Characterization of Galois closed sets using multiway dissimilarities

Jean Diatta

IREMIA
Université de La Réunion
15 avenue René Cassin - BP 7151,
97715 Saint-Denis messag cedex 9, France
(e-mail: jdiatta@univ-reunion.fr)

Abstract. We place ourselves in a so-called meet-closed description context; that is a context consisting of a finite nonempty entity set E whose elements are described in a complete meet-semilattice $\underline{D}$, by means of a descriptor δ. Then we consider multiway quasi-ultrametric dissimilarities on E, a class of multiway dissimilarities that, with their relative k-balls, extend the fundamental in classification bijection between ultrametric dissimilarities and indexed hierarchies. We also consider multiway dissimilarities agreeing with entity descriptions in a quite natural sense called δ-meet compatibility. It turns out that there exists an integer k such that any strictly δ-meet compatible k-way dissimilarity is quasi-ultrametric. On the other hand, the descriptor δ induces a Galois connection between the powerset $\mathcal{P}(E)$ and $\underline{D}$, which, in turn, induces a closure operator, say ϕ_δ, on $\mathcal{P}(E)$. then it is proved that nonempty ϕ_δ-closed subsets of E coincide with k-balls relative to some strictly δ-meet compatible multiway dissimilarities on E.
Keywords: Galois connection, Multiway dissimilarity, Closure operator, Description-meet compatibility, Quasi-ultrametric.

1 Introduction

Multiway dissimilarities are natural extensions of classical pairwise dissimilarities, that allow global comparison of more than two entities. In the last decade, they have been investigated or considered from different approaches in many works among which we just mention [Bandelt and Dress, 1994], [Joly and Le Calvé, 1995], [Daws, 1996] and [Bennani and Heiser, 1997]. In this paper, these approaches are extended onto the so-called meet-closed data description context. A meet-closed description context represents a finite entity set E using a complete meet-semilattice $\underline{D}$. Then we consider multiway quasi-ultrametric dissimilarities on E [Bandelt and Dress, 1994], [Diatta, 1997], a class of multiway dissimilarities that, with their relative k-balls, extend the fundamental in classification bijection between ultrametric dissimilarities and indexed hierarchies [Johnson, 1967]. We also consider multiway dissimilarities agreeing with entity descriptions in a quite natural sense called δ-meet compatibility. It turns out that there exists an integer k such that any strictly δ-meet compatible k-way dissimilarity is quasi-ultrametric.

On the other hand, any descriptor δ induces a Galois connection between the powerset $\mathcal{P}(E)$ and $\underline{D}$, which, in turn, induces a closure operator, say ϕ_δ, on $\mathcal{P}(E)$ [Birkhoff, 1967]. It is proved that nonempty ϕ_δ-closed subsets of E are the k-balls of some strictly δ-meet compatible multiway dissimilarities on E.

2 Multiway dissimilarities

Before introducing multiway dissimilarities, let us first recall the classical pairwise ones. Let E be a finite nonempty set.

A (pairwise) dissimilarity on E is a map $d : E \times E \to \mathbb{R}$ satisfying reflexivity ((R2) $d(x,x) = 0$), non-negativity ((N2) $d(x,y) \geq 0$) and symmetry ((S2) $d(x,y) = d(y,x)$).

Considering maps on $E^3, E^4, \ldots, E^k$, with similar properties, naturally leads to the notion of 3-way, 4-way,..., k-way dissimilarity. For instance, a 3-way dissimilarity on E will be any map $d : E^3 \to \mathbb{R}$ satisfying: (R3) $d(x,x,x) = 0$, (N3) $d(x,y,z) \geq 0$ and (S3) $d(x,y,z) = d(x,z,y) = d(y,x,z) = d(y,z,x) = d(z,x,y) = d(z,y,x)$. The term *multiway* dissimilarity will be used to mean a k-way dissimilarity, for some $k \geq 2$.

Of course, due to the tuple-based definition above, the complexity of expressions related to k-way dissimilarities increases when k grows. Then, for the sake of simplicity, we adopt in the present paper a set-based definition based on the following observation: according to (R2) and (N2), $d(x,x) \leq d(x,y)$ for all x, y. Thus, a dissimilarity on E can be defined as being a nonnegative real valued map d on the set of singletons and pairs of E, satisfying $d(\{x\}) = 0$ and $d(\{x\}) \leq d(\{x,y\})$. This set-based definition makes the symmetry condition implicit. Moreover, for $k \geq 2$, its generalization to k-way dissimilarities involves shortest expressions.

For reasons explained in Remark 3 below, we will drop out the reflexivity condition and thus be rather concerned with so-called (multiway) pseudo-dissimilarities. However, we will still use the term dissimilarity, keeping in mind that the condition $d(\{x\}) = 0$ is not required.

For any set S and any integer $k \geq 1$, $S^*_{\leq k}$ will denote the set of all nonempty subsets of S with at most k elements. Then we formally define multiway dissimilarities as follows.

Definition 1 *A* k-way dissimilarity *on* E *will be any nonnegative real valued and isotone map defined on the set of all nonempty subsets of* E *with at most* k *elements, i.e., any map* $d : E^*_{\leq k} \to \mathbb{R}_+$ *such that* $d(X) \leq d(Y)$ *when* $X \subseteq Y$.

Example 1 *Table 2 presents a dataset, say* $\mathcal{D}$*, about seven market baskets and five items: bread (brd), butter (btr), cheese (chs), eggs (egg), milk (mlk); for instance, the market basket labeled 1 contains bread and cheese. For any* k *such that* $2 \leq k \leq 5$*, a* k*-way dissimilarity on the set of items, can be* *defined by letting* $\mathrm{dis}_k(X)$ *be seven minus the number of baskets that contain each of the items in* X*. Then, for instance,* $\mathrm{dis}_3(\{brd, chs\}) = 4$ *and* $\mathrm{dis}_3(\{brd, btr, chs\}) = 7$*.*

	brd	btr	chs	egg	mlk
1	x		x		
2		x	x		x
3	x		x	x	
4		x		x	x
5	x	x			x
6	x	x		x	
7	x		x		x

Table 1. Example dataset

Remark 1 *For* $\{x, y, z\} \subseteq E$*, we will simply write* $d(x)$ *or* $d(x, y)$ *or* $d(x, y, z)$ *instead of* $d(\{x\})$ *or* $d(\{x, y\})$ *or* $d(\{x, y, z\})$*, respectively. Moreover, as in the tuple-based setting, the notation* $d(x, y)$ *or* $d(x, y, z)$ *will not require* x*,* y *and* z *be distinct.*

3 Quasi-ultrametric multiway dissimilarities

Key notions in the definition of quasi-ultrametrics given below are those of a d-ball, (d, k)-ball and d-diameter, where d is a k-way dissimilarity. To catch their meaning, let us first cast them in the case of a pairwise dissimilarity, say d_2.

The d_2-diameter of a nonempty subset Z of E is the maximum d_2-dissimilarity between elements of Z, i.e.: $\mathrm{diam}_{d_2}(Z) = \max\{d_2(x, y) : x, y \in Z\}$.

Let now x and y be two distinct elements of E and r a nonnegative real number. The d_2-ball of center x and radius r is the set $B^{d_2}(x, r)$ of elements of E whose d_2-dissimilarity degree from x is at most r, i.e., formally, $B^{d_2}(x, r) = \{z \in E : d_2(x, z) \leq r\}$; the $(d_2, 2)$-ball generated by $\{x\}$ is the set $B^{d_2}_x = B^{d_2}(x, d_2(x))$, and the $(d_2, 2)$-ball generated by $\{x, y\}$ is the set $B^{d_2}_{xy} = B^{d_2}(x, d_2(x, y)) \cap B^{d_2}(y, d_2(x, y))$. If $x = y$, $B^{d_2}_{xy} = B^{d_2}_x$.

All these notions have been naturally generalized to multiway dissimilarities in [Diatta, 1997]. For $k \geq 2$, let d_k denote a k-way dissimilarity on E.

The d_k*-diameter* (or, simply, *diameter*) of a nonempty subset Z of E is the maximum d_k-dissimilarity degree between elements of Z, i.e.: $\mathrm{diam}_{d_k}(Z) = \max\{d_k(T) : T \in Z^*_{\leq k}\}$.

Let $X \in E^*_{\leq k-1}$. The d_k*-ball* (or, simply, *ball*) of center X and radius r is the set $B^{d_k}(X, r)$ defined by $B^{d_k}(X, r) = \{y \in E : d_k(X \cup \{y\}) \leq r\}$. If

$X \in E^*_{\leq k}$, then the (d_k, k)-*ball* (or, simply, k-*ball* relative to d_k) generated by X is the set $B_X^{d_k}$ defined by $B_X^{d_k} = B^{d_k}(X, d_k(X))$ when $|X| \leq k-1$, and $B_X^{d_k} = \underset{x \in X}{\cap} B^{d_k}(X \setminus \{x\}, d_k(X))$ otherwise. The superscript d_k may be omitted if there is no risk of confusion.

Before defining quasi-ultrametrics, let us recall a well-known particular case of them, namely ultrametrics. A (2-way) dissimilarity d_2 is said to be *ultrametric* if for all x, y, z:

$$d_2(x, y) \leq \max\{d_2(x, z), d_2(y, z)\}.$$

Next are some characterizations of ultrametric 2-way dissimilarities, which may help in understanding the definition of quasi-ultrametrics given below.

Proposition 1 *[Diatta and Fichet, 1998] For a 2-way dissimilarity d_2 on E, the following assertions are equivalent.*

(i) *d_2 is ultrametric.*
(ii) *for all x, y, z: the greatest two values among $d_2(x, y)$, $d_2(x, z)$ and $d_2(y, z)$ are equal.*
(iii) *for all x, y:* $\mathrm{diam}_{d_2}(B(x, d_2(x, y))) = d_2(x, y)$ *(diameter condition).*
(iv) *for all x, y, u, v: $u, v \in B(x, d_2(x, y))$ implies $B(u, d_2(u, v)) \subseteq B(x, d_2(x, y))$ (inclusion condition).*

Example 2 *Figure 1 presents three dissimilarities d_1, d_1' and d_1'' on the set $\{i, j, k, l\}$. It is easily checked that d_1 satisfies the diameter condition; but d_1 does not satisfy the inclusion condition because $j, k \in B_{jl}^{d_1}$ whereas $i \in B_{jk}^{d_1}$ and $i \notin B_{jl}^{d_1}$. It is also easily checked that d_1' satisfies the inclusion; but d_1' does not satisfy the diameter condition because $i, j \in B_{kl}^{d_1'}$ so that* $\mathrm{diam}_{d_1'}(B_{kl}^{d_1'}) > d_1'(k, l)$. *The dissimilarity d_1'' is clearly quasi-ultrametric since $B_i^{d_1''} = B_j^{d_1''} = B_{ij}^{d_1''} = \{i, j\}$, for $x \neq i, j$, $B_x^{d_1''} = \{x\}$, and for $\{x, y\} \neq \{i, j\}$, $B_{xy}^{d_1''} = \{i, j, k, l\}$.*

d_1:

	i	j	k	l
i	0			
j	1	0		
k	1	1	0	
l	3	2	1	0

d_1':

	i	j	k	l
i	0			
j	3	0		
k	1	1	0	
l	1	1	2	0

d_1'':

	i	j	k	l
i	0			
j	0	0		
k	1	1	0	
l	1	1	1	0

Fig. 1. Three pairwise dissimilarities on the set $\{i, j, k, l\}$: d_1 satisfies the diameter but not the inclusion condition; d_1' satisfies the inclusion but not the diameter condition; d_1'' is quasi-ultrametric.

Conditions (iii) and (iv) of Proposition 1 above can be extended to the case of multiway dissimilarities by replacing balls with k-balls. The two extended conditions define what we call the quasi-ultrametric multiway dissimilarities [Diatta, 1997]:

Definition 2 *A k-way dissimilarity d_k on E is said to*

(i) satisfy the inclusion condition *if for all $X, Y \in E^*_{\leq k}$, $Y \subseteq B^{d_k}_X$ implies $B^{d_k}_Y \subseteq B^{d_k}_X$;*

(ii) satisfy the diameter condition *if for all $X \in E^*_{\leq k}$, $\mathrm{diam}_{d_k}(B^{d_k}_X) = d_k(X)$;*

(iii) be quasi-ultrametric *if it satisfies both of the inclusion and the diameter conditions.*

Example 3 *The reader may check that the 3-way dissimilarity dis_3 defined in Example 1 is quasi-ultrametric. This can also be derived from Theorem 1 below (see Remark 4).*

4 Description-meet compatibility

In this section, we place ourselves in a so-called *meet-closed description context.* That is a context consisting of a finite nonempty entity set E whose elements are described in a complete meet-semilattice $\underline{D}$, by means of a descriptor δ. We will denote such a context as a triple $\mathbb{K} = (E, \underline{D}, \delta)$ where E stands for the entity set, $\underline{D} := (D, \leq)$ the entity description space, and δ the descriptor that associates to each entity $x \in E$ its description $\delta(x)$ in $\underline{D}$.

In all what follows, E will denote a finite nonempty entity set, $\underline{D}$ a complete meet-semilattice, δ a descriptor that maps E into $\underline{D}$, and $\mathbb{K}$ the meet-closed description context $(E, \underline{D}, \delta)$.

Example 4 *Consider Table 4 presenting five visitors of a given Web site, described by three attributes:* LiLo, NoLi, ReSu, *where* $\mathrm{LiLo}(x)$ *is the login-logout time interval of visitor x within the interval* $[0, 24]$, $\mathrm{NoLi}(x)$ *is the number of times visitor x logs in at* $\mathrm{LiLo}(x)$ *interval during a given fixed period, and* $\mathrm{ReSu}(x)$ *is the subjects requested by x during a session; requested subjects are sets of subjects from: Arts & Humanities (AH), Business & Economy (BE), Computers & Internet (CI), News & Media (NM), Recreation & Sports (RS), Science & Health (SH), Society & Culture (SC).*
Then Table 4 can be seen as representing a meet-closed description context $\mathbb{K}_2 := (E_2, \underline{D}_2, \delta_2)$ *where E_2 is the set* $\{1, 2, 3, 4, 5\}$, $\underline{D}_2$ *the direct product of three partially ordered sets (posets): the set* $(\mathrm{FUCI}([0, 24]), \subseteq)$ *of finite unions of closed intervals of* $[0, 24]$ *endowed with the set inclusion order, the set* $([\![30; 40]\!], \leq)$ *of integers from 30 to 40, endowed with the integer usual order, and the powerset* $(\mathcal{P}(S), \subseteq)$ *of the set* $S = \{AH, BE, CI, NM, RS, SC\}$*, endowed with the set inclusion order, and* $\delta_2(x) = (\mathrm{LiLo}(x), \mathrm{NoLi}(x), \mathrm{ReSu}(x))$.

	LiLo	NoLi	ReSu
1	0-2	30	CI,RS
2	21-24	35	AH,NM,SC
3	0-3	40	AH,BE,CI,RS
4	22-24	35	AH,SC
5	12-14	30	BE,NM

Table 2. Example meet-closed description context

The description-meet compatibility defined below has been introduced in [Diatta and Ralambondrainy, 2002] in the case of pairwise dissimilarities. It uses the notion of valuation on a poset.

A *valuation* on a poset $(P, \leq)$ is a map $h : P \to \mathbb{R}_+$ such that $h(x) \leq h(y)$ when $x \leq y$. A *strict* valuation will then be a valuation h such that $x < y$ implies $h(x) < h(y)$.

Before defining the description-meet compatibility, let us introduce a further notation: for any $X \subseteq E$, $\delta(X)$ will denote the set of descriptions of entities belonging to X.

A multiway dissimilarity d on E will be said to be *δ-meet compatible* if there exists a valuation h on $\underline{D}$ with which it is δ-meet compatible, i.e., such that

$$d(X) \leq d(Y) \iff h(\inf \delta(X)) \geq h(\inf \delta(Y)),$$

for $X, Y \subseteq E$. If h is a strict valuation, d will be said to be *strictly* δ-meet compatible.

Remark 2 *The reader may observe that when $\underline{D}$ is a complete join-semilattice, a dual compatibility condition, say δ-join compatibility, can be defined by reversing the right-hand side inequality in the above equivalence and replacing meets by joins.*

Description-meet compatibility is a kind of natural agreement expressing the following fact: the lower the meet of descriptions of entities in X, the larger the dissimilarity degree of X.

Remark 3 *If d is a strictly δ-meet compatible (multiway) dissimilarity, then $\delta(x) < \delta(y)$ implies $d(y) < d(x)$. This is why we drop out the condition $d(x) = 0$, since it is very likely to happen that two entities x and y satisfy $\delta(x) < \delta(y)$.*

Example 5 *Consider the meet-closed description context $\mathbb{K}_2$ defined in Example 4. Define a multiway dissimilarity on E_2 by*

$$\mathrm{dis}'(X) = 47 - (\lambda(\cap_{x \in X} \mathrm{LiLo}(x)) + \min_{x \in X} \mathrm{NoLi}(x) + |\cap_{x \in X} \mathrm{ReSu}(x)|),$$

where $\lambda([\alpha, \beta]) = \beta - \alpha$. For instance, $\mathrm{dis}'(1, 2, 3) = 47 - (\lambda([0, 2] \cap [21, 24] \cap [0, 3]) + \min\{30, 35, 40\} + |\{CI, RS\} \cap \{AH, NM, SC\} \cap \{AH, BE, CI, RS\}|) =$

$47 - (\lambda(\varnothing) + 30 + |\varnothing|) = 47 - (0 + 30 + 0) = 17$. *Then* dis$'$ *is strictly* δ_2*-meet compatible. Indeed,* λ, $x \mapsto x$ *and* $Y \mapsto |Y|$ *are strict valuations on* $(\mathrm{FUCI}([0,24]), \subseteq)$, $([\![30; 40]\!], \leq)$ *and* $(\mathcal{P}, \subseteq)$*, respectively. Thus* h_2 *defined by*

$$h_2(u, v, w) = \lambda(u) + v + |w|$$

is a strict valuation on $\underline{D}_2$*, and the fact that* dis$'$ *is* δ_2*-meet compatible with* h_2 *follows from the fact that* $\mathrm{dis}'(X)$ *is decreasing w.r.t.* $h_2(\inf \delta_2(X))$.

Before outlining the relationship between quasi-ultrametricity and description-meet compatibility, let us recall the following technical notion: the *breadth* of a meet-semilattice $(P, \leq)$ is the least positive integer k such that the meet of any $(k+1)$ elements of P is always the meet of k elements among these $k+1$ [Birkhoff, 1967]. Having noticed this, we agree to say that a subset Q of a meet-semilattice is of breadth k if k is the least positive integer such that for any $(k+1)$-element subset W of Q there is $w \in W$ such that $\inf(W \setminus \{w\}) \leq w$.

Example 6 *Consider the dataset* $\mathcal{D}$ *given in Table 2 as presenting a meet-closed description context* $\mathbb{K}_1 := (E_1, \underline{D}_1, \delta_1)$*, where* E_1 *is the set of five items and* $\underline{D}_1$ *the boolean lattice* $\{0,1\}^7$*; for instance* $\delta_1(brd) = (1,0,1,0,1,1,1)$*. Then* $\delta_1(E_1)$ *is of breadth at least* 3 *since*

$$\inf \delta_1(\{brd, chs, mlk\}) = (0,0,0,0,0,0,1),$$

which is different from either of $\delta_1(brd) \wedge \delta_1(chs) = (1,0,1,0,0,0,1)$*,* $\delta_1(brd) \wedge \delta_1(mlk) = (0,0,0,0,1,0,1)$ *and* $\delta_1(chs) \wedge \delta_1(mlk) = (0,1,0,0,0,0,1)$*. Moreover,*

$$\begin{aligned} \inf \delta_1(\{brd, btr, chs, egg\}) &= \inf \delta_1(\{brd, btr, chs, mlk\}) \\ &= \inf \delta_1(\{brd, btr, chs\}), \end{aligned}$$

$$\begin{aligned} \inf \delta_1(\{brd, btr, egg, mlk\}) &= \inf \delta_1(\{brd, chs, egg, mlk\}) \\ &= \inf \delta_1(\{brd, egg, mlk\}), \end{aligned}$$

and $\inf \delta_1(\{btr, chs, egg, mlk\}) = \inf \delta_1(\{btr, chs, egg\})$*, so that* $\delta_1(E_1)$ *is of breadth* 3.

We now go on stating the result showing the existence of an integer $k \geq 2$ such that any strictly δ-meet compatible k-way dissimilarity on E is quasi-ultrametric.

Theorem 1 *(i) If* $\delta(E)$ *is of breadth one, then every strictly* δ*-meet compatible* 2*-way dissimilarity on* E *is ultrametric.*

(ii) If $\delta(E)$ *is of breadth* $k \geq 2$*, then every strictly* δ*-meet compatible* k*-way dissimilarity on* E *is quasi-ultrametric.*

The converse of Theorem 1 does clearly not hold since, for $k \geq 2$, every constant k-way dissimilarity on E is quasi-ultrametric but never strictly δ-meet compatible, regardless of the descriptor δ. Indeed, otherwise, we would have, for all $x, y \in E$, $\delta(x) = \delta(y)$ so that $\delta(E)$ would be a singleton, hence of breadth one.

Remark 4 *As claimed in Example 3, it follows from Theorem 1 that the* 3*-way dissimilarity* dis_3 *defined in Example 1 is quasi-ultrametric. Indeed, on the one hand, as observed in Example 6,* $\delta_1(E_1)$ *is of breadth* 3*. On the other hand, for each* k *such that* $2 \leq k \leq 5$*,* dis_k *is strictly* δ_1*-meet compatible with the valuation* h_1 *defined on* $\underline{D}_1$ *by letting* $h_1(x)$ *be the number of ones occurring in* x*.*

The entity set E being finite, there is an integer $k \geq 1$ such that k is the breadth of $\delta(E)$. Moreover, as any pairwise ultrametric dissimilarity is quasi-ultrametric, we derive the following from Theorem 1.

Corollary 1 *There is an integer* $k \geq 2$ *such that any strictly* δ*-meet compatible* k*-way dissimilarity on* E *is quasi-ultrametric.*

Following [Diatta, 1997], a k-way dissimilarity d will be said to be ultrametric if for all $X \in E^*_{\leq k}$ and $x \in E$:

$$d(X) \leq \max_{Y \in X^*_{\leq k-1}} d(Y + x).$$

When $\delta(E)$ is of breadth one, Theorem 1 (i) extends to ultrametric multiway dissimilarities:

Theorem 2 *If* $\delta(E)$ *is of breadth one, then for* $k \geq 2$*, every strictly* δ*-meet compatible* k*-way dissimilarity on* E *is ultrametric.*

5 Characterization of Galois closed entity sets

Given the meet-closed description context $\mathbb{K} = (E, \underline{D}, \delta)$, the descriptor δ induces a Galois connection between $(\mathcal{P}(E), \subseteq)$ and $\underline{D}$ by means of the maps

$$f : X \mapsto \inf \{\delta(x) : x \in X\}$$

and

$$g : I \mapsto \{x \in E : I \leq \delta(x)\},$$

for $X \subseteq E$ and $I \in \underline{D}$. Then it is well known that, in these conditions, the map $\phi_\delta := g \circ f$ is a closure operator on $\mathcal{P}(E)$ [Birkhoff, 1967]. That is ϕ_δ is *extensive* ($X \subseteq \phi_\delta(X)$), *isotone* ($X \subseteq Y$ implies $\phi_\delta(X) \subseteq \phi_\delta(Y)$) and *idempotent* ($\phi_\delta(\phi_\delta(X)) = \phi_\delta(X)$). A subset X of E is said to be ϕ_δ-*closed* (or a *Galois closed entity set* of $\mathbb{K}$, relative to ϕ_δ) when $\phi_\delta(X) = X$.

Galois closed entity sets play an important role in classification because they provide easy-to-interpret clusters [Domenach and Leclerc, 2002]. Indeed, if X is a Galois closed entity set, then $f(X)$ is the description of X.

When $\underline{D}$ is a complete join-semilattice, the descriptor δ induces a Galois connection between $(\mathcal{P}(E), \subseteq)$ and the order-dual of $\underline{D}$ by means of the maps

$$f^{\partial} : X \mapsto \sup \{\delta(x) \,:\, x \in X\}$$

and

$$g^{\partial} : I \mapsto \{x \in E \,:\, I \geq \delta(x)\},$$

for $X \subseteq E$ and $I \in \underline{D}$. Similarly, this Galois connection induces the closure operator $\phi_\delta^{\partial} := g^{\partial} \circ f^{\partial}$ on $\mathcal{P}(E)$. Galois closed entity sets relative to ϕ_δ^{∂} have been considered in the framework of symbolic data analysis [Bock and Diday, 2000].

Example 7 *Consider the meet-closed description context $\mathbb{K}_2$ given in Example 4. The pair $\{1,3\}$ is ϕ_δ-closed; but $\{1,2,3\}$ is not ϕ_δ-closed because* $\inf \delta_2(\{1,2,3\}) = (\emptyset, 30, \emptyset) \leq \delta_2(4)$. *On the other hand, the pair $\{4,5\}$ is ϕ_δ^{∂}-closed; but $\{1,2,3\}$ is not ϕ_δ^{∂}-closed because*

$$\delta_2(4) \leq \sup \delta_2(\{1,2,3\}) = ([0,3] \cup [21,24], 40, \{AH, BE, CI, NM, RS, SC\}).$$

The following proposition shows that the ϕ_δ-closure of any subset $X \subseteq E$ is contained in a ball centered in a subset of X and relative to some δ-meet compatible multiway dissimilarity.

Proposition 1 *Let d be a δ-meet compatible k-way dissimilarity measure on E and let $X \in E^*_{\leq k}$. Then for all $Y \in X^*_{\leq k-1}$ and all $y \in B^d(Y, d(X))$, $\phi_\delta(Y + y) \subseteq B^d(Y, d(X))$. Moreover, $\phi_\delta(X) \subseteq B^d(Y, d(X))$.*

The next proposition gives a necessary and sufficient condition for the ϕ_δ-closure of an entity subset X to be a ball (resp. k-ball) relative to some δ-meet compatible multiway dissimilarity.

Proposition 2 *Let d be a δ-meet compatible k-way dissimilarity measure on E. Then, for all $X \in E^*_{\leq k}$ and all $Y \in X^*_{\leq k-1}$:*

(i) $\phi_\delta(X) = B^d(Y, d(X))$ *if and only if* $\inf \delta(B^d(Y, d(X))) = \inf \delta(X)$.
(ii) $\phi_\delta(X) = B^d_X$ *if and only if* $\inf \delta(B^d_X) = \inf \delta(X)$.

We now go on stating the result showing the coincidence between nonempty Galois closed entity sets of a meet-closed description context and k-balls relative to some strictly description-meet compatible multiway dissimilarity.

Theorem 3 *For an integer $p \geq 2$, let d_p be a strictly δ-meet compatible p-way dissimilarity on E.*

(i) *If* $\delta(E)$ *is of breadth one, the set* $\mathcal{F}_\delta{}^*$ *of nonempty* ϕ_δ*-closed subsets of* E *coincides with the set of* $(d_2, 2)$*-balls generated by singletons of* E*.*
(ii) *If* $\delta(E)$ *is of breadth* $k \geq 2$*, then* $\mathcal{F}_\delta{}^*$ *coincides with the set of* (d_k, k)*-balls.*

Finally, as observed above, E being finite, there is an integer $k \geq 1$ such that k is the breadth of $\delta(E)$. Moreover, as any pairwise ultrametric dissimilarity is quasi-ultrametric, we derive the following from theorems 1 and 3.

Corollary 2 *There is an integer* $k \geq 2$ *such that nonempty Galois closed entity subsets of* E *coincide with* k*-balls relative to some* k*-way quasi-ultrametric dissimilarity on* E*.*

It may be noticed that, when $\underline{D}$ is a complete join-semilattice, similar results hold for Galois closed entity sets relative to ϕ_δ^∂, using δ-join compatible multiway dissimilarities.

Acknowledgements: The author is grateful to Boris Mirkin for helpful advise.

References

[Bandelt and Dress, 1994]H-J. Bandelt and A. W. M. Dress. An order theoretic framework for overlapping clustering. *Discrete Mathematics*, 136:21–37, 1994.

[Bennani and Heiser, 1997]M. Bennani and W. J. Heiser. Triadic distance models: axiomatization and least squares representation. *J. Math. Psychology*, 41:189–206, 1997.

[Birkhoff, 1967]G. Birkhoff. *Lattice theory.* 3rd edition, Coll. Publ., XXV. American Mathematical Society, Providence, RI, 1967.

[Bock and Diday, 2000]H.H. Bock and E. Diday, editors. *Analysis of Symbolic Data.* Springer-Verlag, 2000.

[Daws, 1996]J. T. Daws. The analysis of free-sorting data : beyond pairwise cooccurrences. *J. Classification*, 13:57–80, 1996.

[Diatta and Fichet, 1998]J. Diatta and B. Fichet. Quasi-ultrametrics and their 2-ball hypergraphs. *Discrete Mathematics*, 192:87–102, 1998.

[Diatta and Ralambondrainy, 2002]J. Diatta and H. Ralambondrainy. The conceptual weak hierarchy associated with a dissimilarity measure. *Mathematical Social Sciences*, 44:301–319, 2002.

[Diatta, 1997]J. Diatta. Dissimilarités multivoies et généralisations d'hypergraphes sans triangles. *Math. Inf. Sci. hum.*, 138:57–73, 1997.

[Domenach and Leclerc, 2002]F. Domenach and B. Leclerc. On the roles of Galois connections in classification. In O. Opitz M. Schwaiger, editor, *Explanatory Data Analysis in Empirical Research*, pages 31–40. Springer, 2002.

[Johnson, 1967]S. C. Johnson. Hierarchical clustering schemes. *Psychometrika*, 32:241–254, 1967.

[Joly and Le Calvé, 1995]S. Joly and G. Le Calvé. Three-way distances. *J. Classification*, 12:191–205, 1995.