PARTIAL METRIC ON FINITE FORMAL CONTEXTS

Alena Chernikava
Doctoral Degree Programme (4), FEEC BUT
E-mail: xcerni07@stud.feec.vutbr.cz

Supervised by: Martin Kovár
E-mail: kovar@feec.vutbr.cz

Abstract: In this article we propose an interaction of the following areas of mathematics: formal concept analysis and partial metrics. Rudolf Wille in 1982 introduced a formal concept analysis (FCA) as an attempt of restructuring lattice theory [2]. FCA works with data and data is described with a binary relation between an object set and an attribute set. The attribute set, the object set and the relation between them form a triple called a formal context. The formal context forms a mathematical structure that is a source of some information (obvious and hidden). Knowledge could be extracted from formal contexts in many different ways; thus, many different topologies could be defined on a formal context. Because information stored in a formal context mostly is not total information about objects, we could describe a formal context with a partial metric. A partial metric was introduced by Steve Matthews in 1992. The main idea was to generalize metric by “dropping” the first axiom of the metric (zero self-distance). This approach could be used in Computer Science, because at every particular moment of time we know (could calculate) only a part of information about the object. And known part of information at every particular moment of time is finite and could be represented as a finite formal context.

Keywords: Formal context, Partial Metrics, Information

1 INTRODUCTION

Practical applications of formal context analysis were found in different fields including data mining, text mining, machine learning, hierarchical organization of web search results, software development and etc. On the other hand, general topology studies properties of topological spaces. An arbitrary topological space for an arbitrary formal context we can construct in the next way. An arbitrary topological space (X,τ) we can interpret as a formal context $(X,\tau,\in)$ with the object set X, the attribute set τ and the incidence relation $\in$. Many other, more advanced interpretations for a topological space can be found in the literature. On the contrary, from an arbitrary formal context we can construct in natural way several different topologies with a specific properties. How could we compare objects? How could we compare information about object in different moments of time? How is it possible to measure stored information in objects at one moment of time? Because information stored in formal context mostly isn’t a total information about objects we could describe a formal context with a partial metric. The main idea was to generalize metric by “dropping” the first axiom of the metric (zero self-distance). Zero self-distance axiom was replaced by non-negative self distance axiom. What does it mean? A zero-self distance means that all information about this object/structure is already known. But positive self-distance means that we don’t know all the details about the object. The smaller self-distance is the more defined object is. As a simple example we could take a map. If we know were object is situated we could find that point. But if we don’t have enough information we could only find an area where the point is situated. So, from this point of view that point is totally defined object and area is a partial defined object. This approach could be used in Computer Science, because at every particular moment of time we know(could calculate) only a part of information about the object. Also other mathematical structures could be represented as a formal context

2 BASIC DEFINITIONS

Definition 1 *A σ-algebra on a set X is a collection Σ of subsets of a set X that contains $\emptyset$ and X, and is closed under complements, countable unions, and countable intersections.*

Definition 2 *A measure is a countably additive, non-negative, extended real-valued function defined on a σ-algebra.*

Definition 3 *A function $p : X \times X \to \mathbb{R}^+$ is a partial metric if*

(p1) $\forall x,y \in X \quad p(x,y) \geq p(x,x)$

(p2) $\forall x,y \in X \quad p(x,y) = p(y,x)$

(p3) $\forall x,y,z \in X \quad p(x,z) \leq p(x,y) + p(y,z) - p(y,y)$

(p4) $\forall x,y \in X, \quad x = y \quad \Leftrightarrow \quad p(x,y) = p(x,x) = p(y,y)$

Definition 4 *A formal context is a triple $(X,A,\vdash)$ where X, A are sets and $\vdash \subseteq X \times A$ is a binary relation between them.*

In a formal concept analysis, the elements of X are called *objects* and the elements of A are called *attributes* of the context $(X,A,\vdash)$. The binary relation $\vdash$ is called the *incidence* relation. We say x has (the attribute) a or x satisfies a.

Definition 5 *(formal concept, extent, intent) Let $(X,A,\vdash)$ be a formal context, $P \subseteq X$, $F \subseteq A$. We put $P' = \{a | \, a \in A, x \vdash a \text{ for every } x \in P\}$ and $F' = \{x | x \in X, x \vdash a \text{ for every } a \in F\}$. Note: If $P = \{p\}$ is a singleton, we simply write $p' = P'$. Similarly we write $f' = F'$ for $F = \{f\}$. The pair (P,F) is called a formal concept of the context $(X,A,\vdash)$ if $P' = F$ and $F' = P$. The mappings $' : 2^X \to 2^A$ and $' : 2^A \to 2^X$ we would call the derivation operators. P is called the extent and F the intent of the concept (P,F).*

Now we will define the second derivation operator for a context $(X,A,\vdash)$ (by a composition of the first derivation operators):

(1) Map $'' : 2^X \to 2^X$ that for $P \in X$, $P \mapsto P''$,

(2) Map $'' : 2^A \to 2^A$ that for $F \in A$, $F \mapsto F''$.

Definition 6 *A context $(X,A,\vdash)$ is called row-clarified if for each $g,h \in X$ $g' = h'$ implies $g = h$, column-clarified if for each $m,n \in A$ $m' = n'$ implies $n = m$, clarified if it is column- and row-clarified.*

Lemma 1 *Let $(X,A,\vdash)$ be a formal context, τ be its left topology on X, $\leq$ is a preorder of specialization on X equipped with topology τ. The following statements for arbitrary elements $x,y \in X$ are equivalent:*

(1) $x \leq y$,

(2) $x \in \mathrm{cl}\{y\}$,

(3) $x \in y''$,

(4) $y' \subseteq x'$,

(5) $x'' \subseteq y''$,

For other definitions and results we refer to [2].

On the formal context it is possible to generate many different topologies. For example we could generate topologies on the object set, on the attribute set, on the set of all concepts, on the set of extents, on the set on intents. Here we are interested only in one type of the topologies, that could be generated on the object set and attribute set. That topologies behave in the same way, and that why we would pay attention only for one of them. All results could be simply retranslated to the other topology.

Definition 7 *(left and right topologies) Let $(X,A,\vdash)$ be a formal context. The topology τ on X, generated by its closed subbase $\{a'|a \in A\}$ is called the left topology on $(X,A,\vdash)$. Similarly, the right topology on $(X,A,\vdash)$ is the topology on A generated by the family $\{x'|x \in X\}$ used as its subbase for the closed sets.*

The topological closure operator induced by this topology we will denote by cl. All closed sets in the left topology τ denote as $\mathcal{C}$.

Definition 8 *A preorder of specialization on a topological space (X,τ) is the binary relation $\leq$ satisfying the condition $x \leq y \Leftrightarrow x \in \mathrm{cl}\{y\}$. We can rewrite this formula as $\mathrm{cl}\{y\} = \downarrow_{\leq} \{y\}$.*

Other definitions and results the reader could find in the [4].

3 MAIN RESULTS

Let μ be a finite counting measure. A Finite counting measure is an intuitive way to put a measure on any finite set. A measure of a set is taken to be a number of its elements: $\mu(A) = |A|$.

Lemma 2 *Let's take an arbitrary finite set A. Then let's define a finite counting measure μ on this set. A function $p : 2^A \times 2^A \longrightarrow \mathbb{R}^+$ constructed as $p(x,y) = \mu(x \cup y)$,* where *$x,y \subseteq A$ is a partial metric on the set 2^A .*

Proof. We need to check all axioms of a partial metric.

(1) Notice that the measure μ is a monotonic function, so $x \cap y \subseteq y$ implies $\mu(y) \geq \mu(x \cap y)$. Using this fact we can deduce $p(x,y) = \mu(x \cup y) = \mu(x) + \mu(y) - \mu(x \cap y) \geq \mu(x) = p(x,x)$.

(2) Axiom $(p2)$ is clear.

(3) The idea is to divide all sets into disjoint sets. Let's calculate $p(x,z) - [p(x,y) + p(y,z) - p(y,y)] = \mu(x \cup z) - [\mu(x \cup y) + \mu(y \cup z) - \mu(y)] = \mu(x) + \mu(z) - \mu(x \cap z) - [\mu(x) + \mu(y) - \mu(x \cap y) + \mu(y) + \mu(z) - \mu(y \cap z) - \mu(y)] = -\mu(x \cap z) - \mu(y) + \mu(x \cap y) + \mu(y \cap z) = -\mu(x \cap z \cap y) - \mu((x \cap z) \backslash y) - [\mu((y \cap x) \cap z) + \mu((y \cap x) \backslash z) + \mu((y \cap z) \backslash x) + \mu((y \backslash x) \backslash z)] + \mu((x \cap y) \cap z) + \mu((x \cap y) \backslash z) + \mu((y \cap z) \cap x) + \mu((y \cap z) \backslash x) = -\mu((x \cap z) \backslash y) - \mu((y \backslash x) \backslash z)]$. So we have $p(x,z) - [p(x,y) + p(y,z) - p(y,y)] = -\mu((x \cap z) - y) - \mu((y - x) - z)] < 0$ and it means that $p(x,z) < p(x,y) + p(y,z) - p(y,y)$

(4) for a finite case it is obvious (but for infinite sets it doesn't hold).

Lemma 3 *Let's take a finite row-clarified formal context $(X,A,\vdash)$. Given a finite counting measure $\mu : \Sigma \to \mathbb{R}^+$ on the finite set A (where Σ is a σ-algebra on A), define a function $p : X \times X \longrightarrow \mathbb{R}^+$ $p(x,y) = \mu(x' \cup y')$,* where *$x,y \in X$.*

Then p is a partial metric on X.

Proof. Again we need to check all axioms of a partial metric. Doing as in previous Lemma 2 we can easily check axioms $(p1),(p2),(p3)$.
For $(p4)$ we will prove first that for all $x,y \in X, \quad x=y$ implies $p(x,y)=p(x,x)=p(y,y)$. It is obvious(from the definition of the derivation operators on a formal context) that $x=y$ implies $x'=y'$ and then $x'=y'=x'\cup y'$. Then $\mu(x')=\mu(y')=\mu(x'\cup y')$. Thus $p(x,y)=p(x,x)=p(y,y)$.

The most interesting part is to prove that for all $x,y,\in X \quad p(x,y)=p(x,x)=p(y,y)$ implies $x=y$. From $p(x,y)=p(x,x)=p(y,y)$ we can obtain $\mu(x')=\mu(y')=\mu(x'\cup y')$. Thus $\mu(x'\backslash y')=\mu(y'\backslash x')=0$. Then it implies $x'\subseteq y'$ since μ is a finite counting measure. And similarly $y'\subseteq x'$. Thus $x'=y'$. And finally $x=y$ simply because the formal context $(X,A,\vdash)$ is row-clarified. Thus for all $x,y,\in X \quad p(x,y)=p(x,x)=p(y,y)\Rightarrow x=y$.

In the article [1] Matthews introduce partial metrics and topologies defined on it. For readers comfort we remind here the following definition and theorem.

Definition 9 *For each partial metric $p:U^2\to\mathbb{R}$, $\ll_p\subseteq U^2$ is the binary relation such that,*

$$\forall x,y\in U, x\ll_p y \Leftrightarrow p(x,x)=p(x,y).$$

The topology $\tau[\ll_p]$ arises from this binary relation. It is a topology of all upwardly closed sets $\tau[\ll_p]=\{S\subseteq U|\text{for all } x\in S, x\ll_p y\Rightarrow y\in S\}$

Theorem 1 *For each partial metric p, $\ll_p$ is a partial ordering.*

Lemma 4 *Let take a row-clarified context $(X,A,\vdash)$, where A is a finite set. Let's denote as μ a counting finite measure on the set A. Let p be a partial metric on X generated by counting finite measure μ on A. Then $\ll_p=\preceq$, where $\preceq$ is a specialisating preorder for a left topology generated on the context $(X,A,\vdash)$.*

Proof. Suppose that $x\ll_p y$, then by the definition of the binary relation $\ll_p$ we have $p(x,x)=p(x,y)$. The partial metric p is generated by the counting finite measure μ, hence $\mu(x')=\mu(x'\cup y')$. Then $\mu(x')=\mu(x')+\mu(y')-\mu(x'\cap y')$ and then $\mu(x'\cap y')=\mu(y')$. Now we divide the set y' into two disjoint sets $y'=(y'\cap x')\cup(y'\backslash x')$. Then $\mu(y')=\mu(y'\backslash x')+\mu(x'\cap y')$ and it immediately follows $\mu(y'\backslash x')=0$. Because μ is a counting finite measure then $y'\subseteq x'$. And by the Lemma 1 we have $x\preceq y$.
Now let's suppose that $x\preceq y$. By the Lemma 1 this is equivalent to $y'\subseteq x'$. It is obvious that $y'\backslash x'=\emptyset$. Then by the property of measure $\mu(y'\backslash x')=0$. From the $y'=(y'\backslash x')\cup(x'\cap y')$ by the definition of the measure we obtain $\mu(y')=\mu(y'\backslash x')+\mu(x'\cap y')$. Hence $\mu(y')=\mu(x'\cap y')$. By adding a $\mu(x')$ to the every side of equation we have $\mu(x')+\mu(y')=\mu(x')+\mu(x'\cap y')$. Thus $\mu(x')+\mu(y')-\mu(x'\cap y')=\mu(x')$. Then $\mu(x')=\mu(x'\cup y')$. Then we obtain $p(x,x)=p(x,y)$. And $x\ll_p y$ follows by he definition of $\ll_p$ binary relation.

It is necessary to mention, if we take the same subsets $(x=y)$, than a partial metric could be calculated as $p(x,x)=\mu(x\cup x)=\mu(x)$.

As it was stated in [1] the set of all open balls form an open base for some topology denoted as $\tau[p]$. And this topology is T_0 topology.

Definition 10 *An open ball for a partial metric $p:U^2\to\mathbb{R}$ is a set of the form $B^p_\varepsilon(x)=\{y\in U|p(x,y)<\varepsilon\}$ for each $\varepsilon>0$ and $x\in U$.*

The topologies $\tau[\ll_p], \tau[p]$ are the same if for every $x \in X$ there exists $\varepsilon > 0$ that $B^p_\varepsilon = \{p(x,y) < \varepsilon\}$. That was proved by S. Matthews in [1].

Theorem 2 *Let take a row-clarified context $(X,A,\vdash)$, where A is a finite set. Let p be a partial metric on X generated by counting finite measure μ on A. Then $\tau[\ll_p] = \tau[p]$.*

Proof. Let's take an arbitrary $x \in X$. Then as ε we could take a $p(x,x)+0.5$. Then open ball could be represented as $B^P_\varepsilon = \{y | y \in X, p(x,y) < \varepsilon\} = \{y | y \in X, p(x,y) < p(x,x)+0.5\}$. But μ is a finite counting measure, and it implies that the differences between measure of the sets could be only integers. Using that fact and axiom $(p2)$ we get $B^P_\varepsilon = \{y | y \in X, p(x,y) = p(x,x)\}$. Thus $\tau[\ll_p] = \tau[p]$.

That means, that the left topology generated by the finite row-clarified context coincide with the topology generated by the open balls with respect to the partial metric p.

And now we have a good tool for combining partial metrics and left topology generated by a formal context. Let's sum up. Let's take a finite context $(X,A,\vdash)$. In a native way we can introduce a partial metric on the set X. We can easily see that information that is carried in partial-metric function is enough.

Corollary 1 *Let's take a row-clarified context $(X,A,\vdash)$, where A is a finite set. Then the specialization preorder $\preceq$ for a left topology is a partial order.*

Proof. Immediately follows from the previous lemma and theorem 1.

Now we have a possibility for constructing a partial order on set, with help of some additional finite set with defined counting finite measure.

4 CONCLUSION

In this article we proposed a relationship between partial metrics and finite row-clarified contexts. The left topology defined in [4] on a finite row clarified contexts coincide with the open ball topology defined with help of partial metrics in [1]. And it follows, that left topology in this particular case has nice properties described in that article. It is obvious, that formal context is a general mathematical structure and we could represent other mathematical structures as contexts. Such representation we could easy implement in computer. And that why it is a great advantage.

ACKNOWLEDGMENT

This article was supported by the International Visegrad Fund.

REFERENCES

[1] Matthews, S.G., *Partial Metric Topology*, Proc. 8th summer conference on topology and it applications, ed S. Andima et al., Annals of the New York Academy of Science, New York,728, (1994), 183-197.

[2] Wille, R., *Restructuring Lattice Theory: An Approach Based on Hierarchies of Concepts*, I. Rival (ed.), Ordered Sets , Vol. 83 pp. 445-470. Springer Netherlands . ISBN: 978-94-009-7800-3.

[3] Ganter, B., Wille, R., *Formal Concept Analysis*, Springer-Verlag, Berlin (1999), pp.1-285.

[4] Chernikava, A. *Closure Properties of the Formal Contexts*, Proc. 19th conferecne EEICT, Brno University of Technology, Brno, Vol.3 pp.139-144. ISBN 978-80-214-4695-3.