



HAL
open science

A Knowledge Compilation Map for Conditional Preference Statements-based Languages

Hélène Fargier, Jérôme Mengin

► **To cite this version:**

Hélène Fargier, Jérôme Mengin. A Knowledge Compilation Map for Conditional Preference Statements-based Languages. [Research Report] IRIT/RR-2021-02-FR, Institut de Recherche en Informatique de Toulouse (CNRS : UMR5505; INPT de Toulouse; Universités de Toulouse I, II et III). 2021. hal-03133187

HAL Id: hal-03133187

<https://hal.science/hal-03133187>

Submitted on 8 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Knowledge Compilation Map for Conditional Preference Statements-based Languages

Research Report IRIT/RR-2021-02-FR
Institut de Recherche en Informatique de Toulouse

Hélène Fargier
IRIT-CNRS, Université de Toulouse
Toulouse, France
helene.fargier@irit.fr

Jérôme Mengin
IRIT-CNRS, Université de Toulouse
Toulouse, France
jerome.mengin@irit.fr

ABSTRACT

Conditional preference statements have been used to compactly represent preferences over combinatorial domains. They are at the core of CP-nets and their generalizations, and lexicographic preference trees. Several works have addressed the complexity of some queries (optimization, dominance in particular). We extend in this paper some of these results, and study other queries which have not been addressed so far, like equivalence, thereby contributing to a knowledge compilation map for languages based on conditional preference statements. We also introduce a new parameterised family of languages, which enables to balance expressiveness against the complexity of some queries.

1 INTRODUCTION

Preference handling is a key component in several areas of Artificial Intelligence, notably for decision-aid systems. Research in Artificial Intelligence has led to the development of several languages that enable compact representation of preferences over complex, combinatorial domains. Some preference models rank alternatives according to their values given by some multivariate function; this is the case for instance with valued constraints [30], additive utilities and their generalizations [9, 25]. Ordinal models like CP nets and their generalisations [4, 8, 33], or lexicographic preferences and their generalisations [3, 10, 18, 22, 31, 34] use sets of conditional preference statements to represent a pre-order over the set of alternatives.

Many problems of interest, like comparing alternatives or finding optimal alternatives, are NP-hard for many of these models, even PSPACE hard for some models, which makes these representations difficult to use in some decision-aid systems like configurators, where real-time interaction with a decision maker is needed. One approach to tackle this problem is Knowledge Compilation, whereby a model, or a part of it, is *compiled*, off-line, into another representation which enables fast query answering, even if the compiled representation has a much bigger size. This approach has first been studied in propositional logic: [13, 14] compare how various subsets of propositional logic can succinctly, or not, express some propositional knowledge bases, and the complexity of queries of interest. [12] follow a similar approach to compare extensions of propositional logic which associate real values to models of a knowledge base; [19] provide such a map for value function-based models.

The aim of this paper is to initiate such a compilation map for models of preferences based on the language of conditional preference statements. We compare the expressiveness and succinctness of various languages on these conditional preference statements, and the complexity of several queries of interest for these languages.

The next section recalls some basic definitions about combinatorial domains and pre-orders, and introduces notations that will be used throughout. Section 3 gives an overview of various languages based on conditional preference statements that have been studied in the literature. We also introduce a new parameterised family of languages, which enables to balance expressiveness against the complexity of some queries. Section 4 and 5 respectively study expressiveness and succinctness for languages based on conditional preference statements. Sections 6 study the complexity of queries for these languages. Proofs can be found in [20].

2 PRELIMINARIES

2.1 Combinatorial Domain

We consider languages that can be used to represent the preferences of a decision maker over a combinatorial space \mathcal{X} ; here \mathcal{X} is a set of attributes that characterise the possible alternatives, each attribute $X \in \mathcal{X}$ having a finite set of possible values \underline{X} ; we assume that $|\underline{X}| \geq 2$ for every $X \in \mathcal{X}$; then $\underline{\mathcal{X}}$ denotes the cartesian product of the domains of the attributes in \mathcal{X} , its elements are called alternatives. For a binary attribute X , we will often denote by x, \bar{x} its two possible values. In the sequel, n is the number of attributes in \mathcal{X} .

For a subset U of \mathcal{X} , we will denote by \underline{U} the cartesian product of the domains of the attributes in U , called instantiations of U , or partial instantiations (of \mathcal{X}). If v is an instantiation of some $V \subseteq \mathcal{X}$, $v[U]$ denotes the restriction of v to the attributes in $V \cap U$; we say that instantiation $u \in \underline{U}$ and v are compatible if $v[U \cap V] = u[U \cap V]$; if $U \subseteq V$ and $v[U] = u$, we say that v extends u .

Sets of partial instantiations can often be conveniently, and compactly, specified with propositional formulas: the atoms are $X = x$ for every $X \in \mathcal{X}$ and $x \in \underline{X}$, and we use the standard connectives \wedge (conjunction), \vee (disjunction), \rightarrow (implication), \leftrightarrow (equivalence) and \neg (negation); we denote by \top (resp. \perp) the formula always true (resp. false). Implicitly, this propositional logic is equipped with a theory that enforces that every attribute has precisely one value from its domain; so, for two distinct values x, x' of attribute X , the

formula $X = x \wedge X = x'$ is a contradiction; also, the interpretations are thus in one-to-one correspondence with \underline{X} . If α is such a propositional formula over \mathcal{X} and $o \in \underline{X}$, we will write $o \models \alpha$ when o satisfies α , that is when, assigning to every literal $X = x$ that appears in α the value true if $o[X] = x$, and the value false otherwise, makes α true.

Given a formula α , or a partial instantiation u , $\text{Var}(\alpha)$ and $\text{Var}(u)$ denote the set of attributes, the values of which appear in α and u respectively.

When it is not ambiguous, we will use x as a shorthand for the literal $X = x$; also, for a conjunction of such literals, we will omit the \wedge symbol, thus $X = x \wedge Y = \bar{y}$ for instance will be denoted $x\bar{y}$.

2.2 Preference Relations

Depending on the knowledge that we have about a decision maker's preferences, given any pair of distinct alternatives $o, o' \in \underline{X}$, one of the following situations must hold: one may be strictly preferred over the other, or o and o' may be equally preferred, or o and o' may be incomparable.

Assuming that preferences are transitive, such a state of knowledge about the DM's preferences can be characterised by a preorder \geq over \underline{X} : \geq is a binary, reflexive and transitive relation; for alternatives o, o' , we then write $o \geq o'$ when $(o, o') \in \geq$; $o > o'$ when $(o, o') \in \geq$ and $(o', o) \notin \geq$; $o \sim o'$ when $(o, o') \in \geq$ and $(o', o) \in \geq$; $o \bowtie o'$ when $(o, o') \notin \geq$ and $(o', o) \notin \geq$. Note that for any pair of alternatives $o, o' \in \underline{X}$ either $o > o'$, or $o' > o$, or $o \sim o'$ or $o \bowtie o'$.

The relation \sim defined in this way is the *symmetric part* of \geq , it is reflexive and transitive, \bowtie is irreflexive, they are both symmetric. The relation $>$ is the *irreflexive part* of \geq , it is what is usually called a strict partial order: it is irreflexive and transitive.

Terminology and notations. We say that alternative o *dominates* alternative o' (w.r.t. \geq) if and only if $o \geq o'$; if $o > o'$, then we say that o *strictly dominates* o' . We use standard notations for the complements of $>$ and \geq : we write $o \not\geq o'$ when it is not the case that $o \geq o'$, and $o \not> o'$ when it is not the case that $o > o'$.

3 LANGUAGES

3.1 Conditional Preference Statements

A *conditional preference statement* (aka., CP statement) over \mathcal{X} is an expression of the form $\alpha \mid V : w \geq w'$, where α is a propositional formula over $U \subseteq \mathcal{X}$, $w, w' \in \underline{W}$ are such that $w[X] \neq w'[X]$ for every $X \in W$, and U, V, W are disjoint subsets of \mathcal{X} , not necessarily forming a partition of \mathcal{X} . Informally, such a statement represents the piece of knowledge that, when comparing alternatives o, o' that both satisfy α , the one that has values w for W is preferred to the one that has values w' for W , irrespective of the values of the attributes in V , every attribute in $\mathcal{X} \setminus (V \cup W)$ being fixed. We call α the conditioning part of the statement; we call W the swapped attributes, and V the free part.

Example 1 ((Example A in [35], slightly extended)). Consider planning a holiday, with three choices / attributes: wait til next month ($W = w$) or leave now ($W = \bar{w}$), going to city 1, 2 or 3 ($C = c_1, C = c_2$ or $C = c_3$), travelling by plane ($P = p$) or by car ($P = \bar{p}$). I would rather go now, irrespective of the other attributes: $\top \mid \{CP\} : \bar{w} \geq w$.

All else being equal, I prefer to go to city 3, city 1 being my second best choice: $\top \mid \emptyset : c_3 \geq c_1 \geq c_2$. Also, if I go now, I prefer to fly: $\bar{w} \mid \emptyset : p \geq \bar{p}$. Together, the last two statements imply that if I go now, I prefer to go to city 3 by plane than go to city 1 by car; however these statements do not say what I prefer between flying to city 1 or driving to city 3. In fact, I prefer the former, this *tradeoff* can be expressed with the statement $\bar{w} \mid \emptyset : c_1 p \geq c_3 \bar{p}$. Finally, if I go later, I prefer to drive, irrespective of the city: $w \mid \{C\} : \bar{p} \geq p$.

Conditional preference statements have been studied in many works, under various language restrictions. They are the basis for CP-nets [4, 6] and their extensions, and have been studied in a more logic-based fashion by e.g. [24] and [32, 33, 35].¹ They are closely related to *CI-statements* by [7]

For the semantics sets of CP statements, we use the definitions of [35]. Given a statement $\alpha \mid V : w \geq w'$, let $U = \text{Var}(\alpha)$ and $W = \text{Var}(w) = \text{Var}(w')$: a *worsening swap* is any pair of alternatives (o, o') such that $o[U] = o'[U] \models \alpha$, $o[W] = w$ and $o'[W] = w'$, and such that for every attribute $Y \notin U \cup V \cup W$ it holds that $o[Y] = o'[Y]$; we say that $\alpha \mid V : w \geq w'$ *sanctions* (o, o') . For a set of CP-statements φ , let φ^* be the set of all worsening swaps sanctioned by statements of φ , and define \geq_φ to be the reflexive and transitive closure of φ^* . [35] proves that $o \geq_\varphi o'$ if and only if $o = o'$ or φ^* contains a finite sequence of worsening swaps $(o_i, o_{i+1})_{0 \leq i \leq k-1}$ with $o_0 = o$ and $o_k = o'$.²

Example 2 (Example 1, continued). Let $\varphi = \{\top \mid \{CP\} : \bar{w} \geq w, \top \mid \emptyset : c_3 \geq c_1 \geq c_2, n \mid \emptyset : p \geq \bar{p}, \bar{w} \mid \emptyset : c_1 p \geq c_3 \bar{p}, w \mid \{C\} : \bar{p} \geq p\}$. Then $\top \mid \{CP\} : \bar{w} \geq w$ sanctions for instance $(\bar{w}c_2p, wc_3\bar{p})$, so $\bar{w}c_2p \geq_\varphi wc_3\bar{p}$. Also, $\top \mid \emptyset : c_3 \geq c_1 \geq c_2$ sanctions $(\bar{w}c_1p, \bar{w}c_2p)$, $\bar{w} \mid \emptyset : p \geq \bar{p}$ sanctions $(\bar{w}c_2p, \bar{w}c_2\bar{p})$, so, by transitivity, $\bar{w}c_1p \geq_\varphi \bar{w}c_2\bar{p}$. It is not difficult to check that $\bar{w}c_2p \bowtie_\varphi \bar{w}c_1\bar{p}$.

Let us call CP the language where formulas are sets of statements of the general form $\alpha \mid V : w \geq w'$. This language is very expressive: it is possible to represent any preorder “in extension” with preference statements of the form $o \geq o'$ – they all have $W = \mathcal{X}$ as set of swapped attributes, $\alpha = \top$ as condition, and no free attribute.

This expressiveness has a cost: we will see that many queries about pre-orders represented by CP-statements are PSPACE-hard for the language CP. Several restrictions / sub-languages have been studied in the literature, we review them below.

Linearisability. Although the original definition of CP-nets by [6] does not impose it, many works on CP-nets, especially following [4], consider that they are intended to represent a strict partial order, that is, that \geq_φ should be antisymmetric; equivalently, this means that the irreflexive part $>_\varphi$ of \geq_φ can be extended to a linear order. We say that a set φ of CP-statements is *linearisable* in this case.³

¹The formula $u \mid V : x \geq x'$ is written $u : x > x' \mid V$ by [35].

²Actually, [35] proves that (o, o') is in the transitive closure of φ^* if and only there is such a worsening sequence from o to o' , but adding the reflexive closure to this transitive closure does not change the result, since we can add any pair (o, o) to, or remove it from, any sequence of worsening swaps without changing the validity of the sequence.

³Such sets of CP-statements are often called *consistent* in the standard terminology on CP-nets, but we prefer to depart from this definition which only makes sense when one asserts that φ should indeed represent a strict partial order.

Notations. We write $\alpha : w \geq w'$ when V is empty, and $w \geq w'$ when V is empty and $\alpha = \top$. Note that we reserve the symbol \geq for conditional preference statements, whereas “curly” symbols $>$, \neq , \geq , $\not\geq$ are used to represent relations over the set of alternatives.

In the remainder of this section, we present various sublanguages of CP. Some are defined by imposing various simple syntactical restrictions on the formulas, two are languages which have been well studied (CP-nets and lexicographic preference trees); we close the section by introducing a new, parameterised class of sublanguages of CP which have interesting properties, as will be shown in subsequent subsections.

3.2 Statement-wise Restrictions

Some restrictions are on the syntactical form of statements allowed; they bear on the size of the set of free attributes, or on the size of the set of swapped attributes, or on the type of conditioning formulas allowed. Given some language $\mathcal{L} \subseteq \text{CP}$, we define the following restrictions:

- $\mathcal{L}\not\geq$ = only formulas with empty free parts ($V = \emptyset$) for every statement;⁴
- $\mathcal{L}\wedge$ = only formulas where the condition α of every statement is a conjunction of literals;
- $k\text{-}\mathcal{L}$ = only formulas where the set of swapped attributes contains no more than k attributes ($|W| \leq k$) for every statement; in particular, we call elements of 1-CP *unary* statements.

In particular, 1-CP \wedge is the language studied by [35], and 1-CP $\not\geq$ is the language of generalized CP-nets as defined by [24].

3.3 Graphical Restrictions

Given $\varphi \in \text{CP}$ over set of attributes X , we define D_φ as the graph with sets of vertices X , and such that there is an edge (X, Y) if there is $\alpha | V : w \geq w' \in \varphi$ such that $X \in \text{Var}(\alpha)$ and $Y \in \text{Var}(w)$, or $X \in \text{Var}(w)$ and $Y \in V$. We call D_φ the *dependency graph* of φ . Note that D_φ can be computed in polynomial time. This definition, inspired by [35, Def. 15], generalises that of [4], which is restricted to the case where all CP statements are unary and have no free attributes, and that of [8], who study statements with free attributes. Many tractability results on sets of CP statements have been obtained when D_φ has good properties. Given some language $\mathcal{L} \subseteq \text{CP}$, we define:

- $\mathcal{L}\not\geq$ = the restriction of \mathcal{L} to *acyclic* formulas, which are those φ such that D_φ is acyclic;⁵
- $\mathcal{L}\not\geq^{\text{poly}}$ = the restriction of \mathcal{L} to formulas where the dependency graph is a polytree.

[35] also defines a weaker graphical restriction, called “context-uniform conditional acyclicity”, but it turns out that it does give rise to the same complexities as another, weaker restriction called “conditional acyclicity” by [35], which we generalize in section 3.6.

⁴In the literature, the symbol \triangleright is sometimes used to represent an *importance* relation between attributes; and, as explained by [35], statement $\alpha | V : w \geq w'$ is a way to express that attributes in $\text{Var}(w)$ are more important than those in V (when α is true).

⁵This is *full acyclicity* in [35].

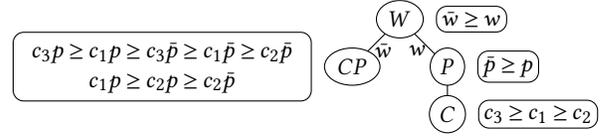


Figure 1: An LP-tree equivalent to the set of CP-statements of Example 2.

3.4 CP-nets

In their seminal work, [4] define a CP-net over a set of attributes X to be composed of two elements:

- (1) a directed graph over X , which should represent *preferential dependencies* between attributes;⁶
- (2) a set of conditional preference tables, one for every attribute X : if U is the set of parents of X in the graph, the conditional preference table for X contains exactly $|\underline{U}|$ rules $u : \geq$, for every $u \in \underline{U}$, where the \geq 's are linear orders over \underline{X} .

Therefore, as shown by [35], CP-nets can be seen as sets of unary CP statements in conjunctive form with no free attribute. Specifically, given a CP-net \mathcal{N} over X , define $\varphi_{\mathcal{N}}$ to be the set of all CP statements $u : x \geq x'$, for every attribute X , every $u \in \underline{U}$ where U is the set of parents of X in the graph, every $x, x' \in \underline{X}$ such that x, x' are consecutive values in the linear order \geq specified by the rule $u : \geq$ of \mathcal{N} . Then the dependency graph of $\varphi_{\mathcal{N}}$, as defined in Section 3.3, coincides with the graph of \mathcal{N} . We call

CPnet = the language that contains all $\varphi_{\mathcal{N}}$, for every CP-net \mathcal{N} .

Note that $\text{CPnet} \subseteq 1\text{-CP}\wedge\not\geq$. For a given $\varphi \in 1\text{-CP}\wedge\not\geq$, being a CP-net necessitates a very strong form of local consistency and completeness: for every attribute X with parents U in D_φ , for every $u \in \underline{U}$, for every $x, x' \in \underline{X}$, φ must explicitly, and uniquely, order ux and ux' .

[8] define TCP-nets as an extension of CP-nets where it is possible to represent tradeoffs, by stating that, under some conditions, some attributes are more important than other ones. [35] describes how TCP-nets can be transformed, in polynomial time, into equivalent sets of 1-CP \wedge statements.

3.5 Lexicographic Preference Trees

LP-trees generalise lexicographic orders, which have been widely studied in decision making – see e.g. [21]. As an inference mechanism, they are equivalent to search trees used by [5], and formalised by [32, 35]. As a preference representation, and elicitation, language, slightly different definitions for LP-trees have been proposed by [3, 10, 18]. We use here a definition which subsumes the others.

An LP-tree that is equivalent to the set of CP-statements of Example 2 is depicted on Figure 1. More generally, an LP-tree over X is a rooted tree with labelled nodes and edges, and a set of preference tables; specifically

- every node N is labelled with a set of attributes, denoted $\text{Var}(N)$;

⁶Given some pre-order \geq over X , attribute X is said to be preferentially dependent on attribute Y if there exist $x, x' \in \underline{X}$, $y, y' \in \underline{Y}$, $z \in \underline{X} \setminus (\{X, Y\})$ such that $xyz \succeq_\varphi x'yz$ but $xy'z \not\succeq_\varphi x'y'z$.

- if N is not a leaf, it can have one child, or $|\text{Var}(N)|$ children;
- in the latter case, the edges that connect N to its children are labelled with the instantiations in $\text{Var}(N)$;
- if N has one child only, the edge that connects N to its child is not labelled: all instantiations in $\text{Var}(N)$ lead to the same subtree;
- we denote by $\text{Anc}(N)$ the set of attributes that appear in the nodes between the root and N (excluding those at N), and by $\text{Inst}(N)$ (resp. $\text{NonInst}(N)$) the set of attributes that appear in the nodes above N that have more than one children (resp. only one child);
- a conditional preference table $\text{CPT}(N)$ is associated with N : it contains local preference rules of the form $\alpha : \geq$, where \geq is a preorder over $\text{Var}(N)$, and α is a propositional formula over some attributes in $\text{NonInst}(N)$.

We assume that the rules in $\text{CPT}(N)$ define their preorder over $\text{Var}(N)$ in extension. Additionally, two constraints guarantee that an LP-tree φ defines a unique preorder over \underline{X} :

- no attribute can appear at more than one node on any branch of φ ; and,
- at every node N of φ , for every $u \in \text{NonInst}(N)$, $\text{CPT}(N)$ must contain exactly one rule $\alpha : \geq$ such that $u \models \alpha$.

Given an LP-tree φ and an alternative $o \in \underline{X}$, there is a unique way to traverse the tree, starting at the root, and along edges that are either not labelled, or labelled with instantiations that agree with o , until a leaf is reached. Now, given two distinct alternatives o, o' , it is possible to traverse the tree along the same edges as long as o and o' agree, until a node N is reached which is labelled with some W such that $o[W] \neq o'[W]$: we say that N decides $\{o, o'\}$.

In order to define \geq_φ for some LP-tree φ , let φ^* be the set of all pairs of distinct alternatives (o, o') such that there is a node N that decides $\{o, o'\}$ and the only rule $\alpha : \geq \in \text{CPT}(N)$ with $o[\text{NonInst}(N)] = o'[\text{NonInst}(N)] \models \alpha$ is such that $o[W] \geq o'[W]$. Then \geq_φ is the reflexive closure of φ^* .

PROPOSITION 1. *Let φ be an LP-tree over \underline{X} , then \geq_φ as defined above is a preorder. Furthermore, \geq_φ is a linear order if and only if 1) every attribute appears on every branch and 2) every preference rule specifies a linear order.*

An LP-tree φ is said to be *complete* if the two conditions in Proposition 1 hold, that is, if \geq_φ is a linear order.

From a semantic point of view, an LP-tree φ is equivalent to the set that contains, for every node N of φ labelled with $W = \text{Var}(N)$, and every rule $\alpha : \geq_N^\alpha$ in $\text{CPT}(N)$, all CP statements of the form $\alpha \wedge u \mid V : w^\# \geq w'^\#$, where

- u is the combination of values given to the attributes in $\text{Inst}(N)$ along the edges between the root and N , and
- $w, w' \in \underline{W}$ such that $w \geq_N^\alpha w'$, and $W^\#$ is the set of attributes on which w and w' have distinct values, and $w^\# = w[W^\#]$, and $w'^\# = w'[W^\#]$; and
- $V = [\underline{X} - (\text{Anc}(N) \cup W)]$.

This set of statements indicate that outcomes that agree on $\text{Anc}(N)$ and satisfy $u \wedge \alpha$, but have different values for $\text{Var}(N)$, should be ordered according to \geq_N^α , whatever their values for attributes in V .

LPT = the language of LP-trees as defined above; we consider that LPT is a subset of CP.⁷

Note that, using the notations defined above, $k\text{-LPT} = \text{LPT} \cap k\text{-CP}$ is the restriction of LPT where every node has at most k attributes, for every $k \in \mathbb{N}$; in particular, 1-LPT is the language of LP-trees with one attribute at each node; and $\text{LPT} \wedge = \text{LPT} \cap \text{CP} \wedge$ is the restriction of LPT where the condition α in every rule at every node is a conjunction of literals. Search trees of [32, 35] and LP-trees as defined by [3, 27] are sublanguages of $1\text{-LPT} \wedge$; LP-trees of [18] and [10] are sublanguages of $\text{LPT} \wedge$.

3.6 Lexico-compatible Formulas

Many graphical restrictions that have been proposed in order to enable polytime answers to some queries are in fact particular cases of a more general property which we study now. We define a new, parameterised family of languages. Given some language $\mathcal{L} \subseteq \text{CP}$ and $k \in \mathbb{N}$, we define:

$\mathcal{L} \mathcal{D}_k^{\text{lex}}$ = the restriction of \mathcal{L} to formulas φ such that there exists some complete LP-tree $\psi \in k\text{-LPT}$ such that \geq_ψ extends \geq_φ .

We say that formulas of $\text{CP} \mathcal{D}_k^{\text{lex}}$ are *k-lexico-compatible*.⁸

[35] proves that acyclic formulas of 1-CP are 1-lexico-compatible when they enjoy some local consistency property; it illustrates that k -lexico-compatibility is indeed a weak form of acyclicity. We will see that k -lexico-compatibility makes some queries tractable.

The next result shows that proving that some $\varphi \in \text{CP}$ is k -lexico-compatible, for a fixed k , is not always easy:

PROPOSITION 2. *For a fixed $k \in \mathbb{N}$, checking if a formula $\varphi \in \text{CP}$ is k -lexico-compatible is coNP-complete.*

Algorithm 1 checks if a given formula is k -lexico-compatible. Given $\varphi \in \text{CP}$, it builds, in a top-down fashion, a complete $\psi \in k\text{-LPT}$ that is compatible with φ . The algorithm is similar to the algorithm proposed by [3] to learn an LP-tree that sanctions a given set of pairs (o, o') . It starts with an empty root node at step 1; then, while there is some empty node, it picks one of them, call it N , and calls at step 2b the function `chooseAttribute` to get a pair (T, \geq) to label N , where T is a set of at most k attributes, none of which appear above N , and \geq is a linear order over \underline{T} ; if no such pair is compatible with φ , in a sense that will be defined shortly, `chooseAttribute` returns failure and the algorithm stops at step 2c; otherwise, if there remain some attributes that do not appear in T nor at any node above N , then the algorithm expands the tree below N at step 2e by creating a branch and a new node for every instantiation $t \in \underline{T}$, and loops.

Note that all edges of the tree built by the algorithm are labelled, so that, at every node N , $\text{NonInst}(N) = \emptyset$, so $\text{CPT}(N)$ must contain only one rule of the form $\top : \geq$, where \top is the formula always true. This is why `chooseAttribute` needs to return one linear order over \underline{T} only, we do not need to specify the trivial condition \top here. There may be a more compact k -LP-tree compatible with φ than the one returned by the above algorithm when it does not fail, but we are only interested here in checking if φ is k -lexico-compatible, and we

⁷Strictly speaking, for $\text{LPT} \subseteq \text{CP}$ to hold, we can add the possibility to augment every formula in CP with a tree structure.

⁸This definition generalises *conditionally acyclic* formulas of [35], which are the formulas of $\text{CP} \mathcal{D}_1^{\text{lex}}$.

Algorithm 1: Build complete LP tree

Input: $\varphi \in \text{CP}$; $k \in \mathbb{N}$;
Output: $\psi \in k\text{-LPT}$, ψ complete, s.t. $\geq_\psi \supseteq \geq_\varphi$, or FAILURE;
(1) $\psi \leftarrow \{\text{an unlabelled root node}\}$;
(2) while ψ contains some unlabelled node:
(a) choose unlabelled node N of ψ ;
(b) $(T, \geq) \leftarrow \text{chooseAttribute}(N, k, \varphi)$;
(c) if $T = \text{FAILURE}$ then STOP and return FAILURE;
(d) label N with (T, \geq) ;
(e) if $\text{Anc}(N) \cup T \neq \mathcal{X}$, for each $t \in \underline{T}$: add new unlabelled node to ψ , attached to N with edge labelled with t ;
(3) return ψ .

have seen that the problem is coNP-complete, so it seems difficult to avoid exploring a tree with size exponential in the size of φ in the worst case. We now specify some condition that chooseAttribute must verify in order for the algorithm to be correct and complete. Given any yet unlabelled node N of the tree being build, let $\varphi(N) = \{\alpha \mid V : w \geq w' \in \varphi \mid \alpha \wedge \text{inst}(N) \not\equiv \perp, W \cap \text{Anc}(N) = \emptyset\}$.

Definition 1. We say that chooseAttribute is φ -compatible if the pair (T, \geq) that chooseAttribute returns at some yet unlabelled node N is such that for every $\alpha \mid V : w \geq w' \in \varphi(N)$: (1) if $\text{Var}(w) \cap T = \emptyset$, then $V \cap T = \emptyset$; (2) if $\text{Var}(w) \cap T \neq \emptyset$, then $t >^N t'$ for every $t, t' \in \underline{T}$ such that $t \wedge w \not\equiv \perp, t' \wedge w' \not\equiv \perp, t[\mathcal{X} \setminus (V \cup W)] = t'[\mathcal{X} \setminus (V \cup W)]$ and $t \wedge \alpha \not\equiv \perp$. If no such pair (T, \geq) can be found, then chooseAttribute must return failure.

Condition (2) guarantees that N will correctly decide every pair of alternatives that is sanctioned by $\alpha \mid V : w \geq w'$ and that will be decided at N . When the entire tree is built in this way, condition ?? guarantees that at every node N , if $\alpha \mid V : w \geq w' \in \varphi(N)$ then $V \cap \text{Anc}(N) = \emptyset$.

PROPOSITION 3. Given $\varphi \in \text{CP}$ and some $k \in \mathbb{N}$, suppose that chooseAttribute is φ -compatible, then $\varphi \in \text{CP}^{\not\equiv}_k^{\text{lex}}$ if and only if the algorithm above returns some $\psi \in k\text{-LPT}$ such that $\geq_\psi \supseteq \geq_\varphi$; otherwise, it returns FAILURE.

Note that chooseAttribute can be implemented to run in polynomial time, for fixed k : there are no more than $\sum_{i=1}^k \binom{n}{i} \leq kn^k$ possibilities for the T it can return, and the number of pairs t, t' that it must check against every statement in $\varphi(N)$ is bounded by $|\underline{T}|^2$, and $|\underline{T}|$ is bounded by d^k , where d is the size of the largest domain of the attributes in \mathcal{X} . Also, each branch of the tree returned by the algorithm, when it succeeds, can have at most n nodes, but the tree can have up to d^n leaves.

4 EXPRESSIVENESS

We detail our results about expressiveness of the various languages studied here in this section, the results about succinctness are in the next section. These results are summarised on Figure 2.

Definition 2. Let \mathcal{L} and \mathcal{L}' be two languages for representing preorders. We say that \mathcal{L} is at least as expressive as \mathcal{L}' , written $\mathcal{L} \sqsupseteq \mathcal{L}'$, if every preorder that can be represented with a formula of \mathcal{L}' can also be represented with a formula of \mathcal{L} ; we write $\mathcal{L} \sqsubset \mathcal{L}'$ if

$\mathcal{L} \sqsupseteq \mathcal{L}'$ but it is not the case that $\mathcal{L}' \sqsupseteq \mathcal{L}$, and say in this case that \mathcal{L} is strictly more expressive than \mathcal{L}' . We write $\mathcal{L} \sqsubseteq \mathcal{L}'$ when the two languages are equally expressive.

We reserve the usual ‘‘rounded’’ symbols \subset and \subseteq for (strict) set inclusion, and \supset and \supseteq for the reverse inclusions. Note that \sqsupseteq is a preorder, and obviously $\mathcal{L} \sqsupseteq \mathcal{L}'$ implies $\mathcal{L} \sqsubseteq \mathcal{L}'$.

Clearly, $\text{CP}^{\not\equiv} \subset \text{CP}$ and $\text{CP}^\wedge \subset \text{CP}$; however, these three languages have the same expressiveness, because of the following:

PROPERTY 4. Given some preorder \geq , define $\varphi \in \{o[\Delta(o, o')] \geq o'[\Delta(o, o')] \mid o \geq o', o \neq o'\}$, where $\Delta(o, o')$ is the set of attributes that have different values in o and o' , then $\varphi \in \text{CP}^{\not\equiv} \cap \text{CP}^\wedge$, and $\geq_\varphi = \geq$.

A large body of works on CP-statements since the seminal paper by [5] concentrate on various subsets of 1-CP. With this strong restriction on the number of swapped attributes, CP-statements have a reduced expressiveness.

Example 3. Consider two binary attributes A and B , with respective domains $\{a, \bar{a}\}$ and $\{b, \bar{b}\}$. Define preorder \geq such that $ab > \bar{a}\bar{b} > \bar{a}b > a\bar{b}$. This can be represented in CP with $\varphi = \{ab \geq \bar{a}\bar{b}, \bar{a}\bar{b} \geq \bar{a}b, \bar{a}b \geq a\bar{b}\}$. But it cannot be represented in 1-CP: $\{b : a \geq \bar{a}, \bar{b} : \bar{a} \geq a, a : b \geq \bar{b}, \bar{a} : \bar{b} \geq b\}^* \subseteq \varphi^*$, but this is not sufficient to compare $\bar{a}b$ with $\bar{a}\bar{b}$. The four remaining formulas of 1-CP over these two attributes are $B : a \geq \bar{a}, B : \bar{a} \geq a, A : b \geq \bar{b}, A : \bar{b} \geq b$, adding any of them to φ yields a preorder which would not be antisymmetric.

Forbidding free parts incurs an additional loss in expressiveness:

Example 4. Consider two binary attributes A and B , with respective domains $\{a, \bar{a}\}$ and $\{b, \bar{b}\}$. Define preorder \geq such that $ab > \bar{a}b > \bar{a}\bar{b}$. This can be represented in 1-CP with $\varphi = \{B : a \geq \bar{a}, b \geq \bar{b}\}$. But the ‘‘tradeoff’’ $\bar{a}b > \bar{a}\bar{b}$ cannot be represented in 1-CP $^{\not\equiv}$, any formula of 1-CP $^{\not\equiv}$ that implies it will put some *intermediate* alternative between $\bar{a}b$ and $\bar{a}\bar{b}$.

However, restricting to conjunctive statements does not incur a loss in expressiveness.

PROPOSITION 5. $\text{CP} = \bigcup_{k \in \mathbb{N}} k\text{-CP}$ and, for every $k \in \mathbb{N}$:

$$\text{CP}^\wedge \sqsubseteq \text{CP}^{\not\equiv} \sqsubseteq \text{CP} \sqsupset k\text{-CP} \sqsubseteq k\text{-CP}^\wedge \sqsupset k\text{-CP}^{\not\equiv} \sqsubseteq k\text{-CP}^\wedge \not\equiv k\text{-CP} \sqsupset (k-1)\text{-CP}.$$

Because an LP-tree can be a single node labelled with \mathcal{X} , and a single preference rule $\top : \geq$ where \geq can be any preorder, LPT can represent any preorder. Limiting to conjunctive conditions in the rules is not restrictive. However, restricting to 1-LPT reduces expressiveness, even if one considers formulas of 1-CP that represent total, linear orders:

Example 5. Let $\varphi = \{a \geq \bar{a}, \bar{c} \mid A : \bar{b} \geq b, \bar{a}c : \bar{b} \geq b, ac : b \geq \bar{b}, a : c \geq \bar{c}, \bar{a} \mid B : \bar{c} \geq c\}$. This yields the following linear order: $abc \geq_\varphi \bar{a}bc \geq_\varphi \bar{a}\bar{b}c \geq_\varphi \bar{a}b\bar{c} \geq_\varphi \bar{a}bc \geq_\varphi abc$. No $\psi \in 1\text{-LPT}$ can represent it: A could not be at the root of such a tree because for instance $\bar{a}\bar{b}c \geq_\varphi \bar{a}b\bar{c}$ and $\bar{a}b\bar{c} \geq_\varphi \bar{a}bc$; neither could C , since $\bar{a}bc \geq_\varphi \bar{a}\bar{b}c$ and $\bar{a}b\bar{c} \geq_\varphi \bar{a}bc$; and finally B could not be at the root either, because $abc \geq_\varphi \bar{a}bc$ and $\bar{a}b\bar{c} \geq_\varphi abc$.

PROPOSITION 6. $\text{LPT} = \bigcup_{k \in \mathbb{N}} k\text{-LPT}$ and, for every $k \in \mathbb{N}$:

$$\text{CP} \sqsubseteq \text{LPT} \sqsubseteq \text{LPT}^\wedge \sqsupset k\text{-LPT} \sqsubseteq k\text{-LPT}^\wedge \sqsupset (k-1)\text{-LPT}.$$

Finally, note that k -lexico-compatibility is a weaker restriction than being a k -LP-tree.

PROPOSITION 7. For every $k \in \mathbb{N}$: $CP\mathcal{D}_k^{\text{lex}} \sqsupseteq CP\mathcal{D}_{k-1}^{\text{lex}}$, and $CP\mathcal{D}_k^{\text{lex}} \sqsupseteq k\text{-LPT}$.

[35] proves that $1\text{-CP}\mathcal{D} \subseteq CP\mathcal{D}_1^{\text{lex}}$. Whether this property can be generalised, with an appropriate definition of k -acyclicity, is left for future work.

5 SUCCINCTNESS

Another criterion is the relative sizes of formulas that can represent the same preorder in different languages. [11] study the space efficiency of various propositional knowledge representation formalisms. An often used definition of succinctness [14, 23] makes it a particular case of expressiveness, which is not a problem when comparing languages of same expressiveness. However, we study here languages with very different expressiveness, so we need a more fine grained definition:

Definition 3. Let \mathcal{L} and \mathcal{L}' be two languages for representing preorders. We say that \mathcal{L} is *at least as succinct as* \mathcal{L}' , written $\mathcal{L} \leq \mathcal{L}'$, if there exists a polynomial p such that for every $\varphi' \in \mathcal{L}'$, there exists $\varphi \in \mathcal{L}$ that represent the same preorder as φ' and such that $|\varphi| < p(|\varphi'|)$.⁹ Moreover, we say that \mathcal{L} is *strictly more succinct than* \mathcal{L}' , written $\mathcal{L} \ll \mathcal{L}'$, if $\mathcal{L} \leq \mathcal{L}'$ and for every polynomial p , there exists $\varphi \in \mathcal{L}$ such that:

- there exists $\varphi' \in \mathcal{L}'$ such that $\geq_{\varphi} = \geq_{\varphi'}$, but
- for every $\varphi' \in \mathcal{L}'$ such that $\geq_{\varphi} = \geq_{\varphi'}$, $|\varphi'| > p(|\varphi|)$.

With this definition, $\mathcal{L} \ll \mathcal{L}'$ if every formula of \mathcal{L}' has an equivalent formula in \mathcal{L} which is “no bigger” (up to some polynomial transformation of the size of φ), and there is at least one sequence of formulas (one formula for every polynomial p) in \mathcal{L} that have equivalent formulas in \mathcal{L}' but necessarily “exponentially bigger”.¹⁰

PROPOSITION 8. The following hold, for languages \mathcal{L} , \mathcal{L}' , \mathcal{L}'' :

- if $\mathcal{L} \supseteq \mathcal{L}'$ then $\mathcal{L} \leq \mathcal{L}'$; and if $\mathcal{L} \leq \mathcal{L}'$, then $\mathcal{L} \supseteq \mathcal{L}'$;
- if $\mathcal{L} \ll \mathcal{L}'$ then $\mathcal{L} \leq \mathcal{L}'$ and $\mathcal{L}' \not\leq \mathcal{L}$;
- if $\mathcal{L} \sqsupseteq \mathcal{L}'$, the reverse implication holds:
if $\mathcal{L} \leq \mathcal{L}'$ and $\mathcal{L}' \not\leq \mathcal{L}$ then $\mathcal{L} \ll \mathcal{L}'$
(otherwise, it might be that $\mathcal{L}' \not\leq \mathcal{L}$ because $\mathcal{L}' \not\supseteq \mathcal{L}$);
- if $\mathcal{L} \supseteq \mathcal{L}'$ and $\mathcal{L}' \ll \mathcal{L}''$, then $\mathcal{L} \ll \mathcal{L}''$.

Restricting the conditioning part of the statements to be conjunctions of literals does induce a loss in succinctness.

Example 6. Consider $2n+1$ binary attributes $X_1, X_2, \dots, X_n, Y_1, Y_2, \dots, Y_n, Z$, and let φ contain $2n+2$ unary CP-statements with no free attribute: $(x_1 \vee y_1) \wedge (x_2 \vee y_2) \wedge \dots \wedge (x_n \vee y_n) : z \geq \bar{z}$, $\neg[(x_1 \vee y_1) \wedge (x_2 \vee y_2) \wedge \dots \wedge (x_n \vee y_n)] : \bar{z} \geq z$ and $\bar{x}_i \geq x_i$ and $\bar{y}_i \geq y_i$ for every $i \in \{1, \dots, n\}$. Then $\varphi \in 1\text{-CP}\mathcal{D}$, but φ is not in conjunctive form. A set of conjunctive CP-statements equivalent to φ has to contain all 2^n statements of the form $\mu_1 \mu_2 \dots \mu_n : z \geq \bar{z}$ with $\mu_i = x_i$ or $\mu_i = y_i$ for every i .

⁹Where $|\varphi| = \sum_{\alpha \vee \beta : w \geq w' \in \varphi} (|\alpha| + |\beta| + 2|\text{Var}(w)|)$, with $|\alpha|$ = the number of connectives plus the number of atoms of α .

¹⁰When \ll is defined as the strict counterpart of \leq , it can happen that $\mathcal{L} \ll \mathcal{L}'$ even if there is no real difference in representation size in the two languages, but $\mathcal{L} \sqsupseteq \mathcal{L}'$.

Also, free attributes enable succinct representation of relative importance of some attributes over others; disabling free attributes thus incurs a loss in succinctness:

Example 7. Consider $n+1$ binary attributes X_1, X_2, \dots, X_n, Y , let $U = \{X_1, X_2, \dots, X_n\}$, and let $\varphi = \{U \mid y \geq \bar{y}\}$. Then $\varphi^* = \{(uy, u'\bar{y}) \mid u, u' \in \underline{U}\}$, and φ^* is equal to its transitive closure, so, if $o \neq o'$, then $o \geq_{\varphi} o'$ if and only if $o[Y] = y$ and $o'[Y] = \bar{y}$. This can be represented, without free attribute, with formula ψ that contains, for every $V \subseteq U$ and every $v \in \underline{V}$, the statement $vy \geq \bar{v}\bar{y}$, where \bar{v} denotes the tuple obtained by inverting all values of v . For every $0 \leq i \leq n$ there are $\binom{n}{i}$ subsets of V of size i , with 2^i ways to choose $v \in \underline{V}$, thus ψ contains $\sum_0^n \binom{n}{i} 2^i = 3^n$ statements.

Restricting to CP-nets induces a further loss in succinctness, as the next example shows:

Example 8. Consider $n+1$ binary attributes X_1, X_2, \dots, X_n, Y , and let φ be the $1\text{-CP}\mathcal{D} \wedge$ formula that contains the following statements: $x_i \geq \bar{x}_i$ for $i = 1, \dots, n$; $x_1 x_2 \dots x_n : y \geq \bar{y}$; $\bar{x}_i : \bar{y} \geq y$ for $i = 1, \dots, n$. The size of φ is linear in n . Because preferences for Y depend on all X_i 's, a CP-net equivalent to φ will contain, in the table for Y , 2^n CP statements.

PROPOSITION 9. The following hold:

- $\mathcal{L} \ll \mathcal{L} \wedge$ for every \mathcal{L} such that $1\text{-CP}\mathcal{D} \subseteq \mathcal{L} \subseteq \text{CP}$;
- $\mathcal{L} \ll \mathcal{L} \mathcal{D}$ for every \mathcal{L} such that $1\text{-CP} \wedge \subseteq \mathcal{L} \subseteq \text{CP}$;
- $1\text{-CP}\mathcal{D} \wedge \ll \text{CPnet}$.

6 QUERIES

Table 1 gives an overview of the tractability of the queries that we study in this section. We begin this section with the two queries that have generated most interest in the literature on CP statements.

Linearisability. Knowing that a given $\varphi \in \text{CP}$ is linearisable (that is, that \geq_{φ} is antisymmetric) is valuable, as it makes several other queries easier. It also gives some interesting insights into the semantics of φ . The following query has been addressed in many works on CP statements:¹¹

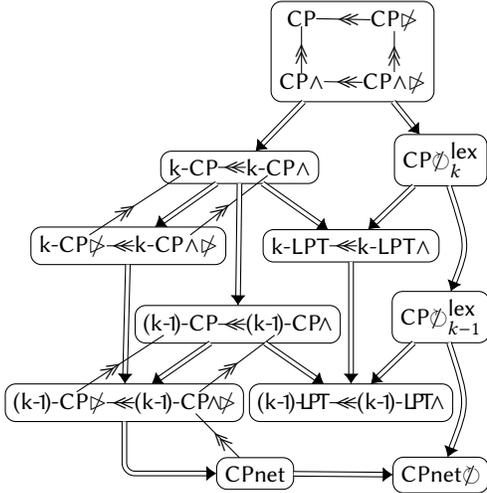
LINEARISABILITY Given φ , is φ linearisable?

[4] prove that when its dependency graph D_{φ} is acyclic, then a CP-net φ is linearisable. This result has been extended by [8, 15, 35], who give weaker, sufficient syntactical conditions that guarantee that a locally consistent set of unary, conjunctive CP statements is linearisable; more generally, by definition of k -lexico-compatibility, every formula of $CP\mathcal{D}_k^{\text{lex}}$ is linearisable (since it is compatible with a complete LP-tree). [24, Theorem 3 and 4] prove that LINEARISABILITY is PSPACE-complete for $1\text{-CP}\mathcal{D} \wedge$.

PROPOSITION 10. LINEARISABILITY can be checked in polynomial time for LPT.

Comparing alternatives. A basic question, given a formula φ and two alternatives o, o' is: how do o and o' compare, according to φ ? Is it the case that $o >_{\varphi} o'$, or $o' >_{\varphi} o$, or $o \bowtie_{\varphi} o'$, or $o \sim_{\varphi} o'$? We define the following query, for any relation $R \in \{>, \geq, \sim, \bowtie\}$:

¹¹This query is often called *consistency*.



$\mathcal{L} \Longrightarrow \mathcal{L}'$: \mathcal{L} is strictly more expressive than \mathcal{L}'
 $\mathcal{L} \Leftarrow \mathcal{L}'$: \mathcal{L} is strictly more succinct than \mathcal{L}'
For $k > 2$. Boxes contain languages that are equally expressive.

Figure 2: Rel. expressiveness and succinctness

R-COMPARISON Given formula φ , alternatives $o \neq o'$, is it the case that $oR_{\varphi}o'$?

For LP-trees, in order to compare alternatives o and o' , one only has to traverse the tree from the root downwards until a node that decides the pair is reached, or up to a leaf if no such node is encountered: in this case o and o' are incomparable. Note that checking if a node decides the pair, and checking if a rule at that nodes applies to order them, can both be done in polynomial time.

PROPOSITION 11. *R-COMPARISON is in P for LPT for $R \in \{>, \geq, \sim, \succ\}$.*

Tractability of comparisons, except in some trivial cases, comes at a heavy price in terms of expressiveness: \geq -COMPARISON is tractable for CP-nets when the dependency graph is a polytree [4, Theorem 14], but [4, Theorems 15, 16] prove that \geq -COMPARISON is already NP-hard for the quite restrictive language of binary-valued, directed-path singly connected CP-nets, which are acyclic. [24, Prop. 7, Corrolary 1] prove that \geq -COMPARISON, $>$ -COMPARISON, \succ -COMPARISON and \sim -COMPARISON are PSPACE complete for 1-CP^{\wedge} and for linearisable, locally complete formulas of 1-CP^{\wedge} . More precise hardness results for acyclic CP-nets are also proved by [28]. Proposition 12 completes the picture.

PROPOSITION 12. *$>$ -COMPARISON and \succ -COMPARISON are NP-hard for the language of fully acyclic CP-nets, and tractable for polytree CP-nets. \sim -COMPARISON is easy for 1-CP^{\wedge} .*

Comparing theories. Checking if two theories yield the same preorder can be useful during the compilation process. We say that two formulas φ and φ' are equivalent if they represent the same preorder, that is, if \geq_{φ} and $\geq_{\varphi'}$ are identical; we then write $\varphi \equiv \varphi'$.

	CP	1-CP^{\wedge}	$1\text{-CP}^{\wedge\wedge}$	CPnet	$\text{CP}^{\wedge\wedge}_{\text{lex}}$	$\text{CP}^{\wedge\wedge}$	$\text{CP}^{\wedge\wedge}_{\text{poly}}$	LPT	LTP w. can. tables
LINEARISABILITY	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$		T	T	T	\checkmark	\checkmark
R-COMPARISON, $R \in \{\geq, >, \succ\}$	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\circ$	$\times\circ$	\times	\checkmark	\checkmark	\checkmark
\sim -COMPARISON	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$		\perp	\perp	\perp	\checkmark	\checkmark
EQUIVALENCE	$\times\mathbf{x}$	$\times\circ$	$\times\circ$	\checkmark	$\times\circ$	\checkmark	\checkmark	$\times\circ$	\checkmark
TOP- p					\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
W. UNDOMINATED \exists	T	T	T	T	T	T	T	T	T
UNDOMINATED \exists	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$		T	T	T	T	T
S. DOM. \exists , DOM. \exists	$\times\mathbf{x}$	$\times\mathbf{x}$				T	T	\checkmark	\checkmark
UNDOM. CHECK, \geq -CUT EXTRACT.	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
S. DOM., DOM., W. UNDOM. CHECK	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$			\checkmark	\checkmark	\checkmark	\checkmark
$>$ -CUT EXTRACTION	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$		\checkmark	\checkmark	\checkmark	\checkmark	\checkmark
$>$ -CUT COUNTING	$\times\mathbf{x}$	$\times\mathbf{x}$	$\times\mathbf{x}$					\checkmark	\checkmark

Each column corresponds to one sub-language of CP. They are sorted in order of decreasing expressiveness from left to right, except when columns are separated by double lines. For each query and sub-language: T = always true for the language; \perp = always false for the language; \checkmark = polytime answer; \times = NP-complete query; $\times\circ$ = NP/coNP-hard query; $\times\mathbf{x}$ = PSPACE-complete query.

Table 1: Complexity of queries.

EQUIVALENCE Given two formulas φ and φ' , are they equivalent?

Consider a formula $\varphi \in \text{CP}$, two alternatives o, o' , and let $\varphi' = \varphi \cup \{o \geq o'\}$; clearly $o \geq_{\varphi'} o'$, thus $\varphi \equiv \varphi'$ if and only if $o \geq_{\varphi} o'$. Therefore, if language \mathcal{L} is such that adding CP statement $o \geq o'$ to any of its formulas yields a formula that is still in \mathcal{L} , then EQUIVALENCE has to be at least as hard as \geq -COMPARISON for \mathcal{L} . This is the case of CP. The problem remains hard for 1-CP^{\wedge} , because it is hard to check the equivalence, in propositional logic, of the conditions of statements that entail a particular swap $x \geq x'$.

Example 9. Consider three attributes A, B and C with respective domains $\{a, \bar{a}\}, \{b, \bar{b}\}$ and $\{c_1, c_2, c_3\}$. Consider two CP statements $s = \bar{a} : c_1 \geq c_2$ and $s' = b : c_2 \geq c_3$, and let $\varphi = \{s, s', a : c_1 \geq c_3\}$. Because of statements s and s' we have $\bar{a}bc_1 \geq_{\varphi} \bar{a}bc_2 \geq_{\varphi} \bar{a}bc_3$; also, $abc_1 \geq_{\varphi} abc_3$ because of statement $a : c_1 \geq c_3$. Hence, for any $u \in \underline{A} \times \underline{B}$, if $u \models a \vee (\bar{a}b)$ then $uc_1 \geq uc_3$. Thus $\varphi \equiv \{s, s'\} \cup \{a \vee (\bar{a}b) : c_1 \geq c_3\} \equiv \varphi \cup \{b : c_1 \geq c_3\}$.

PROPOSITION 13. *EQUIVALENCE is coNP-hard for $1\text{-CP}^{\wedge\wedge}$, and for 1-LPT^{\wedge} , both restricted to binary attributes.*

As usual, comparing two formulas is easier for languages where there exists a canonical form. This is the case of CP-nets, as shown by [26, Lemma 2]; their proof makes it clear that the canonical form of any CP-net φ can be computed in polynomial time. Hence:

PROPOSITION 14. *EQUIVALENCE is in P for CP-net.*

EQUIVALENCE also becomes tractable if some form of canonicity is imposed for the conditions of the rules in an LP-tree; this is because, as with CP-net, it is possible to define a canonical form for the structure, by imposing that the labels of the node be as small as possible – which may lead, in some cases, to splitting some nodes.

Top p alternatives. Given a set of alternatives S and some integer p , we may be interested in finding a subset S' of S that contains p "best" alternatives of S , in the sense that for every $o \in S'$, for every $o' \in S \setminus S'$ it is not the case that $o' \succ_{\varphi} o$. Note that such a set must exist, because \succ_{φ} is acyclic. The TOP- p query is usually defined for totally ordered sets; a definition suited to partial orders is given in [35] (where it is called *ordering*), we adopt this definition here:

TOP- p Given $S \subseteq \underline{X}$, $p < |S|$, and φ , find $o_1, o_2, \dots, o_p \in S$ such that for every $i \in 1, \dots, p$, for every $o' \in S$, if $o' \succ_{\varphi} o_i$ then $o' \in \{o_1, \dots, o_{i-1}\}$.

Note that if o_1, o_2, \dots, o_p is the answer to such query, if $1 \leq i < j \leq p$, then it can be the case that $o_i \preceq o_j$, but it is guaranteed that $o_j \not\preceq o_i$: in the context of a recommender system for instance, where one would expect alternatives to be presented in order of non-increasing preference, o_i could be safely presented before o_j .

[4] prove that TOP- p is tractable for acyclic CP-nets for the specific case where $|S| = 2$. More generally, \succ -COMPARISON queries can be used to compute an answer to a TOP- p query (by asking \succ -COMPARISON queries for every pair of elements of S , the number of such pairs being in $\Theta(|S|^2)$). However, [35] shows that an upper approximation of \succ is sufficient, and proves that such an approximation can be obtained in time polynomial in $|\varphi|$ for a restricted class of lexico-compatible formulas of 1-CP \wedge [35, Th. 5]. We prove that this result does indeed hold for the full class of lexico-compatible formulas of 1-CP \wedge . The TOP- p query is also tractable for LPT.

PROPOSITION 15. TOP- p can be answered in time which is polynomial in the size of φ and the size of S for k -lexico-compatible formulas (for fixed k); and for LPT.

Optimization. Instead of ordering a given set, we may want to find a globally optimal alternative. Following [24], given φ , we say that alternative o is:

- weakly undominated if there is no $o' \in \underline{X}$ such that $o' \succ_{\varphi} o$;
- undominated if there is no $o' \in \underline{X}$, $o' \neq o$, such that $o' \succeq_{\varphi} o$;
- dominating if for every $o' \in \underline{X}$, $o \succeq_{\varphi} o'$;
- strongly dominating if for every $o' \in \underline{X}$ with $o' \neq o$, $o \succ_{\varphi} o'$.

Note that o is strongly dominating if and only if it is dominating and undominated; and that if o is dominating or undominated, then it is weakly undominated. This gives rise to several types of queries:

[w | s] (UNDOMINATED | DOMINATING) EXISTENCE Given φ , is there a [weakly | strongly] (undominated | dominating) alternative?

[w | s] (UNDOMINATED | DOMINATING) CHECKING Given φ , o , is o a [weakly | strongly] (undominated | dominating) alternative?

All these queries are easily shown to be tractable for LPT. The problem UNDOMINATED CHECK has been shown to be tractable for CP-nets [4] and for 1-CP $\not\preceq$ [24]. This can be generalized:

PROPOSITION 16. UNDOMINATED CHECK is in P for CP.

The existence of a weakly undominated alternative is trivially true for CP (in any finite directed graph, at least one vertex has no "strict" predecessor). Linearisability also ensure that there is at least one undominated alternative.

For CP-nets, [4] give a polytime algorithm that computes the only dominating alternative when the dependency graph is acyclic; in this case, this alternative is also the only strongly dominating one and the only undominated one, since the CP-net is linearisable: this implies that DOMINATING \exists , s. DOMINATING \exists , UNDOMINATED \exists , s. DOMINATING CHECK, DOMINATING CHECK and w. UNDOMINATED CHECK are tractable for acyclic CP-nets.

[24, Prop. 8, 9 and 11] prove that w. UNDOMINATED CHECK, DOMINATING CHECK, s. DOMINATING CHECK, DOMINATING \exists and s. DOMINATING \exists are PSPACE-complete for 1-CP $\not\preceq$, and their reductions for proving hardness of w. UNDOMINATED CHECK, DOMINATING CHECK, s. DOMINATING CHECK indeed yield formulas of 1-CP $\not\preceq \wedge$. NP-hardness of UNDOMINATED \exists for 1-CP $\not\preceq \wedge$ is proved by [16],

Cuts. Cuts are sets of alternatives that are at the same "level" with respect to \succeq . For rankings defined with real-valued functions, cuts are defined with respect to possible real values. In the case of pre-orders, we define cuts with respect to some alternative o : given $\varphi \in \text{CP}$, for any $R \in \{\succ, \succeq\}$, for every alternative o , we define

- $\text{CUT}^{R,o}(\varphi) = \{o' \in \underline{X} \mid o' \neq o, o' R_{\varphi} o\}$.

Following [19], we define two families of queries:

R-CUT COUNTING Given φ, o , count the elements of $\text{CUT}^{R,o}(\varphi)$

R-CUT EXTRACTION Given φ, o , return an element of $\text{CUT}^{R,o}(\varphi)$ (or that it is empty)

PROPOSITION 17. \succeq -CUT EXTRACTION is tractable for CP. \succ -CUT COUNTING and \succ -CUT EXTRACTION are PSPACE-hard for 1-CP $\not\preceq \wedge$. For CP $\not\preceq_k^{\text{lex}}$, \succ -CUT EXTRACTION is equivalent to \succeq -CUT EXTRACTION and is tractable. \succ -CUT COUNTING is tractable for LP-trees.

7 CONCLUSION

The literature on languages on CP statements has long focused on statements with unary swaps. Several examples in Section 4 show that this strongly degrades expressiveness. We have introduced a new parameterised family of languages, CP $\not\preceq_k^{\text{lex}}$, which permits to balance expressiveness against query complexity: the lower k is, the less expressive the language is, but the faster answering most queries will be. Table 1 shows that comparison queries seem to resist tractability, even for CP $\not\preceq_k^{\text{lex}}$, but queries like the TOP- p query may be sufficient in many applications. Tractability of the EQUIVALENCE query relies on the existence of canonical form: it is the case when the language enforces a structure like a dependency graph or a tree, and when the conditions of the statements are restricted to some propositional language with a canonical form.

We have not studied here transformations, like conditioning or other forms of projection for instance. Some initial results on projections can be found in [1]. This is an important direction for future work, as well as properties of the various languages studied here with respect to machine learning.

ACKNOWLEDGMENTS

We thank anonymous referees for their valuable comments on previous versions of this paper. This work has benefited from the AI Interdisciplinary Institute ANITI. ANITI is funded by the French "Investing for the Future – PIA3" program under grant agreement ANR-19-PI3A-0004. This work has also been supported by the PING/ACK

project of the French National Agency for Research, grant agreement ANR-18-CE40-0011.

REFERENCES

- [1] Philippe Besnard, Jérôme Lang, and Pierre Marquis. 2005. Variable forgetting in preference relations over combinatorial domains. In *Proceedings of the IJCAI Multidisciplinary Workshop on Advances in Preference Handling (MPREF'05)*.
- [2] Richard Booth, Yann Chevaleyre, Jérôme Lang, Jérôme Mengin, and Chatrarakul Sombaththeera. 2009. *Learning various classes of models of lexicographic orderings*. Rapport de recherche IRIT/RR-2009-21-FR. IRIT, Université Paul Sabatier, Toulouse. http://www.irit.fr/publis/ADRIA/Boothetal_irit09.pdf
- [3] Richard Booth, Yann Chevaleyre, Jérôme Lang, Jérôme Mengin, and Chatrarakul Sombaththeera. 2010. Learning conditionally lexicographic preference relations. In *Proceedings of the 19th European Conference on Artificial Intelligence (ECAI 2010) (Frontiers in Artificial Intelligence and Applications, Vol. 215)*, Helder Coelho, Rudi Studer, and Michael Wooldridge (Eds.). IOS Press, 269–274.
- [4] Craig Boutilier, Ronen I. Brafman, Carmel Domshlak, Holger H. Hoos, and David Poole. 2004. CP-nets: a tool for representing and reasoning with conditional ceteris paribus preference statements. *Journal of Artificial Intelligence Research* 21 (2004), 135–191.
- [5] Craig Boutilier, Ronen I. Brafman, Carmel Domshlak, Holger H. Hoos, and David Poole. 2004. Preference-Based Constrained Optimization with CP-Nets. *Computational Intelligence* 20, 2 (2004), 137–157.
- [6] Craig Boutilier, Ronen I. Brafman, Holger H. Hoos, and David Poole. 1999. Reasoning With Conditional Ceteris Paribus Preference Statements. In *Proceedings of the 15th Annual Conference on Uncertainty in Artificial Intelligence (UAI-99)*, Kathryn B. Laskey and Henri Prade (Eds.). Morgan Kaufmann, 71–80. https://dslpitt.org/uai/displayArticleDetails.jsp?mmnu=1&smnu=2&article_id=155&proceeding_id=21
- [7] Sylvain Bouveret, Ulle Endriss, and Jérôme Lang. 2009. Conditional Importance Networks: A Graphical Language for Representing Ordinal, Monotonic Preferences over Sets of Goods. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI'09)*, Craig Boutilier (Ed.). 67–72. <http://ijcai.org/Proceedings/09/Papers/022.pdf>
- [8] Ronen I. Brafman, Carmel Domshlak, and Solomon E. Shimony. 2006. On graphical modeling of preference and importance. *Journal of Artificial Intelligence Research* 25 (2006), 389–424.
- [9] Darius Brazuinas and Craig Boutilier. 2005. Local Utility Elicitation in GAI Models. In *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence (UAI'05)*, Fahiem Bacchus and Tommi Jaakkola (Eds.). AUAI Press, 42–49.
- [10] Michael Bräuning and Eyke Hüllermeier. 2012. Learning Conditional Lexicographic Preference Trees. In *Preference Learning: Problems and Applications in AI. Proceedings of the ECAI 2012 workshop*, Johannes Fürnkranz and Eyke Hüllermeier (Eds.). 11–15.
- [11] Marco Cadoli, Francesco M. Donini, Paolo Liberatore, and Marco Schaerf. 2000. Space Efficiency of Propositional Knowledge Representation Formalisms. *Journal of Artificial Intelligence Research* 13 (2000), 1–31. <https://doi.org/10.1613/jair.664>
- [12] Sylvie Coste-Marquis, Jérôme Lang, Paolo Liberatore, and Pierre Marquis. 2004. Expressive Power and Succinctness of Propositional Languages for Preference Representation, See [17], 203–212. <http://www.aaai.org/Library/KR/2004/kr04-023.php>
- [13] Adnan Darwiche. 1999. Compiling Knowledge into Decomposable Negation Normal Form. In *Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence (IJCAI 99)*, Thomas Dean (Ed.). Morgan Kaufmann, 284–289. <http://ijcai.org/Proceedings/99-1/Papers/042.pdf>
- [14] Adnan Darwiche and Pierre Marquis. 2002. A Knowledge Compilation Map. *Journal of Artificial Intelligence Research* 17 (2002), 229–264. <https://doi.org/10.1613/jair.989>
- [15] Carmel Domshlak and Ronen I. Brafman. 2002. CP-nets: Reasoning and Consistency Testing. In *Proceedings of the Eighth International Conference on Principles of Knowledge Representation and Reasoning (KR-02)*, Dieter Fensel, Fausto Giunchiglia, Deborah L. McGuinness, and Mary-Anne Williams (Eds.). Morgan Kaufmann, 121–132.
- [16] Carmel Domshlak, Steven David Prestwich, Francesca Rossi, Kristen Brent Venable, and Toby Walsh. 2006. Hard and soft constraints for reasoning about qualitative conditional preferences. *J. Heuristics* 12, 4-5 (2006), 263–285.
- [17] Didier Dubois, Christopher A. Welty, and Mary-Anne Williams (Eds.). 2004. *Proceedings of the Ninth International Conference on the Principles of Knowledge Representation and Reasoning*. AAAI Press.
- [18] Hélène Fargier, Pierre Francois Gimenez, and Jérôme Mengin. 2018. Learning Lexicographic Preference Trees from Positive Examples. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence (AAAI 2018)*, Sheila A. McIlraith and Kilian Q. Weinberger (Eds.). AAAI Press, 2959–2966. <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/17272/16610>
- [19] Hélène Fargier, Pierre Marquis, Alexandre Niveau, and Nicolas Schmidt. 2014. A Knowledge Compilation Map for Ordered Real-Valued Decision Diagrams. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, July 27 -31, 2014, Québec City, Québec, Canada*, Carla E. Brodley and Peter Stone (Eds.). AAAI Press, 1049–1055. <http://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8195>
- [20] Hélène Fargier and Jérôme Mengin. 2021. *A Knowledge Compilation Map for Conditional Preference Statements-based Languages*. Research Report IRIT/RR-2021-02-FR. IRIT - Institut de recherche en informatique de Toulouse.
- [21] Peter C. Fishburn. 1974. Lexicographic Orders, Utilities and Decision Rules: A Survey. *Management Science* 20, 11 (1974), pp. 1442–1471. <http://www.jstor.org/stable/2629975>
- [22] Gerd Gigerenzer and Daniel G. Goldstein. 1996. Reasoning the Fast and Frugal Way: Models of Bounded Rationality. *Psychological Review* 103, 4 (1996), 650–669.
- [23] Goran Gogic, Henry A. Kautz, Christos H. Papadimitriou, and Bart Selman. 1995. The Comparative Linguistics of Knowledge Representation, See [29], 862–869. <http://ijcai.org/Proceedings/95-1/Papers/111.pdf>
- [24] Judy Goldsmith, Jérôme Lang, Mirosław Trzuszczynski, and Nic Wilson. 2008. The Computational Complexity of Dominance and Consistency in CP-nets. *Journal of Artificial Intelligence Research* 33 (2008), 403–432.
- [25] Christophe Gonzales and Patrice Perny. 2004. GAI Networks for Utility Elicitation, See [17], 224–233.
- [26] Frédéric Koriche and Bruno Zanuttini. 2010. Learning conditional preference networks. *Artificial Intelligence* 174, 11 (2010), 685–703. <https://doi.org/10.1016/j.artint.2010.04.019>
- [27] Jérôme Lang, Jérôme Mengin, and Lirong Xia. 2018. Voting on Multi-Issue Domains with Conditionally Lexicographic Preferences. *Artificial Intelligence* 265 (2018), 18–44. <https://doi.org/10.1016/j.artint.2018.05.004>
- [28] Thomas Lukasiewicz and Enrico Malizia. 2019. Complexity results for preference aggregation over (m)CP-nets: Pareto and majority voting. *Artificial Intelligence* 272 (2019), 101–142. <https://doi.org/10.1016/j.artint.2018.12.010>
- [29] Chris S. Mellish (Ed.). 1995. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI 95)*. Morgan Kaufmann. <http://ijcai.org/proceedings/1995-1>
- [30] Thomas Schiex, Hélène Fargier, and Gérard Verfaillie. 1995. Valued Constraint Satisfaction Problems: Hard and Easy Problems, See [29], 631–639. <http://ijcai.org/Proceedings/95-1/Papers/083.pdf>
- [31] Michael Schmitt and Laura Martignou. 2006. On the Complexity of Learning Lexicographic Strategies. *Journal of Machine Learning Research* 7 (2006), 55–83.
- [32] Nic Wilson. 2004. Consistency and Constrained Optimisation for Conditional Preferences. In *Proceedings of the 16th European Conference on Artificial Intelligence (ECAI 2004)*, Ramón López de Mántaras and Lorenza Saitta (Eds.). IOS Press, 888–892.
- [33] Nic Wilson. 2004. Extending CP-Nets with Stronger Conditional Preference Statements. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI'04)*, Deborah L. McGuinness and George Ferguson (Eds.). AAAI Press / The MIT Press, 735–741.
- [34] Nic Wilson. 2006. An Efficient Upper Approximation for Conditional Preference. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI 2006) (Frontiers in Artificial Intelligence and Applications)*, Gerhard Brewka, Silvia Coradeschi, Anna Perini, and Paolo Traverso (Eds.). IOS Press.
- [35] Nic Wilson. 2011. Computational techniques for a simple theory of conditional preferences. *Artificial Intelligence* 175 (2011), 1053–1091. <https://doi.org/10.1016/j.artint.2010.11.018>

A PROOFS

PROPOSITION 1. *Let φ be an LP-tree over \mathcal{X} , then \succeq_φ as defined above is a preorder. Furthermore, \succeq_φ is a linear order if and only if 1) every attribute appears on every branch and 2) every preference rule specifies a linear order.*

PROOF. By definition, \succeq_φ is reflexive. For transitivity, the proof given by [2] is for a restricted family of LP-trees, so we recast it here for our more general family of LP-trees. Suppose that $o \succeq_\varphi o' \succeq_\varphi o''$ and o, o', o'' are distinct. There must be a node N at which $\{o, o'\}$ is decided, let W be the set of attributes that labels N , then $o[\text{Anc}(N)] = o'[\text{Anc}(N)]$, and there is one rule $\alpha : \succeq$ such that $o[\text{NonInst}(N)] = o'[\text{NonInst}(N)] \models \alpha$ and $o[W] \geq o'[W]$. Similarly, let N' be the node at which $\{o', o''\}$ is decided, let W' be the set of attributes that labels N' , then $o[\text{Anc}(N')] = o''[\text{Anc}(N')]$,

and there is one rule $\alpha' : \geq'$ s.t. $o'[\text{NonInst}(N')] = o''[\text{NonInst}(N')] \models \alpha'$ and $o'[W'] \geq' o''[W']$. If $N = N'$, then $o[\text{Anc}(N)] = o'[\text{Anc}(N)] = o''[\text{Anc}(N)]$ and $\geq = \geq'$ is transitive (it is a preorder) thus $o[W] \geq o''[W]$ hence $o \geq_\varphi o''$. If $N \neq N'$, note that both nodes are in the unique branch in φ that corresponds to o' , so one of N, N' must be above the other. Suppose that N is above N' , then, it must be the case that $o'[W] = o''[W]$, and $o[W] \neq o'[W]$, thus N decides $\{o, o''\}$; moreover, since $\text{NonInst}(N) \subseteq \text{NonInst}(N')$, $o'[\text{NonInst}(N)] = o''[\text{NonInst}(N)] \models \alpha$, and $o[W] \geq o'[W] = o''[W]$; hence $o \geq_\varphi o''$. The case where N' is above N is similar.

For the second part of the proposition, suppose first that every attribute appears on every branch and that every preference rule specifies a linear order: we will show that \geq_φ is antisymmetric and connex. For antisymmetry, consider distinct alternatives $o, o' \in \underline{X}$: because every attribute appears on every branch, there must be a node N , labelled with some $W \subseteq \underline{X}$, that decides $\{o, o'\}$, and a unique rule $\alpha : \geq$ at N such that $o[\text{NonInst}(N)] = o'[\text{NonInst}(N)] \models \alpha$; \geq must be a linear order over \underline{W} , so either $o[W] > o'[W]$ and $o >_\varphi o'$, or $o'[W] > o[W]$ and $o' >_\varphi o$: \geq_φ is connex and antisymmetric. For the converse, assuming that either there is some branch where some attribute does not appear, or that there is some rule at some node that does not define a linear order, it is not difficult to define two distinct alternatives that cannot be compared with \geq_φ . \square

PROPOSITION 2. For a fixed $k \in \mathbb{N}$, checking if a formula $\varphi \in \text{CP}$ is k -lexico-compatible is coNP-complete.

PROOF. For membership in coNP: a certificate that some given φ is not lexico-compatible is a branch of a tree built using the algorithm above where failure occurs. coNP-completeness can be proved using the same reduction of 3SAT used by [35, Prop. 24] to prove that checking cuc-acyclicity is coNP-hard. Consider m clauses C_1, \dots, C_m over n binary attributes X_1, \dots, X_n . Let $\underline{X} = \{X_1, \dots, X_n, Y_0, Y_1, \dots, Y_m\}$, where the Y_i s are new binary attributes. Define

$$\varphi = \{l \mid Y_k : y_{k-1} \geq \bar{y}_{k-1} \mid l \in C_k, 1 \leq k \leq m\} \cup \{Y_0 : y_m \geq \bar{y}_m\}.$$

and consider some complete LP tree ψ over set of attributes \underline{X} . Every attribute Y_k appears in the free part of at least one rule of φ , thus cannot be at the root of any complete LP tree ψ that is compatible with φ . On the other hand, any of the X_i 's can be at the root, or at any level, in any branch of ψ . Suppose now that $C_1 \wedge \dots \wedge C_m$ is satisfiable: let u be an instantiation of X_1, \dots, X_n that satisfies this CNF, and consider a branch of ψ where all the X_i s have the same value as in u : at every node N in such a branch, for every Y_k , $\text{inst}(N)$ is consistent with the condition of at least one rule which has Y_k as free part; therefore, no ordering of the Y_k s in such a branch can be compatible with condition 1 in Proposition 18. On the other hand, if $C_1 \wedge \dots \wedge C_m$ is unsatisfiable, then it is not difficult to see that it is possible to build ψ in such a way that the conditions of Proposition 18 are satisfied w.r.t. φ , by taking, for instance, the X_i s for the nodes at the n first levels of ψ : then, since $C_1 \wedge \dots \wedge C_m$ is unsatisfiable, in every branch of the tree there must be one clause C_k that is not satisfied by the corresponding instantiation of the X_i s, so none of the conditions of the corresponding rules $l \mid Y_k : y_{k-1} \geq \bar{y}_{k-1}$ is satisfied; then attribute Y_{k-1} can be chosen for the label of the node at the next level, then Y_{k-1} , and so on. \square

PROPOSITION 3. Given $\varphi \in \text{CP}$ and some $k \in \mathbb{N}$, suppose that chooseAttribute is φ -compatible, then $\varphi \in \text{CP}_k^{\text{lex}}$ if and only if the algorithm above returns some $\psi \in k\text{-LPT}$ such that $\geq_\psi \supseteq \geq_\varphi$; otherwise, it returns FAILURE.

In order to prove Proposition 3, we first state and prove the following result, similar to that of [35, Prop. 13], which shows that it is possible to check in polynomial time, given $\varphi \in \text{CP}$ and some complete $\psi \in \text{LPT}$, if \geq_ψ extends \geq_φ . For ease of presentation, we introduce the following definition:

Definition 4. Statement $\alpha \mid V : w \geq w'$ is *relevant* at node N of some LP-tree ψ if $\text{Var}(w) \cap \text{Anc}(N) = \emptyset$, and $\alpha \wedge \text{inst}(N) \not\models \perp$, and $\text{Var}(w) \cap \text{Var}(N) \neq \emptyset$.

Informally, $\alpha \mid V : w \geq w'$ is relevant at node N if it sanctions some swap $(urvw, urvw')$, with $u \in \text{Var}(\alpha)$ such that $u \models \alpha$ and $v \in \underline{V}$, which may be decided at N .

PROPOSITION 18 (GENERALISATION OF PROP. 13 BY [35]). Let $\varphi \in \text{CP}$ and let ψ be some complete LP-tree. Then $\geq_\psi \supseteq \geq_\varphi$ if and only if for every node $N \in \psi$, for every $\alpha \mid V : w \geq w' \in \varphi$ that is relevant at N , the following two conditions hold:

- (1) $V \cap \text{Anc}(N) = \emptyset$; and
- (2) for every rule $\beta : \geq^\beta \in \text{CPT}(N)$ such that $\alpha \wedge \beta \not\models \perp$, for every $u \in \text{Var}(\alpha)$ such that $u \wedge \alpha \wedge \beta \wedge \text{inst}(N) \not\models \perp$, for every $s \in \underline{\text{Var}(N)} \setminus (\underline{\text{Var}(\alpha)} \cup \underline{V} \cup \underline{\text{Var}(w)})$, for every $v_1, v_2 \in \underline{V}$, it must be the case that $usv_1w[\text{Var}(N)] >^\beta usv_2w'[\text{Var}(N)]$.

PROOF. Let us first prove that conditions 1 and 2 are necessary for every node N of ψ labelled with T and every statement $\alpha \mid V : w \geq w' \in \varphi$ that is relevant at N , assuming that $\geq_\psi \supseteq \geq_\varphi$. Let $R = \underline{X} \setminus (\underline{\text{Var}(\alpha)} \cup \underline{V} \cup \underline{\text{Var}(w)})$. Suppose first that there is some $Y \in V \cap \text{Anc}(N)$: there must be some node N_0 between the root and N in ψ such that Y is in the label S of N_0 . Choose some $u \in \underline{\text{Var}(\alpha)}$, $\bar{v} \in \underline{V} \setminus \{Y\}$, $r \in \underline{R}$ such that $u \models \alpha$ and $u\bar{v}r \wedge \text{inst}(N) \not\models \perp$ (this is possible because $\alpha \wedge \text{inst}(N) \not\models \perp$); pick $y_1, y_2 \in \underline{V}$ with $y_1 \neq y_2$, then φ sanctions $(ury_1\bar{v}w, ury_2\bar{v}w')$ and $(ury_2\bar{v}w, ury_1\bar{v}w')$, so, since $\geq_\psi \supseteq \geq_\varphi$, $ury_1\bar{v}w \geq_\psi ury_2\bar{v}w'$ and $ury_2\bar{v}w \geq_\psi ury_1\bar{v}w'$. Moreover, $\{ury_1\bar{v}w, ury_2\bar{v}w'\}$ and $\{ury_2\bar{v}w, ury_1\bar{v}w'\}$ are both decided in ψ at N_0 because $W \cap \text{Anc}(N_0) \subseteq W \cap \text{Anc}(N) = \emptyset$; let $\beta : \geq^\beta$ be the unique rule in $\text{CPT}(N_0)$ such that $ur\bar{v}[\text{Anc}(N_0)] \models \beta$: then $ury_1\bar{v}[S] \geq^\beta ury_2\bar{v}[S]$ because $ury_1\bar{v}w \geq_\psi ury_2\bar{v}w'$, and $ury_2\bar{v}[S] \geq^\beta ury_1\bar{v}[S]$ because $ury_2\bar{v}w \geq_\psi ury_1\bar{v}w'$: this would imply $ury_2\bar{v}[S] \sim^\beta ury_1\bar{v}[S]$, which is impossible since ψ is a complete LP-tree, and so, by definition, every rule at every node assigns a linear order. This proves that $V \cap \text{Anc}(N) = \emptyset$. Now let $\beta : \geq^\beta \in \text{CPT}(N)$ such that $\alpha \wedge \beta \not\models \perp$: let

- $u \in \underline{\text{Var}(\alpha)}$ such that $u \wedge \alpha \wedge \beta \wedge \text{inst}(N) \not\models \perp$,
- $s \in \underline{\text{Var}(N)} \setminus (\underline{\text{Var}(\alpha)} \cup \underline{V} \cup \underline{\text{Var}(w)})$,
- $v_1, v_2 \in \underline{V}$,
- $r \in \underline{X} \setminus \underline{\text{Var}(N)} \cup \underline{\text{Var}(\alpha)} \cup \underline{V} \cup \underline{\text{Var}(w)}$ such that $r \wedge \beta \not\models \perp$,

then $\alpha \mid V : w \geq w' \in \varphi$ sanctions (uv_1srw, uv_2srw') , so $uv_1srw \geq_\psi uv_2srw'$, and the pair $\{uv_1srw, uv_2srw'\}$ is decided in ψ at N with the rule $\beta : \geq^\beta$, so it must be the case that $uv_1sw[\text{Var}(N)] >^\beta uv_2sw'[\text{Var}(N)]$.

Let us now prove that the condition is sufficient: it is in fact sufficient to prove that, if the condition above holds, then for every statement $\alpha \mid V : w \geq w' \in \varphi$, for every pair of alternatives $\{o, o'\}$ sanctioned by the statement, it is the case that $o \geq_{\psi} o'$. We must have that $o = urv_1w$ and $o' = urv_2w'$ for some $u \in \underline{\text{Var}(\alpha)}$ such that $u \models \alpha$, some $v_1, v_2 \in \underline{V}$, some $r \in \underline{X \setminus (\text{Var}(\alpha) \cup V \cup \text{Var}(w))}$. Let N be the node that decides $\{o, o'\}$ in ψ , let us prove that $\alpha \mid V : w \geq w'$ is relevant at N . That $\text{Var}(w) \cap \text{Anc}(N) = \emptyset$ follows from the fact that $\{o, o'\}$ is not decided at some other node between the root and N ; also, the common part of o and o' must be consistent with $\text{inst}(N)$ (otherwise, the pair would have taken another branch on the tree), so $u \wedge \text{inst}(N) \not\models \perp$, and u is a model of α , so $\alpha \wedge \text{inst}(N) \not\models \perp$; there remains to prove that $\text{Var}(w) \cap \text{Var}(N) \neq \emptyset$. Suppose first that $v_2 = v_1$, then it must be the case that $w[\text{Var}(N)] \neq w'[\text{Var}(N)]$ (since $\{o, o'\}$ is decided at N , so $\text{Var}(w) \cap \text{Var}(N) \neq \emptyset$: $\alpha \mid V : w \geq w'$ is indeed relevant at N , and condition 1 implies that $V \cap \text{Anc}(N) = \emptyset$). Suppose now that $v_2 \neq v_1$, let $o'' = urv_1w'$, let N' be the node that decides $\{o, o''\}$, from what we have just seen, in particular that $V \cap \text{Anc}(N') = \emptyset$ we can conclude that N' also decides $\{o, o'\}$, thus $N = N'$: it follows that $V \cap \text{Anc}(N) = \emptyset$, thus, since N decides $\{o, o'\}$, it must be the case that $\text{Var}(w) \cap \text{Var}(N) \neq \emptyset$: statement $\alpha \mid V : w \geq w'$ is again relevant at N . Consider now the unique rule $\beta : \geq^{\beta} \in \text{CPT}(N)$ such that $ur[\text{Anc}(N)] \models \beta$, then $u \wedge \alpha \wedge \beta \not\models \perp$ since $u \models \alpha$; let $s = r[\text{Var}(N)]$, then $s \in \underline{\text{Var}(N) \setminus (\text{Var}(\alpha) \cup V \cup \text{Var}(w))}$, thus condition 2 implies that $usv_1w[\text{Var}(N)] >^{\beta} usv_2w'[\text{Var}(N)]$; since $\{o, o'\}$ is decided at N , it follows that $o \geq_{\psi} o'$. \square

LEMMA 1. Let $\varphi \in \text{CP}$, suppose that chooseAttribute is φ -compatible, that the algorithm above terminates without failure on φ , k , and returns some complete k -LP-tree ψ , then $\geq_{\varphi} \subseteq \geq_{\psi}$.

PROOF. Let $N \in \psi$, let $\alpha \mid V : w \geq w' \in \varphi$ that is relevant at N , we only need to prove that conditions 1 and 2 in Proposition 18 hold.

Let us first prove that $V \cap \text{Anc}(N) = \emptyset$. Let N' be any node between the root and N , excluding N ; it is labelled with some $T' \subseteq X$ and a linear order $\geq_{N'}$ over T' that have been returned by chooseAttribute when called at N' . We know that $\alpha \wedge \text{inst}(N) \not\models \perp$, so, since $\text{inst}(N)$ extends $\text{inst}(N')$, $\alpha \wedge \text{inst}(N') \not\models \perp$; moreover $\text{Var}(w) \cap \text{Anc}(N) = \emptyset$ because $\alpha \mid V : w \geq w' \in \varphi$ is relevant at N , and $\text{Anc}(N') \subset \text{Anc}(N)$ so $\text{Var}(w) \cap \text{Anc}(N') = \emptyset$, hence $\alpha \mid V : w \geq w' \in \varphi(N')$; since $T' \subseteq \text{Anc}(N)$, and $\text{Var}(w) \cap \text{Anc}(N) = \emptyset$, we also have that $T' \cap \text{Var}(w) = \emptyset$; thus, according to condition ?? that must be verified by chooseAttribute , it must be the case that $V \cap T' = \emptyset$. This holds for every label of every node above N , hence $V \cap \text{Anc}(N) = \emptyset$.

Let us now prove that condition 2 holds. Since the algorithm returns LP-trees with no unlabelled edge, it can be rewritten as follows:

- 2'. for every $u \in \underline{\text{Var}(\alpha)}$ such that $u \models \alpha$ and $u \wedge \text{inst}(N) \not\models \perp$, for every $s \in \underline{\text{Var}(N) \setminus (\text{Var}(\alpha) \cup V \cup \text{Var}(w))}$, for every $v_1, v_2 \in \underline{V}$, it must be the case that $usv_1w[\text{Var}(N)] >^{\beta} usv_2w'[\text{Var}(N)]$.

We know that $\text{Var}(w) \cap T \neq \emptyset$ because $\alpha \mid V : w \geq w' \in \varphi$ is relevant at N . Let $u \in \underline{\text{Var}(\alpha)}$ such that $u \models \alpha$ and $u \wedge \text{inst}(N) \not\models \perp$, let $s \in \underline{\text{Var}(N) \setminus (\text{Var}(\alpha) \cup V \cup \text{Var}(w))}$, let $v_1, v_2 \in \underline{V}$, and let $t =$

$usv_1w[T]$, $t' = usv_2w'[T]$, then clearly $t \wedge w \not\models \perp$, $t' \wedge w' \not\models \perp$, $t[X \setminus (V \cup \text{Var}(w))] = us = t'[X \setminus (V \cup \text{Var}(w))]$, and $t \wedge \alpha \not\models \perp$ (because $u \models \alpha$), thus, according to condition ?? in the specification of chooseAttribute , it must be the case that $t >^N t'$, that is $usv_1w[T] >^N usv_2w'[T]$. \square

LEMMA 2. Let $\varphi \in \text{CP}$, let $k \in \mathbb{N}$, let ψ be some complete k -LP-tree, with no uninstantiated edge, such that $\geq_{\varphi} \subseteq \geq_{\psi}$. Let N be some node of ψ , and let $\psi(N)$ be the subtree of ψ rooted at N . Suppose that, when called at N , chooseAttribute returns a pair (T, \geq) different than the one that labels N . Then $\geq_{\varphi} \in \geq_{\psi'}$, where ψ' is a new k -LP-tree obtained by replacing $\psi(N)$ in ψ with a new subtree as follows:

- N is replaced with a new node N' , labelled with (T, \geq) ; in the sequel, in order to avoid confusion, we denote \geq^N this ordering associated with T returned by chooseAttribute ;
- we create \underline{T} copies of $\psi(N)$, each attached to N' with an edge labelled with some $t \in \underline{T}$;
- for each node M in one of these copies, in the branch below N' corresponding to t , we do as follows, where (S, \geq^M) is the set of attributes and the associated linear ordering that label M :
 - all subtrees below M that correspond to paths incompatible with t are deleted;
 - if $S \subseteq T$, then there remain only one subtree of M , we can safely remove node M ;
 - otherwise, we replace (S, \geq^M) with $(S \setminus T, \geq^M_t)$, where \geq^M_t is the linear order over $S \setminus T$ defined as follows: for every $r, r' \in S \setminus T$, $r \geq^M_t r'$ if and only if $\text{frt}[S] \geq^M r' t[S]$.

PROOF. We prove first that $\psi' \in \text{k-LPT}$:

- the nodes not in $\psi(N)$ are unchanged;
- since T is returned by chooseAttribute , $|T| \leq k$;
- the set of attributes that labels any node below N' is diminished, since the attributes in T are removed.

Thus, since $\psi \in \text{k-LPT}$, $\psi' \in \text{k-LPT}$.

Consider now a swap (o, o') sanctioned by rule $\alpha \mid V : w \geq w' \in \varphi$, and let M be the node in ψ' where $\{o, o'\}$ is decided. If $M \notin \psi(N)$, then $o \geq_{\psi'} o'$ if and only if $o \geq_{\psi} o'$, and \geq_{ψ} extends \geq_{φ} , thus $o \geq_{\psi'} o'$. Let us now consider the case where $M \in \psi(N)$, and let (S, \geq^M) be the label of M . We will consider two subcases. Suppose first that $o[T] = o'[T] = t$, since $o[S] \neq o'[S]$ it implies that $S \not\subseteq T$, thus there is in the subtree that replaces $\psi(N)$ in ψ' a node M' , with $\text{inst}(M')$ compatible with t , obtained from M by removing attributes in $S \cap T$, thus getting new set of attributes $R = S \setminus T$, and by conditioning \geq^M with t . $\{o, o'\}$ is not decided at N' , nor at any other node above M' , it is now decided at M' , and $o[S]$ and $o'[S]$ are of the form $\tilde{t}r$ and $\tilde{t}r'$, where $\tilde{t} = t[S]$; since $o \geq_{\psi'} o'$, it must be the case that $\tilde{t}r >^M \tilde{t}r'$, thus $r >^M_t r'$. Therefore $o \geq_{\psi'} o'$. Let us finally consider the case where $o[T] \neq o'[T]$: $\{o, o'\}$ is now decided at N' . Since $\{o, o'\}$ is decided at M in ψ , we know that $\text{Var}(w) \cap \text{Anc}(M) = \emptyset$ and that $\alpha \wedge \text{inst}(M) \not\models \perp$, thus, since N is above M in ψ , and since $\text{inst}(N) = \text{inst}(N')$, we have that $\alpha \wedge \text{inst}(N') \not\models \perp$. Therefore, $\alpha \mid V : w \geq w' \in \varphi(N')$. Recall that T has been returned by chooseAttribute called at N' : thus, if we had $\text{Var}(w) \cap T = \emptyset$, we would also have $V \cap T = \emptyset$, which is impossible because, since (o, o') is sanctioned by $\alpha \mid V : w \geq w'$, $o[T \setminus (V \cup \text{Var}(w))] = o'[T \setminus (V \cup \text{Var}(w))]$, thus the pair would not

be decided at N' . But, since $\text{Var}(w) \cap T \neq \emptyset$, let $t = o[T]$, $t' = o'[T]$, they are compatible with respectively w and w' (since o and o' respectively extend w and w'), and $t[T \setminus (V \cup \text{Var}(w))] = t'[T \setminus (V \cup \text{Var}(w))]$, thus it must be the case, according to condition ?? in the specification of chooseAttribute, that $t >^{N'} t'$, thus $o >_{\psi} o'$. \square

PROOF OF PROPOSITION 3. Suppose first that the algorithm succeeds and returns a k -LP-tree: lemma 1 proves that $\geq_{\varphi} \subseteq \geq_{\psi}$, and we have mentioned that all labels returned by chooseAttribute have no more than k attributes, so it is in k -LPT, so $\varphi \in \text{CP} \not\subseteq_k^{\text{lex}}$. Now, suppose that $\varphi \in \text{CP} \not\subseteq_k^{\text{lex}}$: there is some k -LP-tree ψ such that $\geq_{\psi} \supseteq \geq_{\varphi}$. Consider the execution of the algorithm with φ and k as inputs. Let N be the first node such that chooseAttribute returns a pair (T, \geq) that is not in ψ at N : then lemma 2 proves that there is some $\psi' \in k$ -LPT that is equal to ψ except for the subtree rooted at N , and that is equal to ψ for the part built before N , and that has (T, \geq) at N : this shows that the algorithm cannot run into a dead end. \square

PROPOSITION 5. $\text{CP} = \bigcup_{k \in \mathbb{N}} k\text{-CP}$ and, for every $k \in \mathbb{N}$:
 $\text{CP} \wedge \sqsupseteq \text{CP} \not\subseteq \sqsupseteq \text{CP} \sqsupseteq k\text{-CP} \sqsupseteq k\text{-CP} \wedge \sqsupseteq k\text{-CP} \not\subseteq \sqsupseteq k\text{-CP} \wedge \not\subseteq$
 $k\text{-CP} \sqsupseteq (k-1)\text{-CP}$.

PROOF. That $\text{CP} \sqsupseteq \text{CP} \wedge \sqsupseteq \text{CP} \wedge \not\subseteq$ follows from property 4.

By definition $\text{CP} \sqsupseteq 1\text{-CP} \sqsupseteq 1\text{-CP} \not\subseteq \sqsupseteq 1\text{-CP} \wedge \not\subseteq$ and $1\text{-CP} \sqsupseteq 1\text{-CP} \wedge \sqsupseteq 1\text{-CP} \wedge \not\subseteq$, thus $\text{CP} \sqsupseteq 1\text{-CP} \sqsupseteq 1\text{-CP} \not\subseteq \sqsupseteq 1\text{-CP} \wedge \not\subseteq$ and $1\text{-CP} \sqsupseteq 1\text{-CP} \wedge \sqsupseteq 1\text{-CP} \wedge \not\subseteq$. Restricting to conjunction of literals does not induce a loss in expressiveness because, given a statement $\alpha \mid V : x \geq x'$, it is possible to compute a DNF logically equivalent to α , and then consider a set of statements, each statement having one disjunct of the DNF as conditioning part. Example 3 prove that $\text{CP} \sqsupseteq 1\text{-CP}$. Example 4 proves that $1\text{-CP} \sqsupseteq 1\text{-CP} \not\subseteq$. \square

PROPOSITION 10. LINEARISABILITY can be checked in polynomial time for LPT.

PROOF. This is because $\varphi \in \text{LPT}$ is linearisable if and only for every rule $\alpha : \geq$ at every node, \geq is antisymmetric. \square

PROPOSITION 12. $>$ -COMPARISON and \bowtie -COMPARISON are NP-hard for the language of fully acyclic CP-nets, and tractable for polytree CP-nets. \sim -COMPARISON is easy for $1\text{-CP} \not\subseteq^{\text{lex}}$.

Proposition 12 is proved using these simple observations:

LEMMA 3. Let $\mathcal{L} \subseteq \text{CP}$.

- (1) If \geq -COMPARISON is tractable for \mathcal{L} , then so are $>$ -COMPARISON, \sim -COMPARISON and \bowtie -COMPARISON.
- (2) If ORDERING is tractable for \mathcal{L} , then \bowtie -COMPARISON is at least as hard as \geq -COMPARISON for \mathcal{L} .

Furthermore, if \mathcal{L} has the property that all its formulas are known to be linearisable, then \geq_{φ} is antisymmetric and:

- (3) \geq -COMPARISON and $>$ -COMPARISON have the same complexity for \mathcal{L} , because $>_{\varphi}$ is the irreflexive part of \geq_{φ} ;
- (4) \sim -COMPARISON is easy for \mathcal{L} , because $o \sim_{\varphi} o'$ holds if and only if $o = o'$.

PROOF. (1) $>$ -COMPARISON, \sim -COMPARISON and \bowtie -COMPARISON queries can be answered with $2 \geq$ -COMPARISON queries. (2) In order to check if $o \geq_{\varphi} o'$, one can ask if $o \bowtie_{\varphi} o'$: if the answer is yes, then $o \not\geq_{\varphi} o'$; if the answer is no, then one can ask to order $\{o, o'\}$: if the answer is that $o \not\geq_{\varphi} o'$, then we have the answer to the initial query; if the answer is that $o' \not\geq_{\varphi} o$, since we know that o and o' are not incomparable, it must be the case that $o \geq_{\varphi} o'$.

The remaining properties are elementary. \square

PROOF OF PROPOSITION 12. That $>$ -COMPARISON is NP hard for the language of fully acyclic CP-nets follows from the fact that \geq -COMPARISON is hard for this language [4, Theorems 15, 16] and from property 3 in lemma 3. Property 2 shows that \bowtie -COMPARISON too is NP hard for this language because ORDERING is easy for this language [4, Theorems 5] and \geq -COMPARISON is hard. These two queries are tractable for polytree CP-nets because \geq -COMPARISON is (property 1 in lemma 3). \sim -COMPARISON is easy for $1\text{-CP} \not\subseteq^{\text{lex}}$ because lexico-compatible formulas are linearisable. \square

PROPOSITION 13. EQUIVALENCE is coNP-hard for $1\text{-CP} \not\subseteq \not\subseteq$, and for $1\text{-LPT} \wedge$, both restricted to binary attributes.

Given a propositional language \mathcal{P} we define \mathcal{P}^{\vee} to be the set of finite disjunctions of formulas in \mathcal{P} , and:

$1\text{-CP} \not\subseteq \mathcal{P}$ is $1\text{-CP} \not\subseteq$ restricted to those statements such that the condition is in \mathcal{P}

$1\text{-LPT} \mathcal{P}$ is 1-LPT restricted to those LP-trees such that the condition of every rule is in \mathcal{P} .

The proof of the proposition is based on the following lemma, which formalizes the intuition suggested by Example 9.

LEMMA 4. Given a propositional language \mathcal{P} closed for conjunction, EQUIVALENCE for \mathcal{P}^{\vee} (in the sense of propositional logic), reduces to EQUIVALENCE for $1\text{-CP} \not\subseteq \mathcal{P}$ restricted to fully acyclic formulas, and to EQUIVALENCE for $1\text{-LPT} \mathcal{P}$.

PROOF. Consider two formulas $\alpha = \bigvee_{i \in I} \alpha_i$ and $\alpha' = \bigvee_{i \in I'} \alpha'_i$ over a set \mathcal{X} of binary attributes, with all α_i 's and α'_i 's in \mathcal{P} ; take some binary attribute $X \notin \mathcal{X}$, with values x and \bar{x} , and let $\varphi = \{\alpha_i : x \geq \bar{x} \mid i \in I\}$ and $\varphi' = \{\alpha'_i : x \geq \bar{x} \mid i \in I'\}$. Note that $\varphi, \varphi' \in 1\text{-CP} \not\subseteq \mathcal{P}$, that they are acyclic, and that they can be computed in time polynomial in $|\alpha| + |\alpha'|$. Then $\varphi^* = \{(ox, o\bar{x}) \mid o \in \underline{\mathcal{X}}, o \models \alpha\}$ and for every $o_1, o_2 \in \underline{\mathcal{X}}$, for every $x_1, x_2 \in \underline{\mathcal{X}}$, $o_1 x_1 \geq_{\varphi} o_2 x_2$ if and only if $o_1 = o_2$, $x_1 = x$, $x_2 = \bar{x}$ and $o_1 \models \alpha$; similarly, $o_1 x_1 \geq_{\varphi'} o_2 x_2$ if and only if $o_1 = o_2$, $x_1 = x$, $x_2 = \bar{x}$ and $o_1 \models \alpha'$. Thus $\alpha \equiv \alpha'$ iff and only if for every $o \in \underline{\mathcal{X}}$, $o \models \alpha \Leftrightarrow o \models \alpha'$, iff for every $o \in \underline{\mathcal{X}}$, $ox \geq_{\varphi} o\bar{x} \Leftrightarrow ox \geq_{\varphi'} o\bar{x}$, if and only if $\varphi \equiv \varphi'$.

Similarly, we can define two linear 1-LP-trees ψ and ψ' as follows: the top $|\mathcal{X}|$ nodes are labelled with attributes from \mathcal{X} , in any order and with no rule; then there is one node labelled with X , and the same preference rules as above. \square

PROPOSITION 15. TOP- p can be answered in time which is polynomial in the size of φ and the size of S for k -lexico-compatible formulas (for fixed k); and for LPT.

PROOF. It suffices to prove the result for the case where $|S| = 2$. The fact that the query is tractable for LPT in this case is a simple consequence of the fact that $>$ -COMPARISON is tractable for LPT.

For k -lexico-compatible formulas, given some $\varphi \in \text{CP}$ known to be k -lexico-compatible, and given a pair of alternatives $\{o, o'\}$, we can run the algorithm proposed in Section 3.6 that builds a complete LP tree ψ that extends φ , but build one branch only, the one that corresponds to the pair $\{o, o'\}$, as long as the chosen attributes have equal values for o and o' : when reaching a node where the chosen set of attributes T is such that $o[T] \neq o'[T]$, the node decides the pair, and o and o' can be ordered accordingly, as in the case of LP trees: if $o[T] > o'[T]$, then $o >_{\psi} o'$, thus $o' \not\prec_{\psi} o$, hence $o' \not\prec_{\varphi} o$; similarly, if $o'[T] > o[T]$, then it cannot be the case that $o >_{\varphi} o'$. The algorithm will not return FAILURE because φ is known to be k -lexico-compatible. The branch has no more than $|X|$ nodes, and there is, for fixed k , a polynomial number of possible labels to try at each node. \square

PROPOSITION 16. *UNDOMINATED CHECK is in P for CP.*

PROOF. Alternative o is not undominated if and only if there is some o' such that $o' \succeq_{\varphi} o$, which can happen if and only if φ contains a statement $\alpha \mid V : w > w'$ with $o \models \alpha$ and $o[\text{Var}(w')] = w'$: this can be checked in polynomial time. \square

PROPOSITION 17. *\succeq -CUT EXTRACTION is tractable for CP. $>$ -CUT COUNTING and $>$ -CUT EXTRACTION are PSPACE-hard for $1\text{-CP} \not\prec \wedge$. For CP_k^{lex} , $>$ -CUT EXTRACTION is equivalent to \succeq -CUT EXTRACTION and is tractable. $>$ -CUT COUNTING is tractable for LP-trees.*

PROOF. \succeq -CUT EXTRACTION is easy for CP: given o and φ , in order to return an element of $\text{CUT}^{\succeq, o}(\varphi)$, it is sufficient to find one statement in φ which sanctions an improving swap for o .

Note that alternative o is weakly undominated iff $\text{CUT}^{>, o}(\varphi) = \emptyset$, iff $|\text{CUT}^{>, o}(\varphi)| = 0$; therefore, $>$ -CUT COUNTING and $>$ -CUT EXTRACTION are at least as hard as WEAKLY UNDOMINATED CHECK, they are therefore PSPACE-hard for $1\text{-CP} \not\prec \wedge$. For acyclic CP-nets, and more generally for cuc-acyclic CP theories, $>$ is the irreflexive part of \succeq , thus $>$ -CUT EXTRACTION is equivalent to \succeq -CUT EXTRACTION and is easy. Finally, $>$ -CUT COUNTING is tractable for LP-trees: given some o , when going down some LP-tree ψ in the branch that corresponds to o , it is possible, at each node N encountered, labelled with T , to count the number of t' in \underline{T} such that $t > o[t']$ (according to the preference rule $\beta : \succeq^{\beta}$ at N such that $o \models \beta$), and to multiply this number by the sizes of the domains of the attributes that have not been encountered yet; adding these sums of products along the branch will give the number of alternatives o such that $o' >_{\psi} o$. \square