows:

$$f(x;\alpha,\beta,\gamma) = \begin{cases} 0, & x \le \alpha; \\ \frac{1}{2}(\frac{x-\alpha}{\beta-\alpha})^2, & \alpha \le x \le \beta; \\ 1 - \frac{1}{2}(\frac{x-\gamma}{\beta-\gamma})^2, & \beta \le x \le \gamma; \\ 1, & x > \gamma. \end{cases}$$



A z-type membership function can be constructed from s-type membership function, for example, mirrored at the line $y = \frac{1}{2}$.

In [13] it is assumed that Large and Small are not necessarily strictly monotonous.

Large and Small can be considered as fuzzy modifiers [34] – operations that modify the meaning of fuzzy values. In our case, they are fuzzy sets D_p^1 and D_p^2. Recently, Novák [25, 26] has introduced a general theory of *evaluating linguistic expressions* that includes fuzzy modifiers as a part. These evaluating linguistic expressions characterize linguistically some value or number [24].

Notice that (21), (22) and (23) representing (17), (18), (16) correspondingly, describe the inference mechanism of Disco in the general fuzzy relation terminology for medical diagnosis. But the structure of this inference is different from those of CADIAG-II and Conorm-CADIAG-II-like systems.

4.3. Results of inference, D_p

Choosing $\max_{d_j} D_p(d_j)$, the corresponding d_j can be considered as the most believable diagnosis for a patient. The word *belief* is used here in the sense of *experience*, and has no connections to the theory of belief functions.

In general, an interpretation of diagnostic results can be done in accordance with the CADIAG-II method (see Section 2.5.): if $D_p(d_j) = 1$, or $D_p(d_j) = 0$, or $0 < D_p(d_j) < 1$, d_j is confirmed, or excluded, or possible, correspondingly. The combination of fuzzy sets to a final result in (23) – the computational schemata – resembles the total degree (13) of Conorm-CADIAG-II. (13) and (23) describe the analogous situation: the confirmation of a diagnosis is increased as the value for the exclusion of this diagnosis decreases. But (13) and (23) use different operations: an Abelian group operation in (13) and arithmetic mean in (23). Arithmetic mean belongs to the class of averaging operations. Moreover, it can be seen that (19) and (21) (or (20) and (22)) can be considered as a composition of fuzzy relations in the form of (9), where instead of t-conorm an averaging operation is used and instead of min the multiplication is used.

A crucial point has to be mentioned here. It is known that the min operator produces the largest fuzzy set from among those produced by all possible t-norms, the max operator produces the smallest one by all possible t-conorms and averaging operations lay between min and max:

$$\overbrace{t\text{-norm} \le \underbrace{\min \le \text{averaging functions} \le \max}_{\text{averaging operations}} \le t\text{-conorms}}^{\text{aggregation operations}}$$

The CADIAG-II-like systems work with aggregation operations from max to the right, whereas Disco takes the place from max to the left:

$$t\text{-norm} \leq \min \leq \text{averaging functions} \leq \boxed{\max} \leq t\text{-conorms}$$

$$\begin{array}{c} | \\ \text{CADIAG-II} \\ \text{Disco} \leftarrow \ | \rightarrow \text{CADIAG-II-like systems} \end{array}$$

Remark 1. In investigations of Disco ([2, 13]) and its current representation in the fuzzy relation framework (19) – (23) the arithmetic mean is used. To apply a different averaging, a particular generalization of Disco is needed. Some additional investigations are still left here.

4.4. Averaging operations in the inference mechanism of Disco

Taking the arithmetic mean M as an averaging operator from equation (3), the diagnostic process in Disco described by (19) and (21) (or (20) and (22)) can be represented by a composition of fuzzy relations as follows: for all $d_j \in \Delta$

$$D^1_p(d_j) =_{\text{def}} \text{M}(\odot(S_p(s_1), R^c_{SD}(s_1, d_j)), \ldots, \odot(S_p(s_m), R^c_{SD}(s_m, d_j))) \tag{24}$$

$$D^2_p(d_j) =_{\text{def}} \text{M}(\odot(S_p(s_1), R^e_{SD}(s_1, d_j)), \ldots, \odot(S_p(s_m), R^e_{SD}(s_m, d_j))) \tag{25}$$

and with final description of D_p from (23) as

$$D_p(d_j) =_{\text{def}} \text{M}(\text{Large}(D^1_p(d_j)), \text{Small}(D^2_p(d_j))) \tag{26}$$

where $\odot$ is an operation of multiplication, $d_j \in \Delta$.

Notice that, for example, an expression $D^1_p(d_j) = \frac{\sum_{s_i} S_p(s_i) \cdot R^c_{SD}(s_i, d_j)}{m}$ is a more clear expression for understanding (24), but we prefer the last one to support the uniqueness in the representation in CADIAG-II-like and Disco systems. We call an inference in (24) and (25) an averaging-multiplication composition, keeping in mind Remark 1.

In particular, the averaging operator (3) applied in (24) – (26) belongs to the class of *generalized means* operators [17]. It can be seen that M (3) is a special instance of OWA according to the Definition 10. As was mentioned in Section 4.2 Large and Small fuzzy sets indicate elements due to their strength for the disease.

5. RELATIONS BETWEEN CADIAG–II–LIKE SYSTEMS AND DISCO

In this section we summarize the main similarities, differences, relations between these systems.

Let us first summarize the description of Disco. The tabular knowledge base of this system represents the occurrence relations, defined numerically. The initial table is transformed into two ones, so called positive and negative discrimination tables. Due to the ability of the fuzzy set LargeRatio to specify the influenced symptoms for the diseases, the elements of the transformed tables can be interpreted as values described "how strong a symptom confirms (disconfirms) a disease". Therefore, the denotation of confirmation and exclusion relations $R^c_{SD}(s_i, d_j)$ and $R^e_{SD}(s_i, d_j)$ from Conorm-CADIAG-II can be used here: as was already told in the Section 4.1 the different knowledge is represented in the same way in CADIAG-II-like and Disco systems.

The new table data are cumulated further. For each disease the mean confirmation and non-confirmation values are calculated with a help of fuzzy sets Large and Small. The value from $[0, 1]$ represents the accumulated belief/experience and is assigned to each disease.

Full confirmation can be, however, very rare in Disco: it is effected only if all the present symptoms fully confirm the diagnosis in question. Similarly the diagnosis is excluded only if all the present symptoms fully exclude the diagnosis (fully confirm other diagnoses). Moreover, all the symptoms included in the system should be examined. Whereas a diagnosis is confirmed in Conorm-CADIAG-II iff at least one of the (fully) present symptoms fully confirms the diagnosis and no (fully) present symptom fully excluded it. Analogously for exclusion.

To underline again, a diagnosis for a patient can be confirmed and excluded in the same time in CADIAG-II and CADIAG-II-like systems, but not in Disco.

Interesting correspondences are to see between min operation of max-min inference and expressions for D^1_p (21) and D^2_p (22). Multiplication of reals is used in the similar way (as min) in Disco. The results are numerically the same, as Disco's inputs are 0s and 1s only (thus, $\min(0, x) = 0 = 0 \cdot x$ and $\min(1, x) = x = 1 \cdot x$ for all $0 \leq x \leq 1$). But this way of comparison can be done only for separate aggregation of positive contributions (or negative contributions) in Conorm-CADIAG-II. In Conorm-CADIAG-II the positive and negative impacts of symptoms are effected almost separately, whereas all the negative impacts of symptoms influence both p_{ij} and π_j with averaging in Disco, and similarly positive impacts of symptoms influence both n_{ij} and ν_j. In Conorm-CADIAG-II the positive and negative impacts of symptoms can be effected under assumption of statistical interpretation of inputs R^c_{SD}. If this interpretation is refused, relation R^c_{SD} can be constructed in such a way that it has always a positive impact to diagnoses and similarly R^e_{SD} which has always negative impact to diagnoses.

The system Disco allows an investigated patient only to have or not to have a symptom ("yes–no" case) without any gradations and all of the symptoms must be examined.

In CADIAG-II-like systems $S_p(s_i)$ is a fuzzy set. If s_i is a combination of symptoms (see Section 2.2.), $S_p(s_i)$ is computed from the values of each symptom occurring in this combination, using truth functions of fuzzy logic.

CADIAG-II-like systems and system Disco can be considered as compositional rule based systems, systems, where effects (contributions) of the rules are *composed*, and a numerical result is attached to the diagnosis. As with all compositional rule-based systems, this *compositionality* [6] is their weak point, too.

As was already told, the knowledge base of CADIAG-II-like systems consists of the IF–THEN rules that are usually given in a form of tables. Also Disco uses a tabular knowledge base (e. g., Table). Thus, both knowledge bases describe relations between symptoms and diseases: both systems can work with a knowledge base represented in the terms of matrices (tables) of frequencies/weights relating symptoms and diagnoses.

Both considered systems use an occurrence relation. CADIAG-II uses an occurrence relation constructed from $f(s_i|d_j)$ only for sure exclusion. On the other hand it uses a confirmation relation as a more important source of knowledge. A confirmation relation R^c in CADIAG-II-like systems is directly built from the initial source of information, from $f(d_j|s_i)$, and is used for confirmation and as an additive rule for sure exclusion.

An occurrence relation is the only source of knowledge for Disco, whereas for CADIAG-II-like systems it is a part of available data.

Negative knowledge was introduced in Conorm-CADIAG-II [6] and its knowledge base then consists of confirmation and exclusion relations. In Conorm-CADIAG-II only relation R^e is used for exclusion.

In the system Disco confirmation and exclusion relations are a result of a transformation of occurrence relations.

The Disco approach has a disadvantage: it cannot distinguish whether result 0.5 is computed because of weak arguments or as arithmetical mean of contradiction of full positive and negative contributions ($\frac{1}{2}(0+1) = 0.5$). This problem is important in medical applications as physicians often prefer three classes of diagnoses: confirmed, excluded and possible.

Total diagnostic contradictions (when diagnoses should be confirmed and excluded in the same time) are underlined in CADIAG's approaches, whereas they are hidden among possible diagnoses in Disco's approach.

The inference mechanism for Disco and Conorm-CADIAG-II can be described by a composition of fuzzy relations.

An averaging operator is idempotent [16], non-associative [17] (except of max, min or some other operations) and produces a fuzzy set that is larger than any fuzzy intersection and smaller than any fuzzy union (2). The general scheme of the inference mechanism (5), applied to positive and negative discrimination tables of Disco (respectively confirmation and exclusion relations), can be described with the arithmetic mean operator (see (24), (25)). Both contributions are computed together in (26). In the case of Conorm-CADIAG-II, (5) takes the form of (10), (11) and results are summarized in (13).

Notice that in CADIAG-II the results of inference are not combined in a total values, they classified into the predefined classes of confirmed, excluded by present symptoms, excluded by absent symptoms and possible diagnoses.

Concerning a combination of the final results in Disco and Conorm-CADIAG-II, another important questionable problem is using of fuzzy modifiers in Disco. In particular, Large and Small are defined only generally in the versions [2, 13] and in the current investigation. Therefore these functions can change the order of resulting values (degrees of confirmation) of diagnoses. Moreover, a possible diagnosis d_i ($D_p(d_i) < D_p(d_j)$) with given fuzzy modifiers, can be confirmed diagnosis, whereas

d_j remain a possible diagnosis when the fuzzy modifiers are changed. This "danger" has to be taken into consideration under the reconstruction of these functions. This "danger" shows that fuzzy modifiers have an important role which has a big influence on results, thus their selection must be careful and reasonable to obtain reasonable results.

These influence of fuzzy modifiers on results deserve further investigations.

Conorm-CADIAG-II is more safe from this point of view.

Another principal difference between Disco and CADIAG-II-like systems is that computation of positive and negative contributions is both theoretically and numerically same in Conorm-CADIAG-II, but it can be significantly different in Disco. If there are the same arguments with the same values both for a diagnosis d_i and against it, i. e., $R^c_{PD}(p, d_i) = R^e_{PD}(p, d_i)$ then resulting $R^{tot}_{PD}(p, d_j) = R^c_{PD}(p, d_j) \oplus -R^e_{PD}(p, d_j) = 0$, when (extended) group operation $\oplus$ is used. It means, in this case the neutral value, neither argument for nor against the diagnosis is obtained.

This does not hold in Disco, where we can obtain $\frac{1}{2}$ corresponding to 0 in CADIAG-II-like systems, or any other value close to 1 or some another close to 0, depending on a selection of fuzzy modifiers Large and Small. I.e., we can obtain anything from high support of d_i to high support that d_i is not occurred at the patient p. This underline again the difference between CADIAG-II-like systems and Disco and the fact, that Large and Small must be selected very carefully.

Let us summarize some important features discussed in this section.

- There is only one source of data in Disco (one table of frequencies $f(s_i|d_j)$), whereas there are two sources both in CADIAG-II (R^c and R^o) and Conorm-CADIAG-II (R^c and R^e).

- A knowledge base Θ_C given by occurrence relation R^o_{SD} of CADIAG-II-like systems is a knowledge base Θ_D of Disco, describing relations between a symptom (combination) s_i and a disease d_j.

- Differently calculated inputs, positive and negative contributions of symptoms to the diagnoses – positive and negative discrimination values in Disco, confirmation and exclusion relations in Conorm-CADIAG-II – can be denoted by the same symbols R^c_{SD} and R^e_{SD}. Thus, a different knowledge can be represented in the same way as it is used in CADIAG-II-like systems.

- The Conorm-CADIAG-II and Disco differ due to the sets of used values. In Disco $[0, 1]$ is used, in Conorm-CADIAG-II $[-1, 1]$. Correspondingly, 0 in Disco is interpreted as "no", whereas 0 in Conorm-CADIAG-II is interpreted as a neutral element.

- Disco and CADIAG-II-like systems can process data in a form of vectors of patients' symptoms value. In particular, in Disco only fully present or fully absent symptoms are considered, $S_p(s_j) = 1$, $S_p(s_j) = 0$, correspondingly, whereas in CADIAG-II-like systems S_p takes values from $[0, 1]$.

- The t-conorm-min composition is used in Conorm-CADIAG-II for aggregation of positive and negative contributions separately whereas in Disco averaging-multiplication is applied. An extended Abelian group operation is used in Conorm-CADIAG-II for combination of positive and negative contributions together, whereas arithmetic mean (or averaging, more generally) is used in Disco.
- Using fuzzy modifiers in Disco is an important difference between Disco and Conorm-CADIAG-II.
- The inference mechanism of Conorm-CADIAG-II and Disco can be described by scheme (5).

6. CONCLUSION

It can be seen from the discussion in the previous sections, the general language for description and comparison of CADIAG-II-like and Disco systems is based on the representation (1). Once again, the components of (1) relate to medical knowledge, patient's information, inference mechanism, diagnosis, correspondingly. We call this scheme (1) a general framework for medical diagnoses based on fuzzy relations. Since (1) describes the natural processes in each decision-support medical system, it can be interpreted in a framework, different from the fuzzy relations framework, for example, in a probabilistic one (conditional probability of symptoms given diseases, a priori probabilities of diseases, Bayes' formula). But for many systems the last approach would not have enough expressive power to represent some aspect of indeterminacy [27] – fuzziness (vagueness) – formalized in such systems. Thus, the general framework (1) based on fuzzy relations can be suitable for description of the systems where vagueness is present in the different form (relations between symptoms and diseases, in symptoms' representation, etc.). Notice that systems like CADIAG-II-like explicitly use the fuzzy relations in their representation and reasoning. Therefore the scheme (1) "one-to-one" corresponds to them. Others, e. g., MYCIN-like and Disco need some transformation to fit the scheme.

In [6] CADIAG-II and MYCIN like systems are compared. In general, the components of the scheme (1) were investigated for both types of systems. There are defined conditions under which it is possible to embed CADIAG-II-like into MYCIN-like systems and vice versa. [6] shows the procedures (and conditions) how a knowledge base of one class of the systems can be transformed into knowledge base of the second class such that both give the same results.

In this paper we have presented Disco and CADIAG-II-like systems in the same fuzzy relation-based framework (1). The systems are different from the theoretical point of view, but it was shown that they can be described in the same general language based on the scheme (1). This "generalization" allows to compare both types of systems on the certain level. Several interpretations of the components of the scheme were deduced during the present investigation. Similarities and differences between two types of systems were analyzed. For example, part of medical knowledge base of CADIAG-II-like systems can be present in Disco as occurrence relations.

Positive and negative contributions to the diagnoses in both systems are calculated due to the (5), but with different interpretations: t-conorm-min in CADIAG-II-like and arithmetic mean-multiplication in Disco. Both contributions are combined also differently: by an Abelian group operation in Conorm-CADIAG-II and averaging in Disco. Linguistic modifiers also distinguish the inference mechanism Disco from CADIAG-II-like systems.

Such a comparison may be interesting not only from the theoretic-academic point of view. Collected during investigation different variations, interpretations of R_{SD}, R_{PS}, $R_{PD} =_{\text{def}} R_{PS} \circ R_{SD}$ can help to build new decision support systems in medicine. Thus from the perspective of applications, the subject of special interest can be systems with practical realizations. CADIAG-II and Disco belong to this category [4, 13].

As well as previous investigations in this direction, our contribution, of course, does not complete the study, but poses new questions, opens new problems and each step nears the far-reach goal.

ACKNOWLEDGEMENT

I would like to express my gratitude to an anonymous referee, who helped to improve the paper considerably, and to Professor Claudio Moraga (University of Dortmund), whose support is always a source of knowledge and inspiration for me.

(Received April 19, 2005.)

REFERENCES

[1] K.-P. Adlassnig and G. Kolarz: CADIAG-II: computer-assisted medical diagnosis using fuzzy subsets. In: Approximate Reasoning in Decision Analysis (M. M. Gupta and E. Sanchez, eds.), North Holland, Amsterdam 1982, pp. 495–505.

[2] J. F. Baldwin, B. W. Pilsworth, and D. Norris: Medical diagnoses from patient records. Fuzzy Sets and Systems *23* (1987), 73–87.

[3] K. Boegl: Design and Implementation of a Web-Based Knowledge Acquisition Toolkit for Medical Expert Consultation Systems. Ph.D. Thesis, Technische Universität Wien, 1997.

[4] K. Boegl, K.-P. Adlassnig, Y. Hayashi, T. Rothenfluh, and H. Leitich: Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system. Artificial Intelligence in Medicine *30* (2004), 1–26.

[5] M. Daniel: Theoretical comparison of inference in CADIAG and MYCIN-like systems. Tatra Mountains Math. Publ. *16* (1999), 255–272.

[6] M. Daniel, P. Hájek, and H. Nguyen: CADIAG-2 and MYCIN-like systems. Artificial Intelligence in Medicine *9* (1997), 241–259.

[7] D. Dubois and H. Prade: Fuzzy Sets and Systems: Theory and Applications. Academic Press, New York 1980.

[8] D. Dubois and H. Prade: A review of fuzzy aggregation connectives. Inform. Sci. *36* (1985), 1–2, 85–121.

[9] R. Fullèr: Introduction to Neuro-fuzzy Systems. Physica-Verlag, Heidelberg 2000.

[10] P. Hájek, T. Havránek, and R. Jiroušek: Uncertain Information Processing in Expert Systems. CRC Press, Boca Raton, Florida 1992.

[11] P. Hájek and J. Valdés: An analysis of MYCIN-like expert systems. Mathware and Soft Computing *1* (1994), 45–68.
[12] F. Hayes-Roth: Expert systems. In: Encyclopedia of Artificial Intelligence – Second edition (S. C. Shapiro, ed.), Volume 1, Wiley, New York 1992, pp. 477–489.
[13] T. Kiseliova: A Computer Assisted Method in Medical Diagnosis Using New Information Technologies. Ph.D. Thesis, Tbilisi State University, Georgia Technical University 1995.
[14] E.-P. Klement, R. Mesiar, and E. Pap: Triangular Norms. Kluwer, Dordrecht 2000.
[15] G. J. Klír, U. H. Clair, and B. Yuan: Fuzzy Sets Theory, Foundations and Applications. Prentice Hall, Upper Saddle River, NJ 1997.
[16] G. J. Klír and M. J. Wierman: Uncertainty–Based Information. Physica–Verlag, Heidelberg 1998.
[17] G. J. Klír and B. Yuan: Fuzzy Sets and Fuzzy Logic. Theory and Applications. Prentice Hall PTR, Upper Saddle River, NJ 1995.
[18] G. Kolousek: The System Architecture of an Integrated Medical Consultation System and Its Implementation Based on Fuzzy Technology. Ph.D. Thesis, Technische Universität Wien 1997.
[19] H. Leitich, K.-P. Adlassnig, and G. Kolarz: Evaluation of two different models of semi-automatic knowledge acquisition for the medical consultant system CADIAG-II/RHEUMA. Artificial Intelligence in Medicine *25* (2002), 215–225.
[20] W. J. V. Melle, E. H. Shortliffe, and B. G. Buchanan: EMYCIN: A knowledge engineer's tool for constructing rule-based expert systems. In: Rule-Based Expert Systems. Addison-Wesley, Reading, MA 1984.
[21] J. M. Mendel: Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions. Prentice Hall PTR, Upper Saddle River 2000.
[22] K. Menger: Statistical metric spaces. Proc. Nat. Acad. Sci. U. S. A. *28* (1942), 535–537.
[23] P. H. Nguyen: Some Contributions to Fuzzy Expert Systems with Emphasis on Medicine. Ph.D. Thesis, Technische Universität Wien 1995.
[24] V. Novák: Fuzzy logic deduction with words applied to ancient sea level estimation. In: Fuzzy Logic in Geology (R. Demicco and G. J. Klír, eds.), Academic Press, Amsterdam 2003, pp. 301–336.
[25] V. Novák: Fuzzy relation equations with words. In: Fuzzy Partial Differential Equations and Relational Equations: Reservoir Characterization and Modeling, Studies in Fuzziness and Soft Computing (M. Nikravesh, L. A. Zadeh, and V. Korotkikh, eds.) Physica–Verlag, Heidelberg 2004, pp. 167–185.
[26] V. Novák and I. Perfilieva: On the semantics of perception-based fuzzy logic deduction. Internat. J. Intelligent Systems *19* (2004), 1007–1031.
[27] V. Novák, I. Perfilieva, and J. Močkoř: Mathematical Principles of Fuzzy Logic. Kluwer, Boston – Dordrecht 1999.
[28] T. E. Rothenfluh, K. Bögl, and K.-P. Adlassnig: Representation and acquisition of knowledge for a fuzzy medical consultation system. In: Fuzzy Systems in Medicine (P. S. Szczepaniak, P. J. G. Lisboa, and J. Kacprzyk, eds.), Physica Verlag, Heidelberg – New York 2000, pp. 636–651.
[29] E. Sanchez: 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.
[30] B. Schweizer and A. Sklar: Probabilistic Metric Spaces. North Holland, Amsterdam 1983.
[31] R. Seising, C. Schuh, and K.-P. Adlassnig: Medical knowledge, fuzzy sets and expert systems. In: Workshop on Intelligent and Adaptive Systems in Medicine, Prague 2003, http://cyber.felk.cvut.cz/EUNITE03-BIO.

[32] E. H. Shortliffe: Computer–Based Medical Consultations, MYCIN. Elsevier, North–Holland, New York 1976.

[33] H. Thiele: Einführung in die Fuzzy-Logik, University of Dortmund. Spezialvorlesung. Script, ftp://ftp.informatik.uni-dortmund.de/pub/ls01.info/fuzzy.logik/skript, 1998.

[34] H. Thiele: Interpreting linguistic hedges by concepts of functional analysis and mathematical logic. In: 6th European Congress on Intelligent Techniques and Soft Computing, EUFIT'98, Aachen 1008, pp. 114–120.

[35] R. Yager: Level sets for membership evaluation of fuzzy subsets. In: Fuzzy Set and Possibility Theory, Recent Developments (R. Yager et al., eds.), Pergamon Press, 1982, pp. 90–97.

[36] R. R. Yager: Fuzzy logic method in recommended systems. Fuzzy Sets and Systems *136* (2003), 133–149.

[37] R. R. Yager and D. P. Filev: Essentials of Fuzzy Modelling and Control. Ketherine Schowalter, 1994.

[38] L. A. Zadeh: Fuzzy sets. Inform. and Control *8* (1965), 338–353.

[39] L. A. Zadeh: The concept of a linguistic variable and its application to approximate reasoning, Part I, Part II, Part III. Inform. Sci. *8, 9* (1975), 199–249, 301–357, 43–80.

[40] X.-H. Zhou, N. A. Obuchowski, and D. K. McClish: Statistical Methods in Diagnostic Medicine. (Wiley Series in Probability and Statistics.) Wiley, New York 2002.

[41] H.-J. Zimmermann: Fuzzy Set Theory and Its Applications. Kluwer Academic Publishers, Dordrecht 1991.

Tatiana Kiseliova, University of Dortmund, Department of Computer Science I, Otto-Hahn-Str. 16, D-44221 Dortmund, Germany and Ivane Javakhishvili Tbilisi State University, University St. 2, 0145, Tbilisi, Georgia.

e-mail: Tatiana.Kiseliova@udo.edu