ELSEVIER SCIENCE B.V. [DTD 4.2.0]
JOURNAL EOR ARTICLE No. 4798
PAGES 1-12 DISPATCH 8 August 2001
EOR 4798 A
PROD. TYPE: FROM DISK

ELSEVIER
European Journal of Operational Research 000 (2001) 000–000
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Theory and Methodology

Electre-like clustering from a pairwise fuzzy proximity index

Raymond Bisdorff

Dpt. des Études en Gestion et en Informatique, Centre Universitaire, 162a, avenue de la Fäiencerie, L-1511 Luxembourg, Luxembourg

Abstract

In this paper, we propose an Electre-like approach for clustering judges from their $\mathscr{L}$-valued pairwise proximities in preference judgements. The approach is based on the extraction of $\mathscr{L}$-valued null kernels from a pairwise $\mathscr{L}$-valued proximity index. A practical application will concern the clustering of movie critics. © 2001 Published by Elsevier Science B.V.

Keywords: Multiple criteria analysis; Fuzzy clustering; Graph theory

1. Introduction

In this paper we propose to apply the concept of $\mathscr{L}$-valued kernels (see [1,2]) to the problem of clustering judges from a pairwise $\mathscr{L}$-valued binary proximity index observed on a set of qualitative preference judgements as encountered in the fuzzy preference modelling context (see [7] for instance).

This work follows two of our papers (see [4,5]) concerning the application of initial and terminal $\mathscr{L}$-valued kernels to bipolar ranking of decision actions from a pairwise fuzzy outranking index as proposed in the Electre decision aid methods (see [8]). Here we propose to apply a same operational technique to construct similarity clusters from a pairwise fuzzy proximity index.

First we introduce the clustering problem, then we briefly sketch the concept of $\mathscr{L}$-valued kernel and show its eventual use in implementing a clustering procedure. In Section 4, we will finally present the application of our method to the clustering of movie critics in Luxembourg. In particular we will discuss how to cope with missing values.

2. Clustering movie critics

In this section, we first present the practical clustering problem that we propose for our investigation. In Section 2.2, we introduce an Electre-based construction of a global proximity index between criteria evaluations (see [8]).

2.1. The movie critics in Luxembourg

The Luxembourg movie magazine “Graffiti” publishes monthly a list of appreciations of currently shown movies in Luxembourg's movie theaters by some well-known local journalists and cinema critics (see Appendix A, Table 8). The

E-mail address: bisdorff@cu.lu (R. Bisdorff).

0377-2217/01/$ - see front matter © 2001 Published by Elsevier Science B.V.
PII: S0377-2217(01)00249-1

Table 1
The movie critics' opinions in Luxembourg

Movies	jpt	cn	pf	vt	jh	mr	...
Courage under Fire	**	**	**	*	**	*	...
Didier	**	**	*	*	***	/	...
Un Eté à la Goulette	***	*	*	**	*	/	...
The First Wives Club	**	0	00	*	**	*	...
Lost Highway	****	****	***	*	*	**	...
...	...	...	...	...	...	...	

evaluation data set, we use in this paper, is collected from the March/April and September 1997 issues of the Graffiti magazine (see Appendix A, Fig. 5). In the extract shown in Table 1 one may notice that critics express their opinions on the basis of an ordinal preference scale ranging from four stars (****) (very much appreciated) to two zeros (00) (very much disliked). A slash (/) indicates missing data, i.e., a critic did not evaluate that movie. In order to clearly separate the positive stars from the negative zeros, we introduce a neutral null point as separator between positive stars and negative 0s, i.e., we extend the original scale to a set of seven ordinal grades $\{-2, -1, 0, 1, 2, 3, 4\}$. For an individual critic, this preference scale gives a complete ordering $\geqslant$ from the best (**** = 4) to the worst (00 = −2) evaluation. For instance, critic jpt evaluates the movie *The First Wives Club* as being rather good (** = 2), whereas critic pf evaluates the same movie as being very bad (00 = −2).

The particular question we are interested in is, to uncover to what extent, these critics express similar opinions or not.

2.2. Constructing a proximity index

Naturally, if one critic expresses exactly the same evaluations as another one, we may easily deduce that the two critics express similar opinions and we cluster them together. Take for instance the movies "Courage under Fire" and "Didier". The evaluations of two critics (jpt and cn) express exactly the same opinion and limited to this sample, our conclusion would be that both critics express a similar opinion and belong in fact to a same cluster.

In general, let C denote the set of considered critics. For each critic $c_i \in C$, let M_i denote the set of movies he has evaluated and for each $m \in M_i$, let $v_i(m) \in \{-2, -1, 0, 1, 2, 3, 4\}$ denote the numeric code of the evaluation he has given.

A natural proximity index s_{ij} logically evaluating the proposition "*critic c_i is expressing similar judgements to critic c_j*" may be computed in the following way:

$$s_{ij} = \frac{|\{m \in M_i \cap M_j : v_i(m) \text{ similar to } v_j(m)\}|}{|M_i \cap M_j|}. \tag{1}$$

We may see in s_{ij} the result of a voting in favor of the proposition"*critic c_i expresses similar opinions to critic c_j*" and we take such a proposition as more or less verified if it is supported by a more or less large majority of the movies the critics have conjointly evaluated. Ideally, only strict equal evaluations, i.e., $v_i(m) = v_j(m)$, should be considered as being similar. But, this kind of index gives in general poor clustering results as practically all critics may easily appear to express in fact different opinions (see Table 2). In our small sample of critics and movies, two clusters ({jpt, cn} and {vt, jh}) appear nevertheless satisfying our strict similarity condition.

We may however progressively soften this strict equality assumption and assume the existence of a similarity threshold, i.e., that a difference in eval-

Table 2
A strict similarity based proximity index

s_{ij}	jpt	cn	pf	vt	jh
jpt	**1**	**.6**	.2	0	0
cn	**.6**	**1**	.4	.2	.4
pf	.2	.4	1	.2	.2
vt	0	.2	.2	**1**	**.6**
jh	0	.4	.2	**.6**	**1**

uation of one, two or even more grades on the preference scale expresses nevertheless a more or less 'similar' qualitative judgement. Formally:

$$\forall (c_i, c_j) \in C \times C, \forall m \in \mid M_i \cap M_j : \\ v_i(m) \text{ similar to } v_j(m) \\ \iff \Delta = |v_i(m) - v_j(m)| \leqslant k \\ \text{with } k = 0, 1, \ldots \quad (2)$$

We obtain thus, by choosing a similarity threshold Δ with larger and larger values k, larger and larger clusters of critics expressing more or less similar opinions. And in the limit, if all judgements are to be considered as "similar" appreciations, all critics are consequently to be seen as expressing same similar opinions, i.e., we gather indeed the whole set C of critics as a global equivalence class. In Table 3, we show on our small sample data extract the proximity index for $\Delta \leqslant 1$ where a difference in evaluation of one grade or less is considered to express a similar opinion. It is worthwhile noticing in Table 3 that we obtain with our proximity index (see Formulas (1) and (2)) in general a strictly symmetric ($s_{ij} = s_{ji}\ \forall i, j = 1..5$) but potentially intransitive proximity relation ($s_{\text{jpt,vt}} > 0.5$ and $s_{\text{vt,jh}} > 0.5$ but $s_{\text{jpt,jh}} < 0.5$ for instance).

If we adopt now a simple majority rule ($s_{ij} > .5$) for fixing a credible proximity, we obtain in our sample data set two overlapping clusters: {jpt, cn, pf, vt} and {cn, pf, vt, jh} as may be seen in Table 3.

To construct formally such similarity clusters from a given proximity index constructed on the whole set of evaluations, we use $\mathscr{L}$-valued kernel constructions (see [1–3]). The reader more interested in the practical clustering results may jump over the following section and later come back to the more formal constructions at the basis of our general clustering approach.

Table 3
Example of relaxed proximity index ($\Delta \leqslant 1$)

s_{ij}	jpt	cn	pf	vt	jh
jpt	**1**	**.6**	**.6**	**.6**	.4
cn	**.6**	**1**	**1**	**.6**	**.6**
pf	**.6**	**1**	**1**	**.8**	**.6**
vt	**.6**	**.6**	**.8**	**1**	**.8**
jh	.4	**.6**	**.6**	**.8**	**1**

3. Computing similarity classes from $\mathcal{L}$-valued proximity relations

In this section, we first briefly present the concept of symmetric or projectively boolean $\mathscr{L}$-valued credibility calculus. In Section 3.2, we then formally introduce $\mathscr{L}$-valued proximity relations and corresponding $\mathscr{L}$-valued similarity classes. In Section 3.3, we introduce kernels on $\mathscr{L}$-valued relations and finally show how to construct associated similarity classes by using null kernels, i.e., conjointly initial and terminal kernel solutions on $\mathscr{L}$-valued proximity relations.

3.1. $\mathscr{L}$-valued credibility calculus

In the truth assessment via the majority rule of the previous similarity assertions, we made a clear semantic distinction between the underlying credibility calculus qualifying the truthfulness of given similarity assertions, and the effective truthfulness of the logical expression involving these assertions. [1]

More formally, let $\mathscr{P}$ represent a set of atomic assertions p to which we may associate a finite rational degree of credibility $r(p) \in [0, 1]$ describing its potential truthfulness. If $r(p) = 1$, assertion p is perceived as *certainly true*, and if $r(p) = 0$, it is perceived as *certainly false*. The complete ordered finite set of involved credibility degrees is denoted V. Their underlying ordering is denoted $(V, \leqslant)$, where $\leqslant$ denotes a complete, reflexive, anti-symmetric and transitive relation.

Let $(\mathscr{P}, r)$ be a set of atomic assertions p associated with corresponding degrees of credibility $r : \mathscr{P} \to V$. Let $\neg$, $\vee$, $\wedge$ and $\Rightarrow$ denote, respectively, negation, disjunction, conjunction and implication of logical expressions. The set $\mathscr{E}$ of all *well-formulated finite expressions* will be generated inductively from the following grammar:

$$\forall p \in \mathscr{P} : p \in \mathscr{E}, \\ \forall x, y \in \mathscr{E} : \neg x \mid (x) \mid x \vee y \mid x \wedge y \mid x \Rightarrow y \in \mathscr{E}.$$

[1] For a more general discussion of this approach see [6].

The unary operator $\neg$ has a higher precedence in the interpretation of a formula, but generally we use bracketing parentheses to control the application range of a given operator and thus to make all formulas have unambiguous semantics.

We extend the credibility calculus on such logical expressions in the following way. Let $\mathscr{E}$ be a set of well-formulated expressions based on $\mathscr{P}$. $\forall x, y \in \mathscr{E}$:

$$r(\neg x) = 1 - r(x), \tag{3}$$

$$r(x \vee y) = \max(r(x), r(y)), \tag{4}$$

$$r(x \wedge y) = \min(r(x), r(y)), \tag{5}$$

$$r(x \Rightarrow y) = \max(r(\neg x), r(y)). \tag{6}$$

From the inductive definition of our well-formulated expressions, we are thus able to compute the credibility of any such formula in what we call a *symmetric evaluation domain* $\mathscr{L} = (V, \leqslant, \neg, \min, \max, 0, \frac{1}{2}, 1)$. The negation operator '$\neg$' implements a strict anti-tonic bijection with credibility $\frac{1}{2}$ acting as negational fix-point. Classic min and max operators capture credibilities of conjunction respective disjunction of formulas. The implication operator follows the classic Kleene–Dienes definition, i.e., $x \Rightarrow y \equiv \neg(x \wedge \neg y)$.

Finally, we denote the couple $(\mathscr{E}, r)$ as $\mathscr{E}^{\mathscr{L}}$ and simply speak of $\mathscr{L}$-valued expressions in the rest of the paper.

Knowing the credibility of any given $\mathscr{L}$-valued expression, we are now able to induce its supposed truthfulness. In classical bi-valued logic, it is usual to work syntactically only on the *truth* point of view of the logic, the *falseness* point of view being redundant through the coercion to the excluded middle. For instance, writing "$(a, b) \in R$" implicitly means assuming that this proposition is actually true and its negation false, otherwise we would write "$(a, b) \notin R$". In our $\mathscr{L}$-valued logic however, each well-formed expression $x \in \mathscr{E}^{\mathscr{L}}$ is associated explicitly with a credibility degree $r(x)$ giving its truth denotation in the following way:

$$x \text{ is } \mathscr{L}\text{-true} \equiv r(x) \geqslant r(\neg x) \iff r(x) > \frac{1}{2}, \tag{7}$$

$$x \text{ is } \mathscr{L}\text{-false} \equiv r(\neg x) \geqslant r(x) \iff r(x) < \frac{1}{2}, \tag{8}$$

$$x \text{ is } \mathscr{L}\text{-undetermined} \equiv r(x) = r(\neg x) \iff r(x) = \frac{1}{2}. \tag{9}$$

Our induced $\mathscr{L}$-valued truth calculus is therefore complete on every set $\mathscr{E}^{\mathscr{L}}$ of well-formulated $\mathscr{L}$-valued expressions, i.e., any expression $x \in \mathscr{E}^{\mathscr{L}}$ is either $\mathscr{L}$-true, $\mathscr{L}$-false or$\mathscr{L}$-undetermined. Furthermore, truthfulness of a given expression x is only defined in case the expression's credibility $r(x)$ exceeds the credibility $r(\neg x)$ of its contradiction $\neg x$.

Concerning implicational expressions of the form '$x \Rightarrow y$', we furthermore impose that to be logically valid, a fact we call $\mathscr{L}$-proper, they must verify the following condition:

$$(x \Rightarrow y) \text{ is called } \mathscr{L}\text{-proper} \equiv r(x \Rightarrow y) \geqslant r(y \Rightarrow x) \iff r(x) \leqslant r(y). \tag{10}$$

An $\mathscr{L}$-implication is called proper [2] iff its credibility is at least as large as the credibility of the converse implication, or iff the credibility of the consequent is at least as large as that of the antecedent. This last condition is of great importance for our clustering approach.

Let us now introduce $\mathscr{L}$-valued proximity relations and corresponding $\mathscr{L}$-valued similarity classes.

3.2. $\mathscr{L}$-valued proximity relations and associated similarity classes

We call $\mathscr{L}$-valued *binary relation* on a finite set C the Cartesian product $\mathrm{S} = C \times C$ evaluated in $\mathscr{L}$. Such an $\mathscr{L}$-valued binary relation S is called a *proximity relation* if it is conjointly $\mathscr{L}$-reflexive and $\mathscr{L}$-symmetric, i.e., $\forall a, b \in C : a\mathrm{S}a$ is $\mathscr{L}$-true and $a\mathrm{S}b$ being $\mathscr{L}$-true implies $b\mathrm{S}a$ being $\mathscr{L}$-true. Tables 2 and 3 illustrate naturally such kind of fuzzy relations.

[2] In a classic Boolean evaluation domain, all implicational expressions are necessarily proper, so that this supplementary condition makes no sense there, contrary to the general $\mathscr{L}$-valued case, where $\mathscr{L}$-proper implications play a central role as will become evident in the $\mathscr{L}$-valued kernel constructions.

Given such an $\mathscr{L}$-valued relation S, we call $\mathscr{L}$-preclass, an $\mathscr{L}$-subset K of C, i.e., membership assertions $(a \in K)$, for all $a \in C$ evaluated in $\mathscr{L}$ such that:

$$\forall a, b \in C \\ : \min\{r(a \in K), r(b \in K)\} \leqslant r(a\mathrm{S}b). \quad (11)$$

The concept of preclass gathers the fact that two critics a and b are in a same preclass K with respect to a similarity relation S only if they are similar under S. Indeed, condition (11) expresses the $\mathscr{L}$-proper implication '$a, b \in K \Rightarrow a\mathrm{S}b$' in terms of the underlying $\mathscr{L}$-valued credibility calculus (see Formulas (5) and (10)).

In our clustering problem, we are naturally interested in particular $\mathscr{L}$-preclasses, namely those that will describe the largest eventual similarity classes we are looking for. Therefore we call $\mathscr{L}$-class, an $\mathscr{L}$-preclass K verifying following supplementary conditions:

$$\exists a_0 \in C : r(a_0 \in K) \geqslant \frac{1}{2}, \quad (12)$$

$$\forall a, b \in C : \min\{r(a \in K), r(a\mathrm{S}b)\} \leqslant r(b \in K). \quad (13)$$

An $\mathscr{L}$-class thus contains always at least one $\mathscr{L}$-true selected element and if a critic a, who is similar to a critic b, is in some class K, then this critic b is also in this same similarity class K. Indeed conditions (11)–(13) conjointly assure that the underlying preclass K is maximal in the sense of $\mathscr{L}$-true inclusion, i.e., gathers a maximum of critics from C.

Finally, we call $\mathscr{L}$-cover a family $\mathbb{K}$ of $\mathscr{L}$-subsets of C verifying the following condition: $\forall a \in C : \max_{K \in \mathbb{K}} r(a \in K) \geqslant \frac{1}{2}$. Examples of such $\mathscr{L}$-covers are given by the rows of Tables 2 and 3 for instance.

In our clustering problem we are now interested in constructing from a given proximity index, modelling in fact an $\mathscr{L}$-valued proximity relation S on the set of critics C, a set $\mathbb{K}$ of $\mathscr{L}$-classes that might give an $\mathscr{L}$-cover of the set of critics C. The operational instrument to do so is given by the $\mathscr{L}$-valued kernel construction (see [1]).

3.3. Initial and terminal kernels on $\mathscr{L}$-valued binary relations

Let R be any $\mathscr{L}$-valued binary relation on a finite set C. A kernel on R represents an $\mathscr{L}$-valued subset K of C which is conjointly maximal interior stable and minimal exterior stable [1]. The interior stability may be expressed in terms of credibility degrees by the following condition:

$$\forall a, b \in C : \min\{r(a \in K), r(b \in K)\} \leqslant 1 - r(a\mathrm{R}b). \quad (14)$$

All nodes $\mathscr{L}$-truly selected in the kernel are therefore mutually R-incomparable. Correspondingly, the exterior stability condition is formulated in terms of credibility degrees as follows:

$$\forall a \in C : \exists b \in C \\ : \min\{r(b \in K), r(b\mathrm{R}a)\} \leqslant 1 - r(a \in K). \quad (15)$$

If b is in the kernel K and b is in relation with a then a is not in the kernel K. We distinguish in general two types of exterior stabilities, initial or terminal ones, depending on the way we consider the relation R in condition (15) (($b\mathrm{R}a$) or ($a\mathrm{R}b$)). Initial kernels correspond to nodes dominating in the sense of R the nodes outside the kernel, and terminal kernels correspond to nodes absorbing in the sense of R the nodes outside the kernel. Condition (15) actually represents the dominating version and therefore formulates an "*initial*" exterior stability condition.

Computing now $\mathscr{L}$-valued kernels from a given $\mathscr{L}$-valued binary relation, is achieved by enumerating, with the help of constraint logic programming, all maximal degrees of credibility of the kernel membership assertions for every $a \in C$ where the stability conditions (14) and (15) are used as propagating mechanisms (see [3]).

In general, we denote $\mathbb{K}^i$ (respectively $\mathbb{K}^t$) the set of all initial (respectively terminal) $\mathscr{L}$-valued kernel solutions computable on a given graph (C, R), i.e., verifying interior and respective exterior $\mathscr{L}$-valued stability conditions.

To illustrate these concepts, we consider a first example of $\mathscr{L}$-valued binary relation (see Table 4) which represents an $\mathscr{L}$-true complete

Table 4
Example of $\mathscr{L}$-valued binary relation

A	*a*	*b*	*c*	*d*
a	1	.8	.8	.8
b	.2	1	.8	.8
c	.2	.2	1	.8
d	.2	.2	.2	1
K^i	**.8**	.2	.2	.2
K^t	.2	.2	.2	**.8**

order relation on A. For such an ordering, the corresponding $\mathscr{L}$-valued initial and terminal kernel solutions are shown in Table 4. The first solution K^i suggests with a credibility of 80%, node a as $\mathscr{L}$-true *initial kernel* and correspondingly the second solution K^t suggests with a similar credibility of 80%, node d as $\mathscr{L}$-true *terminal kernel*.

Let us now consider the special case of $\mathscr{L}$-valued proximity relations.

3.4. Null kernels on $\mathscr{L}$-valued proximity relations

To illustrate this case, let us first consider a sample $\mathscr{L}$-valued proximity relation S shown in Fig. 1 and Table 5. If we apply our kernel construction to the $\mathscr{L}$-complement S^c of this relation S ($aS^cb = \neg(aSb), \forall a, b \in A$), we obtain as conjointly initial and terminal kernel solutions, a set of $\mathscr{L}$-classes, i.e., subsets of similar elements under

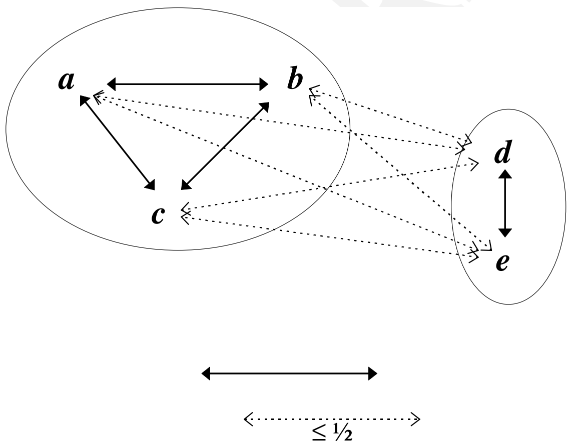

Fig. 1. Clsustering from an $\mathscr{L}$-valued proximity relation.

Table 5
Example of $\mathscr{L}$-valued proximity relation

S	*a*	*b*	*c*	*d*	*e*
a	**1**	**.7**	**.8**	.3	.2
b	**.8**	**1**	**.9**	.2	.1
c	**.7**	**.6**	**1**	.2	.2
d	.4	.2	.2	**1**	**.8**
e	.2	.3	.2	**.8**	**1**

relation S (see Table 6) giving the $\mathscr{L}$-cover we are looking for.

Indeed, initial and terminal kernel constructions do coincide on $\mathscr{L}$-symmetric relations and the sets $\mathbb{K}^i$ and $\mathbb{K}^t$ of initial, respective terminal kernel solutions for this kind of graphs represent the largest subset of nodes being conjointly interiorly stable, i.e., incomparable in the sense of the $\mathscr{L}$-complement of the proximity index, and exteriorly stable, i.e., comparable in the sense of the $\mathscr{L}$-complementary relation with all nodes outside the kernel subset. Or being incomparable (resp. comparable) in the $\mathscr{L}$-complementary relation, means being similar (resp. dissimilar) in the original similarity relation. As the proximity relation S is $\mathscr{L}$-symmetric, both initial and terminal solutions $\mathscr{L}$-truly select the same nodes (see Table 6) and we may speak of $\mathscr{L}$-valued *null kernels*. Furthermore, we notice that the interior stability condition (14) on R^c in fact represents exactly condition (11) of the $\mathscr{L}$-preclass concept. Indeed, let K be a null kernel on S^c:

$$\forall a, b \in C : \min\{r(a \in K), r(b \in K)\} \leqslant (1 - r(aS^cb) = 1 - (1 - r(aSb)) = r(aSb).$$

Table 6
Initial and terminal kernels from the complement of an $\mathscr{L}$-valued proximity relation

S^c	*a*	*b*	*c*	*d*	*e*
a	0	.3	.2	**.7**	**.8**
b	.2	0	.1	**.8**	**.9**
c	.3	.4	0	**.8**	**.8**
d	**.6**	**.8**	**.8**	0	.2
e	**.8**	**.7**	**.8**	.1	0
K_1^i	**.7**	**.6**	**.8**	.2	.2
K_2^i	.2	.2	.2	**.8**	**.8**
K_1^t	**.7**	**.7**	**.6**	.3	.3
K_2^t	.2	.2	.2	**.8**	**.8**

Table 7
Example of null kernel computation

Kernel	a	b	c	d	e
K_1^i	**.7**	**.6**	**.8**	.2	.2
K_1^t	**.7**	**.7**	**.6**	.3	.3
K_1^n	**.7**	**.6**	**.6**	.3	.3
K_2^i	.2	.2	.2	**.8**	**.8**
K_2^t	.2	.2	.2	**.8**	**.8**
K_2^n	.2	.2	.2	**.8**	**.8**

Similarly, exterior stability condition (15) as well implies condition (13) of the $\mathscr{L}$-class concept:

$$\forall a, b \in C : \min\{r(a \in K), (1 - r(a\mathrm{S}^{\mathrm{c}}b))\} = \min\{r(a \in K), r(a\mathrm{S}b)\} \leqslant r(b \in K).$$

In general, let again $\mathbb{K}^i$ and $\mathbb{K}^{\mathrm{t}}$ represent the $\mathscr{L}$-valued sets of all initial (respectively terminal) kernel solutions computable on an $\mathscr{L}$-valued proximity relation S which is $\mathscr{L}$-reflexive and $\mathscr{L}$-symmetric. We know (see [1]) that $\mathscr{L}$-symmetric relations admit the same $\mathscr{L}$-true initial and terminal kernel solutions. Let us identify j couples $(K_j^i, K_j^{\mathrm{t}})$ of such corresponding initial and terminal kernel solutions. From these couples, we construct a set of null kernels $\mathbb{K}_n$ in the following way: $\forall a \in A$ and $\forall (K_j^i, K_j^{\mathrm{t}}) \in \mathbb{K}^i \times \mathbb{K}^{\mathrm{t}}$

$$K_j^n(a) = \begin{cases} \min\left(r(a \in K_j^i), r(a \in K_j^{\mathrm{t}})\right) \\ \quad \iff r(a \in K_j^i) > \frac{1}{2}, \\ \max\left(r(a \in K_j^i), r(a \in K_j^{\mathrm{t}})\right) \\ \quad \iff r(a \in K_j^i) \leqslant \frac{1}{2}. \end{cases} \tag{16}$$

With this construction we assure that corresponding null kernels on S^{c} represent convenient $\mathscr{L}$-classes correctly modelling all similarity classes underlying a given similarity relation S.

As illustration, we may compute on our sample relation in Table 6 the null kernels $\mathbb{K}^n = \{K_1^n, K_2^n\}$ (see Table 7), where the two resulting $\mathscr{L}$-classes define indeed an appropriate $\mathscr{L}$-cover.

We may now come back to our initial practical problem, i.e., clustering the movie critics on the basis of their pairwise proximity index.

4. Clustering the Luxembourg movie critics

As mentioned in the beginning, the complete data set we collected for our proximity calculus is coming from the March/April and September issues of the "Graffiti" magazine. All gathered opinions are expressed by 12 movie critics (see Appendix A, Table 8) upon a list of 57 reference movies (see Appendix A, Fig. 5). The critic's evaluations are, as mentioned before, expressed on a purely ordinal scale numerically coded with seven grades $\{-2, -1, 0, 1, 2, 3, 4\}$; from four stars $(****=4)$ meaning "*very much appreciated*" to two zeros $(00=-2)$ meaning "*very much disliked*".

Unfortunately, our evaluation tableau shown in Fig. 5 contains a high rate of missing values, namely in case a critic has not had the opportunity to evaluate a movie.

We have investigated two possible ways of coping with these missing evaluations.

4.1. Changing missing evaluations into median ones

A first, classical solution consists in considering that all missing evaluations may be assimilated to a median evaluation. With this rule we obtain a strictly symmetric proximity index as shown in Table 9 in Appendix A.

On this proximity index, we obtain 11 null kernels with a strict equality $(\Delta = 0)$ for similar evaluations (see Appendix A, Table 10) and the clusters we deduce from these null kernels are shown in Fig. 2.

One may notice that we naturally obtain overlapping clusters mainly due to the partial $\mathscr{L}$-intransitivity of the underlying proximity relation.

If we relax now the similarity criterion by considering two evaluations, with a difference of up to one grade $(\Delta \leqslant 1)$, as still a more or less similar appreciation, we obtain a unique null kernel giving a complete $\mathscr{L}$-cover modelling the following unique similarity class:

$$K^n = \{\mathrm{jpt}(74), \mathrm{cf}(74), \mathrm{as}(74), \mathrm{rr}(72), \mathrm{vt}(70), \mathrm{mr}(68), \mathrm{RR}(68), \mathrm{pf}(67), \mathrm{dr}(67), \mathrm{jh}(63), \mathrm{rr}(60), \mathrm{cn}(60)\}. \tag{17}$$

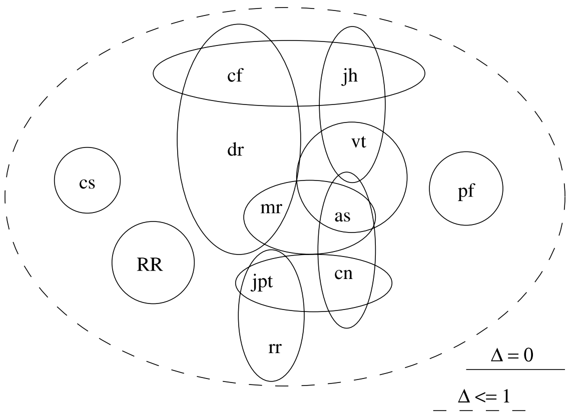

Fig. 2. Clustering the movie critics: solution 1.

Our clustering approach is monotone with respect to the similarity threshold Δ. Indeed, the relative levels of credibility concerning the class membership propositions rise monotonically with the relative frequency of observed similar preference judgements.

It appears unfortunately that replacing missing evaluations with median ones, introduces implicitly a lot of artificially created similarity between the critics' opinions, as for instance in case both missed to evaluate a same movie. Therefore our first clustering result appears not necessarily being very reliable and we propose hereafter an alternative way of coping with the commonly high rate of missing evaluations; a way being more "natural" from an algebraic point of view in the sense of your $\mathcal{L}$-valued credibility calculus.

4.2. "Naturally" taking into account missing evaluations

Our idea here is that in the limit, two critics, who have both seen none of our reference movies, express neither similar nor dissimilar opinions, i.e., the credibility of the proposition that "*the first critic expresses similar or dissimilar opinions compared to the second critic*" must be given an $\mathcal{L}$-undetermined value $\frac{1}{2}$.

Now, the more a critic is missing common evaluations with all the others, the more the proximity of his opinions with respect to all the other's, is tending towards the $\mathcal{L}$-undetermined value $\frac{1}{2}$. Formally, we adjust the former proximity index (see Eq. (2)) as follows.

Let s_{ij} be the original proximity index computed between the evaluations of critic c_i and critic c_j, and let m_{ij} be the ratio of common evaluations with respect to the number of reference movies. Then the proposed rectified proximity index s^r_{ij} is defined in the following way:

$$s^r_{ij} = s_{ij}\, m_{ij} + (1 - m_{ij})\tfrac{1}{2}. \qquad (18)$$

Semantically speaking, we weight the initial proximity index s_{ij} with the relative frequency of common evaluations, and we add halve of the relatively missing evaluations as similar and the other halve as dissimilar proportion. A graphical representation of the transformation may be seen in Fig. 3. In the limit, if m_{ij} approaches 1 (both critics have seen all reference movies), s^r_{ij} remains rather unchanged. On the other hand, if m_{ij} approaches the value 0, (no common evaluations between the critics), s^r_{ij} is more and more restricted to close values around $\frac{1}{2}$.

From a more technical point of view, the above proposed transformation is *natural* (in an algebraic categorical sense) for our kernel construction, in

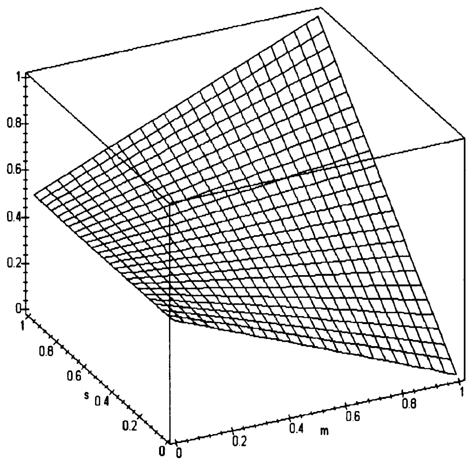

Fig. 3. Naturally taking into account missing evaluations.

the sense that $\mathscr{L}$-true (resp. $\mathscr{L}$-false) similarities remain $\mathscr{L}$-true (resp. $\mathscr{L}$-false) through transformation (18). Thus, the basic structure of the kernel solutions is coherently affected by the modification.

4.3. Clustering the movie critics

Considering now that only equal evaluations are similar ($\Delta = 0$), all similarity classes we obtain, are $\mathscr{L}$-singletons, i.e., one critic is just similar to himself. If we consider however a similarity threshold of one grade ($\Delta \leqslant 1$) we observe the null kernels shown in Fig. 4 (see Appendix A, Table 11).These null kernels model the following clusters:

1. *JP. Thilges* (Revue & Graffiti), *Viviane Thill* (Le Jeudi), Christian Spielmann (Journal), Claude Neu (Luxpost), Joy Hoffmann (Zinemag),
2. *JP. Thilges* (Revue & Graffiti), *Viviane Thill* (Le Jeudi), Christian Spielmann (Journal), Claude Neu (Luxpost), Romain Roll (Zeitung), Raoul Reis (Noticias & Radio ARA),
3. *JP. Thilges* (Revue & Graffiti), Duncan Roberts (Luxembourg News), Christian Spielmann (Journal), Romain Roll (Zeitung) Alain Stevenart (La Meuse),
4. *JP. Thilges* (Revue & Graffiti), *Viviane Thill* (Le Jeudi), Alain Stevenart (La Meuse),
5. *Viviane Thill* (Le Jeudi), Peter Feist (Grëngespoun),
6. Martine Reuter (Tageblatt & RTL Radio Lëtzebuerg),
7. Claude François (Luxemburger Wort & Télécran & DNR).

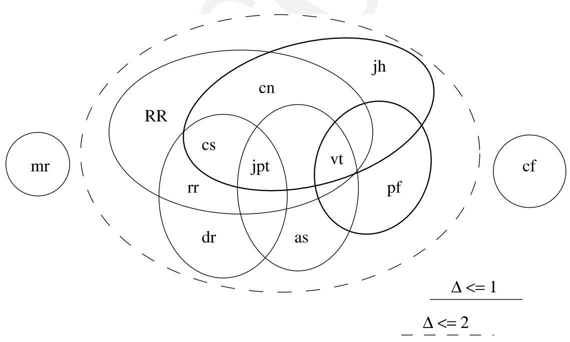

Fig. 4. Clustering the movie critics with missing values.

It is worth noticing, that our clusters partly overlap and therefore propose a rich interpretation for the media sociologist. JP. Thilges (jpt) for instance, as editor of the Graffiti magazine he apparently takes a central position by expressing at the same time and in some particular sense, similar opinions to different subsets of critics, either more marginal or more popular press oriented ones. But also, Viviane Thill (*vt*), one of the outstanding movie critics in Luxembourg, clearly appears as a leading opinion maker. Furthermore, we notice that both isolated critics, M. Reuter (mr) and Cl. François (cf) have missed a lot of evaluations and it appears quite natural with our approach, that they therefore do not compare well with all the other critics.

Finally, if we furthermore relax our similarity condition in considering a difference of up to two grades ($\Delta \leqslant 2$) as still being 'insignificant', we obtain one big cluster gathering all critics except both previous journalists, who remain all the same incomparable.

5. Conclusion

In this paper, we propose an innovative method for constructing fuzzy similarity classes from $\mathscr{L}$-valued proximity relations. First we have introduced the practical concern of our investigation, namely clustering a set of movie critics from a given set of evaluations on a reference set of movies. The formal problem of constructing clusters on this kind of data is operationally solved with the help of $\mathscr{L}$-valued null kernels, i.e., kernels being conjointly initial and terminal. Finally, an original method for dealing with numerous missing evaluations has been developed and discussed.

Appendix A

See Tables 8–11 and Fig. 5.

Table 8
The Luxembourg Movie Critics in our data sets

Identifier	Name	Press affiliation
jpt	JP Thilges	Revue and Graffiti
cn	Claude Neu	Luxpost
mr	Martine Reuter	Tageblatt and RTL Radio Lëtzebuerg
as	Alain Stevenart	La Meuse
pf	Peter Feist	Grengespoun
vt	Viviane Thill	Le Jeudi
dr	Duncan Roberts and Luxembourg News	
jh	Joy Hoffmann	Zinemag
rr	Romain Roll	Zeitung
RR	Raoul Reis	Noticias & Radio Ara
cs	Christian Spielman	Journal
cf	Claude Francois	LW and Telecran and DNR

Table 9
Proximity index with strict similarity and without missing values

$s_{ij/\Delta=0}$	jpt	cn	mr	as	pf	vt	dr	jh	rr	RR	cs	cf
jpt	100	53	37	37	42	30	46	39	53	40	37	47
cn		100	53	46	42	40	49	44	49	44	32	47
mr			100	67	44	47	60	46	39	46	42	56
as				100	44	51	49	40	35	39	26	44
pf					100	30	44	32	39	30	35	46
vt						100	33	56	33	30	21	46
dr							100	46	47	33	44	56
jh								100	35	39	28	58
rr									100	42	42	46
RR										100	32	44
cs											100	40
cf												100

Table 10
Null kernels on s_{ij} with a strict similarity ($\Delta = 0$)

$s_{ij/\Delta=0}$	jpt	cn	mr	as	pf	vt	dr	jh	rr	RR	cs	cf
K_1^n	**53**	**51**	47	47	47	47	47	47	49	47	47	47
K_2^n	**53**	49	47	47	47	47	47	47	**51**	47	47	47
K_3^n	49	49	49	**51**	49	**51**	49	49	49	49	49	49
K_4^n	47	**51**	**53**	49	47	47	49	47	47	47	47	49
K_5^n	47	49	**53**	**51**	47	47	49	47	47	47	47	49
K_6^n	47	47	**53**	47	47	47	**53**	47	47	47	47	53
K_7^n	46	46	46	46	**54**	46	46	46	46	46	46	46
K_8^n	46	46	46	46	46	**54**	46	**54**	46	46	46	46
K_9^n	46	46	46	46	46	46	46	**54**	46	46	46	**54**
K_{10}^n	46	46	46	46	46	46	46	46	46	**54**	46	46
K_{11}^n	44	44	44	44	44	44	44	44	44	44	**56**	44

Movies	jpt	cn	mr	as	pf	vt	dr	jh	rr	RR	cs	cf
Courage under Fire	**	**	*	/	**	*	*	*	**	*	*	*
Didier	**	**	/	**	*	*	/	***	**	**	***	/
Un Été à la Goulette	***	*	/	***	*	**	/	*	/	*	*	/
The First Wives Club	**	o	*	*	oo	o	*	o	**	*	*	**
The Frighteners	*	**	**	/	/	*	**	**	***	***	**	**
Lost Highway	****	****	***	***	***	*	**	*	****	****	**	**
The Mirror has two Faces	**	/	*	/	/	*	/	*	*	/	**	**
The Pillow Book	***	***	***	/	***	**	**	*	***	***	*	/
Portrait of a Lady	***	***	**	*	***	**	*	/	/	**	o	**
La Promesse	***	****	**	****	***	****	**	***	***	*	***	***
Ransom	**	/	*	o	/	***	*	***	**	**	**	*
The Relic	**	**	/	/	/	/	*	*	***	o	*	*
Salut Cousin	***	**	**	**	**	***	**	**	***	/	*	**
Tout doit disparaître	*	/	/	/	/	/	/	oo	/	/	*	*
Truth about Cats and Dogs	**	*	*	*	*	***	*	***	***	**	**	**
Blood and Wine	**	*	*	**	***	**	*	**	**	**	*	**
The Cricible	**	*	/	*	*	*	***	*	**	*	***	*
Dante's Peak	**	/	*	/	o	*	*	**	*	*	o	o
The Devil's Own	**	**	o	*	/	*	o	*	*	**	*	*
The Empire strikes back	***	***	**	**	*	***	/	/	***	**	****	***
Everyone says I love you	****	****	***	****	**	****	***	****	/	****	o	**
The Fan	*	*	/	o	o	o	*	/	*	*	*	*
Jerry Maguire	**	*	*	oo	o	oo	**	oo	**	***	**	**
Le Jour et la Nuit	oo	/	/	/	o	oo	/	oo	/	/	o	/
Jude	***	***	***	***	****	***	**	***	****	*	*	**
Kleines Arschloch	*	**	/	/	*	**	/	***	/	o	*	o
Michael Collins	***	***	**	**	**	***	***	***	****	**	**	/
101 Damatiens	*	**	o	o	/	o	*	/	*	**	**	*
Shine	***	***	**	**	***	**	***	***	***	**	**	****
Les Soeurs Soleil	/	oo	/	oo	/	/	/	oo	/	/	o	/
Star Wars	****	***	**	**	**	**	***	/	****	***	****	***
Troublemakers	**	**	**	/	**	**	**	**	***	*	*	**
Absolute Power	***	**	**	***	/	*	*	*	**	***	**	o
Antonia's Line	****	***	**	*	****	***	***	**	***	**	***	/
Arlette	*	o	/	o	*	o	/	o	/	oo	*	o
Assassin(s)	oo	***	*	o	*	**	*	**	*	**	*	**
Balto	**	/	/	/	/	/	/	**	*	*	**	/
Beavis & Butthead do America	/	/	oo	oo	/	oo	**	oo	**	**	**	/
Big Night	***	**	/	***	***	**	***	**	/	/	*	***
Carla's Song	***	**	/	**	*	*	***	*	**	*	*	/
Donnie Brasco	**	**	*	*	**	***	***	***	***	**	***	***
The Fith Element	**	*	*	*	o	*	**	**	***	**	***	**
The Funeral	***	****	/	**	**	***	**	****	*	****	*	/
Funny Boys	**	**	/	/	**	*	**	o	**	/	**	**
Grace of my Heart	**	**	*	***	/	*	**	o	/	*	**	/
Lorenz im Land der Lügner	/	/	/	/	/	/	/	/	*	/	o	**
Michael	*	/	/	*	o	/	*	**	*	o	*	/
Palookville	**	**	/	***	**	***	**	**	/	*	**	**
Return of the Jedi	***	***	**	**	**	**	/	**	***	***	***	**
Romeo & Juliet	***	o	**	***	***	**	***	/	***	***	**	**
Smilla's Sense of Snow	*	**	*	*	o	/	o	/	**	**	*	*
Some Mother's Son	***	***	/	***	**	**	***	***	**	o	****	/
Tenue Correcte Exigée	*	*	/	/	/	**	/	**	/	o	*	**
Turbulence	*	/	o	/	oo	o	*	/	o	/	*	o
Unstrung Heroes	***	**	**	**	***	***	**	****	**	***	**	***
When we were Kings	/	**	/	****	****	***	****	**	***	**	**	/
Y-aura-t-il de la Neige à Noël ?	/	***	**	*	***	***	/	**	/	/	**	/

Fig. 5. The complete data set (*Source*: Graffiti March/April and September 1997).

Table 11
Null kernels on s^r_{ij} with relaxed similarity ($\Delta \leqslant 1$)

$s^r_{ij/\Delta \leqslant 1}$	jpt	cn	mr	as	pf	vt	dr	jh	rr	RR	cs	cf
K^n_1	**52**	**52**	46	46	46	**51**	48	49	**51**	**51**	**54**	46
K^n_2	**52**	**52**	46	46	46	**52**	48	**52**	48	48	**54**	46
K^n_3	**51**	49	49	**51**	49	**51**	49	49	49	49	49	49
K^n_4	**52**	49	46	46	46	49	**51**	49	**51**	49	**54**	46
K^n_5	49	49	**51**	49	49	49	49	49	49	49	49	49
K^n_6	48	48	48	48	**52**	**52**	48	48	48	48	48	48
K^n_7	47	47	47	47	47	47	47	47	47	47	47	**53**

References

[1] R. Bisdorff, M. Roubens, On defining fuzzy kernels from $\mathscr{L}$-valued simple graphs, in: Proceedings Information Processing and Management of Uncertainty, IPMU'96, Granada, July 1996, pp. 593–599.

[2] R. Bisdorff, M. Roubens, On defining and computing fuzzy kernels from $\mathscr{L}$-valued simple graphs, in: Da Ruan, et al. (Eds.), Intelligent Systems and Soft Computing for Nuclear Science and Industry, FLINS'96 Workshop, World Scientific, Singapore, 1996, pp. 113–123.

[3] R. Bisdorff, On computing kernels from $\mathscr{L}$-valued simple graphs, in: Proceedings Fifth European Congress on Intelligent Techniques and Soft Computing EUFIT'97, Aachen, September 1, 1997, pp. 97–103.

[4] R. Bisdorff, Bi-pole ranking from pairwise comparisons by using initial and terminal $\mathscr{L}$-valued kernels, in: Proceedings of the conference IPMU'98 (Editions E.D.K., Paris), 1998, pp. 318–323.

[5] R. Bisdorff, Bi-pole ranking from pairwise fuzzy outranking, Belgian Journal of Operations Research Statistics and Computer Science, JORBEL 37 (4) 97 (1999) 53–70.

[6] R. Bisdorff, Logical foundation of fuzzy preferential systems with application to the Electre decision aid methods, Computers and Operations Research 27 (2000) 673–687.

[7] J. Fodor, M. Roubens, Fuzzy Preference Modelling and Multi-Criteria Decision Support, Kluwer Academic Publishers, Dordrecht, 1994.

[8] B. Roy, D. Bouyssou, Aide multicritère à la décision: Méhodes et cas, Economica, Paris, 1993.