
Peer reviewed version

License (if available):
Other

Link to published version (if available):
10.1109/SMC.2018.00443

Link to publication record in Explore Bristol Research
PDF-document

This is the accepted author manuscript (AAM). The final published version (version of record) is available online via OSA at https://doi.org/10.1364/OL.44.000275. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/
Graded Concepts for Collaborative Intelligence

T P Martin *, 1, B Azvine * 2

* Machine Intelligence and Uncertainty Lab, Engineering Maths, University of Bristol, Bristol, BS8 1UB, UK
1 trevor.martin@bristol.ac.uk

* Security Futures Practice, BT Research & Innovation, Adastral Park, Ipswich, IP5 3RE, UK
2 ben.azvine@bt.com

Abstract — Collaborative intelligence involves a combination of human and machine-based analysis, in which humans focus on higher-level tasks involving insight and understanding, whilst machines deal with gathering, filtering and processing data into a convenient and understandable form. We have proposed the use of graded concept lattices as a representation for exchanging information between machine and human in a collaborative intelligent system. Graded concepts allow summarization at multiple levels of discernibility (granularity). In this paper, we outline a new interpretation of fuzzy concept lattices as graded sets of crisp concepts. In addition, we prove equivalence between graded (fuzzy) formal concept analysis and the standard crisp framework. Consequently, any software tools developed for crisp data can be extended to the graded case without change.

I. INTRODUCTION

Collaborative intelligence is a term used to describe any system where processing is shared between humans and machines. It can be traced back to multi-agent systems where computational tasks were distributed amongst processors - either in a heterogeneous manner where each component had a specific “expertise” or in a homogeneous fashion, where a problem was divided into similar, but smaller, sub-problems, processed separately. Many systems are founded (formally or informally) on the idea of collaborative intelligence - an obvious example is the idea of querying large datasets to gain some insight into a topic, then refining queries once further understanding is gained. This is an iterative process - the machine contributes speed of searching/indexing, while the human contributes insight and conceptual organisation.

With the increasing volume of data generated by online systems (internet logs, transaction records, communication records, transport networks, sensor networks, etc.), machine assistance is needed in filtering, fusing and finding causal and other relations in data flows. Collaborative intelligence can be used to monitor and proactively control the behaviour of a complex system such as a computer network, and includes prediction, perception and comprehension of data (in a broad sense) on a fluid basis that evolves over time. Collaborative intelligence relies on simple, effective and efficient communication of information between the “processing components” i.e. computers and human analysts. A key feature in enhancing human understanding of large scale data is the notion of summarisation. By combining many values (or objects) into a few entities, concise summaries enable analysts to gain insight into bulk patterns and focus on the mechanisms underlying relations in the data, leading to greater understanding of current data and prediction of future data.

A. Knowledge Representation for Collaborative Intelligence

To function effectively, the "components" of a collaborative intelligent system must be able to communicate efficiently. However, there is a fundamental difference in the knowledge representations used. Machine processing is generally centred on well-defined entities and relations, ranging from the flat table structures of database systems through graph-based representations and up to ontological approaches involving formal logics. On the other hand, human language and communication is based on a degree of vagueness and ambiguity that leads to an efficient transmission of information between humans without the need for precise definition of every term used. Even quantities that can be measured precisely (height of a person or building, volume of a sound, amount of rainfall, colour of an object, etc.) are usually described in non-precise terms such as tall, loud, quite heavy, dark green, etc. More abstract properties such as beautiful landscape, delicious food, pleasant weather, clear documentation, corporate social responsibility, are essentially ill-defined, whether they are based on a holistic assessment or reduced to a combination of lower-level, measurable quantities.

Zadeh’s initial formulation of fuzzy sets [1] was inspired primarily by the flexibility of definitions in natural language. He argued that most natural language terms (concepts) admit “graded membership”, in that it is possible to compare two objects and to say whether or not one belongs more strongly to the concept. Clearly in the case of an elementary quantity such as height, we are generally able to say that person-1 satisfies the concept “tall” better than person-2 (or that they satisfy the concept equally well). Such gradation can be confirmed by measurement, if necessary. However, it is also valid to speak of graded membership in the case of more complex concepts such as those listed above. We can generally rank the membership of different objects in the set representing the concept extension - in other words, the concept extension can be modelled as a fuzzy set. The interval [0, 1] is a convenient range for the membership function. It maps naturally to a scale where definite membership can be represented by 1, non-membership by 0, with intermediate values used to reflect the lack of a precise border between the two extremes. However, the fundamental idea is that membership is ordered, not the precise membership on a scale.

In addition to the use of flexible terms, human reasoning is characterised by an ability to switch between different levels of granularity when dealing with a problem - for example, at
Formal Concept Analysis (FCA) extracts the implicit structure in data using an approach based on lattice theory. The basic method uses data in binary object-attribute-value form and groups objects into sets which have shared attributes. These sets form a lattice, which can be used for association rule analysis, ontology creation, knowledge representation and other tasks in data mining. The novel contributions of this paper are twofold. We introduce a new interpretation of fuzzy formal concept analysis, in which a fuzzy lattice is regarded as a fuzzy set of crisp lattices. The main result is to show that a fuzzy version of formal concept analysis can be reduced to a set of crisp calculations. The approach converts each fuzzy set of objects to a crisp set (of object-membership pairs), and we prove that a crisp concept lattice derived in this way is isomorphic to the fuzzy concept lattice derived by using our previous formulation [5]. By converting the fuzzy problem to a crisp equivalent, we enable existing software and tools to be adapted for fuzzy use without any change in the underlying code.

II. BACKGROUND

A. Formal Concept Analysis

Formal concept analysis (FCA) extracts the implicit structure in data using an approach based on lattice theory. The basic method uses data in binary object-attribute-value form and groups objects into sets which have shared attributes. These sets form a lattice, which can be used for association rule analysis, ontology creation, knowledge representation and other tasks in data mining. The standard formulation starts with a formal context \((O, A, R)\), where

- \(O\) is a set of objects,
- \(A\) is a set of attributes and
- \(R\) is a relation.

\[
R \subseteq O \times A
\]  

(1)

For \(X\), a subset of objects and \(Y\), a subset of the attributes \(X \subseteq O\) and \(Y \subseteq A\),

the operators \(\uparrow\) and \(\downarrow\) are defined as follows:

\[
X\uparrow = \{y \in Y | \forall x \in X : (x,y) \in R\}
\]

\[
Y\downarrow = \{x \in X | \forall y \in Y : (x,y) \in R\}
\]

(2)

Any pair \((X, Y)\) such that \(X\uparrow = Y\) and \(Y\downarrow = X\) is a formal concept with extension \(X\) (the set of objects belonging to the concept) and intension \(Y\) (the set of properties defining the concept). For example, Table I shows a relation on the set of objects \(\{o1, o2, o3\}\) and attributes \(\{a1, a2, a3\}\) which has the formal concepts:

<table>
<thead>
<tr>
<th></th>
<th>a1</th>
<th>a2</th>
<th>a3</th>
</tr>
</thead>
<tbody>
<tr>
<td>o1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>o2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>o3</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Using the second and fourth concepts for illustration, the object \(o2\) is the only one to have both attributes \(a1\) and \(a2\), objects \(o1\) and \(o2\) are the only objects to have attribute \(a1\) etc. Fig 1 shows the lattice of formal concepts derived from the relation.

![Fig. 1.](image)

To simplify the lattice view of formal concepts, it is normal to adopt reduced labelling [2] as shown in Fig 1. Instead of listing the entire extension and intension of each concept, we adopt the convention that the intension of a concept in the lattice is the attribute label attached to the concept AND all attributes attached to its super-concepts. In an analogous manner, the extension of a concept is the object label attached to the concept and all objects attached to its sub-concepts. Each object and attribute thus appears exactly once, at the greatest lower bound of concepts whose extension contains an object or least upper bound of concepts whose intension contains an attribute respectively. The lattice reflects the natural partial ordering on concepts arising from subset (resp. superset) relations between concept extensions (resp. intensions). The approach is closely linked to frequent itemset discovery, often used as the basis of association rule-mining [6] - for instance, in Fig 1 we see that all objects having attribute \(a2\) also have attribute \(a1\).

Formal concept analysis is extended to non-binary attributes by scaling, i.e. creation of one or more binary
attributes which discretise or summarise the underlying values. For example, an attribute taking numerical values in the range 0-10 might be replaced by two mutually exclusive binary attributes $A_{0.5}$ and $A_{5.10}$ (or a finer partition if required).

### B. Fuzzy Formal Concept Analysis

The extension of FCA based on a propositional table to the fuzzy propositional case is straightforward - instead of a table entry indicating that an object has (1) or does not have (0) an attribute, intermediate values are allowed to indicate the degree to which an attribute is applicable - in other words, the relation $R$ in eq (1) is generalised from a binary to a fuzzy relation with membership function

$$\mu_R:O \times A \rightarrow [0,1]$$

Formal treatments such as [7, 8] generalise both concept extension and intension to fuzzy sets. In previous work [5, 9, 10] we have adopted a formulation of fuzzy FCA in which a fuzzy formal concept is defined as a pair $(X, Y)$ where $X$ is a fuzzy set of objects and $Y$ is a crisp set of attributes. As in the crisp case, concepts satisfy

$$X^\uparrow = Y \quad \text{and} \quad Y^\downarrow = X$$

where we adopt the usual definition of equality for fuzzy sets (any element has identical membership in both sets) and

$$X^\uparrow = \{ y \in Y \mid \forall x \in X : \mu_x(x, y) \geq \mu_x(x) \}$$

$$Y^\downarrow = \{ x \mid \mu_x(x) = \min_{y \in Y} \mu_x(x, y) \}$$

(crisp sets of objects and fuzzy sets of attributes arise from a dual of these operators). An efficient algorithm for calculating concepts from a fuzzy context table is presented in [5].

### C. Simple Example

Table II shows a small illustrative fuzzy context, with four objects $\{o_1, o_2, o_3, o_4\}$ and four attributes $\{a_1, a_2, a_3, a_4\}$. The associated lattice is shown in Fig 2 with full labelling, and in Fig. 3 with reduced fuzzy labelling.

#### Table II. A Fuzzy Context

<table>
<thead>
<tr>
<th></th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$o_1$</td>
<td>1</td>
<td>0.8</td>
<td>0</td>
<td>0.8</td>
</tr>
<tr>
<td>$o_2$</td>
<td>1</td>
<td>0</td>
<td>0.7</td>
<td>0</td>
</tr>
<tr>
<td>$o_3$</td>
<td>0</td>
<td>0</td>
<td>0.5</td>
<td>1</td>
</tr>
<tr>
<td>$o_4$</td>
<td>0.4</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
</tbody>
</table>

The reduced labelling for the fuzzy lattice (Fig 3) shows each element/membership pair exactly once. The extension of a concept is the fuzzy union of the object label attached to a concept and the extensions of all its sub-concepts.

The reduced labelling for a fuzzy lattice is defined as follows. For any concept $C = (ext(C), int(C))$, we associate the attribute label $a$ with the concept if

$$a \in int(C) \land \forall C, C > C : a \notin int(C)$$

and we associate the object label and membership $o / \mu_{ext(o)}(o)$ with the concept if

$$\forall o \in O, \forall C \subset C : \mu_{ext(C)}(o) < \mu_{ext(C)}(o)$$

for non-zero memberships. Note that in some cases, this leads to nodes with no associated label, as shown in Fig. 3.
III. A GRADED INTERPRETATION OF FUZZY FCA

More recent work [11] has combined the notions of formal concept analysis with graded tolerance relations. A concept is a set of objects, which cannot be distinguished on the basis of the describing attributes (intension). Given a partition of attribute values, equivalence classes naturally divide the objects into non-overlapping sets, each of which contains objects that are indiscernible on the basis of the attribute. For graded partitions, we obtain a nested sequence of lattices. We form a concept lattice using standard techniques - however, the lattice varies according to the membership grade used to create the partition. For example, Fig 4 shows the concept lattices at \( \alpha=0.7 \) and \( \alpha=1 \). We note that using graded equivalence relations to define concepts automatically ensures that the graded concepts are nested.

The fuzzy FCA lattices obtained using eqs (3) and (4) are often complex and difficult to interpret visually. This problem is shared with crisp FCA, but is more pronounced because each object gives rise to several labels (one label for each distinct membership level for the object in the table).

In the graded view proposed here, we treat the concept lattice as a crisp, membership-dependent structure. The conversion of a general function using fuzzy objects (such as sets, intervals, etc.) to a crisp function with an explicit dependence on membership is described in [12] and builds on our earlier work[13] and (in part) the work of [14, 15].

In this interpretation, we view the conventional FCA process of lattice construction as a mapping

\[
C : R(O, A) \rightarrow (L : P(O) \rightarrow P(A))
\]

where \( R \) is the relation as defined in eq (1), \( P(O), P(A) \) are the powersets of objects \( O \) and attributes \( A \) respectively, and \( L \) is an anti-monotonic 1-1 function (usually a partial function) representing the lattice.

\[
L(X) = \begin{cases} 
X^1 & \text{if } X^1 = X \\
\text{undefined} & \text{otherwise}
\end{cases}
\]  

(5)

For example, in Fig 2,

\[
L(\{o1, o2\}) = \{a1\}, \quad L(\{o1\}) = \{a1, a3\}, \quad \text{etc.}
\]

Similarly, for a fuzzy relation \( R_f \) we can write

\[
C_f : R_f(O, A) \rightarrow (L_f : \tilde{P}(O) \rightarrow P(A))
\]

where \( \tilde{P}(O) \) is the set of fuzzy subsets of \( O \) and \( L_f \) is a fuzzy lattice, defined by eqs (3, 4).

Adopting the \( X\mu \) approach [12], we have a mapping

\[
C_x : R_x(O, A) \rightarrow (X : (0,1) \rightarrow (L : P(O) \rightarrow P(A)))
\]

i.e. from a fuzzy context (table) to a function, \( X \), which takes us from membership (alpha-cut) to a crisp lattice. In other words if \( X(\alpha) \) maps to \( L_{\alpha} \) representing the set of concepts at membership grade \( \alpha \) and \( L_{\alpha} \) is the same at membership grade \( \alpha \) such that \( \alpha_1 > \alpha \), then each object concept \( (X, Y) \) in \( L_{\alpha} \) has a parent object concept \( (X_p, Y_p) \) in \( L_{\alpha_1} \), such that \( X_p \subseteq X \) and each object concept \( (X_p, Y_p) \) in \( L_{\alpha_1} \) has one or more child object concepts in \( L_{\alpha} \). We illustrate how this can be calculated and related to the fuzzy lattice, although we do not present a formal treatment of the problem.

The fuzzy concept lattice is treated as a graded set of crisp lattices, where for any membership value in \((0,1] \) we have a corresponding crisp lattice as shown in Fig. 4. These lattices correspond to membership grades of at least 0.7 and 1 respectively. Clearly the lattices are related to each other (for clarity, the lattice diagrams use the same position for corresponding concepts). They are also related to the full fuzzy concept lattice in Figs 2 and 3 which can be generated from the graded lattices by taking the union of all graded lattices with the appropriate memberships. This approach fits well with incremental methods for generating concept lattices such as [17], as the core lattice an be created quickly, with new concepts introduced as the alpha-threshold is reduced.

![Fig. 4. Graded formal concept lattices from the extended context in Table III at membership ≥ 0.7 (top) and membership 1 (bottom)](image)

IV. EXTENDED CRISP CONTEXTS

When converting a fuzzy concept lattice to its crisp equivalent, we derive a crisp lattice on an extended set of objects which is formed by considering the set of pairs \( (o,m) \) where \( o \) is an object and \( m \) is in \( \Lambda \), the level set of the fuzzy relation \( R_f \)

\[
\Lambda = \{ \alpha | \mu_X(x,y) = \alpha \text{ for some } (x,y) \in O \times A \}
\]
In this case we have the lattice construction operation:

\[ C_e : R_e(O, A) \rightarrow (L_e : P(O \times A) \rightarrow P(A)) \]

| TABLE III. AN EXTENDED CONTEXT CORRESPONDING TO TABLE II |
|------------------|------------------|------------------|------------------|
| o1, 1            | a1               | 1                | 0                |
| o1, 0.8          | a2               | 1                | 1                |
| o2, 1            | a3               | 0                | 0                |
| o2, 0.7          | a4               | 1                | 0                |
| o3, 1            | a5               | 0                | 0                |
| o3, 0.5          | a6               | 0                | 1                |
| o4, 1            | a7               | 0                | 1                |
| o4, 0.5          | a8               | 1                | 1                |
| o4, 0.4          | a9               | 1                | 1                |

Fig. 5. Formal concept lattice from the extended context in Table III

In this formulation, introduced in [5], we replace the fuzzy extension with a crisp extension based on element-membeship pairs as follows. As before, let \( O \) be a set of objects, \( A \) be a set of attributes, and define a relation \( R \) as in eq (1). \( X \), \( Y \) are subsets of \( O, A \) respectively. We use subscripts \( e \) and \( f \) to indicate the extended and fuzzy cases respectively.

Instead of the objects \( O \), the extended context is defined on a set of pairs:

\[ O_e = \{(o_i, m_i) \mid \exists a_j : \mu_e(o_i, a_j) = \mu_e(o_i, a_j) \land m_i \leq \mu_e(o_i, a_j)\} \]  

\[ R_{ext} = \{((o_i, m_i), a_j) \mid \exists a_j : \mu_e(o_i, a_j) \land m_i \leq \mu_e(o_i, a_j)\} \]  

For example, Table III shows the extended fuzzy context corresponding to Table II. There is at least one row in Table III for each row in Table II, i.e. there is at least one element in the extended context for each \( i \) where there is a \( j \) such that \( \mu_e(o_i, a_j) > 0 \) in the fuzzy context. Also

\[ \mu_e(o_i, a_j) = \max_i \{m_i \mid ((o_i, m_i), a_j) \in R_e\} \]  

(7)

Applying the crisp definitions to the extended context:

\[ X \subseteq O_e \]
\[ Y \subseteq A \]

we have

\[ X^f = \\{y \in A \mid \forall (x, m) \in X : ((x, m), y) \in R_e\} \]  

(8)

\[ Y^f = \{((x, m), y) \in O \mid \forall y \in Y : ((x, m), y) \in R_e\} \]  

(9)

These operators applied to the extended context (Table III) do not give an identical set of concepts to those derived from the fuzzy context (Table II) with the fuzzy operators (eqs 3, 4). However, the intensions of the concepts are identical, leading to isomorphic lattices. In the next section, we prove this claim.

A. Equivalence of Extended and Fuzzy Contexts

**Theorem**: the lattice of extended concepts is isomorphic to the lattice of fuzzy concepts i.e. there is a bijection between fuzzy and extended concepts preserving the concept order.

Since the set of attributes, \( A \), is the same in the fuzzy and extended contexts, we prove that

\[ (Y^f, Y) \]

is an extended concept iff

\[ (Y^i, Y) \]

is a fuzzy concept (we use \( f \) to indicate that the operation is with respect to the fuzzy context, \( e \) for the extended context).

**Proof**: let \( (Y^i, Y) \) be a fuzzy concept. By definition,

\[ Y^i \uparrow Y = Y \]  

(10)

Now consider the closure of \( Y \) with respect to the extended context and use proof by contradiction to show that \( Y \) is the intension of a formal concept in the extended context.

Assume \( (Y^i, Y) \) is not a concept, i.e. \( Y \neq Y^i \) so that there is some \( y^* \in Y^i \) where \( y^* \notin Y \) Then

\[ \forall y \in Y : ((x, m), y) \in R_e \implies ((x, m), y^*) \in R_e \]  

by 8,9

i.e. any \( (x, m) \) pair with all attributes in \( Y \) also has attribute \( y^* \) Hence

\[ \forall o \in O : \mu_e(o, y^*) \geq \min \max \{m_{i} \mid ((o, m), y) \in R_e\} \]  

by (7)

where

\[ \min(\phi) = 0 \]

Hence \( y^* \in Y^i \) by (4)

and so \( y^* \in Y \) by (10)

This contradicts our assumption, hence \( Y = Y^i \) so that if \( (Y^i, Y) \) is a fuzzy concept then \( Y^i \) is an extended concept.
A similar argument can be used to show that the reverse implication holds, hence

\[
(Y^+, Y) \text{ is an extended concept iff } (Y^+, Y) \text{ is a fuzzy concept.}
\]

Therefore the lattices are isomorphic. Note that only the intensions are identical. Fig. 5 shows the lattice derived from the extended concept derived from the example fuzzy context in Table II. The lattice in Fig 5 has the same concept intensions as the fuzzy lattice in Fig 3 (and 2), although the concept extensions are obviously different.

For a given concept intension, the extension of the fuzzy concept can be derived straightforwardly from the extension of the extended concept - for example, considering the concept with intension \( a_3 \) we have concept extensions

\[
\{o2/0.7, o3/0.5, o4/1\} \quad \text{in the fuzzy case and}
\]

\[
\{ (o2,0.7), (o3,0.5), (o4,0.4), (o4,0.5), (o4,1) \}
\]

in the extended case. We perform a simple transformation of a fuzzy context, use any standard (crisp) software package to create the concept lattice and, if necessary, transform these back to fuzzy concepts.

We also note that the lattice in Fig 5 has the same concept intensions as the union of lattices shown in Fig 4. It is possible to prove equivalence of the \( X-\mu \) and extended approach.

V. APPLICATION

A specific application area is in the analysis of cyber attacks where we use an ontology such as https://capec.mitre.org/ to categorise events into hierarchical classes. The category labels are fuzzy (in the sense that events may belong more or less strongly to a category, and the approach described in this paper allows us to extend existing software to the case of fuzzy membership in categories, without needing to modify the underlying software. Fig 6 shows a fragment of the concept lattice used.

VI. SUMMARY

Fuzzy FCA is an enabler for the exchange of information between humans and machines in a collaborative intelligent system. It can model many of the soft definitions used in natural language and also the different levels of granularity used. In this paper, we have outlined two ways in which standard (crisp) FCA algorithms can be extended to fuzzy contexts. They key result is that any software designed for crisp (non-fuzzy) data can be used for fuzzy data without modification. On-going research in this area is examining the extension to triadic concept lattices [18], allowing relational tables rather than the scaled propositional tables used here.

REFERENCES


Fig 6 - Fragment of a formal concept lattice for analysis of cyberattacks