Interval Computations with BK-Products of Fuzzy Relations in Diagnostic Knowledge-Based System CLINAID

LADISLAV J. KOHOUT and ISABEL STABILE* ,
Dept. of Computer Science, Florida State University, Tallahassee, Florida 32306-4530, USA.
*Department of Anatomy, University of Malta, Republic of Malta, Europe

Abstract: – CLINAID is a medical knowledge-based system that uses fuzzy relational structures for both knowledge representation and inference. The system can deal with multiple body systems. Interval-based fuzzy logics employed in CLINAID make it possible to deal efficiently with multiplicity of contexts that appear in medical decision making involving risk and uncertainty. A particular emphasis is placed on the description of the involvement of fuzzy triangle and square relational products that play a significant role in our approach.

Keywords: – Diagnostic knowledge-based systems, interval computing, fuzzy relations, BK-products of relations, fuzzy interval logics, medicine, soft computing

1 Clinaid Generic Architecture

The *Clinaid* approach significantly differs from more conventional AI approaches in its potential capabilities of dealing with incomplete information. This is done by combining fuzzy relational inference with fuzzy interval logics. Furthermore, the majority of extant medical expert systems deal with a limited medical context, the largest domain of knowledge being just a single medical field, e.g. internal medicine or cardiac medicine, etc. The inherent limitation of such knowledge based systems is conceptual and logical: their knowledge bases and inference engines cannot easily mix the knowledge from several fields without adverse effects. CLINAID successfully deals with this problem by introducing a *multi-center architecture* in the Diagnostic Unit by which the multiple contexts are handled concurrently by means of interval relational computations. Further advantage of using fuzzy relations for knowledge representation and inference is it unifying power. Relational inference can be realized by different architectures: it can be equally well embedded in to a logic symbolic architecture or fuzzy neural network architecture [16], according to the implementation needs.

2 Knowledge Representation and Inference in Clinaid

These sections deal with relational approach to KR and Inference.

2.1 Knowledge Representation

B .... a set of body systems.
G .... sets of general diseases.
D .... sets of specific diseases.
I .... a set of investigations.
P .... a set of patients.
S .... a set of symptoms and physical signs.
Y .... sets of syndromes.

These sets enter into the following relations:

- ‘symptoms and signs to body systems’
  $SB \in \mathcal{F}(S \rightsquigarrow B)$
- ‘symptoms and signs to syndromes’
  $SY_i \in \mathcal{F}(S \rightsquigarrow Y)$
- ‘symptoms and signs to syndromes’
  $SY_i \in \mathcal{F}(S \rightsquigarrow Y)$
- ‘syndromes to general diseases’
  $YG_i \in \mathcal{F}(Y \rightsquigarrow G)$
- ‘general diseases G to investigations I’
  $GI \in \mathcal{F}(G \rightsquigarrow I)$
- ‘investigations results J to specific diseases D
  JD $\in \mathcal{F}(J \rightsquigarrow D)$

The index i above refers to *i-th relation* of a family of relations. So we have a family of relations of signs and symptoms, a family of relations of syndromes, a family of relations of general diseases in order to cover all the body systems. There are 11 body systems altogether [8]. We have explored so far the following 5 body systems in CLINAID in some detail: cardio-vascular, endocrine, respiratory, muscular, blood and reticulo-endothelial systems. We have also started work on acquisition of medical knowledge for another body system, namely *reproductive*.

2.2 Dealing With the Inferential Context

The diagnostic unit of CLINAID deals with a *multiplicity* of contexts provided by the several body systems and other factors. We distinguish: a) the *context of different body systems*; b) the *context of different diagnostic levels* within the super-context of a specific body system.

The strategy of clinical inference used over this diagnostic hierarchy is realized by means of fuzzy triangle products [4]. The relations are created, represented and computed dynamically. Only the items of information which are relevant for reaching the conclusion of a specific level are utilized when each level is activated. The activation of levels within the specific body system represented in a center of the diagnostic unit is ordered by the following sequence:

(body system) $\Rightarrow$ *syndromes* $\Rightarrow$ *general diseases* $\Rightarrow$ *working diagnoses.*

Signs and symptoms of a patient entered into the system are matched with the knowledge structures distributed within the individual levels of a specific center. The stream of computations passes through the levels of the above sequence, in an iterative manner.

2.3 Relational Inference

All the versions of CLINAID have used BK-relational products [10],[1],[2],[4],[15],[5]. In addition to the standard fuzzy composition $(R \circ S)_{ik} = \bigvee_j (R_{ij} \bigwedge S_{jk})$ we distinguish the following BK–product types.

2.3.1 Mathematical definitions of BK-Products

Where R is a relation from X to Y, and S a relation from Y to Z, a *product relation* $R@S$ is a relation from X to Z, determined by R and S such that $@ \in \{\triangleleft, \triangleright, \square\}$. In the following definitions of the products, R_{ij}, S_{jk} represent the fuzzy degrees to which the respective statements $x_i R y_j$, $y_j S z_k$ are true.

PRODUCT TYPE	MANY-VAL. LOGIC DEFIN.
Subproduct	$(R \triangleleft S)_{ik} = \mathcal{Q}_j(R_{ij} \rightarrow S_{jk})$
Superproduct	$(R \triangleright S)_{ik} = \mathcal{Q}_j(R_{ij} \leftarrow S_{jk})$
Square product	$(R \square S)_{ik} = \mathcal{Q}_j(R_{ij} \equiv S_{jk})$

The symbols for the both *implications* and the *equivalence* in the formulas shown above represent connectives of some many-valued logic, **chosen** according to the logic properties of the products required. The generic formula

$$(R@S)_{ik} := \mathcal{Q}_j(R_{ij} \# S_{jk}),$$

yields two generic types of fuzzy relational products. We can replace the generalized quantifier $\mathcal{Q}$ with *infimum* or with $\frac{1}{|J|}\sum$:

$(R@S)_{ik} := \inf_j (R_{ij} \# S_{jk})$: *Harsh* product,

$(R@S)_{ik} := \dfrac{1}{|J|}\sum_j (R_{ij} \# S_{jk})$: *Mean* product.

where $@ \in \{\triangleleft, \triangleright, \square\}$ and $\# \in \{\rightarrow, \leftarrow, \equiv\}$.

The details of choice of the appropriate many-valued connectives are discussed in [1],[2],[3],[14].

The knowledge representation scheme and inference algorithms uses not only 2-argument but also 3-arguments (ternary) relations. The computation of BK-products is not different from 2-argument relations shown above. Thus given two ternary relations, say R and S the a subproduct for example, will be computed by the formula

$$(R \triangleleft S)_{iklm} = \mathcal{Q}_j(R_{ijk} \rightarrow S_{jlm})$$

matching the elements of R with elements of S over the index j by the implication operator. The resulting relation is a quaternary relation $(R \triangleleft S)$.

The matching indexes need not be the same. For example, one can compute the following product:

$$(R \triangleleft T)_{ijlm} = \mathcal{Q}_{(k=n)}(R_{ijk} \rightarrow T_{lmn})$$

matching the elements of R with index k with elements of S with index n by the implication operator. Which indexes have to be matched is indicated in the quantifier by the subscript equality $k = n$. The result is again a quaternary relation.

2.3.2 Fuzzy Interval Inference

An important extension of the reasoning capabilities of CLINAID is *interval-based* inferential structure [17], [8],[7]. The semantic justification of logics used for this purpose is provided by a theoretical device called the *checklist paradigm* [1],[7]. This paradigm not only gives a sound epistemological justification for interval-valued inference, but also supports knowledge and data retrieval techniques utilizing information structures with interval credibility weights.

The inequality restricting the possible values of measure m(F) expressing the logical values that fall within the interval, is written in its general form as: $contop \geq m(F) \geq conbot$. In this paper we employ $m(F) = m_1(F) = 1 - (\alpha_{10}/n)$. This choice yields the bounds [1],[7]:

$$min(1, 1 - a + b) \geq m_1(\rightarrow) \geq max(1 - a, b),$$

Łukasiewicz and Kleene-Dienes implication operators, respectively. These implication operators were used in the examples in the sequel. For $\equiv$ we have used

$TOPCON: \quad a \equiv b = min(1-a+b, 1-b+a).$

$BOTCON: \quad a \equiv b = max(a+b-1, 1-(a+b)).$

There is a deeper justification for the choice of connectives used for interval logic reasoning in CLINAID. The checklist paradigm semantics reveals that structural interdependencies between connectives of fuzzy interval logics used here are captured by the subgroups of the 8 element symmetric group $\mathcal{S}_{2\times2\times2}$ of logic transformations. The space restriction does not allow us to discuss the underlying theory here. Detailed presentation of these issues, however, can be found in [11],[7].

2.3.3 Example of Fuzzy Interval-Based Inference in the Cardiac Body System

Global inferential activity of CLINAID consists of matching the relevant empirical findings with medical theories and normative facts ([12] Chapter 10 section 4, pp. 181-184).

Some relational formulas employed in CLINAID are listed and their conceptual meaning explained below.

Name:		A set of:
B	....	body systems.
D	....	specific diseases.
V	....	variability index of observation events.
P	....	patients.
S	....	symptoms and physical signs.

Relation $PVS \in \mathcal{R}(P \times V \times S)$ is a relation between *patients, variability index (time) of observation events* and *signs and symptoms*. $VBS \in \mathcal{R}(V \times B \times S)$ is between *variability index of observation events* and *signs and symptoms*. $PS \in \mathcal{R}(P \rightsquigarrow S)$ relates *patients* to *symptoms and physical signs*. $PD \in \mathcal{R}(P \rightsquigarrow D)$ is a relation from *patients* to *specific diseases*. $SB\mathcal{R}(S \rightsquigarrow B)$ relates *symptoms and physical signs* to *body systems*. Finally, $SD \in \mathcal{R}(S \rightsquigarrow D)$ is a relation from *symptoms and physical signs* to *specific diseases*.

Given are two ternary relations, PVS and VBS, from which the composed relations PBS and PB are computed. PVS is a ternary (3-place) relation relating the set of patients P with the set of all time instances V at which the observations of relevant signs and symptoms S are made. This relation belongs to the family of relations capturing the relevant empirical findings. VBS is a ternary relation relating the variability index of observational events V with the set of body systems B and the set of signs and symptoms S. The VBS relation captures the relationship between signs and symptoms and body systems. The actual observations of patients captured by the PVS relation have to be matched with the knowledge contained in the relation VBS over the set V. This yields the computation

$$PBS_{imk} = (PVS \lhd VBS)_{imk} = \\ (1/cardV) \textstyle\sum_{j=l} (PVS_{ijk} \rightarrow VBS_{lmk})$$

$$PB_{im} = (PBS \square^{*} PBS)_{im} = \\ (1/cardS) \textstyle\sum_{k} (PBS_{imk} \wedge (PBS_{imk} \equiv PBS_{imk}))$$

where *card V, card S* are the cardinalities (sizes) of the sets of relevant elements, over which the relational composition (matching) is performed. The second product yielding the relation PB computes the degree of involvement of each body system relative to the composition over available symptoms and signs (and generally also over other observables) characterizing individual patients.

We present a sample result of working diagnoses computed from real clinical data [8].

Let us assume that our clinical case is a patient that exhibits the following set S of signs and symptoms. The set S is a fuzzy set consisting of $s_1 = SOB/.8$ (Shortness of breath), $s_2 = AS/.6$ (Ankle swelling), $s_3 = HM/.7$ (Hepatomegaly), $s_4 = ANX/.6$ (Anxiety), $s_5 = JVPH/.8$ (Jugular venous pressure high), $s_6 = CMRM/.7$ (Cardiac murmurs), $s_7 = PLP/.7$ (Palpitations) and $s_8 = RVEP/.75$ (Right ventricular enlargement).

From this set, and using the clinical knowledge structure of the Diagnostic unit one computes the relation *PB* that suggests the involvement of individual body systems. This relation is computed by the products $PVS \lhd VBS$ and $PBS \square^{*} PBS$ listed above. The computation shows that the top candidate is the *cardio-vascular* body system. Further details of selecting the relevant body systems will be presented in Sec. 3.2 below.

In this section we proceed with demonstration of inference assuming that the cardiac system has already been selected as the most relevant, given the above listed patient's signs and symptoms used as the input data into relations PVS and P.

The relevant syndromes in a specific body system are computed by the fuzzy triangle product

$PY = PS \lhd SY$ from patients to syndromes given by the formula

$$PS_{ik} = (PS \lhd SY)_{ik} = \\ = (1/cardS) \sum_j (PS_{ij} \to SY_{jk})$$

The following syndromes within the cardiovascular system are identified as the most plausible for patient p_1: *Acute cardiac failure (ACF), Left sided heart failure (LHF), Right sided heart failure (RHF), Murmurs (MRM)* and *Dysrythmia (DSR)*. The intervals of the possibility of their occurrence are listed in decreasing order as the fuzzy interval set of syndromes Y:

Y={ACF/[.72, .57], DSR/[.63, .43], MRM/[.56, .48], RHF/[.54, .47], LHF/[.38, .36]}.

The fuzzy triangle product $PYD = PS \lhd SYD$ from patients to diseases (indexed by syndromes) is given by the formula

$$PYD_{ikm} = (PS \lhd SYD)_{ikm} = \\ (1/cardS) \sum_j (PS_{ij} \to SYD_{jkm})$$

It provides the final working diagnosis. This is listed in Table 1 below. *Congestive Heart Failure* is the correct diagnosis.

Table 1: Working Diagnoses for syndrome Acute Cardiac Failure ACF/[0.72, 0.57]

Disease	Possibility of occurrence
Congestive Heart Failure	[1.0, 0.76]
Heart Block	[0.89, 0.64]
Myocardial Infarction	[0.73, 0.59]
Malignant Hypertension	[0.72, 0.59]
Cardiomyopathy	[0.73, 0.58]

3 The Integrating Role of Fuzzy Relational Products

There were two basic problems that precluded us from adapting conventional techniques used in the expert systems field, namely:

1. The large number of signs, symptoms and diseases used in CLINAID.

2. Each body system of CLINAID deals with a different medical field.

The first problem noted above is resolved by adequate parallel relational structures. The second one is more serious, as it is a problem that is both logical and conceptual: no amount of computational power can avoid problems caused by the *mixing of several distinct contexts*. Only a good, well structured design of a Knowledge-Based System based on logics with well defined semantics and epistemology can provide a remedy.

For these reasons, we have radically departed from established practice and used fuzzy relational structures in most parts of CLINAID, not only for the processes of inference and knowledge representation, but have also included fuzzy relational dynamic protection structures in the data bases [6] and fuzzy relational methods of retrieval of data base objects [13].

3.1 Multi-Center Approach

CLINAID chooses to use fuzzy relations to capture this large amount of clinical knowledge in a structured way, and to make it easier to deal with parallelism that is essential for a multi-center approach. From the point of view of system architecture, our approach is grounded on two essential features that are built into the system:

1. Plurality of control and inferential centers.

2. Inference via triangular relational products.

The plurality of centers requires their *mutual co-operation*. Briefly, the plural centers co-operate as follows:

1. Each center is adapted to dealing with a particular portion of the common task.

2. Each center filters out different aspects of data, according to the context or to what its special competence is.

The power of this way of organizing things comes jointly from the *specialization* of the former and the *simplification* of the latter. How the *common task* is *subdivided* and how the centers are co-ordinated are of course delicate and crucial questions [9], but assuming that they can be achieved successfully then the advantages of subdivision are evident

1. Each center deals only with a small portion of systemic knowledge.

2. The centers can operate in parallel.

These provide a major weapon against combinatorial explosion, with the following desirable consequences:

1. Inference and meta-inference are both expressible as relations.

2. Inference is computable in parallel.

3.2 The Context of Body Systems

In the previous sections we have shown how inference is made within a specific body system, namely a cardiac one. Each body system represents a different context, determined by a particular medical specialty. So, the CLINAID has to be capable of deciding which contexts are relevant with respect to given specific input data, and which contexts ought to be eliminated form further inference.

Table 2 below presents a small fragment of the knowledge structure of CLINAID, capturing the relationship of signs/symptoms to several body systems. We shall use this table to demonstrate, how CLINAID selects relevant body system(s).

Let us assume that our clinical case is a patient that exhibits the following set S of signs and symptoms. The set S is a fuzzy set consisting of $s_1 = SOB/.8$ (Shortness of breath), $s_2 = AS/.6$ (Ankle swelling), $s_3 = HM/.7$ (Hepatomegaly), $s_4 = ANX/.6$ (Anxiety), $s_5 = JVPH/.8$ (Jugular venous pressure high), $s_6 = CMRM/.7$ (Cardiac murmurs), $s_7 = PLP/.7$ (Palpitations) and $s_8 = RVEP/.75$ (Right ventricular enlargement).

From this set, and using the clinical knowledge structure of Table 2 one computes the relation *PB* that suggests the involvement of individual body systems.

Table 2: Relation of Signs/symptoms to Body systems

Signs/symptoms	*Body systems*			
	RS	CVS	MS	HEM
SOB	.9	.9	.3	.6
PND	.1	.9	.1	.6
Cyanosis	.8	.8	.3	.3
Tachycardia	.5	.9	.2	.6
Fine basal creps	.2	.9	.2	.6
Ankle swelling	.3	.9	.2	.6
Hepatomegaly	.3	.7	.1	.6

This relation is computed by the products $PVS \triangleleft VBS$ and $PBS \square^* PBS$ listed in Sec. 2.3.3 above. The results of this computation are shown in Table 3.

It can be seen from Table 3 that there is *high* positive evidence for the involvement of cardiovascular system CVS, *low*, negligible positive and negative evidence (around .5) for respiratory system RS and *substantial* negative evidence for the involvement of the muscular system MS.

Table 3: Interval-Based Diagnosis: Identification of Possible Body Systems of Patient p_1

Body Syst.	top$^+$	bot$^+$	top$^-$	bot$^-$
RS	.5	.43	.57	.5
CVS	1.0	.67	.33	0
MS	.5	.33	.67	.5
HEM	.9	.2	.8	.1
Grounds for:	*Acceptance*		*Rejection*	

There is, however, great uncertainty in indicating the possibility of the involvement of the hematological system HEM. From the clinical point of view, the wide interval [.2, .9] indicates that the given signs and symptoms are not as specific for HEM as they are for the cardiovascular system CVS. Hence, the CVS is the winning candidate body system.

4 Conclusion

Fuzzy relations present a unifying framework for integrating a number of different techniques [16]. Namely, fuzzy approximate reasoning techniques, many-valued logic based interval computations, symbolic logic computations based on non-associative compositions of relations, semiotic descriptors based handling of linguistic statements, and diffuse distributed computations in neural networks. We have shown how this inegration has been done in a medical multi-centre, multicontext KBS CLINAID. Designing CLINAID as a fuzzy relational architecture allows us to employ fuzzy information retrieval methods as the means for unified coordination of information flow between the individual communicating centers of CLINAID, be it relational symbolic computation modules, interval computation modules or neural networks.

References

[1] W. Bandler and L.J. Kohout. Semantics of implication operators and fuzzy relational products. *Internat. Journal of Man-Machine Studies*, 12:89–116, 1980. Reprinted in Mamdani, E.H. and Gaines, B.R. eds. *Fuzzy Reasoning and its Applications*. Academic Press, London, 1981, pages 219-246.

[2] W. Bandler and L.J. Kohout. A survey of fuzzy relational products in their applicability to medicine and clinical psychology. In L.J. Kohout and W. Bandler, editors, *Knowledge Representation in Medicine and Clinical Behavioural Science*, pages 107–118. Gordon and Breach Publ., London and New York, 1986.

[3] W. Bandler and L.J. Kohout. Fuzzy implication operators. In M.G. Singh, editor, *Systems and Control Encyclopedia*, pages 1806–1810. Pergamon Press, Oxford, 1987.

[4] W. Bandler and L.J. Kohout. Relations, mathematical. In M.G. Singh, editor, *Systems and Control Encyclopedia*, pages 4000 – 4008. Pergamon Press, Oxford, 1987.

[5] P. Hájek. A remark on Bandler-Kohout products of relations. *Internat. Journal of General Systems*, 25(2):165–166, 1996.

[6] V. Kaliappan, L.J. Kohout, and J. Anderson. Design of a well-protected patient record unit for multi-centre knowledge-based system clinaid. In K.-P. Adlasnik, G. Grabner, S. Bengtsson, and R. Hansen, editors, *MIE91 – 10th International Congress (Lecture Notes in Medical Informatics)*, pages 309–313, Berlin, August 19-22 1991. European Federation for Medical Informatics, Springer Verlag.

[7] L. J. Kohout and W. Bandler. Interval-valued systems for approximate reasoning based on the checklist paradigm. In P. Wang, Paul, editor, *Advances in Fuzzy Theory and Technology, vol. 1*, pages 167–193. Bookwrights Press, Durham, N.C., 1993.

[8] L. J. Kohout, I. Stabile, W. Bandler, and J. Anderson. CLINAID: Medical knowledge-based system based on fuzzy relational structures. In M. Cohen and D. Hudson, editors, *Comparative Approaches in Medical Reasoning*, pages 1–25. World Scientific, 1995.

[9] L.J. Kohout. *A Perspective on Intelligent Systems: A Framework for Analysis and Design*. Chapman and Hall & Van Nostrand, London & New York, 1990. A Scientific Monograph, 255 pages. In 1991, received an international prize from The International Institute for Advanced Studies in Systems Research: *"The best book of the year in the area of AI Systems"*.

[10] L.J. Kohout. Boolean and fuzzy relations. In P.M. Pardalos and C.A. Floudas, editors, *The Encyclopedia of Optimization*, pages 189–202. Kluwer, Boston, 2001. vol.I, A-D.

[11] L.J. Kohout. Checklist paradigm semantics for fuzzy logics. In P.M. Pardalos and C.A. Floudas, editors, *The Encyclopedia of Optimization*, pages 237–246. Kluwer, Boston, 2001. vol.I, A-D.

[12] L.J. Kohout, J. Anderson, and W. et al. Bandler. *Knowledge-Based Systems for Multiple Environments*. Ashgate Publ. (Gower), Aldershot, U.K., 1992. A Scientific Monograph, 382 pages. Written by the principal author in collaboration with others. Awarded *"Outstanding Scholarly Contribution Award"* by the Systems Research Foundation in 1993.

[13] L.J. Kohout and W. Bandler. The use of fuzzy information retrieval techniques in construction of multi-centre knowledge-based systems. In B. Bouchon and R.R. Yager, editors, *Uncertainty in Knowledge-Based Systems (Lecture Notes in Computer Science vol. 286)*, pages 257–264. Springer Verlag, Berlin, 1987.

[14] L.J. Kohout and W. Bandler. Fuzzy relational products in knowledge engineering. In V. Novák et al., editor, *Fuzzy Approach to Reasoning and Decision Making*, pages 51–66. Academia and Kluwer, Prague and Dordrecht, 1992.

[15] L.J. Kohout and W. Bandler. Use of fuzzy relations in knowledge representation, acquisition and processing. In L.A. Zadeh and J. Kacprzyk, editors, *Fuzzy Logic for the Management of Uncertainty*, pages 415–435. John Wiley, New York, 1992.

[16] L.J. Kohout and E. Kim. The role of BK–products of relations in soft computing. *Soft Computing*, 6(2):87–91, 2002.

[17] L.J. Kohout and I. Stabile. Interval-valued inference in medical knowledge-based system Clinaid. *Interval Computations*, 2(3):88–115, 1993.