

# Concepts in Fuzzy Scaling Theory: Order and Granularity

Karl Erich Wolff

Fachbereich Mathematik und Naturwissenschaften, Fachhochschule Darmstadt,  
Schöfferstr. 3, D-64295 Darmstadt, Germany  
Phone: +49-6151-16-8677, Fax: +49-6151-16-8975  
email: [wolff@mathematik.tu-darmstadt.de](mailto:wolff@mathematik.tu-darmstadt.de)

**ABSTRACT:** This paper explains mainly by examples the mathematical foundation of Fuzzy Scaling Theory described in Wolff (1998). It illustrates the solution of fundamental problems in Fuzzy Theory, especially the problem of the pragmatic meaning of ordered sets as logics, the problem of the meaning of Fuzzy implications and the problem of the construction of direct products of linguistic variables. The key to the solution of these problems is the description of membership functions by cut contexts and of linguistic variables by conceptual scales. This renders the translation of Conceptual Scaling Theory into Fuzzy Scaling Theory and the generalization of the latter to L-Fuzzy Scaling Theory for an arbitrary ordered set  $L$  as logic.

This paper bridges the gap between two scientific communities both using "graded concepts" as tools to handle granularity: it is shown that L-Fuzzy Scaling Theory is equivalent to Conceptual Scaling Theory.

**KEYWORDS:** Fuzzy Scaling Theory, linguistic variables, Formal Concept Analysis, Fuzzy implications

## 1 INTRODUCTION: Problems in Fuzzy Theory

Fuzzy Theory introduced by Zadeh (1965) got its huge influence by the very convincing idea of representing "graded concepts" by "a theory in which everything is a matter of degree, or to put it figuratively, everything has elasticity" (Zadeh in the foreword of Zimmermann 1994). The formal description of this "elasticity" replaces the classical "crisp" sets by Fuzzy sets, represented by membership functions mapping each object of a set  $X$  into the real unit interval  $[0,1]$ . From a mathematical point of view the definition of a membership function can be generalized by replacing the unit interval by more general structures like L-Fuzzy algebras in the sense of Wechler (1978) or complete lattices with a semigroup structure (Goguen 1967). Omitting the algebraic structure one can clearly choose even an arbitrary ordered set as the "Fuzzy logic" of the membership functions.

But these generalizations did not become really valuable for applications. Why? I believe that the meaning of these general Fuzzy logics in practice is not understood:

- How to apply a general Fuzzy logic in practice?
- Is there a canonical Fuzzy logic for all applications?
- What about "elasticity" in discrete logics?

But also the meaning of classical Fuzzy logic, the unit interval  $[0,1]$  is not well understood:

- How to define the conceptual meaning of Fuzzy implications of the form "If  $X$  is  $A$ , then  $Y$  is  $B$ " ?
- What is the formal role of objects and measurement values in Fuzzy Theory?
- What is the right membership value of a missing value of some measurement?

The reader should be able to solve these problems after having studied the basic concepts and examples in this paper and the mathematical foundation of it in Wolff (1998).

## 2 CLASSICAL MEMBERSHIP FUNCTIONS: The unit interval is too narrow

It is obvious that membership functions play a crucial role in Fuzzy Theory. Why? What is the central idea behind the technical definition of a membership function? A classical membership function  $f$  maps an arbitrary crisp set  $X$  into the unit interval,  $f: X \rightarrow [0,1]$ . The elements of  $X$  are usually interpreted as objects, but sometimes as measurement values of objects which will be discussed later. The unit interval is used not just as a set without structure, but mainly as an ordered set which allows to represent the "gradedness" and "elasticity" of concepts. In this paper we concentrate only on this ordinal point of view and do not discuss the algebraic structure of the unit interval.

The advantage of the unit interval lies in the fact that any chain can be order embedded in  $[0,1]$ , its disadvantage is that *only* chains can be represented by an order embedding in  $[0,1]$ .

The following example concerning the fuzzy construct "confidence for investors" shows a typical multidimensional structure, far away from the simple idea of a chain from low to high confidence even though (or better: since) it is constructed from many chains.

### Example 1: Confidence for investors

This example is taken from Jambu (1991, p.236, 450):

"With a view to advising their customers, a group of banks asked some economists to study the risk of their investment abroad; 43 countries were selected according to 15 confidence criteria. Each of the economists gave a mark between 0 (the maximum of risk) and 4 (the minimum of risk)."

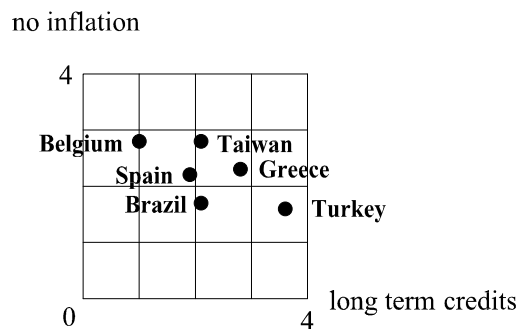
For this paper we study all 43 countries but only 5 confidence criteria. The mean values of the marks are shown in the following Table 1 for only 6 countries.

Table 1: Confidence marks

	long term credits	marketing services	economical increasing	no inflation	political stability
Belgium	1.0	<b>3.3</b>	<b>2.6</b>	<b>2.8</b>	<b>3.4</b>
Brazil	<b>2.1</b>	<b>2.4</b>	<b>3.2</b>	1.7	<b>2.8</b>
Greece	2.8	2.3	2.0	<b>2.3</b>	<b>2.1</b>
Spain	1.9	2.0	<b>2.6</b>	<b>2.2</b>	<b>2.4</b>
Taiwan	<b>2.1</b>	<b>2.8</b>	<b>2.7</b>	<b>2.8</b>	2.0
Turkey	<b>3.6</b>	1.9	2.0	1.6	1.9

By construction the values of each confidence criterion can be arranged in a chain from 0 to 4. Hence each country can be represented as a point in the direct product of 5 chains, each chain a copy of  $[0,4]$ . To visualize a 2-dimensional projection of this 5-dimensional space we draw the usual scatterplot in the direct product  $[0,4] \times [0,4]$  and embed the 6 countries of Table 1 at the places of their value pairs for "long term credits" and "no inflation", say.

Figure 1: A scatterplot



Such direct products of two (or more) chains are the standard representation frame for metric and for ordinal data – but they are no longer chains and therefore not order embeddable into a chain. We will see later that this is the central obstacle for the development of classical Fuzzy Theory since it prohibits the construction of the direct product of two linguistic variables as long as the logics are restricted to be chains.

Very similar but much more general than this classical representation of objects in direct product of chains is the conceptual representation of objects in direct products of concept lattices of conceptual scales.

### 3 FORMAL CONCEPT ANALYSIS: Concepts and Order

Formal Concept Analysis (FCA) was introduced by Wille (1982). The interested reader is also referred to Ganter, Wille (1999), Wille (1996) and, for a short introduction, to Wolff (1994). The basic tool is the notion of a *formal context*  $(G,M,I)$  where  $I$  is a binary relation between a set  $G$  (of "objects") and a set  $M$  of "attributes", i.e.  $I \subseteq G \times M$ . A formal context can be represented by its *concept lattice*. Using conceptual scaling techniques data tables with arbitrary values can be represented in concept lattices of suitable granularity.

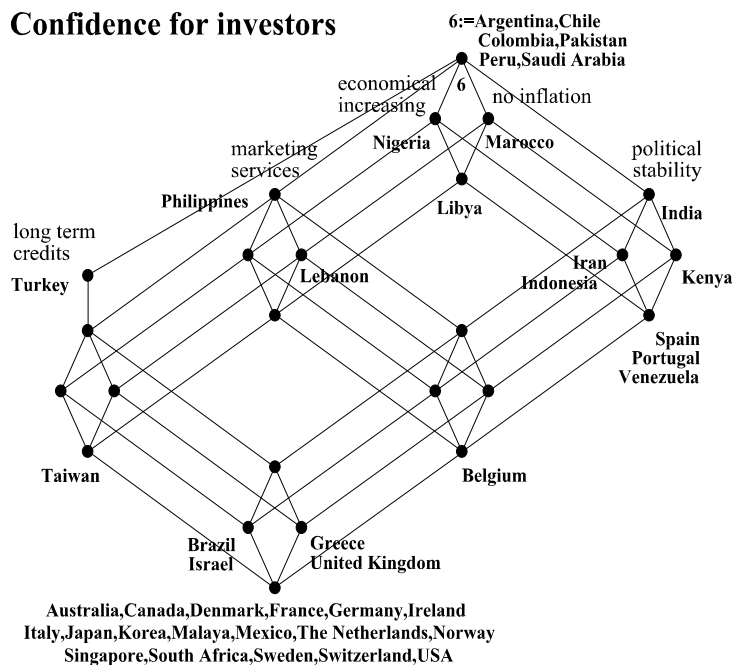
Now we demonstrate a typical application of FCA using a very simple conceptual scaling procedure in the example above. At first we study the distribution of all 43 countries over the five confidence criteria of Table 1 looking very roughly only whether a value is greater than 2 (printed **bold** in Table 1). This yields a 43x5-table from which only the 6 rows corresponding to Table 1 are shown in Table 2.

Table 2:

	long term credits	marketing services	economical increasing	no inflation	political stability
Belgium		×	×	×	×
Brazil	×	×	×		×
Greece				×	×
Spain			×	×	×
Taiwan	×	×	×	×	
Turkey	×				

The whole 43x5-table can be described as a formal context  $\mathbf{K} = (G,M,I)$  where  $G$  is the set of all 43 chosen countries,  $M$  is the set of the 5 confidence criteria in Table 2 and the incidence relation  $I$  is defined by  $g I m : \Leftrightarrow m(g) > 2$ , where  $m(g)$  denotes the value of the confidence criterion  $m$  at the country  $g$ . The information of this formal context  $\mathbf{K}$  is represented in the line diagram in Figure 2.

Figure 2:



How to read this line diagram? First of all, it does not represent an undirected graph but a directed one, more precisely a conceptual hierarchy, called the concept lattice  $\mathbf{B}(\mathbf{K})$  of the given formal context  $\mathbf{K}$ . The elements of the concept lattice are called (formal) concepts. They are represented graphically by small circles. A formal concept of the context  $(G,M,I)$  is by definition a pair  $(A,B)$  where  $A \subseteq G$  and  $B \subseteq M$ , such that  $A = B^\downarrow$ , the set of all objects having all attributes of  $B$ ,

and  $B = A^\uparrow$ , the set of all attributes which hold for all objects of  $A$ .  $A$  is called the extent and  $B$  the intent of  $(A, B)$ . The hierarchy between concepts is defined by the set inclusion of their extents:

$(A_1, B_1) \leq (A_2, B_2) : \Leftrightarrow A_1 \subseteq A_2$  (which is equivalent to  $B_2 \subseteq B_1$ ).

A concept  $c$  is called a lower neighbour of a concept  $d$ , iff  $c < d$  and there is no concept  $x \in \mathbf{B}(\mathbf{K})$  such that  $c < x < d$ . The lower neighbourhood relation is represented graphically by upwards leading lines. Hence upward leading paths represent the hierarchy of concepts.

The bottom circle in Figure 2 represents the smallest concept  $(M^\downarrow, M)$ . In our example  $M^\downarrow$  is the set of those countries fulfilling all 5 confidence criteria. The top circle represents the biggest concept  $(G, \emptyset)$  consisting of all objects and (in this example) the empty set. The labelling of the line diagram is explained now.

Each object  $g$  determines the concept

$\gamma g := (\{g\}^\uparrow, \{g\}^\downarrow)$ , called the *object concept of g*,

while for each attribute  $m$  the concept

$\mu m := (\{m\}^\downarrow, \{m\}^\uparrow)$  is called the *attribute concept of m*.

In a line diagram the circle of an object concept  $c$  is labelled by the names of those objects having  $c$  as object concept. Objects having the same object concept are called equivalent, they have exactly the same attributes. The classes of this equivalence relation are called the *(object) contingency classes*. Analogously the circles of the attribute concepts are labelled by the names of equivalent attributes.

The reading rule for line diagrams says: An object  $g$  has an attribute  $m$  if and only if there is an upwards leading path from the circle of  $g$  to the circle of  $m$ . Example: From the circle of Spain (Portugal and Venezuela) there are upwards leading paths to the circles of "economical increasing", "no inflation", "political stability", but not to the circles of the other attributes. Hence the formal context is represented without any loss of information in the line diagram.

Finally we mention that concept lattices are (up to isomorphism) just the complete lattices (where any subset has an infimum and a supremum), the main examples are all finite lattices and all closed real intervals, especially the unit interval.

Now we are ready to discuss the central theme in all knowledge representation theories, namely granularity, restricted to its formal representation in Fuzzy Theory and FCA.

## 4 SCALING THEORY: Working with granularity

With respect to a given purpose we have to choose an appropriate granularity of the structure under consideration. This leads to granularity tools in all theories. Scaling Theory is understood here as the theory of granularity. The central aim is the development of flexible tools for the generation of purpose oriented granularity.

In this section we compare both Conceptual and Fuzzy Scaling Theory. The main granularity generating tools in these theories are the conceptual scales and the linguistic variables.

### 4.1 CONCEPTUAL SCALING THEORY

Conceptual Scaling Theory was developed by Ganter and Wille (1989). The general process in conceptual scaling starts with the representation of knowledge in a data table with arbitrary values and possibly missing values. These data tables are formally described by *many-valued contexts*  $(G, M, W, I)$ , where  $G$  is a set of "objects",  $M$  is a set of "many-valued attributes",  $W$  is a set of "values" and  $I$  is a ternary relation,  $I \subseteq G \times M \times W$ , such that for any  $g \in G$ ,  $m \in M$  there is at most one value  $w$  satisfying  $(g, m, w) \in I$ . Therefore a many-valued attribute  $m$  is usually interpreted as a (partial) measurement function and we write  $m(g) = w$  iff  $(g, m, w) \in I$ . An example is given in Table 1.

The central granularity choosing process in conceptual scaling theory is the construction of a formal context  $\mathbf{S}_m = (W_m, M_m, I_m)$  for each  $m \in M$  such that  $W_m \supseteq mG := \{m(g) \mid g \in G\}$ . Such formal contexts, called *conceptual scales*, represent a contextual language about the set of values of  $m$ . Usually one chooses  $W_m$  as the set of all "possible" values of  $m$  with respect to some purpose. Each attribute  $n \in M_m$  is called a scale attribute. The set  $n^\downarrow = \{w \mid w I_m n\}$  is the extent of the attribute concept of  $n$  in the scale  $\mathbf{S}_m$ . Hence the choice of a scale induces a selection of subsets of  $W_m$  - describing the granularity of the contextual language about the possible values. The set of all intersections of these subsets constitute just the closure system of all extents of the concept lattice of  $\mathbf{S}_m$ .

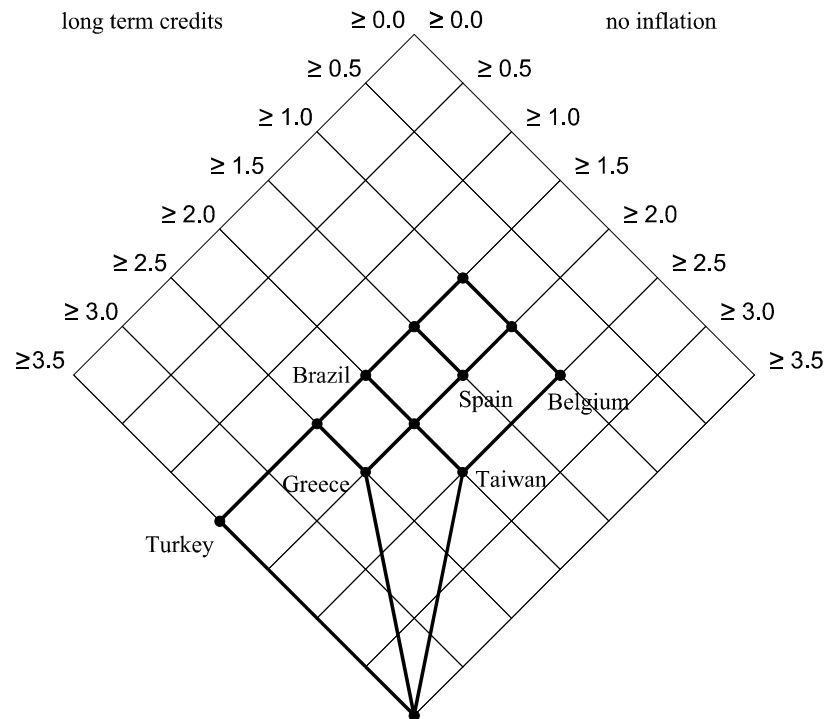
The granularity of the language about the possible values of  $m$  induces in a natural way a granularity on the set  $G$  of objects of the given many-valued context since each object  $g$  is mapped via  $m$  onto its value  $m(g)$  and  $m(g)$  is mapped via the object concept mapping  $\gamma_m$  of  $\mathbf{S}_m$  onto  $\gamma_m(m(g))$ :  $g \rightarrow m(g) \rightarrow \gamma_m(m(g))$ .

Hence the set of all object concepts of  $\mathbf{S}_m$  plays the role of a frame into which each object of  $G$  can be embedded. For two attributes  $m, m' \in M$  each object  $g$  can be mapped onto the corresponding pairs:

$g \rightarrow (m(g), m'(g)) \rightarrow (\gamma_m(m(g)), \gamma_{m'}(m'(g))) \in \mathbf{B}(\mathbf{S}_m) \times \mathbf{B}(\mathbf{S}_{m'})$ .

As a simple example we scale each of the two many-valued attributes used in Figure 2 by a scale whose concept lattice is a chain of length 8 and get the line diagram in Figure 3.

Figure 3:



Reading example: Brazil has the attribute (long term credits,  $\geq 2.0$ ) but not the attribute (long term credits,  $\geq 2.5$ ) reflecting the fact that the original value (2.1) for long term credits for Brazil satisfies  $2.0 \leq 2.1 < 2.5$ .

This embedding of objects into the direct product of the concept lattices of two (or more) scales (not necessarily chains!) is the fundamental conceptual scaling technique.

The contextual connection between the objects of the given many-valued context and the scale attributes of the scales  $S_m$  ( $m \in M$ ) is described in the *derived context*  $K = (G, \{(m,n) \mid m \in M, n \in M_m\}, J)$  defined by

$$g J (m,n) :\Leftrightarrow m(g) I_m n \quad (\text{for } g \in G, m \in M, n \in M_m).$$

In the example above, Brazil has the attribute (long term credits,  $\geq 2.0$ ) since the value "long term credits(Brazil)" has the scale attribute " $\geq 2.0$ ".

The derived context has the property that its concept lattice  $B(K)$  can be (supremum-)embedded into the direct product of the concept lattices of the chosen scales. Figure 3 shows the supremum-embedded concept lattice in the direct product of two chains. This is the discrete analogon of the scatterplot in Figure 1. The corresponding theory is described in Ganter, Wille (1989, 1999).

Another typical example is the concept lattice in Figure 2. Each many-valued attribute  $m$  of the many-valued context in Figure 1 is scaled by the scale  $S_m = ([0,4], \{n\}, I_m)$ , where  $w I_m n :\Leftrightarrow w > 2$  (for  $w \in [0,4]$ ). Hence the extent of the only scale attribute is the set  $\{w \in [0,4] \mid w > 2\}$ . Therefore  $B(S_m)$  is a chain with exactly 2 concepts. The direct product of these 5 chains is a lattice which is isomorphic to the 5-dimensional Boolean lattice with 32 elements. The derived context for these scales is the formal context  $K$  represented by the line diagram of Figure 2.

#### 4.2 CLASSICAL FUZZY SCALING THEORY: Linguistic variables

Classical Fuzzy Theory has a powerful tool to manage granularity, namely linguistic variables, introduced by Zadeh (1975):

“By a *linguistic variable* we mean a variable whose values are words or sentences in a natural or artificial language. For example, *Age* is a linguistic variable if its values are linguistic rather than numerical, i.e., *young, not young, very young, quite young, old, not very old and not very young*, etc., rather than 20, 21, 22, 23,....”

For the description of linguistic values Fuzzy sets are used, formally described as membership functions. The set of all membership functions (or the set of all Fuzzy sets) on a set  $X$  is denoted by  $F(X, [0,1])$ . Hence a classical linguistic variable can be defined as follows:

**Definition:**

A *classical linguistic variable* is a quintupel  $(X, V, \mu, [0,1], \leq)$ , where  $X$  is a set (called the ‘*domain*’),  $V$  is a set (of ‘*linguistic values*’),  $([0,1], \leq)$  is the real unit interval with its usual ordering and  $\mu$  is a mapping  $\mu: V \rightarrow F(X, [0,1])$  which represents each linguistic value  $v$  by a membership function  $\mu_v := \mu(v)$  on  $X$ .

To understand the notion of a linguistic variable from a conceptual point of view we study the fundamental connection between Fuzzy sets and crisp sets – namely  $\alpha$ -cuts:

**Definition:**

For  $f \in F(X, [0,1])$  and  $\alpha \in [0,1]$  the crisp set  $f_\alpha := \{x \in X \mid f(x) \geq \alpha\}$  is called the  *$\alpha$ -cut of  $f$* .

The well-known representation theorem (cf. Böhme 1993) states that each membership function  $f$  can be represented by the family  $(f_\alpha \mid \alpha \in [0,1])$  of its  $\alpha$ -cuts. But what about the relation between several membership functions of a linguistic variable?

We start with a small example:

Table 3: A classical linguistic variable

<b>X</b>	<b>young</b>	<b>old</b>
<b>20</b>	0.8	0.2
<b>40</b>	0.6	0.4
<b>60</b>	0.4	0.6
<b>80</b>	0.2	0.8
<b>100</b>	0	1.0

Each membership function  $f$  can be described much coarser marking only the values  $x$  with  $f(x) \geq \alpha$ , say  $\alpha = 0.6$ . This clearly yields a conceptual scale.

Table 4:

<b>X</b>	<b>(young, 0.6)</b>	<b>(old, 0.6)</b>
<b>20</b>	×	
<b>40</b>	×	
<b>60</b>		×
<b>80</b>		×
<b>100</b>		×

This Fuzzy-to-crisp transformation can be viewed as a quite rough conceptual scaling of the many-valued context in Table 3, using for each of the many-valued attributes "young" and "old" the same "threshold" scale with one scale attribute describing the property " $\geq 0.6$ ", namely the scale  $([0,1], \{0.6\}, \geq)$ . In the derived context (Table 4) the extent of the attribute concept of (young, 0.6) is just the  $\alpha$ -cut of the membership function of "young" where  $\alpha = 0.6$ , in this example the set  $\{20,40\}$ .

Since a membership function  $f$  is determined by the family of its  $\alpha$ -cuts we now apply the very fine scale

$S_f := ([0,1], fX, \geq)$  to Table 3 and get the derived context shown in Table 5.

Table 5: The context of a linguistic variable

X	young					old				
	0.0	0.2	0.4	0.6	0.8	0.2	0.4	0.6	0.8	1
20	×	×	×	×	×	×				
40	×	×	×	×		×	×			
60	×	×	×			×	×	×		
80	×	×				×	×	×	×	
100	×					×	×	×	×	×

Reading example: In the "young" part of Table 5 the second row contains the elements of  $fX$  and for each  $\alpha \in fX$  we see crosses in the  $\alpha$ -column at the elements of the extent of the attribute concept of (young,  $\alpha$ ) which is just the  $\alpha$ -cut of the membership function  $f$  of "young".

Reading this part row-wise we see that the row of  $x \in X$  has crosses for all  $\alpha \leq f(x)$ . Roughly speaking: There are crosses for all values  $\leq f(x)$ .

**Definition:**

For each membership function  $f: X \rightarrow [0,1]$  we introduce the contexts:

$K_f := ([0,1], X, I_f)$  where  $\alpha I_f x :\Leftrightarrow f(x) \geq \alpha$  (for  $x \in X, \alpha \in [0,1]$ ), called the *cut context of f* (Wolff 1998),

$C_f := (X, [0,1], J_f)$  where  $x J_f \alpha :\Leftrightarrow f(x) \geq \alpha$  (for  $x \in X, \alpha \in [0,1]$ ), called *the context of f* and its subcontext

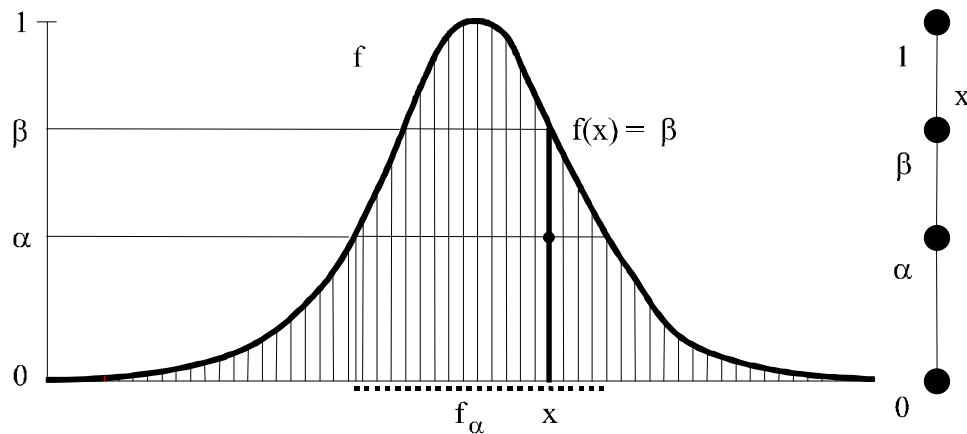
$D_f := (X, fX, J_f \cap (X \times fX))$ , called the *subdirect context of f*.

Clearly  $C_f$  is the dual of  $K_f$  and from each of these contexts the membership function  $f$  can be reconstructed since it can be reconstructed from  $D_f$  because  $f(x) = \sup\{\alpha \in fX \mid x J_f \alpha\}$ .

In Table 5 the subdirect contexts of the two membership functions are represented.

The incidence relation  $J_f$  is just the "set under the curve  $f$ " visualized in Figure 4.

Figure 4: A classical Fuzzy set and some cuts - the concept lattice of its cut context is a chain



The concept lattice  $B(C_f)$  is a chain (hence also  $B(D_f)$  and  $B(K_f)$ ) which follows from the fact that the extents of any two attributes  $\alpha, \beta \in [0,1]$  are comparable. Proof: since  $\alpha$  and  $\beta$  are comparable, say  $\alpha \leq \beta$ ,  $f(x) \geq \beta$  implies  $f(x) \geq \alpha$ , hence the  $f_\beta \subseteq f_\alpha$ . Remark: In the general situation of arbitrary ordered sets as logics (instead of the unit interval) the corresponding concept lattices are no longer chains – and this leads to the desired freedom in applications.

In our example the two chains  $B(D_f)$  corresponding to "young" and "old" are drawn in Figure 5.

Figure 5:

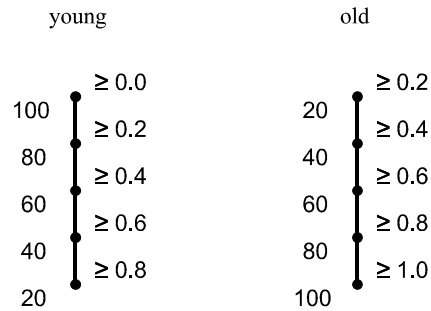
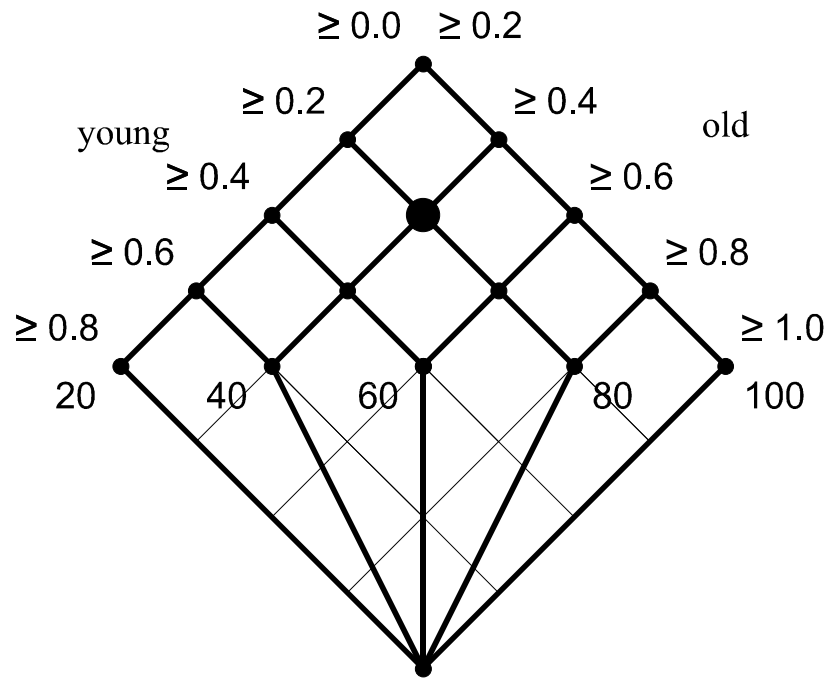


Figure 5 shows that high values of the membership functions correspond to small concepts which follows from the mentioned fact that  $\alpha \leq \beta$  implies  $f_\beta \subseteq f_\alpha$ . This "handstand effect" can be reversed by dualization: in the concept lattice of the cut context small values correspond to small concepts as shown in the right part of Figure 4. The concept lattice of the context in Table 5 is shown in Figure 6.

Figure 6:



This line diagram shows that the union of the scale attributes for "young" and "old" now combine to a new language about the given age values allowing to express all "intervals" constructed from these age values, for example the interval  $\{40, 60, 80\}$  in the "middle" of these age values which can be described as the intersection of the extent of the attribute concepts of "(young,  $\geq 0.2$ )" and "(old,  $\geq 0.4$ )" – and this intersection is itself an extent of a concept – graphically represented by the big circle – namely the infimum of these two attribute concepts. Now new concepts, namely those without attribute labels arise and may be used to describe the meaning of for example "quite young" in the framework of the chosen language.

Now it is obvious (and this was proved in Wolff 1998)

- that each linguistic variable can be described without any loss of information by its context (which is just the apposition of the (subdirect) contexts of its membership functions).

This crucial result might be irritating on the first view. How can the continuity of the real unit interval be represented with discrete structures like formal contexts? Answer: it is right that finite formal contexts are discrete but an arbitrary context need not to be discrete. For example the concept lattice of the infinite context  $([0,1], [0,1], \leq)$  is isomorphic to

the real unit interval with its usual real ordering. This shows that "conceptual" does not imply "discrete". But one might ask: "Where exactly does finiteness come in?" It is clear that a context  $(G, M, I)$  has only finitely many concepts if  $G$  or  $M$  is finite. Hence the context  $\mathbf{D}_f := (X, fX, J_f \cap (X \times fX))$  has finitely many concepts if  $fX$  is finite which is clearly the case if  $X$  is finite. As in Table 3 one often wishes to have a finite set  $X$ . But even if  $X$  is not finite we have in any data table in practice only a finite number of objects and therefore only a finite number of actually occurring measurement values.

## 5 MEASUREMENTS: Objects, implications and realized linguistic variables

### 5.1 "IF OBJECTS COMES IN..."

The examples above demonstrate two different roles of formal objects in knowledge processing. In the "confidence" example of Table 1 countries are the objects while in the "linguistic variable" example of Table 3 values of age play the role of formal objects but we didn't mention up to now the persons we would like to describe by their age values.

In conceptual knowledge processing an object  $g$  of a given many-valued context, its value  $m(g)$  of a many-valued attribute  $m$  and the scale attributes  $(m, n)$  of the scale  $S_m$  are glued together in the derived context in a very simple way:

$g$  has an attribute  $(m, n)$  in the derived context if and only if there is a value  $w \in W$  such that  $m(g) = w$  and  $w I_m n$ .

Example: ADAM has the attribute (age, young), if the age of ADAM has the attribute young in the scale for age.

As far as I know there is no corresponding construction in Fuzzy Theory which describes formally this fundamental relation between the data (given by a measurement function  $m$ ) and a linguistic variable describing the Fuzzy language about the possible values of  $m$ . But clearly this process is done in practice and is usually described by words like "If an input-value comes in..." (Driankov 1993, p.90).

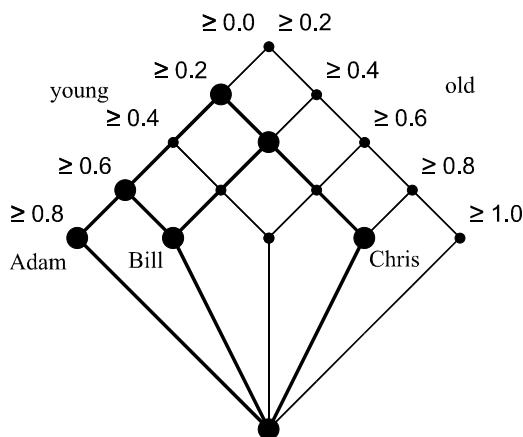
As a small example let us take a part of a questionnaire given in Table 6.

Table 6:

	<b>age</b>
<b>Adam</b>	20
<b>Bill</b>	40
<b>Chris</b>	80

Since we are interested to understand how the persons are classified according to their age we use the context in Table 5 as a conceptual scale for the many-valued attribute "age". The corresponding derived context (now called the *realized scale*) obtained from this second (!) scaling procedure has a concept lattice which is drawn in Figure 7 using bold lines indicating the supremum-embedding into the concept lattice of this scale.

Figure 7: A realized scale



This diagram will be interpreted in the next section.

## 5.2 ...THEN IMPLICATIONS FOLLOW"

After having introduced the objects of our primary interest we study implications (in the sense of the following definition) as one of the most important tools for the description of object distributions in conceptual structures.

### Definition:

Let  $\mathbf{K} = (G, M, I)$  be a context and  $P, C \subseteq M$  ( $P$  stands for "premise" and  $C$  for "conclusion").

We say that *the implication  $(P, C)$  holds (in  $\mathbf{K}$ )*, written  $P \Rightarrow C$  (in  $\mathbf{K}$ ), if  $P^\downarrow \subseteq C^\downarrow$ , in words: For all objects  $g \in G$ : if  $g$  has all attributes of  $P$ , then  $g$  has all attributes of  $C$ .

Interpreting Figure 7 we may say: "All persons which are not old are quite young" understood as "All persons with the attribute (young,  $\geq 0.4$ ) have the attribute (young,  $\geq 0.6$ )", shortly described as the implication "(young,  $\geq 0.4$ )  $\Rightarrow$  (young,  $\geq 0.6$ )". Also the inverse implication "(young,  $\geq 0.6$ )  $\Rightarrow$  (young,  $\geq 0.4$ )" holds in this context but this one holds independently of the given object distribution in the "logic" of this scale! But it holds even in the logic of the scale of this scale – namely the chain of "young" in Figure 5. Why does it hold there? Since the implication " $\geq 0.6 \Rightarrow \geq 0.4$ " holds in the context  $([0,1], [0,1], \geq)$ . Clearly this implication should be written precisely " $0.6 \Rightarrow 0.4$ " which means "For all  $x \in [0,1]$ : if  $x \geq 0.6$  then  $x \geq 0.4$ ". And this is equivalent to " $0.6 \geq 0.4$ " in the ordered set  $([0,1], \geq)$ .

This example shows in the language of implications a sequence of nested structures representing a "flow of knowledge" from basic abstract knowledge to meaningful knowledge.

This example gives also a rough idea how conceptual scaling can be used to understand information flow from one contextual logic to another. Indeed, there is a nice equivalence between conceptual scaling and information channels in the sense of Barwise and Seligman (1997) described in Wolff (1999a).

## 5.3 EXTENDING CLASSICAL FUZZY THEORY: Realized linguistic variables

Conceptual scales are modeled in Fuzzy Theory by linguistic variables, but the connection of the measurement and the linguistic variable, conceptually described by:  $g \rightarrow m(g) \rightarrow \gamma_m(m(g))$

is not described theoretically in Fuzzy Theory. Both situations, the scale and the realized scale can be described by linguistic variables on the object sets  $X$  respectively  $G$ . I believe that the similarity of both situations led to an ambiguity in many books including the above mentioned definition of a linguistic variable by Zadeh (1975).

What should be developed? Clearly a combination of the measurement and the linguistic variable. Therefore we recall (at first only for the case of the classical logic  $([0,1], \geq)$ ) the definition of a realized linguistic variable introduced in Wolff (1998).

### Definition:

Let  $\lambda := (X, V, \mu, [0,1], \leq)$  be a *classical linguistic variable* and  $m: G \rightarrow X$  be a function from an arbitrary set  $G$  into  $X$ . Then the tuple  $(G, m, \lambda) := (G, m, X, V, \mu, [0,1], \leq)$  is called a *classical realized linguistic variable of  $\lambda$*  and  $\lambda$  *the linguistic variable of  $(G, m, \lambda)$* . The function  $m$  is called a *measurement function from the set  $G$  of objects into the domain  $X$  of the linguistic variable  $\lambda$* .

Now the context of a realized linguistic variable with respect to the subdirect scales can be defined.

### Definition:

Let  $(G, m, \lambda) := (G, m, X, V, \mu, [0,1], \leq)$  be a classical realized linguistic variable. Then the context  $\mathbf{D}(G, m, \lambda)$  is defined to be the derived context of the many-valued context  $(G, \{m\}, mG, \{(g, m, m(g)) \mid g \in G\})$  using the scale  $\mathbf{D}_\mu$  defined as the apposition of the contexts  $\mathbf{D}_{\mu(v)}$  where  $v \in V$ .

This is the main tool for the investigation of problems in Fuzzy reasoning.

## 5.4 PROBLEMS IN FUZZY REASONING: The unit interval is too narrow!

What is the relation between Fuzzy reasoning and the above mentioned logical conceptual aspects of Fuzzy scaling? For a detailed discussion of this relation the reader is referred to Wolff (1999b). We just sketch the main arguments.

In my opinion Fuzzy reasoning (as described in Fuzzy Theory until now) suffers mainly from two facts:

Fact 1: The unit interval is too narrow for the construction of direct products of linguistic variables.

Fact 2: The different roles of objects and values in the scaling process were not understood.

Now I explain in which sense the unit interval is too narrow:

Let  $\rho := (G, m, X, V, \mu, [0,1], \leq)$  and  $\rho' := (G, m', X', V', \mu', [0,1], \leq)$  be two classical realized linguistic variables.

If we try to combine them into a single classical realized linguistic variable which contains the whole information of both we could try to build a classical realized linguistic variable of the form

$$(G, m \times m', X \times X', V \times V', \mu^*, [0,1], \leq),$$

where the mapping  $m \times m' : G \rightarrow X \times X'$  is defined by  $(m \times m')(g) := (m(g), m'(g))$ . But how to define a mapping

$\mu^* : V \times V' \rightarrow F(X \times X', [0,1])$  which preserves the information given by  $\mu$  and  $\mu'$ ?

For  $(v, v') \in V \times V'$   $\mu^*(v, v') \in F(X \times X', [0,1])$ , hence it is a Fuzzy relation which should be constructed from the given membership functions  $\mu_v$  and  $\mu'_{v'}$ . Therefore many Fuzzy researchers tried to find meaningful constructions of Fuzzy relations  $\in F(X \times X', [0,1])$  from two membership functions on  $X$  and on  $X'$ . In the literature they are called "implications", for example the Gaines-Rescher implication, Goguen i., Gödel i., Larsen i., Lukasiewicz i., Kleene-Dienes i., Mamdani i., Reichenbach i., stochastic i. and the Zadeh implication (Driankov 1993). As "the most important implication" (Driankov 1993, p.89) we mention the Mamdani implication:

$$\mu^*_{(v,v')}(x, x') := \min(\mu_v(x), \mu'_{v'}(x')).$$

What is the meaning of this construction? It has at least a certain effect of "information flow" similar to the modus ponens rule for the classical implication: If the first value is big and the minimum of both is big, then the second value is big. But it destroys information irreversibly. All of these implications have some further disadvantages described by Hellendoorn (1990, 1992).

On the basis of the examples in this paper it is quite easy to understand that there is no reason to represent knowledge necessarily in a chain. Therefore we take the direct product  $[0,1] \times [0,1]$  as the natural logic for the product of two classical linguistic variables.

To think of this direct product as a logic is quite hard for a classical Fuzzy theorist but it was prepared by Goguen (1967) to use as logics not only chains but lattices or even arbitrary ordered sets. The development of Formal Concept Analysis and its intimate relations to logics give us now the possibility to work with arbitrary ordered sets as logics using conceptual scaling theory which leads to deep insights into several difficult problems in contextual logics.

To finish the construction of the product of two classical realized linguistic variables we take

$$(G, m \times m', X \times X', V \times V', \mu \times \mu', [0,1] \times [0,1], \leq_x),$$

where the mapping  $\mu \times \mu' : V \times V' \rightarrow F(X \times X', [0,1] \times [0,1])$  is defined by  $(\mu \times \mu')(v, v') := \mu_v \times \mu'_{v'}$ , where

$$(\mu_v \times \mu'_{v'})(x, x') := (\mu_v(x), \mu'_{v'}(x')) \in [0,1] \times [0,1] \text{ for all } (x, x') \in X \times X' \text{ and}$$

$\leq_x$  means the usual product of the real ordering with itself:

$$(a, a') \leq_x (b, b') : \Leftrightarrow a \leq b \text{ and } a' \leq b' \text{ (for all } a, b, a', b' \in [0,1]).$$

It should be clear from the given examples that we can now construct the context of this product of two classical realized linguistic variables whose concept lattice can be embedded into the direct product of two chains in the same way as it is shown in Figure 3.

But clearly the general theory has to start now with an arbitrary ordered set, say  $(L, \leq)$  instead of the unit interval. This L-Fuzzy Scaling Theory was described by Wolff (1998) and will be discussed in the next section.

## 6 GENERAL FUZZY SCALING THEORY: Embeddings into meaningful contextual logics

Independently of the purely mathematical generalization argument of Goguen (1967) there is a way to understand the effect of the substitution of the unit interval by an arbitrary ordered set.

The central observation is that the object concept function  $\gamma : G \rightarrow \mathbf{B}(\mathbf{K})$  maps objects of a formal context into a complete lattice, just like, but more general than a classical membership function  $f : X \rightarrow [0,1]$ . The concept lattices of the "confidence" example demonstrate that membership in the general sense can be described conceptually very well in a concept lattice allowing "graded concepts" in the sense of the conceptual hierarchy, allowing even incomparability of concepts which often occurs in practice and a great flexibility in the choice of granularity resulting from the flexibility in choosing scales. Hence the idea of replacing the narrow unit interval by an arbitrary concept lattice seems very successful. It has not only the advantage to have all complete lattices as possible logics but also to use the short description of concept lattices by formal contexts and their general applicability and interpretability in many situations. Since an arbitrary ordered set can always be embedded into the smallest complete lattice containing it, namely its Dedekind-MacNeille completion (Davey, Priestley 1990; Ganter, Wille 1999), the advantages of concept lattices can also be used in the general case of arbitrary ordered sets. Clearly this ordered set can again be generalized to an arbitrary context and then the concept lattice of his context represents its contextual logic.

## 6.1 BASIC NOTIONS: L-Fuzzy Sets and $\alpha$ -cuts

Let  $X$  be a set and  $(L, \leq)$  an ordered set, i.e.  $\leq$  is a reflexive, antisymmetric and transitive relation on the set  $L$ . The set  $\mathbf{F}(X,L) := \{ f \mid f: X \rightarrow L \}$  is called the set of all *L-Fuzzy sets* (or *L-membership functions*) on  $X$ .

The definition of  $\alpha$ -cuts can be generalized easily:

For each L-Fuzzy set  $f \in \mathbf{F}(X,L)$  and each  $\alpha \in L$  the  *$\alpha$ -cut* of  $f$  is defined to be the set  $f_\alpha := \{x \in X \mid f(x) \geq \alpha\}$ .

## 6.2 L-FUZZY SCALING THEORY: Linguistic variables (over L)

### Definition:

A *linguistic variable* is a quintuple  $(X, V, \mu, L, \leq)$ , where  $X$  is a set (called the '*domain*'),  $V$  is a set (of '*linguistic values*'),  $(L, \leq)$  is an ordered set and  $\mu$  is a mapping  $\mu: V \rightarrow \mathbf{F}(X, L)$  which represents each linguistic value  $v$  by an L-membership function  $\mu_v := \mu(v)$  on  $X$ .

### Definition:

Let  $\lambda := (X, V, \mu, L, \leq)$  be a *linguistic variable* and  $m: G \rightarrow X$  be a function from an arbitrary set  $G$  into  $X$ . Then the tuple  $(G, m, \lambda) := (G, m, X, V, \mu, L, \leq)$  is called a *realized linguistic variable of  $\lambda$*  and  $\lambda$  the *linguistic variable of  $(G, m, \lambda)$* . The function  $m$  is called a *measurement function from the set  $G$  of objects into the domain  $X$  of the linguistic variable  $\lambda$* .

### Definition:

Let  $\rho := (G, m, X, V, \mu, L, \leq)$  and  $\rho' := (G, m', X', V', \mu', L', \leq')$  be two realized linguistic variables on the same set  $G$  of objects. The mapping  $m \times m': G \rightarrow X \times X'$  which is defined by  $(m \times m')(g) := (m(g), m'(g))$  is called the *product of the two measurement functions  $m$  and  $m'$* .

The mapping  $\mu \times \mu': V \times V' \rightarrow \mathbf{F}(X \times X', L \times L')$  is defined by  $(\mu \times \mu')(v, v') := \mu_v \times \mu'_{v'}$ , where  $(\mu_v \times \mu'_{v'})(x, x') := (\mu_v(x), \mu'_{v'}(x')) \in L \times L'$  for all  $(x, x') \in X \times X'$ .

$(L \times L', \leq_x)$  is the usual product of the ordered sets  $(L, \leq)$ ,  $(L', \leq')$ , where

$(a, a') \leq_x (b, b') : \Leftrightarrow a \leq b \text{ and } a' \leq' b'$  (for all  $a, b \in L, a', b' \in L'$ ).

Then the following tuple  $\rho \times \rho' := (G, m \times m', X \times X', V \times V', \mu \times \mu', L \times L', \leq_x)$

is a realized linguistic variable on the product  $(L \times L', \leq_x)$ , called the *product of  $\rho$  and  $\rho'$* .

The quintuple  $\lambda \times \lambda' := (X \times X', V \times V', \mu \times \mu', L \times L', \leq_x)$  is called the *product of the corresponding linguistic variables  $\lambda$  and  $\lambda'$* .

### Discussion of this definition:

1. The product  $\rho \times \rho'$  of two realized linguistic variables  $\rho$  and  $\rho'$  clearly contains all the information about  $\rho$  and  $\rho'$ .

2. If one tries to define another product of two *classical* linguistic variables  $\rho$  and  $\rho'$  which

- is again a *classical* linguistic variable and
- contains all the information about  $\rho$  and  $\rho'$

then the information about  $\mu$  and  $\mu'$  represented in the product  $[0,1] \times [0,1]$  has to be represented in  $[0,1]$ . As far as I see there is no natural way to do this. This difficulty does not occur in the definition above since we used an arbitrary ordered set instead of the usual ordering on the unit interval – and the direct product of two ordered sets is an ordered set while the direct product of two intervals is not an interval.

## 6.2 TRANSLATION INTO CONTEXTS: Cut context and its dual

### Definition:

For each L-Fuzzy set  $f \in \mathbf{F}(X,L)$  we introduce the contexts:

$\mathbf{K}_f := (L, X, I_f)$  where  $\alpha I_f x : \Leftrightarrow f(x) \geq \alpha$  (for  $x \in X, \alpha \in L$ ), called the *cut context of  $f$*  (Wolff 1998),

$\mathbf{C}_f := (X, L, J_f)$  where  $x J_f \alpha : \Leftrightarrow f(x) \geq \alpha$  (for  $x \in X, \alpha \in L$ ), called *the context of  $f$*  and its subcontext

$\mathbf{D}_f := (X, fX, J_f \cap (X \times fX))$ , called the *subdirect context of  $f$* .

The subdirect context  $\mathbf{D}_f$  (and therefore also  $\mathbf{K}_f$  and  $\mathbf{C}_f$ ) determines  $f$  since  $f(x) = \sup\{\alpha \in fX \mid x J_f \alpha\}$ . Clearly the concept lattices of these contexts are not necessarily chains.

As in the examples, contexts describing the (realized) linguistic variables can be defined such that the (realized) linguistic variable can be reconstructed from its context. For the mathematical details and an example showing the new

freedom of L-Fuzzy Scaling Theory even in the case of the existence of missing values the reader is referred to Wolff (1998).

### 6.3 L-FUZZY REASONING: Implications and Dependencies

Fuzzy implications of the form "If X is A, then Y is B" can be introduced easily syntactically – but what is its semantic and what its pragmatic meaning? Using realized linguistic variables we define Fuzzy implications syntactically (in its most simple version) in the following conceptual way:

Let  $\rho := (G, m, X, V, \mu, L, \leq)$  and  $\rho' := (G, m', X', V', \mu', L', \leq')$  be two realized linguistic variables on the same set  $G$  of objects. We introduce a *Fuzzy implication between  $\rho$  and  $\rho'$*  as a pair  $((m, v), (m', v'))$  where  $v \in V$  and  $v' \in V'$  and refer to it by the phrase "If  $m$  is  $v$ , then  $m'$  is  $v'$ ".

As a first attempt we define a semantic meaning of a Fuzzy implication by the validity of an implication of a suitable derived context:

**Definition:**

For  $(\alpha, \alpha') \in L \times L'$  the Fuzzy implication  $(m, v, m', v')$  between  $\rho$  and  $\rho'$  is called *valid with respect to  $(\mathbf{a}, \mathbf{a}')$* , if  $\forall g \in G (\alpha \leq \mu_v(m(g)) \Rightarrow \alpha' \leq \mu'_{v'}(m'(g)))$ .

That the Fuzzy implication  $(m, v, m', v')$  is valid with respect to  $(\alpha, \alpha')$  is equivalent to the validity of the implication  $((v, \alpha) \Rightarrow (v', \alpha'))$  in the apposition  $\mathbf{K}_\rho \mid \mathbf{K}_{\rho'}$  of the derived contexts of  $\rho$  and  $\rho'$  (Wolff 1998).

Example: We interpret Figure 3 as the line diagram of the apposition of the derived contexts of the linguistic variable  $\rho$  for  $m :=$  "long term credits" and the linguistic variable  $\rho'$  for  $m' :=$  "no inflation" both defined on the same set  $G$  of the six countries as in Table 1.  $X := X' := [0,4]$ ,  $V := \{m\}$ ,  $V' := \{m'\}$ ,  $L := L' := [0,4]$ ,  $\mu_m(x) := \mu'_{m'}(x) :=$  the result of rounding off  $x$  to the next half, for example  $\mu_m(2.8) = 2.5$ . From Figure 3 we see that the following implication holds in this context: "long term credits  $\geq 1.5$  and no inflation  $\geq 2.5$  implies long term credits  $\geq 2.0$ ."

But the implications of a formal context do not contain the whole information of the context. They just represent the closure system of the intents. They do not represent the distribution of the objects. If we wish to use also the object distribution we leave the classical logical field which is mainly concerned with the attribute logic and the structure of the concept lattice. But in our understanding the object distribution belongs essentially to the data which is obvious from the use of frequency distributions given for example in histograms.

Tools for the investigation of the frequency distribution have been developed in Statistics and Correspondence Analysis especially in the case of the joint frequency distribution in contingency tables (which just represent direct products of the object concepts of the used scaled many-valued attributes). A useful new graphical representation of dependencies between two scaled many-valued attributes was introduced by the author and first published in the diploma thesis of Wehrle (1997). Connections to several dependency notions and "soft implications" will be described in Wolff (1999b).

## 7 CONCLUSIONS: L-Fuzzy Scaling Theory is equivalent to Conceptual Scaling Theory

We have seen that Fuzzy Scaling Theory can be generalized to L-Fuzzy Scaling Theory by the introduction of (realized) linguistic variables over an arbitrary ordered set  $L$ . This surmounts the two central obstacles in classical Fuzzy Theory namely the narrowness of the unit interval and the lack of a scaling procedure. These obstacles prohibited in classical Fuzzy Theory the investigation of some difficult relations between contextual logics occurring in linguistic variables. The key to the solution of these problems was the description of the central tool in Fuzzy Theory namely the membership function by the cut context which represents the given membership function without any loss of information - independently whether these contexts are finite or not. In practice finiteness of the data implies finiteness of the corresponding derived concept lattices. Hence L-Fuzzy Scaling Theory can be described by Conceptual Scaling Theory. But clearly each formal context is a special (realized) linguistic variable where the characteristic functions of the extents of the attribute concepts are the membership functions. This demonstrates the main idea how to prove that both theories are equivalent.

The comparison of two theories has many aspects hence a formal equivalence does not mean that the theories are used in the same way. But it is a hint for both communities that they should try to understand one another.

## References

- Barwise, J., Seligman, J.: Information Flow – The Logic of Distributed Systems. Cambridge Tracts in Theoretical Computer Science 44, 1997.
- Böhme, G.; Fuzzy-Logik - Einführung in die algebraischen und logischen Grundlagen. Springer-Verlag, 1993.
- Davey, B.A., Priestley, H.A.; Introduction to lattices and order. Cambridge University Press 1990.
- Driankov, D., Hellendoorn, H., Reinfrank, M.; An introduction to Fuzzy Control. Springer-Verlag, 1993.
- Dubois, D., Esteva, F., Garcia, P., Godo, L., de Mantaras, R.L., Prade, H. 1997. Fuzzy Rule-Based Models for Case-Based Reasoning. In: Proceedings of EUFIT'97, 5<sup>th</sup> European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany, pp. 923-927.
- Ganter, B., R. Wille; Conceptual Scaling. In: F.Roberts (Hrsg.) Applications of combinatorics and graph theory to the biological and social sciences,139-167. Springer Verlag, New York, 1989.
- Ganter, B., R. Wille; Formale Begriffsanalyse: Mathematische Grundlagen. Springer-Verlag, Berlin-Heidelberg 1996.
- Ganter, B., R. Wille; Formal Concept Analysis: mathematical foundations. (translated from the German by Cornelia Franzke) Springer-Verlag, Berlin-Heidelberg 1999.
- Ganter, B., R. Wille: Contextual Attribute Logic. Submitted for publication in the Proceedings of the 7<sup>th</sup> International Conference On Conceptual Structures ICCS '99, (1999a).
- Ganter, B., Wille, R., Wolff, K.E. (Hrsg.); Beiträge zur Begriffsanalyse. B.I.-Wissenschaftsverlag, Mannheim, 1987.
- Goguen, J.A.; L-Fuzzy sets. Journal of Mathematical Analysis and Applications **18** (1967), 145-174.
- Hellendoorn, H.; Reasoning with Fuzzy Logic. Ph.D. thesis, Delft University of Technology, Delft, 1990.
- Hellendoorn, H.; The Generalized Modus Ponens Considered as a Fuzzy Relation. Fuzzy Sets and Systems, **46** (1), 1992, pp.29-48.
- Hirota, K. (ed.); Industrial Applications of Fuzzy Technology. Springer-Verlag Tokyo, 1993.
- Jambu, M.; Exploratory and Multivariate Data Analysis. Academic Press, San Diego, 1991.
- Kruse, R., Gebhardt, F., Klawonn, F.; Fuzzy-Systeme. B.G. Teubner, Stuttgart 1993.
- Lin, T.Y.; Granular Fuzzy Sets. In: Zimmermann, H.J.: EUFIT '98, 6<sup>th</sup> European Congress on Intelligent Techniques and Soft Computing, Aachen 1998, Vol. I, 94-98.
- Luxenburger, M.; Partielle Implikationen und partielle Abhängigkeiten zwischen Merkmalen. Diplomarbeit, TH Darmstadt, 1988.
- Luxenburger, M.; Dependencies between Many-valued Attributes. In: Ewa Orłowska (ed.); Incomplete Information: Rough Set Analysis. Physika-Verlag, Heidelberg, 1998, p.316-343.
- Pollandt, S.; Fuzzy-Begriffe: Formale Begriffsanalyse unscharfer Daten. Springer-Verlag, Berlin-Heidelberg 1996.
- Umbreit, S.; Formale Begriffsanalyse mit unscharfen Begriffen. Dissertation. Martin-Luther-Universität Halle-Wittenberg, 1995.
- Vogt, F., R.Wille; TOSCANA - a graphical tool for analyzing and exploring data. In: R.Tamassia, I.G.Tollis (eds.); Graph Drawing. Springer-Verlag, Heidelberg, 1994, 193-205.
- Wechler, W.; The Concept of Fuzziness in Automata and Language Theory. Akademie Verlag, Berlin, 1978.
- Wehrle, C.; Abhängigkeitsuntersuchungen in mehrwertigen Kontexten. Diplomarbeit am Fachbereich Informatik der Fachhochschule Darmstadt, 1997.
- Wille, R.; Restructuring lattice theory: an approach based on hierarchies of concepts. In: Ordered Sets (ed. I.Rival). Reidel, Dordrecht-Boston 1982, 445-470.
- Wille, R.; Line diagrams of hierarchical concept systems. International Classification 11, 1984, 77-86.
- Wille, R.; Lattices in data analysis: how to draw them with a computer. In: I.Rival (ed.): Algorithms and order. Kluwer, Dordrecht-Boston 1989, 33-58.
- Wille, R., Zickwolff, M. (Hrsg.); Begriffliche Wissensverarbeitung - Grundfragen und Aufgaben. B.I.-Wissenschaftsverlag Mannheim, 1994.
- Wolff, K.E.; Begriffsanalytische Visualisierung mehrdimensionaler Daten aus einer SPIEGEL-Umfrage an westdeutschen Universitäten. Preprint Fachhochschule Darmstadt, 1992.
- Wolff, K.E.; A first course in Formal Concept Analysis - How to understand line diagrams. In: Faulbaum, F. (ed.): SoftStat '93, Advances in Statistical Software 4, Gustav Fischer Verlag, Stuttgart 1994, 429-438.
- Wolff, K.E.; Comparison of graphical data analysis methods. In: Faulbaum, F. & Bandilla, W. SoftStat '95 Advances in Statistical Software 5, Lucius&Lucius, Stuttgart 1996, 139-151.
- Wolff, K.E.; Conceptual Interpretation of Fuzzy Theory. In: Zimmermann, H.J.: EUFIT '98, 6<sup>th</sup> European Congress on Intelligent Techniques and Soft Computing, Aachen 1998, Vol. I, 555-562.
- Wolff, K.E.; Information Flow and Conceptual Scaling. Preprint Fachhochschule Darmstadt, 1999a.
- Wolff, K.E.; Implications and Soft Implications – Tools for the Analysis of Dependencies. Preprint Fachhochschule Darmstadt, 1999b.
- Yager, R.R., Zadeh, L.A.; An Introduction to Fuzzy Logic Applications in Intelligent Systems. Kluwer Academic Publishers, 1992. Third printing 1995.

Zadeh, L.A.; Fuzzy sets. *Information and Control* **8**, 1965, 338 - 353.

Zadeh, L.A.; Fuzzy algorithms. *Information and Control* **19**, 1969, 94 - 102.

Zadeh, L.A.; The concept of a linguistic variable and its application to approximate reasoning. Memorandum ERL-M 411 Berkeley, October 1973.

Zadeh, L.A.; The concept of a linguistic variable and its application to approximate reasoning. Part I: *Inf. Science* 8,199-249; Part II: *Inf. Science* 8, 301-357; Part III: *Inf. Science* 9, 43-80, 1975.

Zadeh, L.A.; Possibility theory and soft data analysis. In: L.Cobb, R.M.Thrall (eds.): *Mathematical Frontiers in the Social and Policy Sciences*. Boulder, CO, 69 - 129.

Zadeh, L.A.; Knowledge representation in Fuzzy Logic. In: Yager, R.R., Zadeh, L.A. (eds.): *An Introduction to Fuzzy Logic Applications in Intelligent Systems*. Kluwer Academic Publishers 1992, 1-25.

Zimmermann, H.J.; *Fuzzy Sets, Decision Making, and Expert Systems*. Boston, Dordrecht, Lancaster 1987.

Zimmermann, H.J.; *Fuzzy Set Theory - and Its Applications*. Kluwer Academic Publishers,1991. Seventh printing 1994.