# Approximating Infinite Solution Sets by Discretization of the Scales of Truth Degrees


Radim Belohlavek
Dept. Systems Science and Industrial Engineering
Binghamton University—SUNY
Binghamton, NY, 13902, U. S. A.,
rbelohla@binghamton.edu

Michal Krupka and Vilem Vychodil
Dept. Computer Science
Palacky University, Olomouc
Tomkova 40, Olomouc 77900, Czech Republic
michal.krupka@upol.cz, vilem.vychodil@upol.cz



***Abstract*— The present paper discusses the problem of approximating possibly infinite sets of solutions by finite sets of solutions via discretization of scales of truth degrees. Infinite sets of solutions we have in mind in this paper typically appear in constraint-based problems such as "find all collections in a given finite universe satisfying constraint $C$". In crisp setting, i.e. when collections are conceived as crisp sets, the set of all such collections is finite and often computationally tractable. In fuzzy setting, i.e. when collections are conceived as fuzzy sets, the set of all such collections may be infinite and, *ipso facto,* computationally intractable when one uses the unit interval $[0, 1]$ as the scale of membership degrees. A natural solution to this problem is to uses, instead of $[0, 1]$, a finite subset $K$ of $[0, 1]$ which approximates $[0, 1]$ to a satisfactory degree. This idea is pursued in the present paper. To be sufficiently specific, we illustrate the idea on a particular method, namely, on formal concept analysis. We present several results including estimation of degrees of similarity of the finitary approximation to the possibly infinite original case by means of the degree of approximation of $K$ of $[0, 1]$.**


## I. Introduction and problem setting

Using the real unit interval $[0, 1]$ as a scale of truth degrees is the most common choice in fuzzy logic applications. The aim of our paper is to bring to the attention one aspect of using $[0, 1]$. Namely, given a universe set $X$, the set $[0, 1]^X$ of all fuzzy sets in $X$ is infinite even if $X$ is finite. While this fact is advantageous from the point of view of representation capability of fuzzy sets in $X$, there are apparent disadvantages of using $[0, 1]$ as well. Namely, using $[0, 1]$ can lead to problems which are computationally not feasible even if the corresponding crisp problems, i.e., problems with $[0, 1]$ replaced by $\{0, 1\}$, are computationally tractable. Examples of these problems can be drawn from data mining tasks where one tries to extract all collections of elements of some set, say $X$, which satisfy certain constraint $C$. When "collection" is understood as a set, the search space is $2^X$ and the solution set $\{A \in 2^X \mid A \text{ satisfies } C\}$ can be efficiently computable. On the other hand, when "collection" is understood as a fuzzy set, both the search space $[0, 1]^X$ and the solution set $\{A \in [0, 1]^X \mid A \text{ satisfies } C\}$ can be infinite. In such a case, problem of computing the solution set is not feasible in principle. A natural idea in this case is to consider a finite subset $K$ of $[0, 1]$ which is a "good approximation" of $[0, 1]$. Then both $K^X$ and the solution set $\{A \in K^X \mid A \text{ satisfies } C\}$ are finite and the solution set can be a good approximation of the original infinite solution set $\{A \in [0, 1]^X \mid A \text{ satisfies } C\}$. In our paper, we attempt to formalize the above considerations. We provide general results which address several issues of the idea of approximating an infinite solution set by a finite one. In particular, we obtain approximation formulas and a result which enables us to infer, given a required approximation level, how to select a $K$ which guarantees the required approximation level.

We demonstrate our ability to approximate infinite solution sets by several examples. For illustration, we consider formal concept analysis (FCA, see [6]) which is a particular method of knowledge extraction. FCA deals with object-attribute data describing relationship between objects and attributes. In more detail: an input for FCA is a data table with rows corresponding to objects, columns corresponding to attributes, and table entries containing degrees to which objects have attributes. In its basic setting, FCA considers degrees 0 and 1 only, meaning that each object has/does not have an attribute. The output of FCA is a hierarchically ordered set of conceptual clusters extracted from the data. Since it is often the case that attributes are fuzzy rather than bivalent (attributes apply to objects to various degrees), one can consider an extension of FCA using arbitrary truth degrees from the real unit interval $[0, 1]$ as table entries, see [2], [3]. The output of FCA in this case is again a hierarchy of clusters. It can happen, however, that the hierarchy is infinite due to the fact that we have shifted from a finite scale $\{0, 1\}$ to the interval $[0, 1]$. This infinite scale is in fact an infinite solution to the clustering problem which we want to approximate by a finite one. Using general results proposed in this paper, we are able to replace $[0, 1]$ by a suitable finite scale $K \subseteq [0, 1]$ which, being used as a structure of truth degrees, produces a finite solution (hierarchy with finitely many clusters) which is computationally tractable and approximates well the infinite one.

Section II presents preliminaries. Section III present our approach and results. In Section IV, we outline some further issues connected to the present problem.

## II. Preliminaries

The basic concept in fuzzy logic is that of a structure of truth degrees which represents a set of truth degrees we use

1-4244-1214-5/07/$25.00 ©2007 IEEE

to describe graded truth of propositions (e.g., graded relationship of objects, graded properties of objects, similarities of values, etc.) plus logical connectives (e.g., conjunction, implications, ...) which are used to calculate truth degrees from other truth degrees. In this paper we are going to use so-called complete residuated lattices as our structures of truth degrees. This choice is general enough because complete residuated lattices include the popular t-norm-based structures of truth degrees as well as finite structures of truth degrees which we use to approximate the infinite ones. The rest of this section presents an introduction to the complete residuated lattices and derived notions we will need in the sequel. Further details can be found e.g. in [2], [7], [8], a good introduction to fuzzy logic and fuzzy sets is presented in [9].

A complete residuated lattice [8] is an algebra

$$\mathbf{L} = \langle L, \wedge, \vee, \otimes, \rightarrow, 0, 1\rangle \tag{1}$$

such that $\langle L, \wedge, \vee, 0, 1\rangle$ is a complete lattice with 0 and 1 being the least and greatest element of $L$, respectively; $\langle L, \otimes, 1\rangle$ is a commutative monoid (i.e. $\otimes$ is commutative, associative, and $a \otimes 1 = 1 \otimes a = a$ for each $a \in L$); $\otimes$ and $\rightarrow$ satisfy so-called adjointness property:

$$a \otimes b \leq c \quad \text{iff} \quad a \leq b \rightarrow c \tag{2}$$

for each $a, b, c \in L$. Elements $a$ of $L$ are called truth degrees. $\otimes$ and $\rightarrow$ are (truth functions of) "fuzzy conjunction" and "fuzzy implication". For each complete residuated lattice (1) we consider a derived (truth function of) logical connective $\leftrightarrow$ ("fuzzy equivalence") defined by $a \leftrightarrow b = (a \rightarrow b) \wedge (b \rightarrow a)$. By $a^n$ we denote $a \otimes \cdots \otimes a$ ($n$-times).

A common choice of $\mathbf{L}$ is a structure with $L = [0, 1]$ (unit interval), $\wedge$ and $\vee$ being minimum and maximum, $\otimes$ being a left-continuous t-norm with the corresponding $\rightarrow$. Three most important pairs of adjoint operations on the unit interval are:

$$\text{Łukasiewicz:} \quad \begin{aligned} a \otimes b &= \max(a + b - 1, 0), \\ a \rightarrow b &= \min(1 - a + b, 1), \end{aligned} \tag{3}$$

$$\text{Gödel:} \quad \begin{aligned} a \otimes b &= \min(a, b), \\ a \rightarrow b &= \begin{cases} 1 & \text{if } a \leq b, \\ b & \text{otherwise}, \end{cases} \end{aligned} \tag{4}$$

$$\text{Goguen (product):} \quad \begin{aligned} a \otimes b &= a \cdot b, \\ a \rightarrow b &= \begin{cases} 1 & \text{if } a \leq b, \\ \frac{b}{a} & \text{otherwise}. \end{cases} \end{aligned} \tag{5}$$

Complete residuated lattices on $[0, 1]$ given by (3), (4), and (5) are called standard Łukasiewicz, Gödel, Goguen (product) algebras, respectively. The class of complete residuated lattices include finite structures as well. For instance, one can put

$$L = \{a_0 = 0, a_1, \ldots, a_n = 1\} \subseteq [0, 1], \tag{6}$$

where $a_0 < \cdots < a_n$ and with $\otimes$ and $\rightarrow$ given by

$$a_k \otimes a_l = a_{\max(k+l-n,0)}, \tag{7}$$

$$a_k \rightarrow a_l = a_{\min(n-k+l,n)}. \tag{8}$$

Such an $\mathbf{L}$ is called a finite Łukasiewicz chain. If in addition $\{a_0, \ldots, a_n\} \subseteq [0, 1]$ are equidistant, in which case (7) and (8) are restrictions of the operations from (3), then $\mathbf{L}$ is called an equidistant Łukasiewicz chain. For instance,

$$\begin{aligned} &L_3 = \{0, 0.5, 1\}, \\ &L_4 = \{0, \tfrac{1}{3}, \tfrac{2}{3}, 1\}, \\ &L_5 = \{0, 0.25, 0.5, 0.75, 1\}, \ \ldots \end{aligned}$$

equipped with operations defined by (7) and (8) are equidistant Łukasiewicz chains. Another class of complete residuated lattices defined on finite subsets of $[0, 1]$ is the class of finite Gödel chains, where subsets of $L \subseteq [0, 1]$ are equipped with restrictions of Gödel operations (4) on $[0, 1]$ to $L$. A special case of a complete residuated lattice is the two-element Boolean algebra $\langle\{0, 1\}, \wedge, \vee, \otimes, \rightarrow, 0, 1\rangle$, denoted by $\mathbf{2}$, which is the structure of truth degrees of the classical logic. That is, the operations $\wedge, \vee, \otimes, \rightarrow$ of $\mathbf{2}$ are the truth functions (interpretations) of the corresponding logical connectives of the classical logic.

With $\mathbf{L}$ taken as a structure of truth degrees, we use the following notions: an $\mathbf{L}$-set (fuzzy set) $A$ in universe $U$ is a mapping $A \colon U \rightarrow L$, $A(u)$ being interpreted as "the degree to which $u$ belongs to $A$". Let $\mathbf{L}^U$ denote the collection of all $\mathbf{L}$-sets in $U$. Operations with $\mathbf{L}$-sets are defined componentwise. For instance, the intersection of $\mathbf{L}$-sets $A, B \in \mathbf{L}^U$ is an $\mathbf{L}$-set $A \cap B$ in $U$ such that $(A \cap B)(u) = A(u) \wedge B(u)$ for each $u \in U$, etc. Binary $\mathbf{L}$-relations (binary fuzzy relations) between $X$ and $Y$ can be thought of as $\mathbf{L}$-sets in the universe $X \times Y$. That is, a binary $\mathbf{L}$-relation $I \in \mathbf{L}^{X \times Y}$ between a set $X$ and a set $Y$ is a mapping assigning to each $x \in X$ and each $y \in Y$ a truth degree $I(x, y) \in L$ (a degree to which $x$ and $y$ are related by $I$). For any $A, B \in \mathbf{L}^U$, we define a similarity degree

$$A \approx B = \textstyle\bigwedge_{u \in U}\bigl(A(u) \leftrightarrow B(u)\bigr), \tag{9}$$

which expresses a degree to which fuzzy sets $A$ and $B$ are similar. In particular, we write $A = B$ if $A \approx B = 1$. We have $A = B$ (i.e., $A \approx B = 1$) iff $A(u) = B(u)$ ($u \in U$).

In the following we use well-known properties of residuated lattices and fuzzy structures which can be found in monographs [2], [8]. Throughout the rest of the paper, $\mathbf{L}$ denotes an arbitrary complete residuated lattice.

## III. Discretization in Formal Concept Analysis

### *A. Introduction to FCA*

We first present a brief account on formal concept analysis of data with fuzzy attributes, see e.g. [6] and [2], [3]. The input data to FCA consists of a data table describing a relationship between objects and attributes. The output of FCA consists of a hierarchically ordered collection of clusters. The clusters are called formal concepts and can be seen as natural concepts well-understandable and interpretable by humans. A data table with fuzzy attributes can be represented by a triplet $\langle X, Y, I\rangle$, called an $\mathbf{L}$-context, where $X$ and $Y$ are a non-empty sets of objects (table rows) and attributes (table columns), and $I$ :

$X \times Y \to L$ is an $\mathbf{L}$-relation with $I(x,y)$ representing the degree to which object $x \in X$ has attribute $y \in Y$ (table entry corresponding to row $x$ and column $y$). For $A \in \mathbf{L}^X$, $B \in \mathbf{L}^Y$ (i.e. $A$ is a fuzzy set of objects, $B$ is a fuzzy set of attributes), we define fuzzy sets $A^{\uparrow_I} \in \mathbf{L}^Y$ (fuzzy set of attributes), $B^{\downarrow_I} \in \mathbf{L}^X$ (fuzzy set of objects) by

$$A^{\uparrow_I}(y) = \bigwedge_{x\in X}\big(A(x) \to I(x,y)\big),$$
$$B^{\downarrow_I}(x) = \bigwedge_{y\in Y}\big(B(y) \to I(x,y)\big).$$

We put

$$\mathcal{B}(X,Y,I) = \{\langle A,B\rangle \in \mathbf{L}^X \times \mathbf{L}^Y \mid A^{\uparrow} = B,\ B^{\downarrow} = A\}$$

and define for $\langle A_1,B_1\rangle, \langle A_2,B_2\rangle \in \mathcal{B}(X,Y,I)$ a partial order $\leq$ by $\langle A_1,B_1\rangle \leq \langle A_2,B_2\rangle$ iff $A_1 \subseteq A_2$ (or, iff $B_2 \subseteq B_1$; both ways are equivalent). $\langle \mathcal{B}(X,Y,I), \leq\rangle$ is called a fuzzy concept lattice associated to $\langle X,Y,I\rangle$. Elements $\langle A,B\rangle$ of $\mathcal{B}(X,Y,I)$ are naturally interpreted as concepts (clusters) hidden in the input data represented by $I$. Namely, $A^{\uparrow_I} = B$ and $B^{\downarrow_I} = A$ say that $B$ is the collection of all attributes shared by all objects from $A$, and $A$ is the collection of all objects sharing all attributes from $B$. Note that these conditions represent exactly the definition of a concept as developed in the so-called Port-Royal logic; $A$ and $B$ are called the extent and the intent of the concept $\langle A,B\rangle$, respectively, and represent the collection of all objects and all attributes covered by the particular concept. Furthermore, $\leq$ models the natural subconcept-superconcept hierarchy—concept $\langle A_1,B_1\rangle$ is a subconcept of $\langle A_2,B_2\rangle$ iff each object from $A_1$ belongs to $A_2$ (dually for attributes).

We can see that formal concepts are just pairs $\langle A,B\rangle$ of fuzzy sets satisfying constraint $A^{\uparrow_I} = B$ and $B^{\downarrow_I} = A$. While in crisp case, i.e. $L = \{0,1\}$, $\mathcal{B}(X,Y,I)$ is finite and can be efficiently computed, it can be infinite in fuzzy setting, e.g. with $L = [0,1]$ with Łukasiewicz operations.

For two fuzzy concept lattices $\mathcal{B}_1 = \mathcal{B}(X,Y,I_1)$ and $\mathcal{B}_2 = \mathcal{B}(X,Y,I_2)$, we define a degree $\mathcal{B}_1 \approx_{\text{Ext}} \mathcal{B}_2$ to which $\mathcal{B}(X,Y,I_1)$ and $\mathcal{B}(X,Y,I_2)$ are similar via their extents by

$$\mathcal{B}_1 \approx_{\text{Ext}} \mathcal{B}_2 = \Big(\bigwedge_{\langle A_1,B_1\rangle\in\mathcal{B}_1} \bigvee_{\langle A_2,B_2\rangle\in\mathcal{B}_2} (A_1 \approx A_2)\Big) \wedge$$
$$\wedge \Big(\bigwedge_{\langle A_2,B_2\rangle\in\mathcal{B}_2} \bigvee_{\langle A_1,B_1\rangle\in\mathcal{B}_1} (A_1 \approx A_2)\Big).$$

In an analogous way, one can define a degree $\mathcal{B}_1 \approx_{\text{Int}} \mathcal{B}_2$ to which $\mathcal{B}(X,Y,I_1)$ and $\mathcal{B}(X,Y,I_2)$ are similar via their intents.

### *B. Replacing infinite structures of truth degrees by finite scales*

Suppose $\mathbf{L}_1 = \langle L_1, \wedge, \vee, \otimes, \to, 0, 1\rangle$ is a complete residuated lattice and $L_2 \subseteq L_1$. If $L_2$ is a nonempty subuniverse of $L_2$, i.e., if $L_2$ is nonempty and it is closed under all operations from $\mathbf{L}_1$, then $L_2$ equipped with restriction of the operations from $\mathbf{L}_1$ is a complete residuated lattice which is a subalgebra of $\mathbf{L}_1$. In such a case, we denote the subalgebra by $\mathbf{L}_2$. The new structure $\mathbf{L}_2$ can be seen as an approximation of the original structure of truth degrees $\mathbf{L}_1$. For example, an equidistant five-element Łukasiewicz chain (see Section II) can be seen as an approximation of the standard Łukasiewicz algebra defined on the real unit interval. Finite substructures seem to be good candidates for replacing infinite structures of truth degrees. One issue connected with moving from infinite structures to finite ones is our ability to estimate quality of approximation of the infinite structure by the finite one. Therefore, for a complete residuated lattice $\mathbf{L}_1$ and its subalgebra $\mathbf{L}_2$, we introduce the following degree of approximation:

$$\text{appr}(\mathbf{L}_1, \mathbf{L}_2) = \bigwedge_{a\in L_1} \bigvee_{b\in L_2} \big(a \leftrightarrow b\big). \qquad (10)$$

Using standard rules of fuzzy logic, one can see that formula (10) represents a degree to which it is true that "for each truth degree $a$ from $L_1$ there is a truth degree $b$ in $L_2$ which is equivalent to $a$". Thus, $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)$ can be understood as a degree to which $\mathbf{L}_2$ is a faithful approximation of $\mathbf{L}_1$.

*Example 1:* Consider the standard Łukasiewicz algebra (denote it by $\mathbf{L}_1$) and its equidistant substructure with $L_2 = \{0, 0.25, 0.5, 0.75, 1\}$. Then $\text{appr}(\mathbf{L}_1, \mathbf{L}_2) = 0.75$. More generally, for

$$L_2 = \{0 = \tfrac{0}{n}, \tfrac{1}{n}, \dots, \tfrac{n-1}{n}, \tfrac{n}{n} = 1\}, \qquad (11)$$

one can see that $\text{appr}(\mathbf{L}_1, \mathbf{L}_2) = 1 - \frac{1}{n}$.

Consider now a complete residuated lattice $\mathbf{L}_1$ and an $\mathbf{L}_1$-context $\langle X,Y,I_1\rangle$. If we choose a finite substructure $\mathbf{L}_2$ of $\mathbf{L}_1$, we might consider transforming $\langle X,Y,I_1\rangle$ into an $\mathbf{L}_2$-context $\langle X,Y,I_2\rangle$ so that the formal concepts present in $\langle X,Y,I_2\rangle$ are good approximations of the formal concepts presented in $\langle X,Y,I_1\rangle$. We can think of $\mathcal{B}(X,Y,I_1)$ as the original solution set which may be infinite and $\mathcal{B}(X,Y,I_2)$ as its finite approximation. The first step in the process is to find a suitable $\mathbf{L}_2$-context $\langle X,Y,I_2\rangle$. To achieve this goal, we define a mapping assigning to each truth degree from $L_1$ its approximation in $L_2$. For simplicity, we restrict ourselves to cases where both $\mathbf{L}_1$ and $\mathbf{L}_2$ are linearly ordered and $\mathbf{L}_2$ is finite only. This is sufficient for the goal of approximating solution sets over $[0,1]$ by solutions sets over finite subsets of $[0,1]$.

Define $\text{disc}(L_1, L_2)\colon L_1 \to L_2$ as a mapping satisfying

$$(\text{disc}(L_1,L_2))(a) = b \quad \text{iff} \quad a \leftrightarrow b = \bigvee_{c\in L_2}\big(a \leftrightarrow c\big). \qquad (12)$$

Since $L_2$ is a finite chain, there is always at least one mapping satisfying (12). Note that $\text{disc}(L_1, L_2)$ is not given uniquely in general. The mapping satisfying (12) will be called a *discretization function.* For a fuzzy set $A \in L_1^U$ we define a fuzzy set $(\text{disc}(L_1,L_2))(A) \in L_2^U$ by a componentwise application of $\text{disc}(L_1, L_2)$ as follows:

$$\big((\text{disc}(L_1,L_2))(A)\big)(u) = (\text{disc}(L_1,L_2))(A(u)). \qquad (13)$$

The definition (13) can also be introduced for binary fuzzy relations $A \in L_1^{U\times V}$.

Now, given $\langle X,Y,I_1\rangle$ where $I_1 : X \times Y \to L_1$ we can consider $I_2\colon X\times Y \to L_2$ which equals to $(\text{disc}(L_1,L_2))(I_1)$. Hence, $I_2$ is a discretization of $I_1$ which is induced by mapping $\text{disc}(L_1, L_2)$, see (12) and (13).

*Example 2:* Let $\mathbf{L}_1$ be the finite Łukasiewicz chain and consider an $\mathbf{L}_1$-context which is given by the table

| $I_1$ | $y_1$ | $y_2$ | $y_3$ |
|---|---|---|---|
| $x_1$ | 0.00 | 0.27 | 0.52 |
| $x_2$ | 0.56 | 1.00 | 0.68 |
| $x_3$ | 0.34 | 0.73 | 1.00 |

Furthermore, consider $L_2 = \{0, 0.25, 0.5, 0.75, 1\}$ and a discretization function $\mathrm{disc}(L_1, L_2)$ defined by

$$(\mathrm{disc}(L_1, L_2))(a) = \begin{cases} 0 & \text{if } a \in [0, 0.125), \\ 0.25 & \text{if } a \in [0.125, 0.375), \\ 0.5 & \text{if } a \in [0.375, 0.625), \\ 0.75 & \text{if } a \in [0.625, 0.875), \\ 1 & \text{if } a \in [0.875, 1]. \end{cases}$$

Then, the induced $\mathbf{L}_2$-context $I_2 = (\mathrm{disc}(L_1, L_2))(I_1)$ will be the following:

| $I_2$ | $y_1$ | $y_2$ | $y_3$ |
|---|---|---|---|
| $x_1$ | 0.00 | 0.25 | 0.50 |
| $x_2$ | 0.50 | 1.00 | 0.75 |
| $x_3$ | 0.25 | 0.75 | 1.00 |

Intuitively, if the discretization of the context looks similar to the original context (i.e., the truth degrees contained in the new context are similar to the corresponding degrees from the original table), then the two contexts should induce similar concepts. This is indeed so, as we are going to show in the sequel.

The following assertion says that every fuzzy set is similar to its discretization at least to a degree given by (10).

*Lemma 1:* Let $\mathbf{L}_1$ and $\mathbf{L}_2$ be linear complete residuated lattices such that $\mathbf{L}_2$ is a finite substructure of $\mathbf{L}_1$. Then, for each $A_1 \in L_1^U$, we have:

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) \leq A_1 \approx (\mathrm{disc}(L_1, L_2))(A_1). \tag{14}$$

*Proof:* We show that for each $u \in U$,

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) \leq A_1(u) \leftrightarrow ((\mathrm{disc}(L_1, L_2))(A_1))(u).$$

Using (10), (12) and (13), we get

$$\begin{aligned} \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) &\leq \textstyle\bigvee_{c \in L_2} \big(A_1(u) \leftrightarrow c\big) = \\ &= A_1(u) \leftrightarrow ((\mathrm{disc}(L_1, L_2))(A_1(u)) = \\ &= A_1(u) \leftrightarrow ((\mathrm{disc}(L_1, L_2))(A_1))(u). \end{aligned}$$

which finishes the proof. ■

The following assertion shows that the derivation operators ${}^{\downarrow}, {}^{\uparrow}$ used in discretized contexts with discretized fuzzy sets of objects or attributes yield similar results as the original operators in the original contexts. The similarity of results is bounded from below by the degree of the approximation of the original structure of truth degrees by the finite one.

*Lemma 2:* Let $\mathbf{L}_1$ and $\mathbf{L}_2$ be linear complete residuated lattices such that $\mathbf{L}_2$ is a finite substructure of $\mathbf{L}_1$. Moreover, let $\langle X, Y, I_1 \rangle$ be an $\mathbf{L}_1$-context and $\langle X, Y, I_2 \rangle$ be an $\mathbf{L}_2$-context such that $I_2 = (\mathrm{disc}(L_1, L_2))(I_1)$. Then, for each $A_1 \in L_1^X$, $B_1 \in L_1^Y$, we have

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq A_1^{\uparrow_{I_1}} \approx (\mathrm{disc}(L_1, L_2))(A_1)^{\uparrow_{I_2}}, \tag{15}$$

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq B_1^{\downarrow_{I_1}} \approx (\mathrm{disc}(L_1, L_2))(B_1)^{\downarrow_{I_2}}, \tag{16}$$

where $\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2$ denotes $\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)$.

*Proof:* Due to the limited scope of this paper, we present only a sketch of the proof. The full version of the proof is postponed to the full version of the paper.

We prove only (15) since the proof of (16) is symmetrical. Denote $\big(\mathrm{disc}(L_1, L_2)\big)(A_1)$ by $A_2$. Using adjointness and properties of residuated lattices, one can show that (15) is true iff for each $x \in X$ and $y \in Y$, the following inequalities are satisfied:

$$A_2(x) \otimes A_1^{\uparrow_{I_1}}(y) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq I_2(x, y), \tag{17}$$

$$A_1(x) \otimes A_2^{\uparrow_{I_2}}(y) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq I_1(x, y). \tag{18}$$

In order to prove (17), observe that

$$\begin{aligned} &A_2(x) \otimes A_1^{\uparrow_{I_1}}(y) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \\ &\leq A_2(x) \otimes (A_1(x) \to I_1(x, y)) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2. \end{aligned}$$

Using Lemma 1, we get

$$\begin{aligned} \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) &\leq A_2(x) \to A_1(x), \\ \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) &\leq I_1(x, y) \to I_2(x, y), \end{aligned}$$

i.e. by adjointness,

$$\begin{aligned} A_2(x) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) &\leq A_1(x), \\ I_1(x, y) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) &\leq I_2(x, y), \end{aligned}$$

from which we get

$$\begin{aligned} &A_2(x) \otimes (A_1(x) \to I_1(x, y)) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \\ &\leq A_1(x) \otimes (A_1(x) \to I_1(x, y)) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) \leq \\ &\leq I_1(x, y) \otimes \mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2) \leq I_2(x, y), \end{aligned}$$

which proves (17); the proof of (18) is symmetrical. ■

The following theorem shows to what degree the concepts present in the discretized data are similar to the concepts present in the original data.

*Theorem 1:* Let $\mathbf{L}_1$ and $\mathbf{L}_2$ be linear complete residuated lattices such that $\mathbf{L}_2$ is a finite substructure of $\mathbf{L}_1$. Moreover, let $\langle X, Y, I_1 \rangle$ be an $\mathbf{L}_1$-context and $\langle X, Y, I_2 \rangle$ be an $\mathbf{L}_2$-context such that $I_2 = (\mathrm{disc}(L_1, L_2))(I_1)$. Then,

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \mathcal{B}(X, Y, I_1) \approx_{\mathrm{Ext}} \mathcal{B}(X, Y, I_2), \tag{19}$$

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \mathcal{B}(X, Y, I_1) \approx_{\mathrm{Int}} \mathcal{B}(X, Y, I_2). \tag{20}$$

*Proof:* We present a sketch of the proof only. We focus on proving (19) because (20) will then be a consequence of results from [1]. Denote by $E_1$ and $E_2$ the sets of all extents of $\mathcal{B}(X, Y, I_1)$ and $\mathcal{B}(X, Y, I_2)$, respectively. It suffices to check

$$\mathrm{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \textstyle\bigwedge_{A_1 \in E_1} \bigvee_{A_2 \in E_2} A_1 \approx A_2, \tag{21}$$

and

$$\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \bigwedge_{A_2 \in E_2} \bigvee_{A_1 \in E_1} A_1 \approx A_2. \qquad (22)$$

The inequality (21) is true iff, for each $A_1 \in E_1$,

$$\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq \bigvee_{A_2 \in E_2} A_1 \approx A_2. \qquad (23)$$

Using Lemma 2, for $A_2 = \big((\text{disc}(L_1, L_2))(A_1^{\uparrow_{I_1}})\big)^{\downarrow_{I_2}}$, we get

$$\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq A_1^{\uparrow_{I_1}\downarrow_{I_1}} \approx \big((\text{disc}(L_1, L_2))(A_1^{\uparrow_{I_1}})\big)^{\downarrow_{I_2}},$$

which proves (23) because $A_1 = A_1^{\uparrow_{I_1}\downarrow_{I_1}}$ and $A_2 \in E_2$, i.e. the inequality (21) is true.

We now prove (22) by showing that for each $A_2 \in E_2$ there is $A_1 \in E_1$ such that $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq A_1 \approx A_2$. Take $A_2 \in E_2$ and put $A_1 = A_2^{\uparrow_{I_2}\downarrow_{I_1}}$. Lemma 2 yields

$$\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq A_1 \approx \big((\text{disc}(L_1, L_2))(A_2^{\uparrow_{I_2}})\big)^{\downarrow_{I_2}}.$$

Observe that for each $a \in L_2$, we have $(\text{disc}(L_1, L_2))(a) = a$. Therefore, $(\text{disc}(L_1, L_2))(A_2^{\uparrow_{I_2}}) = A_2^{\uparrow_{I_2}}$. Since $A_2 = A_2^{\uparrow_{I_2}\downarrow_{I_2}}$, we get $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \leq A_1 \approx A_2^{\uparrow_{I_2}\downarrow_{I_2}} = A_1 \approx A_2$. In addition to that, $A_1 \in E_1$, showing that (22) is true. ■

*Remark 1:* Let us see what the foregoing results say. Technically, they provide estimations of similarity degrees on the right hand side in terms of the degree $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)$ of approximation of $\mathbf{L}_1$ by $\mathbf{L}_2$. For instance, Theorem 1 says that when we use $\mathbf{L}_2$ instead of $\mathbf{L}_1$ and transform $\langle X, Y, I_1 \rangle$ to $\langle X, Y, I_2 \rangle$, then the degree

$$\mathcal{B}(X, Y, I_1) \approx_{\text{Ext}} \mathcal{B}(X, Y, I_2)$$

to which $\mathcal{B}(X, Y, I_1)$ is similar to $\mathcal{B}(X, Y, I_2)$ is at least $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2$. Two aspects of a result of this type need to be mentioned. First, computing the estimation $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2$ is easy. When devising $\mathbf{L}_2$, computing the estimation $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2$ enables us to see how well $\mathcal{B}(X, Y, I_1)$ is approximated by $\mathcal{B}(X, Y, I_2)$. Second, suppose we want to see what kind of approximation we need to use in order to have $\mathcal{B}(X, Y, I_1) \approx_{\text{Ext}} \mathcal{B}(X, Y, I_2)$ at least as high as a prescribed level $a$ of similarity. Then the result tells us that we need to choose $\mathbf{L}_2$ such that $\text{appr}(\mathbf{L}_1, \mathbf{L}_2)^2 \geq a$, i.e. we know how fine the discretization $\mathbf{L}_2$ of $\mathbf{L}_1$ needs to be.

*Example 3:* If $\mathbf{L}_1$ is the standard Łukasiewicz algebra and $\mathbf{L}_2$ is its equidistant substructure with $L_2$ being (11) then in order to approximate the original concepts (computed using $\mathbf{L}_1$) at least to degree $0 \leq a < 1$, we need to take $\mathbf{L}_2$ so that

$$n \geq -\frac{2}{a-1},$$

e.g., for $a = 0.9$, we must take $n \geq 20$, i.e. $\mathbf{L}_2$ must contain 21 truth degrees to achieve the desired logical precision.

*Remark 2:* Using a finite scale instead of the infinite one is beneficial not only from the computational point of view. The discretization of structures of truth degrees can also be seen as a way of reducing the size of concept lattices. Concept lattices generated from contexts using t-norm based structures of truth degrees are usually too large or even infinite. Therefore, there is an effort to reduce the generated concepts lattices in that they contain just some (interesting) concepts (see, e.g., [4], [5]). Concept lattices generated using large scales of truth degrees often contain a large number of concepts which are similar to high degrees so that they are virtually indistinguishable for users. A reasonable choice of a finite scale of truth degrees can reduce the vast amount of concepts to a few representatives only. This is illustrated by the next example.

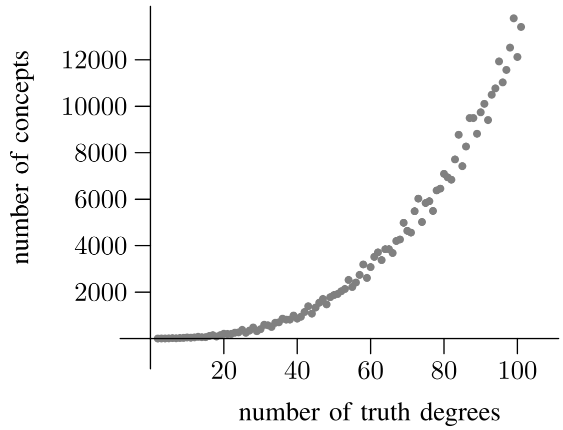


Fig. 1. Dependency between the size of the discrete scale of truth degrees and the number of extracted concepts

*Example 4:* Let $\mathbf{L}_1$ be the finite Łukasiewicz chain and consider again the $\mathbf{L}_1$-context

| $I_1$ | $y_1$ | $y_2$ | $y_3$ |
|---|---|---|---|
| $x_1$ | 0.00 | 0.27 | 0.52 |
| $x_2$ | 0.56 | 1.00 | 0.68 |
| $x_3$ | 0.34 | 0.73 | 1.00 |

If we take $L_2 = \{\frac{n}{100} \,|\, 0 \leq n \leq 100\}$ equipped with Łukasiewicz operations, then $I_1 = (\text{disc}(L_1, L_2))(I_1)$, i.e. the discretized version of $I_1$, will coincide with $I_1$ itself. The concept lattice generated using $\mathbf{L}_2$ contains 13415 concepts which can be seen as not natural because the data table contains just three objects and three attributes. With smaller equidistant scales, we obtain smaller concept lattices generated from the data. The situation for $I_1$ and its discretizations using 2 up to 101 truth degrees is depicted in Fig. 1.

## IV. General Approach and Further Issues

The present approach, which we illustrated on the case of formal concept analysis of data with fuzzy attributes, can be obviously generalized to other situations. A general framework, to which we will generalize our results in future work, is that of a predicate fuzzy logic with the assumption that our constraints are expressed by first-order formulas. Our present example and other examples then become a particular case of this general framework.

Moreover, the present idea of approximating a large (possibly infinite) scale $L$ of truth degrees by a smaller set $K$

leads to related problems such as the one we are now going to outline. Suppose $A : X \to L$ is a fuzzy set taking values in $L$ such that $M = \{A(x) \,|\, x \in X\}$ is the set of all degrees "used by $A$". One might wish to replace $A$ by a different fuzzy set $B$ which approximates $A$ well and "uses" as small a set $K$ of truth degrees as possible. Denote $K = \{B(x) \,|\, x \in X\}$. Then, given a similarity threshold $e$, our problem is to find $B$ with

$$\text{appr}(M, K) \geq e. \tag{24}$$

such that $K$ is minimal in terms of the number of its elements. Here, appr$(M, K)$ is defined as earlier in our paper, see (10). A feasible approach to solve this problem is the following. Denote by $\sim_e$ the tolerance on $L$ defined by

$$a \sim_e b \quad \text{iff} \quad a \leftrightarrow b \geq e \otimes e$$

and for any maximal block $B \subseteq L$ of $\sim_e$ set

$$c(B) = e \to \bigwedge B.$$

One can see that $\bigwedge B \in B$. Finally, for any $c \in L$ set

$$B(c) = \{a \in L \mid e \leq a \leftrightarrow c\}.$$

Then, one can show that for any maximal block $B \subseteq L$ of $\sim_e$ it holds $B = B(c(B))$, i.e. elements of maximal block $B$ can be approximated by $c(B)$ with precision given by the threshold $e$. Moreover, one can prove the following theorem (its proof will be presented in the full version of our paper):

*Theorem 2:* Let $\Omega \subseteq L/{\sim_e}$ be a set of maximal blocks of the tolerance $\sim_e$ such that $M \subseteq \bigcup\Omega$ and $K = \{c(B) \mid B \in \Omega\}$. Then (24) is satisfied. Conversely, if $K \subseteq L$ satisfies (24) and for any $a \in K$, $B(a)$ is a maximal block of $\sim_e$ then there is a covering $\Omega \subseteq L/{\sim_e}$ of $M$ such that $K = \{c(B) \mid B \in \Omega\}$.

The theorem tells how to find set $K$ in terms of maximal blocks of of $\sim_e$ and shows universality of such approach. Moreover, an algorithm can be devised which finds a required $K$ for a given $M$. Details of the just described method will be presented in our next paper.

## Acknowledgment

Supported by grant No. 1ET101370417 of GA AV ČR, by grant No. 201/05/0079 of the Czech Science Foundation, and by institutional support, research plan MSM 6198959214.

## References


[1] Belohlavek R.: Similarity relations in concept lattices. *J. Logic Comput.* 10(6):823–845, 2000.

[2] Belohlavek R.: *Fuzzy Relational Systems: Foundations and Principles*. Kluwer, Academic/Plenum Publishers, New York, 2002.

[3] Belohlavek R.: Concept lattices and order in fuzzy logic. *Annals of Pure and Applied Logic* **128**(1–3)(2004), 277–298.

[4] Belohlavek R., Vychodil V.: Reducing the size of fuzzy concept lattices by hedges. In: FUZZ-IEEE 2005, The IEEE International Conference on Fuzzy Systems, May 22–25, 2005, Reno (Nevada, USA), pp. 663–668.

[5] Belohlavek R., Vychodil V.: Reducing the size of fuzzy concept lattices by fuzzy closure operators. Proc. SCIS & ISIS 2006, pp. 309–314. Tokyo Institute of Technology, Japan, Sep 20–24, 2006, ISSN 1880–3741.

[6] Ganter B., Wille R.: *Formal Concept Analysis. Mathematical Foundations.* Springer, Berlin, 1999.

[7] Gerla G.: *Fuzzy Logic. Mathematical Tools for Approximate Reasoning.* Kluwer, Dordrecht, 2001.

[8] Hájek P.: *Metamathematics of Fuzzy Logic.* Kluwer, Dordrecht, 1998.

[9] Klir G. J., Yuan B.: *Fuzzy Sets and Fuzzy Logic. Theory and Applications.* Prentice-Hall, 1995.