Learning Preferences in Lexicographic Choice Logic

Karima Sedki[1,2][a], Nada Boudegzdame[1,2][b], Jean Baptiste Lamy[1,2][c] and Rosy Tsopra[3,4][d]

[1]*LIMICS (INSERM UMRS 1142), Université Paris 13, Sorbonne Paris Cité, 93017 Bobigny, France*

[2]*UPMC Université Paris 6, Sorbonne Universités, Paris, France*

[3]*INSERM, Université de Paris, Sorbonne Université, Centre de Recherche des Cordeliers,*
Information Sciences to Support Personalized Medicine, F-75006 Paris, France

[4]*Inria Paris, 75012 Paris Department of Medical Informatics, Hopital European Georges-Pompidou, AP-HP, Paris, France*

Keywords: Preferences Learning, Lexicographic Choice Logic.

Abstract: Lexicographic Choice Logic (*LCL*) is a variant of Qualitative Choice Logic which is a logic-based formalism for preference handling. The *LCL* logic extends the propositional logic with a new connective ($\vec{\diamond}$) to express preferences. Given a preference $x\vec{\diamond}y$, satisfying both x and y is the best option, the second best option is to satisfy only x, and satisfying only y is the third best option. Satisfying neither x nor y is not acceptable. In this paper, we propose a method for learning preferences in the context of *LCL*. The method is based on an adaptation of association rules based on the APRIORI algorithm. The adaptation consists essentially of using variations of the support and confidence measures that are suitable for *LCL* semantic.

1 INTRODUCTION

Preferences can be obtained in two ways: i) by elicitation from the user, through a sequence of queries/answers or ii) directly learning them from data. However, even if powerful formalisms have been proposed, preference elicitation is in general not an easy task especially when there are too many outcomes. It is then more appealing to learn preferences from data which is easy to collect. Preference learning (Johannes and Hüllermeier, 2010) has recently received a considerable attention in many disciplines. It aims to learn a preference model from observed preference information. There are three preference learning problems (Johannes and Hüllermeier, 2010): i) Object ranking problem (Waegeman and De Baets, 2010; Joachims et al., 2005), ii) Label ranking (Hüllermeier et al., 2008; Vembu and Gärtner, 2010) , and iii) Instance ranking problem (Cohen et al., 2011).

The purpose of this paper is to learn preferences in the context of Lexicographic Choice Logic (*LCL*) (Bernreiter et al., 2022). *LCL* is a variant of the well-known preference formalism Qualitative Choice Logic (Brewka et al., 2004). In *QCL*, to express preferences, an ordered disjunction connective is added to propositional logic. Intuitively, if x and y are two options then $x\vec{\times}y$ means: "if possible x, but if x is impossible then at least y". *LCL* allows to encode lexicographic ordering over variables by using the logical connective $\vec{\diamond}$. Given a preference $x\vec{\diamond}y$, satisfying both x and y is the best option, satisfying only x is the second best option, and the third best option is to satisfy only y. Satisfying neither x nor y is not acceptable.

The proposed method consists in an adaptation of association rules based on the APRIORI algorithm (Agrawal et al., 1993). The adaptation consists essentially of using variations of the support and confidence measures according to the semantic of *LCL*. In previous work (Sedki et al., 2022), we proposed a method for learning preferences in the context of *QCL*. It is also based on the adaptation of association rules as for *LCL*. However, the method for learning preferences in *LCL* requires different definitions than those of *QCL*, particularly the support, the confidence and the length of the learned formulas.

The paper is organized as follows: We start with some useful notations, then we present a description of some important elements of *LCL*. The fourth section describes the proposed method for learning *LCL* preferences. In Section 5 we present a case study in

[a] https://orcid.org/0000-0002-2712-5431
[b] https://orcid.org/0000-0003-1409-6560
[c] https://orcid.org/0000-0002-6078-0905
[d] https://orcid.org/0000-0002-9406-5547

Sedki, K., Boudegzdame, N., Lamy, J. and Tsopra, R.
Learning Preferences in Lexicographic Choice Logic.
DOI: 10.5220/0011891300003393
In *Proceedings of the 15th International Conference on Agents and Artificial Intelligence (ICAART 2023) - Volume 3*, pages 1012-1019
ISBN: 978-989-758-623-1; ISSN: 2184-433X

the front of medical domain, particularly in antibiotics prescription where we aim to learn an *LCL* preference model of medical experts that provide recommendations of antibiotics. Finally, we conclude the paper.

2 NOTATIONS

Let $\mathcal{V}$ be a finite set of propositional variables. An interpretation I is defined as a set of propositional variables such that $v \in I$ if and only if v is set to true by I. If I satisfies a formula ϕ, we write $I \models \phi$, otherwise, we write $I \not\models \phi$. A model of a formula ϕ is an interpretation I that satisfies the formula. $var(\phi)$ denotes the variables of a formula ϕ. $|.|$ denotes the cardinality of a set. Let us give the following definition:

***Definition* 1.** *Let $\mathcal{V}$ be a set of variables, then:*

- *Each single option is built on the set of variables $\mathcal{V}$ and the connective $\neg$.*
- *Each conjunctive option is built on the set of single options and the connective $\wedge$.*
- *The set of single and conjunctive options is denoted by X.*

v_1, $\neg v_2$ are examples of single options, $v_1 \wedge \neg v_2$ is a conjunctive option.

3 LEXICOGRAPHIC CHOICE LOGIC (*LCL*)

Lexicographic Choice Logic (Bernreiter et al., 2022) is a variant of Qualitative Choice Logic (Brewka et al., 2004). It has two types of connectives: classical connectives (here we use $\neg$, $\vee$, and $\wedge$) and a new connective $\vec{\diamond}$, used to encode lexicographic ordering over variables. Given a preference $x \vec{\diamond} y$, satisfying both x and y is the best option, satisfying only x is the second best option, and the third best option is to satisfy only y. Satisfying neither x nor y is not acceptable. In this paper, we do not use the unified language proposed in (Bernreiter et al., 2022) since we focus only on *LCL* and not on the other choice logics such as *QCL* (Brewka et al., 2004), *PQCL* (Benferhat and Sedki, 2008) and *CCL* (Boudjelida and Benferhat, 2016). Note that the connective $\vec{\diamond}$ of *LCL* is not associative (see (Bernreiter et al., 2022) for more details). Here, we follow the presentation given in (Bernreiter et al., 2022), where we consider a *BCF* formula ϕ presented as follows: $(x_1 \vec{\diamond} (x_2 \vec{\diamond} (\dots (x_{n-1} \vec{\diamond} x_n))))$.
As for *QCL*, the *LCL* language is composed of three types of formulas defined in the following.

***Definition* 2.** *Each propositional formula is built on the set of variables $\mathcal{V}$ and the connectives $\wedge$, $\vee$, $\neg$. $PROP_{\mathcal{V}}$ denotes the language of propositional formulas.*

***Definition* 3.** *Basic choice formulas (BCF) allow the expression of simple preferences. $BCF_{\mathcal{V}}$ denotes the language of BCF formulas and defined as follows:*

a) If ϕ, $\psi \in PROP_{\mathcal{V}}$ then $\phi \vec{\diamond} \psi \in BCF_{\mathcal{V}}$.

b) Every BCF formula is only obtained by applying the item a) above a finite number of times.

***Definition* 4.** *General choice formulas (GCF) can be obtained from $\mathcal{V}$ using connectives $\vec{\diamond}$, $\wedge$, $\vee$, $\neg$. The language composed of GCF formulas, denoted $GCF_{\mathcal{V}}$, is defined as follows:*

c) If ϕ, $\psi \in BCF_{\mathcal{V}}$ then $(\phi \wedge \psi)$, $\neg(\psi)$, $(\phi \vee \psi)$, $(\phi \vec{\diamond} \psi)$ $\in GCF_{\mathcal{V}}$.

d) The language of $GCF_{\mathcal{V}}$ is obtained by applying the item c) a finite number of times.

***Example* 1.** *$\phi_1 = a \vee b$ is an example of propositional formula. $\phi_2 = a \vec{\diamond} b \vec{\diamond} c$ is a BCF formula, $\phi_3 = (a \vec{\diamond} b) \wedge (c \vec{\diamond} d)$ is a GCF formula.*

3.1 Semantics and syntax of *LCL*

The semantics of an *LCL* formula is based on the degree of satisfaction of a formula in a particular interpretation I. The satisfaction degree of a formula given an interpretation is a positive natural number when a formula is satisfied by that interpretation or ∞ otherwise. The higher this degree, the less preferable the interpretation. Unacceptable interpretations have a degree of ∞. The set of satisfaction degrees is denoted by D. Let us define firstly the notions of optionality and length.

The optionality of a formula ϕ is a function that assigns to ϕ a strictly positive integer. It corresponds to the greatest satisfaction degree d ($d \neq \infty$) of all possible degrees of ϕ. The definition of optionality of *LCL* formulas is given in the following.

***Definition* 5.** *The optionality in LCL is defined as follows:*

1. *$opt(v)$=1, for every v in $\mathcal{V}$.*
2. *$opt(\phi \vec{\diamond} \psi) = (opt(\phi) + 1) \times (opt(\psi) + 1) - 1$*
3. *$opt(\phi \wedge \psi) = max(opt(\phi), opt(\psi))$.*
4. *$opt(\phi \vee \psi) = max(opt(\phi), opt(\psi))$.*
5. *$opt(\neg\phi) = 1$.*

We can observe that from Definition 5, the optionality of a *BCF* formula $\phi = (x_1 \vec{\diamond} (x_2 \vec{\diamond} (\dots (x_{n-1} \vec{\diamond} x_n))))$, $opt(\phi) = 2^n - 1$. It corresponds to the degree ascribed by the last preferred interpretation to ϕ.

***Example* 2.** *Let us consider the BCF formula* ϕ= *(a* $\vec{\diamond}$ *(b* $\vec{\diamond}$ *c)). From Item 2 of Definition 5,* $opt(\phi) = 7$*. This means that there an interpretation that ascribes a degree* 7 *to* ϕ *but there is no interpretation that ascribes a degree greater than* 7 *to* ϕ*.*

Let us give in the following the definition of length of *LCL* formulas.

***Definition* 6.** *The length of an LCL formula* ϕ*, denoted by* $len(\phi)$ *corresponds to the number of options that* ϕ *contains.*

1. $len(v)$=1*, for every v in* $\mathcal{V}$*.*
2. $len(\phi \vec{\diamond} \psi)$= $len(\phi)$ + $len(\psi)$
3. $len(\phi \wedge \psi)$= $max(len(\phi), len(\psi))$.
4. $len(\phi \vee \psi)$= $max(len(\phi), len(\psi))$.
5. $len(\neg\phi)$ = 1.

***Example* 3.** *Let us consider the BCF formula* ϕ= *a* $\vec{\diamond}$ *b* $\vec{\diamond}$ *c. From Item 2 of Definition 6,* $len(\phi) = 3$*. This means that* ϕ *contains* 3 *options (a, b, and c).*

Let us now define the inference relation of *LCL* formulas.

***Definition* 7.** *Let v be a propositional atom in* $PROP_{\mathcal{V}}$*,* ϕ *and* ψ *be two LCL formulas, I be an interpretation. The satisfaction degree of a formula* ϕ *under an interpretation I is denoted by* $deg(I, \phi)$*.*

1. $deg(I,v) = \begin{cases} 1 & if \quad v \in I \\ \infty & if \quad v \notin I \end{cases}$

2. $deg(I,\neg\phi) = \begin{cases} 1 & if \quad I \not\models \phi \\ \infty & otherwise \end{cases}$

3. $deg(I,\phi\vec{\diamond}\psi) = \begin{cases} (m-1) \times opt(\psi) + n \\ if \quad I \models_m \phi, I \models_n \psi \\ opt(\phi) \times opt(\psi) + m \\ if \quad I \models_m \phi, I \not\models \psi \\ opt(\phi) \times opt(\psi) + opt(\phi) + n \\ if \quad I \not\models \phi, I \models_n \psi \\ \infty \quad otherwise \end{cases}$

4. $deg(I,\phi \wedge \psi) =$
$\begin{cases} max(m,n) & if \quad I \models_m \phi \quad and \quad I \models_n \psi \\ \infty & otherwise \end{cases}$

5. $deg(I,\phi \vee \psi) = \begin{cases} min(m,n) & if \quad I \models_m \phi \quad or \quad I \models_n \psi \\ \infty & if \quad I \not\models \phi \quad and \quad I \not\models \psi \end{cases}$

For any propositional formula ϕ, there is only one degree of satisfaction (namely 1) obtained when ϕ is satisfied by *I*. Namely, if a propositional formula is satisfied, then it can only be to a degree of 1, otherwise, the degree is ∞. The formula $\phi \wedge \psi$ is assigned the maximum degree of ϕ and ψ because both formulas need to be satisfied. $\phi \vee \psi$ is assigned the minimum degree since it is sufficient to satisfy either ϕ or ψ. Regarding a *BCF* formula ϕ=$(x_1\vec{\diamond}(x_2\vec{\diamond}(...(x_{n-1}\vec{\diamond}x_n)))$, an interpretation *I* satisfies ϕ to a degree 1, if it satisfies the *n* options of ϕ, namely $I \models x_1 \wedge \ldots \wedge x_n$, an interpretation *I* satisfies ϕ to a degree 2, if it satisfies the $n-1$ options of ϕ, namely $I \models x_1 \wedge \ldots \wedge x_{n-1}$ and so on. The formula $\neg(x_1\vec{\diamond}x_2\vec{\diamond}\ldots\vec{\diamond}x_n)$ is equivalent to the propositional formula $\neg x_1 \wedge \neg x_2 \wedge ... \wedge \neg x_n$. So, its degree of satisfaction is 1 if ϕ is not satisfied, ∞ otherwise.

***Example* 4.** *Let us consider the LCL formula* ϕ = *(a* $\vec{\diamond}$ *(b* $\vec{\diamond}$ *c)). The satisfaction degree of* ϕ *for each interpretation is given in Table 1.*

Table 1: The *LCL* inference relation of ϕ.

Interpretations	a	b	c	$deg(I,\phi)$
I_1	F	F	F	∞
I_2	F	F	T	7
I_3	F	T	F	6
I_4	F	T	T	5
I_5	T	F	F	4
I_6	T	F	T	3
I_7	T	T	F	2
I_8	T	T	T	1

With respect to the lexicographic orderings of the variables a, b, and c in ϕ*, the interpretation* I_8=*{a, b, c} ascribes a degree of* 1 *to* ϕ*, the interpretation* I_7=*{a, b} ascribes a degree of* 2 *to* ϕ_1*, and so on. The interpretation* $\emptyset$ *ascribes a degree of* ∞ *to* ϕ*.*

4 LEARNING PREFERENCES IN *LCL*

We consider the problem of learning *LCL* preferences from a preference database $\mathcal{P}$ containing a set of interpretations described with a set of variables and each one associating with a satisfaction degree indicating the degree of preference for the user. In this paper, we are restricted for learning *BCF* formulas. Learning *GCF* formulas follows the same method since each *GCF* formula can be transformed into its equivalent *BCF* formula, the same thing for propositional formula which corresponds to a *BCF* formula with only one option. In addition, for lack of space, we limited to single and conjunctive options. Namely, *BCF* formulas contains only single or conjunctive options instead of a propositional formulas.

***Definition* 8.** *(Bernreiter et al., 2022) Let* $\mathcal{P}$ *be a preference database, I be an interpretation in* $\mathcal{P}$*, D be a set of satisfaction degrees s.t. I is assigned with a degree in D. A degree d is LCL-obtainable from D iff there exists an interpretation I and an LCL formula* ϕ *s.t.* $deg(I, \phi) = d$*. The set of all degrees obtainable from* $\mathcal{P}$ *is denoted by* D_{LCL}*.*

From Definition 8, $\forall\ \phi \in PROP_{\mathcal{V}}$, we have D_{LCL}=$\{1,\infty\}$, $\forall\ \phi \in BCF_{\mathcal{V}}$, we have D_{LCL}=$\mathbb{N}$.

***Proposition* 1.** *Let $\mathcal{P}$ be a preference database, D be a set of satisfaction degrees in $\mathcal{P}$, $d = max(D)$ s.t. $d \neq \infty$. Then, for a smallest positive natural number n verifying that $d < 2^n$ it holds that there exists an LCL formula ϕ s.t. $len(\phi)$ = n.*

Proof. Assume that $\mathcal{P}$ contains a set of interpretations where each one is assigned with one satisfaction degree d in D. Assume that D is *LCL*-obtainable from D. From Definition 8, there exists a formula ϕ s.t. $deg(I,\phi) \in D_{LCL}$.

- If $\phi\in PROP_{\mathcal{V}}$, then from Definition 6, $len(\phi) = 1$. From Definition 8, D_{LCL}=$\{1,\infty\}$. Thus, $d = max(D) = 1$. For a smallest number $n = 1$, we have $d = 1 < 2^{n=1}$.
- If $\phi\in BCF_{\mathcal{V}}$ s.t. ϕ=$(x_1\vec{\diamond}(x_2\vec{\diamond}(\ldots(x_{n-1}\vec{\diamond}x_n)))$, then from Definition 6, $len(\phi) = n$. From Definition 8, D_{LCL}=$\mathbb{N}$. According to the lexicographic ordering over x_i, we have $\forall d \in D_{LCL}$ ($d \neq \infty$), we have $d \leq 2^n - 1$ which means that if $d \geq 2^n$, then there is no *LCL* formula s.t. $len(\phi) = n$ so that d be *LCL*-obtainable from D.

□

For example, assume that D=$\{1,2,3,4,\infty\}$. $\forall\ d\in D$, d is *LCL*-obtainable iff $d = 4$ verifies Proposition 1. A smallest positive natural number that verifies Proposition 1 is $n = 3$. So, $d = 4$ be *LCL*-obtainable, iff there exists an *LCL* formula ϕ s.t. $len(\phi) = 3$. Thus, D_{LCL}=$\{1,2,3,4,\infty\}$ is *LCL*-obtainable from D since there is an *LCL* formula ϕ s.t. $len(\phi)$=3. It is clear that if we consider ψ s.t. $len(\psi) = 2$, D_{LCL}=$\{1,2,3,4,\infty\}$ is not *LCL*-obtainable from D since even if each of $d = 1$ (resp. 2, 3, ∞) is *LCL*-obtainable with ψ s.t. $len(\psi) = 2$, $d = 4$ is not.

We aim to learn a preference model $\mathcal{M}_{LCL}$ defined as follows.

***Definition* 9.** *Let $\mathcal{P}$ be a preference data base, D be the set of satisfaction degrees, then*

$$\mathcal{M}_{LCL} = \begin{cases} \phi = (x_1\vec{\diamond}(x_2\vec{\diamond}(\ldots(x_{n-1}\vec{\diamond}x_n))) \\ x_i \in X \quad and \quad len(\phi) = n \\ according \quad to \quad Proposition \quad 1 \end{cases}$$

Definition 9 states that the learned preference model contains a *BCF* formula. Its length is equal to n. So it contains n options which are single or conjunctive. The length of the formula is determined from the set D according to Proposition 1.

The question we addressed in the following sections is among all possible options in X, what are the options x_1, x_2, $\ldots$, x_n that compose ϕ such that ϕ predicts the correct satisfaction degree of each interpretation in $\mathcal{P}$.
Our aim is to learn an *LCL* preference model maximizing accuracy measure *Acc* with respect to $\mathcal{P}$. We choose the following simple measure that computes the proportion of interpretations that kept their degree of satisfaction with the learned model $\mathcal{M}_{LCL}$ that contains a formula ϕ. The degree of interpretation I in $\mathcal{P}$ is denoted by $d(I)$, and the degree that I ascribes to the learned formula ϕ is denoted by $deg(I,\phi)$.

$$Acc(\mathcal{P},\mathcal{M}_{LCL}) = \frac{|\{I \in \mathcal{P} \mid d(I) = deg(I,\phi)\}|}{|\mathcal{P}|} \quad (1)$$

4.1 Learning *LCL* formulas

The learned *LCL* model contains a formula in the form of ϕ=$(x_1\vec{\diamond}(x_2\vec{\diamond}(\ldots(x_{n-1}\vec{\diamond}x_n)))$ is a single or a conjunctive option. For generating best options of ϕ, our proposed method is inspired from Apriori algorithm (Agrawal et al., 1993) for generating frequent item-sets. Instead of generating all possible options which can be very large, we generate only frequent ones which correspond to those exceeding a minimal fixed support and confidence. Let's first introduce the following proposition before defining the support and confidence of options in *LCL*.

We first explain what we consider lexicographically preferred interpretation. We say that an interpretation I is lexicographically preferred than another interpretation I' with respect to the *LCL* formula ϕ=$(x_1\vec{\diamond}(x_2\vec{\diamond}(\ldots(x_{n-1}\vec{\diamond}x_n)))$ where $x_1 > x_2 > \ldots > x_n$ if there is $j \in \{1,\ldots,n\}$ s.t. $x_i \in I$, $x_i \in I'$ for all $i < j$ and $x_j \in I$ but $x_j \notin I'$. For example, given a formula ϕ=$x_1\vec{\diamond}x_2$, with respect to the orderings of the variables x_1, x_2 in ϕ, the lexicographically first preferred interpretation is $\{x_1,x_2\}$ and it ascribes a degree of 1 to ϕ, the second lexicographically preferred is $\{x_1\}$ and it ascribes a degree of 2 to ϕ, the third preferred is $\{x_2\}$ and it ascribes a degree of 3 to ϕ. The fourth interpretation is $\emptyset$, it is unacceptable and it ascribes a degree of ∞ to ϕ.

***Proposition* 2.** *Let $\mathcal{P}$ be a preference database, D be a set of satisfaction degrees and ϕ=$(x_1\vec{\diamond}(x_2\vec{\diamond}(\ldots(x_{n-1}\vec{\diamond}x_n)))$ be the LCL formula to be learned, it holds that 2^{n-1} lexicographically preferred interpretations are sufficient for learning the options $x_{i=1,\ldots,n}$.*

Proof. Let us consider ϕ=$(x_1\vec{\diamond}(x_2\vec{\diamond}(\ldots(x_{n-1}\vec{\diamond}x_n)))$ the *LCL* formula to be learned. Thus there are 2^n possible interpretations s.t. following the lexicographic ordering on the variables x_1, x_2, ... ,x_n, the interpretation $\emptyset$ ascribes a degree of ∞ to ϕ and the remaining $2^n - 1$ interpretations ascribe to ϕ a degree of 1 to $2^n - 1$.

- x_1 is the preferred variable of ϕ, so from the possible 2^n interpretations, there are 2^{n-1} interpretations that satisfy x_1. From these interpretations, the most preferred one is $\{x_1, x_2, \ldots, x_n\}$ and it ascribes a degree of 1 to ϕ, the second preferred interpretation is $\{x_1, x_2, \ldots, x_{n-1}\}$ and it ascribes a degree of 2 to ϕ and the last preferred interpretation is $\{x_1\}$ and it ascribes a degree of 2^{n-1} to ϕ. Thus, the best option for x_1 is the one that is satisfied by the interpretations I that ascribe a degree of $d(I) = 1, 2, \ldots, 2^{n-1}$ to ϕ.
- x_2 is the second preferred variable of ϕ. Thus, from the 2^{n-1} interpretations that satisfy x_1, we have 2^{n-2} interpretations that satisfy x_1 and x_2. From the 2^{n-2} interpretations, the last preferred interpretation that satisfies x_1 and x_2 ascribes a degree of 2^{n-2} to ϕ. Thus, the best option for x_2 is the one verifying that the interpretations I ascribing the degree of $d(I) = 1, 2, \ldots, 2^{n-2}$ to ϕ contain x_1 and x_2 (i.e., $x_1 \wedge x_2$ is satisfied by these interpretations).
- The unique interpretation that satisfies $x_1, \ldots, x_n$ is the one ascribed the degree 1 to ϕ. So, the best option for x_n is the one verifying that the interpretations I ascribing a degree of $d(I) = 1$ to ϕ contain $x_1, \ldots, x_n$ (i.e., $x_1 \wedge \ldots \wedge x_n$ is satisfied by these interpretations).

Thus, we have 2^{n-1} lexicographically preferred interpretations that are necessary for learning x_1. From the 2^{n-1} interpretations, 2^{n-2} are necessary for learning x_2 and so on. So, 2^{n-1} interpretations are necessary for learning the n options of ϕ. To generalize, the best option $x_{i=1,\ldots,n}$ is the one verifying that $x_1 \wedge \ldots \wedge x_i$ is satisfied by interpretations I ascribed $d(I) = 1, 2, \ldots, 2^{n-i}$. □

From the result given in Proposition 2, the support and confidence in *LCL* are defined as follows:

***Definition* 10** (Support)**.** *Let $\mathcal{P}$ be a preference database, I be an interpretation in $\mathcal{P}$, D be the set of satisfaction degrees, ϕ be an LCL formula to be learned s.t. $len(\phi)$=n according to Proposition 1. The support of an option $x_{i=1,\ldots,n} = x$ for interpretations $I \in \mathcal{P}$ is defined as:*

$$Supp(x_i) = \frac{\left|\left\{I \mid I \models x_1 \wedge \ldots \wedge x_i \quad \wedge \quad d(I) = 1, \ldots, 2^{n-i}\right\}\right|}{\left|\{I \mid d(I) = 1, \ldots, 2^{n-i}\}\right|}$$

***Definition* 11** (Confidence)**.** *Let $\mathcal{P}$ be a preference database, I be an interpretation in $\mathcal{P}$, D be the set of satisfaction degrees s.t. $d(I) \in D$, ϕ be an LCL formula to be learned s.t. $len(\phi)$=n according to Proposition 1. The confidence of an option $x_{i=1,\ldots,n} = x$ for interpretations $I \in \mathcal{P}$ is defined as:*

$$Conf(x_i) = \frac{|\{I \mid I \models (x_1 \wedge \ldots \wedge x_i \wedge d(I) = 1, \ldots, 2^{n-i} \vee (I \not\models x_1 \wedge \ldots \wedge x_i \wedge d(I) \neq 1, \ldots, 2^{n-i}\}|}{|\mathcal{P}|}$$

***Example* 5.** *Let us consider in Table 2, 6 possible configurations for a futur shopping center. The considered services are parking (p), shopping (s), and restaurants (r). The global evaluation of each configuration represents the satisfaction of users,* 1 *for high satisfaction,* 2 *for medium satisfaction,* 3 *for low satisfaction and* ∞ *for unacceptable configurations.*

Table 2: A simple example of user's preferences.

configurations	*p*	*s*	*r*	User's satisfaction (*D*)
I_1	0	0	1	3
I_2	0	1	0	∞
I_3	1	0	0	∞
I_4	1	0	1	3
I_5	1	1	0	2
I_6	1	1	1	1

$\mathcal{P}$=$\{I_1, \ldots, I_6\}$, D =$\{1, 2, 3, \infty\}$. We have $len(\phi) =$ 2. We aim to learn a preference model $\mathcal{M}_{LCL}$ s.t.

$$\mathcal{M}_{LCL} = \begin{cases} \phi \quad s.t. \quad \phi = x_1 \vec{\diamond} x_2, x_{i=1,\ldots,2} \in X \\ and \quad x_1 \neq x_2 \end{cases}$$

The learned model contains a BCF formula ϕ that contains 2 options, each one is a single or conjunctive option built on $\mathcal{V}$=$\{p, s, r\}$.

Let us consider the option $x_1 = p$. From data of Table 2, we have $Supp(x_1 = p)$=1 (p is satisfied by I_6 having d=1 and satisfied also by I_5 having $d = 2$). $Conf(x_1 = p)$=4/6. We have $I_3 \models p$ with $d = \infty$ and $I_4 \models p$ with $d = 3$). So, these two interpretations do not verify the definition of confidence. This means that p is not the best first option of ϕ since we aim to learn an option with maximum support and confidence (ideally 1).

The method for learning the frequent best options is defined in the following.

4.2 Generation of Options of the *LCL* Formula

For generating frequent options, we adapt the approach of association rules (Agrawal et al., 1993).

The idea is to start with all single options, count their support and find all single frequent options, combine them to form *candidate* 2-conjunctive options, go through data and count their support and find all *frequent* 2-conjunctive options, combine them to form *candidate* 3-conjunctive options and so on. More precisely, two principal steps are applied: i) *Join Step* where candidates' conjunctive options (CC_k) are generated by joining frequents ones (FC_{k-1}), ii) *Prune Step* where any (k-1)-option that is not frequent cannot be a subset of a frequent k-option. Once frequent conjunctive options are generated for each $x_{i=1,\dots,n}$, we return only those exceeding a minimal confidence θ, called final frequent options ($FinalF_{x_i}$). Algorithm 1 summarizes these steps.

Algorithm 1: Final frequent options.

```
Data: The preference database P, the set of
      satisfaction degrees D, d∈D,
      ϕ=(x1 ⃗◇(x2 ⃗◇(...(x_{n−1} ⃗◇ x_n)))) the LCL
      formula to be learned, a minimal support σ
      and a minimal confidence θ for x_{i=1...n}
FC_j: Frequent conjunctive options of size j
FC_1: Frequent single options
FC_{x_i}: Final options for x_{i=1,...,n}
Result: FinalF_{x_{i=1,...,n}}: Final frequent options for
        each x_{i=1,...,n} with support and confidence
        exceeding σ and θ
for each x_{i=1,...,n} do
    for j = 1; j ≠ ∅; j++ do
        CC_j: candidate options of size j;
        for each interpretation I having
          d(I) = 1, ..., 2^{n−i} do
            for each x ∈ CC_j do
                Compute Supp(x) according to
                 Definition 10
            end
        end
        FC_j={x ∈ CC_j | Supp(x) ≥ σ}
    end
    FC_{x_i} = ∪_j FC_j
    FinalF_{x_i} = ∅
    for each x ∈ FC_{x_i} do
        if Conf(x) ≥ θ and x is minimal
          according to Definition 12 then
            FinalF_{x_i} = FinalF_{x_i} ∪ x
        end
    end
    Return FinalF_{x_i}
end
```

Many frequent options can be learned for a given option of the *BCF* formula. To deal with this problem, we define in the following the notion of minimal frequent option.

***Definition* 12.** *Let $FC_{x_{i=1,\dots,n}}$ be a set of frequent options for x_i, and x in FC_{x_i}. Then, x is a minimal frequent option for x_i iff there is no frequent option x' in FC_{x_i} s.t. i) $Supp(x_i = x') = Supp(x_i = x)$, ii) $Conf(x_i = x') = Conf(x_i = x)$ and iii) $x' \subseteq x$.*

From Table 2, the option $x = p \wedge s$ in $FinalF_{x_1}$ is a minimal final frequent option.

As each final frequent option has its support and confidence, we propose to order them as follows.

***Definition* 13.** *Let $FinalF_{x_{i=1,\dots,n}}$ be a set of final frequent options for $x_{i=1\dots n}$. Given x, $y \in FinalF_{x_i}$. $x \succ_{x_i} y$ iff*

- *$Conf(x_i = x,) > Conf(x_i = y)$ or,*
- *$Conf(x_i = x)$=$Conf(x_i = y)$ and $Supp(x_i = x) > Supp(x_i = y)$.*

For comparing the final frequent options for a degree x_i, we consider the confidence as the most important criterion and the support the second one. This allows to guarantee selecting interesting options for each x_i.

***Example* 6** (Example 5 continued)**.** *Let us consider the data in Table 2. Table 3 gives final frequent options ($x \in X$) for the option x_1 of ϕ exceeding minimal support $\sigma = 0.9$. FC_{x_1}=$FC_1 \cup FC_2$={p, s, $p \wedge s$}. With minimal confidence θ=0.9, $FinalF_{x_1}$={$p \wedge s$}.*

Table 3: Final frequent options for x_1.

The set CC_1 for x_1	
x	$Supp(x_1 = x)$
p	1
$\neg p$	0
s	1
$\neg s$	0
r	0.5
$\neg r$	0.5

The set FC_1 for x_1	
x	$Supp(x_1 = x)$
p	1
s	1

The set CC_2 for x_1	
x	$Supp(x_1 = x)$
$p \wedge s$	1

The set FC_2 for x_1	
x	$Supp(x_1 = x)$
$p \wedge s$	1

FC_{x_1}=$FC_1 \cup FC_2$		
x	$Supp(x_1 = x)$	$Conf(x_1 = x)$
p	1	0.66
s	1	0.83
$p \wedge s$	1	1

$FinalF_{x_1}$ with $\theta = 0.9$		
x	$Supp(x_1 = x)$	$Conf(x_1 = x)$
$p \wedge s$	1	1

Table 4 gives final frequent options for the option x_2 of ϕ exceeding minimal support $\sigma = 0.9$. The set CC_1 contains all possible single options and their support for interpretations with degree $d = 1$. For example, $supp(x_2 = r)$ is computed according to Definition

10 as follows: We have $p \wedge s$ as the final frequent option for x_1. So,

$$supp(x_2 = r) = \frac{|\{I \mid I \models p \wedge s \wedge r \quad \wedge \quad d(I) = 1\}|}{|\{I \mid d(I) = 1\}|} = 1$$

The set FC_1 contains all frequent single options obtained from CC_1 exceeding minimal support $\sigma = 0.9$. CC_2 contains conjunctive options of size 2 that are composed from the set FC_1. The set FC_2 is composed from CC_2, it contains frequent conjunctive options of size 2 ($p \wedge s$ is removed from CC_2 since it is the unique final frequent option for x_1, it can not be a final frequent option for x_2). The set FC_3 is composed from CC_3, it contains frequent conjunctive options of size 3. FC_{x_2}=$FC_1 \cup FC_2 \cup FC_3$. With minimal confidence θ=0.9 and applying Definition 12 to return only minimal frequent options, $FinalF_{x_2}$=$\{r\}$.

Table 4: Final frequent options for x_2.

The set CC_1 for x_2	
x	$Supp(x_2 = x)$
p	1
$\neg p$	0
s	1
$\neg s$	0
r	1
$\neg r$	0

The set FC_1 for x_2	
x	$Supp(x_2 = x)$
p	1
s	1
r	1

The set CC_2 for x_2	
x	$Supp(x_2 = x)$
$p \wedge r$	1
$s \wedge r$	1

The set FC_2 for x_2	
x	$Supp(x_2 = x)$
$p \wedge r$	1
$s \wedge r$	1

The set CC_3 for x_2	
x	$Supp(x_2 = x)$
$p \wedge s \wedge r$	1

The set FC_3 for x_2	
x	$Supp(x_2 = x)$
$p \wedge s \wedge r$	1

FC_{x_2}=$FC_1 \cup FC_2 \cup FC_3$		
x	$Supp(x_2 = x)$	$Conf(x_2 = x)$
p	1	0.83
s	1	0.83
r	1	1
$p \wedge r$	1	1
$r \wedge s$	1	1
$p \wedge s \wedge r$	1	1

$FinalF_{x_2}$ with $\theta = 0.9$		
x	$Supp(x_2 = x)$	$Conf(x_2 = x)$
r	1	1

Given $FinalF_{x_{i=1,\dots,n}}$ be a set of final frequent options for each $x_{i=1,\dots,n}$ ordered following Definition 13, the *LCL* learned preference model is:

$$\mathcal{M}_{LCL} = \begin{cases} \phi = (x_1 \vec{\diamond} (x_2 \vec{\diamond} (\dots (x_{n-1} \vec{\diamond} x_n))) \\ \quad s.t. \quad x_i \in FinalF_{x_{i=1,\dots,n}} \\ \quad and \quad x_1 \neq x_2 \neq \dots \neq x_n. \end{cases}$$

Thus, the best *LCL* learned model contains a formula which is composed of preferred final frequent options for each option. The accuracy of each model is computed by applying Equation 1. The best preference model is the one that contains formula with the greatest accuracy (ideally 1). To define the satisfaction degree of new interpretations by the learned *LCL* preference model that contains a formula ϕ, we apply Definition 7.

Example 7. *Let us continue Example 6. The LCL preference model learned from data of Table 2 is*

$$\mathcal{M}_{LCL} = \{ \ \phi = (p \wedge s) \vec{\diamond} r$$

The accuracy of $\mathcal{M}_{LCL}$ is 1. Let us consider the following two new interpretations I=$\{$p=0, s=0, r=0$\}$ and I'=$\{$p=0, s=1, r=1$\}$, then deg(I,ϕ)=∞ and deg(I',ϕ)=3.

5 EXPERIMENTAL RESULTS

To further test our method, we provide a case study in the context of antibiotics prescription. The database used here contains a list of antibiotics, each one is described with some features (here we use 7 binary features), and a rank of recommendation as defined in Clinical Practice Guidelines (CPGs). It should be noted that the complete dataset was tested using Algorithm 1 implementation in Python3.

Antibiotics with recommendation rank 1 are recommended in first intention, those having a recommendation rank 2 are recommended in second intention and those having recommendation rank 3 are recommended in third intention. Antibiotics having a recommendation rank 0 are not recommended. So, antibiotics having rank 1 are preferred to those having rank 2 which are also preferred to those having rank 3. Antibiotics having rank 0 are unacceptable since they can not be prescribed for the patient. Thus, in the context of *LCL*, we have *D*=$\{1, 2, 3, \infty\}$.

Our aim is to determine what are the features so that an antibiotic has a given recommendation rank. Thus, we apply our method for learning an *LCL* preference model from the antibiotic database. The antibiotic's features are: Convenient protocol (*proto*), Non precious class (*Precious*), serious side effects (*SideEff*), High level of efficacy (*Efficacy*), Narrow antibacterial spectrum (*Spect*), ecological adverse effects (*RiskResi*) and Taste (*Taste*). The learned *LCL* preference model for pharyngitis is given in Table 5.

Table 5: $\mathcal{M}_{LCL}$ in pharyngitis clinical situation.

$\mathcal{M}_{LCL}$	Accuracy
$(Proto \wedge \neg SideEff) \vec{\diamond} (Proto \wedge \neg RiskResi)$	0.91

Let us explain results of Table 5. $\mathcal{M}_{LCL}$ contains a *BCF* formula that has 2 options. The accuracy of the model is higher which means that *LCL* is fully adapted for modeling experts' reasoning for providing recommendations of antibiotics. The model does not predict correctly some antibiotics. The reason is that it is not possible to obtain a lexicographic ordering over the two options of the *BCF* formula. In fact, the learned model should verify the following conditions: i) only antibiotics with rank 1 satisfy the two options of the formula, ii) only antibiotics with rank 2 satisfy the first option and not satisfy the second option, and iii) only antibiotics with rank 3 satisfy the second option and falsify the first option. There is no model that verifies these conditions in the antibiotic database for the considered clinical situation. The model given here does not verify the condition iii). This is due certainly to some inconsistencies in the database (Tsopra et al., 2018).

6 CONCLUSION

We proposed a method for learning preferences in the context of a logic-based preference formalism, *LCL*. The method is based on an adaptation of association rules based on the Apriori algorithm. The *LCL* learned model is qualitative and easily interpretable for the user. To the best of our knowledge, this is the first proposition for learning preferences in the context of *LCL*.

The choice of train data plays an important role of the learned *LCL* model. It can be different following the considered train data. For example, if we consider a preference database with $D=\{1, 2, 3, \infty\}$, then the learned model will contain an *LCL* formula with 2 options. However, if we consider a preference database with $D=\{1, 2, 3, 4, \infty\}$, then the learned model will contain an *LCL* formula with 3 options. The formula to be learned from $D=\{3, 4, \infty\}$ will be certainly different from the one learned from $D=\{1, 2, 3, 4, \infty\}$. The problem of learning *LCL* preferences is considered as an instance ranking problem where the set of satisfaction degrees corresponds to the set of labels and the set of outcomes corresponds to the set of interpretations. In future work, we would perform some evaluations to compare our method with other preference learning methods, particularly those of instance ranking problem.

REFERENCES

Agrawal, R., Imieliński, T., and Swami, A. (1993). Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, SIGMOD '93, pages 207–216, New York, NY, USA. ACM.

Benferhat, S. and Sedki, K. (2008). Two alternatives for handling preferences in qualitative choice logic. *Fuzzy Sets and Systems*, 159(15):1889–1912.

Bernreiter, M., Maly, J., and Woltran, S. (2022). Choice logics and their computational properties. *Artif. Intell.*, 311:103755.

Boudjelida, A. and Benferhat, S. (2016). Conjunctive choice logic. In *International Symposium on Artificial Intelligence and Mathematics, ISAIM 2016, Fort Lauderdale, Florida, USA, January 4-6, 2016*.

Brewka, G., Benferhat, S., and Berre, D. L. (2004). Qualitative choice logic. *Artif. Intell.*, 157(1-2):203–237.

Cohen, W. W., Schapire, R. E., and Singer, Y. (2011). Learning to order things. *CoRR*, abs/1105.5464.

Hüllermeier, E., Fürnkranz, J., Cheng, W., and Brinker, K. (2008). Label ranking by learning pairwise preferences. *Artif. Intell.*, 172(16-17):1897–1916.

Joachims, T., Granka, L., Pan, B., Hembrooke, H., and Gay, G. (2005). Accurately interpreting clickthrough data as implicit feedback. In *SIGIR'05: Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 154–161, New York, NY, USA. ACM Press.

Johannes, F. and Hüllermeier, E. (2010). *Preference Learning*. Springer-Verlag New York, Inc., New York, NY, USA, 1st edition.

Sedki, K., Lamy, J., and Tsopra, R. (2022). Qualitative choice logic for modeling experts recommendations of antibiotics. In *Proceedings of the Thirty-Fifth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2022*.

Tsopra, R., Lamy, J., and Sedki, K. (2018). Using preference learning for detecting inconsistencies in clinical practice guidelines: Methods and application to antibiotherapy. *Artificial Intelligence in Medicine*, 89:24–33.

Vembu, S. and Gärtner, T. (2010). Label ranking algorithms: A survey. In *Preference Learning.*, pages 45–64.

Waegeman, W. and De Baets, B. (2010). A survey on roc-based ordinal regression learning. In Fürnkranz, J. and Hüllermeier, E., editors, *Preference learning*, pages 127–154. Springer.