Adapting Fuzzy Formal Concept Analysis for Fuzzy Description Logics

Felix Distel

Institute of Theoretical Computer Science, Technische Universität Dresden, Dresden, Germany, felix@tcs.inf.tu-dresden.de

Abstract. Fuzzy Logics have been applied successfully within both Formal Concept Analysis and Description Logics. Especially in the latter field, Fuzzy Logics have been gaining significant momentum during the last two years. Unfortunately, the research on fuzzy logics within the two communities has been conducted independently from each other, leading to different approaches being pursued. We show that if we look at a restricted variant of fuzzy formal concept analysis, then the differences between the two approaches can be reconciled. Moreover, an implicational base can be computed even when the identity hedge is used.

1 Introduction

In many applications one is forced to deal with vague knowledge, knowledge that does not fit into the binary world of classical logics. Questions such as whether a country is large, or whether two cities are close to each other are difficult to answer with true or false. There are various degrees of size and proximity. Fuzzy Logics has successfully proposed to use a scale of truth degrees to describe vague knowledge. It has first been formalized for propositional logic [1] and has since been applied to many other logics and logic related formalisms. Among them are Formal Concept Analysis (FCA) [2] and Description Logics (DL) [3].

Whenever one applies Fuzzy Logics to an existing formalism, one is faced with several choices: Should the real unit interval be used for the set of truth degrees or a more complex lattice of truth degrees? How should the semantics of the conjunction be defined? Which parts of the existing theory should be replaced by their fuzzy counterparts and which should remain unchanged? These decisions have been made independently for fuzzy FCA and fuzzy DL.

In the past, a number of works have used FCA methods in DL. Some use it as a tool for efficiently computing concept hierarchies [4]. Others use it for ontology completion [5] and for exploring and learning from graph data [6, 7]. This work has been possible due to the close ties between FCA and DL. For example, in FCA objects can be described using sets of attributes, and in DL individuals can be described using concept descriptions, in the easiest case conjunctions over concept names. Sets of attributes in FCA and conjunctions over concept names in DL share essentially the same semantics.

While in the crisp case the similarities between fuzzy DL and fuzzy FCA are prominent, the situation is not so clear in the fuzzy variants of the respective theories. In fuzzy FCA one is allowed to use fuzzy sets of attributes. In fuzzy DL the same concept descriptions as in crisp DL are used. They are not fuzzy, only their semantics are. In Section 3 we identify such differences, that hinder the close cooperation that exists between the crisp variants of the two fields. We propose simple adjustments to avoid them. Generally speaking, one can say that the FCA community has been more ambitious and applied Fuzzy Logics to much larger parts of the original theory than fuzzy DL. Unfortunately, for this reason implication bases, which play an important role in the cooperation between crisp DL and crisp FCA, can no longer effectively be computed in the general case [8, 9].[1] We shall see in Section 4 that if we restrict expressivity of fuzzy FCA by considering only crisp sets of attributes, we can effectively compute bases.[2] Moreover, if the Gödel t-norm is used, this restricted version of fuzzy FCA is exactly the segment of fuzzy FCA whose semantics overlaps with fuzzy DL, presumably allowing synergies as in the crisp case.

The restriction to the Gödel t-norm is necessary, since fuzzy FCA uses weak conjunction for the semantics of attribute sets, while DL uses strong conjunction for its semantics. Weak conjunction and strong conjunction coincide only for the Gödel t-norm. From a current DL viewpoint, this is not a severe restriction, since up to now the Gödel t-norm is the only t-norm for which the standard DL reasoning tasks are know to be decidable [10].

2 Preliminaries

2.1 T-Norms, Hedges and Fuzzy Sets

Fuzzy Logics represent vague data while maintaining a well-defined semantics. Instead of using only the two values true and false a scale of *truth degrees* is used. In this work we consider only the most typical choice where truth degrees are values from the real unit interval [0, 1].

Fuzzy Logics provide several operators to define its semantics. A *t-norm* $\otimes$ is a binary operator $\otimes\colon [0,1] \times [0,1] \to [0,1]$ that is associative, commutative, monotone and has 1 as its unit. Every continuous t-norm gives rise to a binary operator $\Rightarrow\colon [0,1] \times [0,1] \to [0,1]$ that is the unique operator satisfying for all $z \in [0,1]$

$$z \leq x \Rightarrow y \text{ iff } x \otimes z \leq y. \tag{1}$$

The intuition is that the t-norm and the residuum can be used to interpret conjunction and implication, respectively. Among the many continuous t-norms perhaps the simplest one, and the one we shall be interested in, is the Gödel t-norm. It is defined as $x \otimes y = \min\{x, y\}$ and its corresponding residuum is

$$x \Rightarrow y = \begin{cases} 1 & \text{if } x \leq y \\ y & \text{otherwise.} \end{cases}$$

[1] They can still be computed if the globalization hedge is used. However, hedges do not exist in fuzzy DL and would even be problematic, as we shall see later.

[2] even if the globalization hedge is not used

A *hedge* $\cdot^*$ is a unary operator that is idempotent and satisfies $1^* = 1$, $a^* \leq a$, and $(a \Rightarrow b)^* \leq a^* \Rightarrow b^*$ for all $a, b \in [0,1]$. It is used for truth-stressing, i.e. to increase the contrast between 1 and the smaller truth values. A simple hedge is the *globalization*, defined as $1^* = 1$ and $a^* = 0$ for $a \neq 0$.

Fuzzy sets are a central idea of Fuzzy Logics. Given a set M a *fuzzy (sub-)set* T of M is a function $T\colon M \to [0,1]$, that maps each element of M to its membership degree in T. The *cardinality* of a fuzzy set T is defined as the cardinality of its support $\{x \in M \mid T(x) > 0\}$. Two fuzzy sets T_1 and T_2 can be compared pointwise by defining $T_1 \subseteq T_2$ iff $T_1(x) \leq T_2(x)$ for all $x \in M$. Alternatively, one can associate a subsethood degree with T_1 and T_2 by defining

$$S(T_1, T_2) = \inf_{x \in M} T_1(x) \Rightarrow T_2(x).$$

For finite fuzzy sets we use notation such as $\{{}^{0.5}/_{a}, {}^{1}/_{b}\}$ to denote the set that contains a with degree 0.5 and b with degree 1.

2.2 Formal Concept Analysis

The crisp setting We introduce crisp FCA in addition to fuzzy FCA, as we shall need the crisp version of the Duquenne-Guigues Base in the later sections. In crisp FCA [11], data is typically represented in the form of cross tables such as the one in Table 1. More formally, a *formal context* is a triple $\mathbb{K} = (G, M, I)$ where G is a set, called the set of *objects*, M is a set, called the set of *attributes*, and $I \subseteq G \times M$ is a binary relation, called the *incidence relation*. For sets $A \subseteq G$ and $B \subseteq M$ the derivation operators are defined as

$$A^{\uparrow} = \{m \in M \mid \forall g \in A\colon (g,m) \in I\}, \quad B^{\downarrow} = \{g \in G \mid \forall m \in B\colon (g,m) \in I\}. \tag{2}$$

The two derivation operators $\cdot^{\uparrow}$ and $\cdot^{\downarrow}$ form an antitone Galois-connection. An *implication* $A \to B$, where $A, B \subseteq M$, is said to hold in the context $\mathbb{K}$ if $A^{\downarrow} \subseteq B^{\downarrow}$. A set of attributes $U \subseteq M$ *respects* $A \to B$ iff $A \not\subseteq U$ or $B \subseteq U$. $A \to B$ *follows* from a set of implications $\mathcal{L}$ iff every set U that respects all implications from $\mathcal{L}$ also respects $A \to B$.

One way to structure the data in a formal context $\mathbb{K}$ is the *Duquenne-Guigues base* $\mathcal{DG}(\mathbb{K})$ [12]. $\mathcal{DG}(\mathbb{K})$ is a set of implication that is *sound* for $\mathbb{K}$, i.e. every implication from $\mathcal{DG}(\mathbb{K})$ holds in $\mathbb{K}$, *complete* for $\mathbb{K}$, i.e. every implication that holds in $\mathbb{K}$ follows from $\mathcal{DG}(\mathbb{K})$, and has *minimal cardinality* among all sound and complete sets of implications. A version that can handle background knowledge has been introduced in [13]. Given a sound set of implication $\mathcal{S}$ (the background knowledge) the $\mathcal{S}$*-Duquenne-Guigues base* $\mathcal{DG}_{\mathcal{S}}(\mathbb{K})$ is a set of implications such that $\mathcal{DG}_{\mathcal{S}}(\mathbb{K})$ is sound for $\mathbb{K}$, $\mathcal{S} \cup \mathcal{DG}_{\mathcal{S}}(\mathbb{K})$ is complete for $\mathbb{K}$ and $\mathcal{DG}_{\mathcal{S}}(\mathbb{K})$ has minimal cardinality [7]. The underlying mathematics of $\mathcal{DG}(\mathbb{K})$ and $\mathcal{DG}_{\mathcal{S}}(\mathbb{K})$ are not relevant for this work. It is, however, important, that both bases can effectively be computed for every finite context $\mathbb{K}$.

The Fuzzy Setting [2] In a fuzzy context $\mathbb{K} = (G, M, I)$ the incidence relation I is a fuzzy relation, i.e. a fuzzy subset of $G \times M$. The derivation operators are defined for fuzzy subsets A of G and fuzzy subsets B of M as follows:

$$A^{\uparrow}(m) = \inf_{g \in G} \big(A(g)^* \Rightarrow I(g,m)\big), \qquad B^{\downarrow}(g) = \inf_{m \in M} \big(B(m) \Rightarrow I(g,m)\big). \tag{3}$$

Notice, that the hedge $\cdot^*$ is used only for the derivation of fuzzy sets of objects. The operators $\cdot^{\uparrow}$ and $\cdot^{\downarrow}$ form a Galois connection with hedges.

In fuzzy FCA the implications are also allowed to be fuzzy. A *fuzzy implication* is a pair written as $A \to B$ where A and B are fuzzy subsets of M. Let U be a fuzzy subset of M. The *degree to which $A \to B$ holds in U* is defined as

$$\|A \to B\|_U = S(A, U)^* \Rightarrow S(B, U) \tag{4}$$

The *degree to which $A \to B$ holds in $\mathbb{K}$* is defined as $\|A \to B\|_{\mathbb{K}} = \min_{g \in G} \|A \to B\|_{I_g}$, where I_g is the fuzzy set to which each $m \in M$ belongs with degree $I(g, m)$. Let $\mathcal{L}$ be a fuzzy set of fuzzy implications. A set $U \subseteq M$ is called a model of $\mathcal{L}$ if $\|A \to B\|_U \geq \mathcal{L}(A \to B)$ holds for every fuzzy implication $A \to B$. We say that $A \to B$ *follows from $\mathcal{L}$ to degree q* if $\|A \to B\|_U \geq q$ for all models U of $\mathcal{L}$. There have been several works where the existence of bases for fuzzy implications has been considered [8, 9]. We shall not go into details, however, we would like to point out two things. First, it can be shown that it suffices to consider *crisp* sets $\mathcal{L}$ that contain *fuzzy* GCIs [14]. Second, in this setting an effective algorithm for computing a base is known only when globalization is used as the hedge [8].

2.3 Fuzzy Description Logics

For DL we only introduce the fuzzy version. The crisp version only occurs in a high-level description in Section 3.1. For a formal introduction of crisp DL we refer to [15]. DL is not just one formalism, but a family of many knowledge representation formalisms. The observations in this work hold for any fuzzy DL that provides for conjunction, i.e. virtually all of them. For brevity we only introduce the lightweight DL called $\mathcal{EL}$. In fuzzy $\mathcal{EL}$ (exactly like in crisp $\mathcal{EL}$) *concept descriptions* can be formed from a set of *concept names* $\mathcal{N}_C$ and a set of *role names* $\mathcal{N}_R$ using the constructors $\top$, $\sqcap$ and $\exists$. More formally, $\top$ and all concept names are concept descriptions, and if C and D are concept description and r is a role name then $C \sqcap D$ and $\exists r.C$ are also concept descriptions.

In fuzzy $\mathcal{EL}$ (in contrast to crisp $\mathcal{EL}$) fuzzy sets are used to interpret both concepts and roles. A fuzzy interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ satisfies

$$A^{\mathcal{I}} \colon \Delta^{\mathcal{I}} \to [0,1], \qquad\qquad r^{\mathcal{I}} \colon \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \to [0,1].$$

for all $A \in \mathcal{N}_C$ and all $r \in \mathcal{N}_R$. Fuzzy interpretations $\mathcal{I}$ are extended to complex concept descriptions by defining $\top^{\mathcal{I}}(x) = 1$ and

$$(C \sqcap D)^{\mathcal{I}}(x) = C^{\mathcal{I}}(x) \otimes D^{\mathcal{I}}(x), \qquad (\exists s.C)^{\mathcal{I}}(x) = \sup_{z \in \Delta^{\mathcal{I}}} s^{\mathcal{I}}(x,z) \otimes C^{\mathcal{I}}(z) \tag{5}$$

for all $x \in \Delta^{\mathcal{I}}$. Fuzzy GCIs are typically written as $\langle C \sqsubseteq D, q \rangle$, where C, D are concept descriptions and $q \in [0, 1]$. The fuzzy GCI $\langle C \sqsubseteq D, q \rangle$ holds in the fuzzy interpretation $\mathcal{I}$ if all $x \in \Delta^{\mathcal{I}}$ satisfy $C^{\mathcal{I}}(x) \Rightarrow D^{\mathcal{I}}(x) \geq q$. The fuzzy interpretation $\mathcal{I}$ is a *model* of the set of fuzzy GCIs $\mathcal{T}$ if all fuzzy GCIs from $\mathcal{T}$ hold in $\mathcal{I}$. $\langle C \sqsubseteq D, q \rangle$ *is entailed by* $\mathcal{T}$ if it holds in all models of $\mathcal{T}$.

3 Comparison of the Two Formalisms

3.1 The Crisp Setting

Most existing works at the intersection of FCA and DL have in common that they associate FCA attributes and DL concept names. The objects are usually chosen to be domain elements of an interpretation [6, 7]. Other choices, such as selecting ABox individuals, usually require extending FCA theory, e.g. to allow for partial knowledge [5]. These choices are motivated by the following observation. Whether we compute the interpretation of the concept description $\mathsf{Large} \sqcap \mathsf{Populous} \sqcap \mathsf{Asian}$ or compute the derivation of the set of attributes $\{\mathsf{Large}, \mathsf{Populous}, \mathsf{Asian}\}$, the intuition in both cases is that we want to know which countries are large *and* populous *and* Asian.

To formalize this connection, for every interpretation $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$ one can define its *induced context* $\mathbb{K}_{\mathcal{I}}$ whose set of objects is $\Delta^{\mathcal{I}}$, whose attributes are the concept names and where $x \in \Delta^{\mathcal{I}}$ and $A \in \mathcal{N}_C$ are incident iff $x \in A^{\mathcal{I}}$. For example, we can think of the context in Table 1 as being induced by an interpretation $\mathcal{I}$ whose domain are the world's 8 most populous countries, and where the concept names $\mathsf{Populous}$, Large and Asian are interpreted as $\mathsf{Populous}^{\mathcal{I}} = \{\mathsf{China}, \mathsf{India}\}$, $\mathsf{Large}^{\mathcal{I}} = \{\mathsf{China}, \mathsf{Russia}, \mathsf{US}\}$, and $\mathsf{Asian}^{\mathcal{I}} = \{\mathsf{China}, \mathsf{India}, \mathsf{Indonesia}, \mathsf{Pakistan}, \mathsf{Russia}\}$.

In the induced context it holds for all sets $U \subseteq \mathcal{N}_C$ that $U^{\downarrow} = (\bigsqcap U)^{\mathcal{I}}$, further supporting the intuition that sets of attributes are treated like conjunctions over attributes. Similarly, for two sets of concept names $U, V \subseteq \mathcal{N}_C$ the GCI $\bigsqcap U \sqsubseteq \bigsqcap V$ holds in the interpretation $\mathcal{I}$ iff the implication $U \to V$ holds in the induced context of $\mathcal{I}$. Hence, the notions of dependencies also coincide in crisp FCA and crisp DL. One could even go so far to say that standard formal contexts and the very simple DL that only allows for conjunction are syntactic variants of each other.

3.2 The Fuzzy Setting

In this section, we analyze the differences between fuzzy FCA and fuzzy DL that hinder a close cooperation like it exists in the crisp setting. First, *the semantics of fuzzy FCA do not treat sets of attributes like conjunctions over concept names.* Remember that in the crisp setting the exact same semantics are used to compute the derivation of a set of attributes or the interpretation of a conjunction of concept names. This is not true in the fuzzy setting because the infimum (or minimum in the case of finite contexts) is used to interpret attribute sets (2) while

Table 1. Induced Context

	Large	Populous	Asian
Brazil			
China	×	×	×
India		×	×
Indonesia			×
Nigeria			
Pakistan			×
Russia	×		×
US	×		

Table 2. Induced Fuzzy Context

	Large	Populous	Asian
Brazil	0.5	0.14	0.0
China	0.56	1.0	1.0
India	0.19	0.9	1.0
Indonesia	0.11	0.18	0.76
Nigeria	0.05	0.13	0.0
Pakistan	0.05	0.13	1.0
Russia	1.0	0.11	0.75
US	0.58	0.23	0.0

the t-norm is used to interpret conjunctions (5).[3] In the case of the Gödel t-norm this is completely harmless, as the Gödel t-norm coincides with the minimum. For the other t-norms the difference is relevant.

As an example, assume that Table 2 is obtained from a fuzzy interpretation $\mathcal{I}$ by using the domain as objects, concept names as attributes and defining $I(x, A) = A^{\mathcal{I}}(x)$ (We could call it the *induced fuzzy context* of $\mathcal{I}$). If we use the Łukasiewicz t-norm, which is defined as $x \otimes y = \max\{0, x + y - 1\}$, then

$$(\mathsf{Populous} \sqcap \mathsf{Large})^{\mathcal{I}}(\mathsf{US}) = 0.23 \otimes 0.58 = 0$$

However, in fuzzy FCA with the Łukasiewicz t-norm we obtain

$$\{{}^{1}/\mathsf{Populous}, {}^{1}/\mathsf{Large}\}^{\downarrow}(\mathsf{US}) = \min\{1 \Rightarrow 0.23, 1 \Rightarrow 0.58\} = 0.23.$$

Thus, unlike in the crisp case, the fuzzy semantics differ even for crisp sets of attributes such as $\{{}^{1}/\mathsf{Populous}, {}^{1}/\mathsf{Large}\}$.

Second, we can observe that *in fuzzy DL concept descriptions on their own are not fuzzy.* The interpretations are fuzzy, the axioms are fuzzy, but the concept descriptions themselves are not. By contrast, in fuzzy FCA it is possible to use a fuzzy set of attributes to describe a class of objects.

To describe all countries that are (completely) huge and somewhat Asian, one can use a fuzzy set of attributes $\{{}^{1}/\mathsf{Large}, {}^{0.5}/\mathsf{Asian}\}$. Then in Table 2 the membership of Russia in the derivation $\{{}^{1}/\mathsf{Large}, {}^{0.5}/\mathsf{Asian}\}^{\downarrow}$ is 1. By contrast, in DL it is not possible to associate a truth degree with the concepts in a conjunction. The best approximation of the above attribute set in DL is the simple conjunction $\mathsf{Large} \sqcap \mathsf{Asian}$, which has, of course, a different semantics. In fact, Russia belongs to $(\mathsf{Large} \sqcap \mathsf{Asian})^{\mathcal{I}}$ only with degree 0.75.[4] In this respect fuzzy FCA is more expressive than fuzzy DL.

Finally, *fuzzy FCA typically uses hedges and fuzzy DL does not.* In principle, fuzzy FCA is more general here, since one could treat fuzzy DL as the special

[3] Some authors use two types of conjunction: a strong conjunction interpreted by the t-norm and a weak conjunction interpreted by the minimum. In this terminology, we could write that fuzzy DL uses strong conjunction while fuzzy FCA uses weak conjunction.

[4] The Gödel t-norm is used to emphasize that this is independent of the first problem.

case where identity is used as the hedge. In practice, if identity is used as the hedge, one cannot effectively compute a base in fuzzy FCA, at least not in the settings that have previously been considered.

On the other hand, using globalization in combination with crisp sets of attributes has practical limitations. Consider a fuzzy implication $A \to B$. Using the globalization as the hedge means, that all those counterexamples $g \in G$ are ignored that do not satisfy $S(A, I_g) = 1$. This is particularly problematic, if we only consider crisp left-hand sides A, since then

$$S(A, I_g) = \min_{m \in A}(1 \Rightarrow I(g, m)) = \min_{m \in A} I(g, m). \tag{6}$$

If for just one $m \in A$ the value $I(g, m)$ is not 1 then $S(A, I_g) < 1$ holds and the object g is ignored. For example, in Table 2 if we consider $A = \{1/\mathsf{Large}, 1/\mathsf{Populous}\}$ then all objects are ignored, i.e. any implication with A as its left-hand side holds. Presumably, in many applications values that differ from 1 are the rule rather than the exception, meaning that almost all objects will be ignored.

4 Bridging the Gap

In the previous section we have identified the three aspects in which fuzzy DL and fuzzy FCA disagree. We shall now consider a restricted subset of fuzzy FCA for which the semantics agree. Unfortunately, it is not possible to use strong conjunction instead of weak conjunction in fuzzy FCA, since the derivation operators would no longer form a Galois-connection. Instead, we caution that the following theory can only be applied directly in fuzzy DL with Gödel t-norm.

Since truly fuzzy implications have no equivalent in fuzzy DL, we only consider implications $A \to B$, where both sets A and B are crisp (from now on called *crisp implications*). A similar idea has been proposed under the name of "one-sided fuzzyness" in [16], with respect to concept lattices, not with respect to bases. Instead of trying to compute a base that is complete for all fuzzy implications we try to find a base that is complete only for crisp implications. In standard fuzzy FCA there is a result, that allows one to consider only crisp sets of fuzzy implications when searching for a base (Lemma 1 in [14]). Unfortunately, this result cannot be applied in our restricted setting. Instead of computing a crisp set containing fuzzy implications we compute a fuzzy set containing crisp implications:

Problem 1. Given a fuzzy context $\mathbb{K} = (G, M, I)$ compute a fuzzy subset T of $\{A \to B \mid A, B \subseteq M\}$ that is

- complete, i.e. for every implication $A \to B$, $A, B \subseteq M$, $\|A \to B\|_{\mathbb{K}} = q$ implies that $A \to B$ follows from T with degree q,
- sound, i.e. every implication $A \to B$ holds in $\mathbb{K}$ with degree at least $T(A \to B)$, and
- irredundant, i.e. no fuzzy set $U \subsetneq T$ is complete.

Furthermore, we use identity as the hedge, thereby ensuring both compatibility with DL and the use of all objects as potential counterexamples. These three restrictions – Gödel t-norm, identity as the hedge, only crisp implications – guarantee that $A \rightarrow B$ holds in the fuzzy induced context $\mathbb{K}$ of $\mathcal{I}$ to degree q iff $\langle \bigsqcap A \sqsubseteq \bigsqcap B, q\rangle$ holds in $\mathcal{I}$. This is analogous to the crisp case.

4.1 Axiomatization

In [14] an axiomatic system is presented that can be used to infer all fuzzy implications that follow from a crisp set of fuzzy implications. We present a similar system of deduction rules, which can be used to infer for each crisp implication the degree to which it follows from a fuzzy set of crisp implications.

Let $\mathcal{L}$ be a fuzzy subset of $\{A \rightarrow B \mid A, B \subseteq M\}$. Our axiomatic system consists of the following deduction rules, where q_1, q_2 are positive truth values. In each deduction step a new fuzzy subset $\mathcal{L}_{i+1}$ is obtained from the previous set $\mathcal{L}_i$, where $\mathcal{L}_0 = \mathcal{L}$. For all implications $E \rightarrow F$ we define $\mathcal{L}_{i+1}(E \rightarrow F) = \mathcal{L}_i(E \rightarrow F)$ unless mentioned otherwise in the rules.

(Refl) From $A \subseteq B$ and $\mathcal{L}_i(A \rightarrow B) < 1$ infer $\mathcal{L}_{i+1}(A \rightarrow B) = 1$
(Union) From $\mathcal{L}_i(A \rightarrow B) = q_1$, $\mathcal{L}_i(A \rightarrow C) = q_2$ and $\mathcal{L}_i(A \rightarrow B \cup C) < \min\{q_1, q_2\}$ infer $\mathcal{L}_{i+1}(A \rightarrow B \cup C) = \min\{q_1, q_2\}$
(Trans) From $\mathcal{L}_i(A \rightarrow B) = q_1$, $\mathcal{L}_i(B \rightarrow C) = q_2$ and $\mathcal{L}_i(A \rightarrow C) < q_1 \otimes q_2$ infer $\mathcal{L}_{i+1}(A \rightarrow C) = q_1 \otimes q_2$.

In each of the three rules the inferred implication obtains a membership degree that is smaller or equal to the membership degrees of the rules in the precondition. Since a rule can only be applied if the degree of the inferred implication strictly increases, no implication can ever be used in its own deduction implicitly or explicitly. There are only finitely many crisp implications and therefore the deduction process must terminate. We now want to show that the deduction system is *sound*, in the sense that if after a finite number k of deduction steps we can deduce $\mathcal{L}_k(A \rightarrow B) = q$ then $A \rightarrow B$ follows from $\mathcal{L}$ with at least degree q, and *complete* in the sense that if $A \rightarrow B$ follows from $\mathcal{L}$ with degree q then $\mathcal{L}_k(A \rightarrow B) = q$ can be deduced.

Lemma 1. *(Refl)–(Trans) is a sound and complete system of deduction rules.*

Proof. To prove soundness, we prove that each rule application does not change the models, i.e. that every model U of $\mathcal{L}_i$ is a model of $\mathcal{L}_{i+1}$. The converse that every model of $\mathcal{L}_{i+1}$ is a model of $\mathcal{L}_i$ is trivial, since $\mathcal{L}_i \subseteq \mathcal{L}_{i+1}$. Soundness of (Refl) is also trivial. *Soundness of (Union):* Assume that U is a model of $\mathcal{L}_i$. We define $\alpha = \min_{m \in A} U(m)$, $\beta = \min_{m \in B} U(m)$ and $\gamma = \min_{m \in C} U(m)$. From (6) and (4) we obtain

$$\|A \rightarrow B\|_U = \alpha \Rightarrow \beta,\ \|A \rightarrow C\|_U = \alpha \Rightarrow \gamma,\ \|A \rightarrow B \cup C\|_U = \alpha \Rightarrow \min\{\beta, \gamma\}.$$

Monotonicity of the residuum yields $\|A \rightarrow B \cup C\|_U = \min\{\alpha \Rightarrow \beta, \alpha \Rightarrow \gamma\} = \min\{q_1, q_2\}$. This proves that U is also a model of $\mathcal{L}_{i+1}$, which suffices to prove soundness of (Union). *Soundness of (Trans):* The preconditions can be rewritten as $\alpha \Rightarrow \beta = q_1$ and $\beta \Rightarrow \gamma = q_2$. Using (1) we obtain $\alpha \otimes q_1 \leq \beta$ and $\beta \otimes q_2 \leq \gamma$. From monotonicity of the t-norm we obtain $\alpha \otimes (q_1 \otimes q_2) \leq \gamma$. Using (1) again we get $q_1 \otimes q_2 \leq \alpha \Rightarrow \gamma = \|A \to C\|_U$. Hence, U is a model of $\mathcal{L}_{i+1}$, which proves soundness of (Trans).

Completeness: Let $X \to Y$ be an implication that follows (semantically) from $\mathcal{L}$ to degree q. Let $\mathcal{L}_k$ be the fuzzy set of implications obtained after exhaustively applying the deduction rules. To prove completeness it suffices to show that that $\mathcal{L}_k(X \to Y) \geq q$.

As a preliminary step, let us define the following fuzzy set $X^+(m) = \mathcal{L}_k(X \to \{m\})$ and show that it is a model of $\mathcal{L}$. Assume that X^+ is not a model of $\mathcal{L}$, i.e. $\|A \to B\|_{X^+} < \mathcal{L}(A \to B)$ for some implication $A \to B$. We use the notation $\alpha = \min_{m \in A} X^+(m)$ and $\beta = \min_{m \in B} X^+(m)$. Then $\|A \to B\|_{X^+} = \alpha \Rightarrow \beta < \mathcal{L}(A \to B)$, or equivalently by (1)

$$\alpha \otimes \mathcal{L}(A \to B) > \beta. \tag{7}$$

On the other hand $X^+(a) = \mathcal{L}_k(X \to \{a\}) \geq \alpha$ holds for all $a \in A$. Because the rules have been applied exhaustively to obtain $\mathcal{L}_k$ (Union) is not applicable to $\mathcal{L}_k$ and therefore $\mathcal{L}_k(X \to A) \geq \alpha$. Using a similar argument for (Trans) we obtain

$$\mathcal{L}_k(X \to B) \geq \mathcal{L}_k(X \to A) \otimes \mathcal{L}_k(A \to B) \geq \alpha \otimes \mathcal{L}(A \to B),$$

where we have exploited the fact that truth values can only increase when a rule is applied and therefore $\mathcal{L}(A \to B) \leq \mathcal{L}_k(A \to B)$. Finally, using (Refl) and (Trans) it follows that $\mathcal{L}_k(X \to \{b\}) \geq \alpha \otimes \mathcal{L}(A \to B)$ for all $b \in B$. This contradicts (7) and thus X^+ must be a model of $\mathcal{L}$.

Since $X \to Y$ follows from $\mathcal{L}$ to degree q it must hold that

$$q \leq \|X \to Y\|_{X^+} = \big(\min_{x \in X} X^+(x) \Rightarrow \min_{y \in Y} X^+(y)\big) = \min_{y \in Y} X^+(y).$$

Therefore $\mathcal{L}_k(X \to \{y\}) = X^+(y) \geq q$ for all $y \in Y$. Since (Union) cannot be applied to $\mathcal{L}_k$ we obtain $\mathcal{L}_k(X \to Y) \geq q$ which proves completeness. $\square$

4.2 Stem Base

We now provide a practical approach for computing a finite base for the Gödel t-norm, the only t-norm for which the semantics of fuzzy FCA and fuzzy DL coincide. Assume that we are given a finite fuzzy context $\mathbb{K} = (G, M, I)$. Let $Q_{\mathbb{K}}$ be the set containing 1 and all truth degrees that occur in $\mathbb{K}$. Let $q_0 \in [0, 1]$ be a fixed truth degree. We define a crisp context $\mathbb{K}_{q_0} = (G_{q_0}, M, I_{q_0})$ as follows. For each $g \in G$ and each $q \in Q_{\mathbb{K}}$ with $q < q_0$ the set G_{q_0} contains an object g_q with

$$\{g_q\}' = \{m \in M \mid I(g, m) > q\},$$

i.e. g_q has exactly those attributes that g has with degree higher than q. As an example, consider a context $\mathbb{K}$ of South American Countries (Table 3).[5]

[5] The value for HighGDP is the fraction of the country's GDP per capita and the GDP per capita of Chile, the largest in South America. Similarly for the other values.

Algorithm 1 Computing a Minimal Base with Gödel t-Norm

$\mathcal{B} = \mathcal{L} = \emptyset$
for all $q \in Q_{\mathbb{K}}$ in decreasing order **do**
  $\mathcal{D} = \mathcal{DG}_{\mathcal{B}}(\mathbb{K}_q)$
  $\mathcal{B} = \mathcal{B} \cup \mathcal{D}$
  $\mathcal{L} = \mathcal{L} \cup \{{}^{q}/_{A \to B} \mid A \to B \in \mathcal{D}\}$
end for
return $\mathcal{L}$

Lemma 2. *$A \to B$ holds in $\mathbb{K}$ with at least degree q_0 iff $A \to B$ holds in $\mathbb{K}_{q_0}$.*

Proof. Assume that $A \to B$ holds in $\mathbb{K}$ with degree less than q_0. According to (4) and the definition of the Gödel-residuum this is equivalent to

$$\begin{aligned}
&\min_{g \in G} \|A \to B\|_{I_g} < q_0 \\
\iff\ & \exists g \in G\colon \left(\min_{a \in A} I(g,a) \Rightarrow \min_{b \in B} I(g,b)\right) < q_0 \\
\iff\ & \exists g \in G\colon \min_{b \in B} I(g,b) < q_0 \text{ and } \min_{a \in A} I(g,a) > \min_{b \in B} I(g,b) \\
\iff\ & \exists g \in G\colon \exists b \in B\colon I(g,b) < q_0 \text{ and } \forall a \in A\colon I(g,a) > I(g,b).
\end{aligned} \tag{8}$$

$I(g,b)$ is a truth degree from $Q_{\mathbb{K}}$ and $g_{I(g,b)}$ satisfies $(g_{I(g,b)}, b) \notin I_{q_0}$ and $(g_{I(g,b)}, a) \in I_{q_0}$ for all $a \in A$. Therefore, $\mathbb{K}_{q_0}$ contains a counterexample to $A \to B$, hence $A \to B$ does not hold in $\mathbb{K}_{q_0}$.

On the other hand if $A \to B$ does not hold in $\mathbb{K}_{q_0}$ then there must be some g_q, $q < q_0$ such that $A \subseteq \{g_q\}'$ and $B \not\subseteq \{g_q\}'$. By definition of $\{g_q\}'$ this is equivalent to $I(g_q, a) > q$ for all $a \in A$ and $I(g_q, b) \leq q$ for some $b \in B$. Since $q < q_0$ it holds that for this value b in particular $I(g_q, b) < q_0$ and $I(g_q, a) > I(g_q, b)$ for all $a \in A$. It then follows from (8) that $A \to B$ does not hold in $\mathbb{K}$ with at least degree q_0. □

Notice, that if $q_1 < q_0$ then $G_{q_1} \subseteq G_{q_0}$ and $I_{q_1} \subseteq I_{q_0}$. This observation, together with Lemma 2, suggests a levelwise approach as sketched in Algorithm 1. One starts with the largest value $q_{\max}$ in $Q_{\mathbb{K}}$ and computes the Duquenne-Guigues Base for $\mathbb{K}_{q_{\max}}$. The base serves two purposes. Its implications are added to the fuzzy set of implications $\mathcal{L}$ with degree q, and it serves as background knowledge in the next iteration. For the context from Table 3 Algorithm 1 yields the base $\{{}^{1}/_{\{\mathsf{Populous},\mathsf{Small}\} \to \{\mathsf{HighGDP}\}}, {}^{0.9}/_{\{\mathsf{Populous}\} \to \{\mathsf{HighGDP}\}}, {}^{0.2}/_{\emptyset \to \{\mathsf{HighGDP}\}}\}$.

Lemma 3. *Upon termination Algorithm 1 returns a fuzzy set of crisp implications $\mathcal{L}$ that is sound and complete for $\mathbb{K}_q$ and has minimal cardinality among all such sets.*

Proof. Soundness follows immediately from Lemma 8. To prove completeness, assume that $U \to V$ holds in $\mathbb{K}$ with degree $q \in Q_{\mathbb{K}}$ (notice that for the Gödel t-norm it always holds that $\|U \to V\|_{\mathbb{K}} \in Q_{\mathbb{K}}$). Then by Lemma 8 $U \to V$ holds

Table 3. South American Countries

	Populous	HighGDP	Small
Argentina	0.2	0.8	0.6
Bolivia	0.1	0.2	0.9
Brazil	1.0	0.9	0.0
Chile	0.1	1.0	0.9
Colombia	0.2	0.5	0.9
Ecuador	0.1	0.3	1.0
Guyana	0.0	0.2	1.0
Paraguay	0.0	0.2	1.0
Suriname	0.0	0.5	1.0
Uruguay	0.0	1.0	1.0
Venezuela	0.1	0.7	0.9

Table 4. $\mathbb{K}_1$

	Populous	HighGDP	Small
Argentina$_{0.6}$		×	
Argentina$_{0.2}$		×	×
Bolivia$_{0.2}$			×
Bolivia$_{0.1}$		×	×
Brazil$_{0.9}$	×		
Brazil$_{0.0}$	×	×	
Chile$_{0.9}$		×	
Chile$_{0.1}$		×	×
Colombia$_{0.5}$			×
Colombia$_{0.2}$		×	×
Ecuador$_{0.3}$			×
Ecuador$_{0.1}$		×	×
Guyana$_{0.2}$			×
Guyana$_{0.0}$		×	×
Paraguay$_{0.2}$			×
Paraguay$_{0.0}$		×	×
Suriname$_{0.5}$			×
Suriname$_{0.0}$		×	×
Uruguay$_{0.0}$		×	×
Venezuela$_{0.7}$			×
Venezuela$_{0.1}$		×	×

in $\mathbb{K}_q$. Consider $\mathcal{D}$, $\mathcal{V}$ and $\mathcal{L}$ after the iteration for q in Algorithm 1. Since $\mathcal{D}\cup\mathcal{B}$ is complete for $\mathbb{K}_q$ and $U \to V$ holds in $\mathbb{K}_q$ the implication $U \to V$ follows from $\mathcal{D}\cup\mathcal{B} = \{A \to B \mid \mathcal{L}(A \to B) \geq q\}$ in the crisp setting.

We show that then $U \to V$ follows to degree q from $\mathcal{L}$ in the fuzzy setting. Assume the contrary, i.e. that there exists a context $\bar{K}$ for which $\mathcal{L}$ is sound, but in which $U \to V$ does not hold to degree at least q. By Lemma 8 this yields that $U \to V$ does not hold in $\bar{\mathbb{K}}_q$ while all implications from $\mathcal{D}\cup\mathcal{B} = \{A \to B \mid \mathcal{L}(A \to B) \geq q\}$ do hold in $\bar{\mathbb{K}}_q$. Because we have shown that $U \to V$ follows from $\mathcal{D}\cup\mathcal{B}$ in the crisp setting this is a contradiction. Hence $U \to V$ follows to degree q from $\mathcal{L}$, which proves completeness.

Assume that $\bar{\mathcal{L}}$ is another sound and complete fuzzy set of implications for $\mathbb{K}$. Then by Lemma 8 for each $q \in Q_{\mathbb{K}}$ the crisp set

$$\bar{\mathcal{L}}_q = \{A \to B \mid \mathcal{L}(A \to B) \geq q\}$$

must be sound and complete for $\mathbb{K}_q$. A simple induction over $q \in Q_{\mathbb{K}}$ can be used to show that $|\bar{\mathcal{L}}_q| \geq |\mathcal{L}_q|$ for all $q \in Q_{\mathbb{K}}$. For q maximal in $Q_{\mathbb{K}}$ the claim follows directly from minimality of the Duquenne-Guigues Base. For the induction step let $q \in Q_{\mathbb{K}}$ where $|\bar{\mathcal{L}}_{\bar{q}}| \geq |\mathcal{L}_{\bar{q}}|$ holds for the next larger value $\bar{q} \in Q_{\mathbb{K}}$. Both $\mathcal{L}_{\bar{q}}$ and $\bar{\mathcal{L}}_{\bar{q}}$ are sound and complete for $\mathbb{K}_{\bar{q}}$, in particular they have the same models. Thus, $\bar{\mathcal{L}}_q = \big(\bar{\mathcal{L}}_q \setminus \bar{\mathcal{L}}_{\bar{q}}\big) \cup \bar{\mathcal{L}}_{\bar{q}}$ and $\big(\bar{\mathcal{L}}_q \setminus \bar{\mathcal{L}}_{\bar{q}}\big) \cup \mathcal{L}_{\bar{q}}$ also have the same models, and are thus both sound and complete for $\mathbb{K}_q$. Minimality of the $\mathcal{L}_{\bar{q}}$-Duquenne-Guigues base implies $|\bar{\mathcal{L}}_q \setminus \bar{\mathcal{L}}_{\bar{q}}| \geq |\mathcal{DG}_{\mathcal{L}_{\bar{q}}}(\mathbb{K}_q)|$. This proves

$$|\bar{\mathcal{L}}_q| = |\big(\bar{\mathcal{L}}_q \setminus \bar{\mathcal{L}}_{\bar{q}}\big)| + |\bar{\mathcal{L}}_{\bar{q}}| \geq |\mathcal{DG}_{\mathcal{L}_{\bar{q}}}(\mathbb{K}_q)| + |\mathcal{L}_{\bar{q}}| = |\mathcal{L}_q|.$$

Since this holds for all $q \in Q_{\mathbb{K}}$ we get that $\mathcal{L}$ has minimal cardinality among all bases. □

5 Conclusion

We have restricted fuzzy FCA by allowing only crisp sets of attributes in the implications and using identity as the hedge. We have presented a sound and complete set of deduction rules for this restricted setting. For the Gödel t-norm the restricted setting corresponds semantically to fuzzy DL. Furthermore, we have presented a simple algorithm for computing a minimal base for the restricted setting. In the general setting this is only possible with globalization.

We do not claim, that this restriction of expressivity is the only feasible approach for reconciling the differences between the two fields. In future work it would be interesting to look at a kind of weighted conjunction in DL (imitating the semantics of fuzzy attribute sets). It would also be interesting to consider a version of fuzzy FCA that uses strong conjunction.

References

1. Hájek, P.: Metamathematics of Fuzzy Logic (Trends in Logic). Springer (2001)
2. Belohlavek, R.: Fuzzy Relational Systems: Foundations and Principles. Kluwer (2002)
3. García-Cerdaña, Á., Armengol, E., Esteva, F.: Fuzzy description logics and t-norm based fuzzy logics. Int. J. of Approx. Reasoning **51** (2010) 632–655
4. Sertkaya, B.: Computing the hierarchy of conjunctions of concept names and their negations in a description logic knowledge base using formal concept analysis. In: Cont. to ICFCA'06, Dresden, Germany (2006) 73–86
5. Baader, F., Ganter, B., Sattler, U., Sertkaya, B.: Completing Description Logic knowledge bases using Formal Concept Analysis. In: IJCAI'07. (2007)
6. Rudolph, S.: Exploring relational structures via FLE. In: Conceptual Structures at Work. Springer (2004) 233–233
7. Distel, F.: Learning Description Logic Knowledge Bases from Data Using Methods from Formal Concept Analysis. PhD thesis, TU Dresden (2011)
8. Belohlavek, R., Vychodil, V.: Attribute implications in a fuzzy setting. In: ICFCA'06. Springer (2006) 45–60
9. Belohlavek, R., Chlupova, M., Vychodil, V.: Implications from data with fuzzy attributes. In: AISTA'04. (2004)
10. Bobillo, F., Delgado, M., Gómez-Romero, J., Straccia, U.: Fuzzy description logics under Gödel semantics. Int. J. of Approx. Reasoning **50**(3) (2009) 494–514
11. Ganter, B., Wille, R.: Formal Concept Analysis: Mathematical Foundations. Springer, New York (1997)
12. Guigues, J.L., Duquenne, V.: Familles minimales d'implications informatives résultant d'un tableau de données binaires. Math. Sci. Humaines **95** (1986) 5–18
13. Stumme, G.: Attribute exploration with background implications and exceptions. In: Data Analysis and Inf. Sys., Berlin, Springer (1996) 457ff
14. Belohlavek, R., Vychodil, V.: Axiomatizations of fuzzy attribute logic. In: IICAI'05. (2005)
15. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.F., eds.: The Description Logic Handbook: Theory, Implementation, and Applications. Cambridge University Press (2003)
16. Krajči S.: Cluster based efficient generation of fuzzy concepts. Neural Network World **5** (2003), 521–530