

# Efficient Inference and Computation of Optimal Alternatives for Preference Languages Based On Lexicographic Models

Nic Wilson, Anne-Marie George

Insight Centre for Data Analytics, School of Computer Science and IT  
 University College Cork, Ireland  
 {nic.wilson, annemarie.george}@insight-centre.org

## Abstract

We analyse preference inference, through consistency, for general preference languages based on lexicographic models. We identify a property, which we call *strong compositionality*, that applies for many natural kinds of preference statement, and that allows a greedy algorithm for determining consistency of a set of preference statements. We also consider different natural definitions of optimality, and their relations to each other, for general preference languages based on lexicographic models. Based on our framework, we show that testing consistency, and thus inference, is polynomial for a specific preference language  $\mathcal{L}'_{pqT}$ , which allows strict and non-strict statements, comparisons between outcomes and between partial tuples, both *ceteris paribus* and strong statements, and their combination. Computing different kinds of optimal sets is also shown to be polynomial; this is backed up by our experimental results.

## 1 Introduction

Preferences are considered in many different fields such as recommender systems and human-computer interaction [Chen and Pu, 2007; Trabelsi *et al.*, 2013], databases [Agrawal and Wimmers, 2000; Kießling, 2002] and multi-objective decision making and social choice [Arrow and Raynaud, 1986; Lang, 2002; Sandholm and Boutilier, 2006]. They are used to, e.g., find optimal decisions, solutions or items that satisfy the wishes or needs of a user or a group. In order to reason with preferences, typically assumptions are made on the form of the relation used to model the user preferences. In this paper we restrict our considerations to lexicographic models, and we consider preference inference, which involves reasoning about the the whole set of preference models that are consistent with the input preferences, as in [Kohli and Jedidi, 2007; Wilson, 2014]; this contrasts with work that focuses on learning one lexicographic model that fits best with the given preferences [Booth *et al.*, 2010; Bräuning and Hüllermeier, 2012; Dombi *et al.*, 2007; Yaman *et al.*, 2010].

Preference inference aims to overcome gaps of knowledge in the user preferences by analysing the given preferences

and deciding whether another preference statement can be deduced. Since in this paper we allow negations of statements, preference inference can be reduced to testing consistency of a set of preference statements. We define a property, strong compositionality, that is satisfied by many natural types of preference statements and enables a simple greedy algorithm for testing consistency.

Fast computation is essential in many applications; we give a concrete instance  $\mathcal{L}'_{pqT}$  of a preference language that allows polynomial computation. This can represent a relatively expressive form of preference input. It can express preferences between a pair of complete assignments to a set of variables, and, more generally, between partial assignments to just a subset of the variables, allowing *ceteris paribus* assumptions or stronger implications on the values of other variables. We also allow non-strict statements, two forms of strict preference statements and can represent certain negated statements.

Previous work on preference inference based on standard lexicographic models have considered more restricted preference languages. Wilson [2014] considered only non-negated non-strict statements, which can only express that one assignment is at least as good as another (or equivalent). Kohli and Jedidi [2007] considers only non-negated strict statements, which can only express that one complete assignment is strictly better than another. In [Wilson *et al.*, 2015a], preference statements are comparisons of complete assignments, and fixed value orders for the variables are given.

Other work on preference inference includes that based on weighted sum models [Wilson *et al.*, 2015b; Montazery and Wilson, 2016], on hierarchical models [Wilson *et al.*, 2015a], conditional lexicographic models [Wilson, 2009]; Pareto orders [George and Wilson, 2016], and on general strict total orders, as in e.g., work on conditional preference structures such as [Boutilier *et al.*, 2004]. While lexicographic models prevent tradeoffs between variables, they usually allow better complexity results for general preference languages.

Given a set of preference inputs and a set of alternatives, there are several natural notions of optimal set; for instance being undominated with respect to the induced preference relation; or being possibly optimal, i.e., optimal in at least one model. We establish relationships between the different notions of optimality. For preference language  $\mathcal{L}'_{pqT}$ , we analyse the running times and set relations theoretically and experimentally, illustrating the efficiency of the algorithms.

The paper is organised as follows: Section 2 gives basic definitions. The concepts of inference, consistency and strongly compositional statements are investigated in Section 3. Section 4 introduces languages  $\mathcal{L}_{pqT}$  and  $\mathcal{L}'_{pqT}$  and an algorithm to decide consistency in this context. In Section 5, different notions of optimality are compared for general preference languages together with a discussion of computational methods and complexity results for  $\mathcal{L}'_{pqT}$ . The last section concludes. A longer version of this paper including proofs and additional comments can be found under [Wilson and George, 2017].

## 2 Lex Models, Composition and Extension

In this section, we give some basic definitions; in particular, we define lexicographic models, along with a natural composition operation and extension relation, that are important in our approach. A lexicographic model has an associated ordering of (some of the) variables, along with value orderings for each such variable. This generates a (lexicographic) ordering relation, by first comparing outcomes  $\alpha$  and  $\beta$  on the first (i.e., most important) variable  $Y$ ; only if  $\alpha(Y) = \beta(Y)$  is the next most important variable considered.

Throughout the paper we consider a fixed set  $V$  of  $n$  variables, where for each  $X \in V$ ,  $D(X)$  is the set of possible values of  $X$ . For subset of variables  $A \subseteq V$  let  $\underline{A} = \prod_{X \in A} D(X)$  be the set of possible assignments to  $A$ . For  $X \in V$ , we abbreviate  $\{X\}$  to  $\underline{X}$  (which is essentially  $D(X)$ ). An *outcome* is an element of  $\underline{V}$ , i.e., an assignment to all the variables. For outcome  $\alpha$  and subset  $U$  of  $V$ , we define  $\alpha(U)$  to be the tuple in  $\underline{U}$  generated by projecting (i.e., restricting)  $\alpha$  to  $U$ . These definitions describe a decision scenario over a set of alternatives given by different features. Consider for example the choice between different flight connections. A set of variables to describe the flight connections could be  $V = \{\text{airline}, \text{time}, \text{class}\}$ . The variable domains are then given by  $D(\text{airline}) = \{\text{KLM}, \text{LAN}\}$ ,  $D(\text{time}) = \{\text{day}, \text{night}\}$  and  $D(\text{class}) = \{\text{economy}, \text{business}\}$ . Then projection of the outcome  $\alpha = (\text{KLM}, \text{day}, \text{economy})$  to the variables  $U = \{\text{airline}, \text{class}\}$  is given by  $\alpha(U) = (\text{KLM}, \text{economy})$ .

Define  $\mathcal{G}$  to be the set of lexicographic models (over the set of variables  $V$ ); a *lexicographic model* or *lex model*,  $\pi$  (over  $V$ ), is defined to be a (possibly empty) sequence of the form  $(Y_1, \geq_{Y_1}), \dots, (Y_k, \geq_{Y_k})$ , where  $Y_i$  ( $i = 1, \dots, k$ ) are different variables in  $V$ , and each  $\geq_{Y_i}$  is a total order on  $\underline{Y_i}$ . The associated relation  $\succ_{\pi} \subseteq \underline{V} \times \underline{V}$  is defined as follows: for outcomes  $\alpha$  and  $\beta$ ,  $\alpha \succ_{\pi} \beta$  if and only if either (i) for all  $i = 1, \dots, k$ ,  $\alpha(Y_i) = \beta(Y_i)$ ; or (ii) there exists  $i \in \{1, \dots, k\}$  such that for all  $j < i$ ,  $\alpha(Y_j) = \beta(Y_j)$  and  $\alpha(Y_i) \succ_{Y_i} \beta(Y_i)$  (i.e.,  $\alpha(Y_i) \geq_{Y_i} \beta(Y_i)$  and  $\alpha(Y_i) \neq \beta(Y_i)$ ). Thus  $\succ_{\pi}$  is a total pre-order on  $\underline{V}$ , which is a total order if  $k = n = |V|$ .

The corresponding strict relation  $\succ_{\pi}$  is given by  $\alpha \succ_{\pi} \beta$  if and only if there exists  $i \in \{1, \dots, k\}$  such that  $\alpha(Y_i) \succ_{Y_i} \beta(Y_i)$  and for all  $j < i$ ,  $\alpha(Y_j) = \beta(Y_j)$ . The corresponding equivalence relation  $\equiv_{\pi}$  is given by  $\alpha \equiv_{\pi} \beta$  if and only if for all  $i = 1, \dots, k$ ,  $\alpha(Y_i) = \beta(Y_i)$ . Thus,  $\alpha \equiv_{\pi} \beta$  if and only if  $\alpha(V_{\pi}) = \beta(V_{\pi})$ , where  $V_{\pi} = \{Y_1, \dots, Y_k\}$  is set of the variables involved in  $\pi$ . Consider lexicographic models  $\pi =$

$(\text{airline}, \text{KLM} > \text{LAN}), (\text{time}, \text{day} > \text{night})$  and  $\pi' = (\text{class}, \text{economy} > \text{business}), (\text{time}, \text{night} > \text{day})$  for the previous example of flight connections. In  $\pi$  the choice of airline decides which connection is preferred; only if these are the same, is the flight time considered.  $\pi'$  first compares the classes; only if these are the same, are the flight times considered. For outcomes  $\alpha = (\text{KLM}, \text{day}, \text{economy})$ ,  $\beta = (\text{KLM}, \text{night}, \text{business})$  and  $\gamma = (\text{LAN}, \text{day}, \text{economy})$  we have  $\alpha \succ_{\pi} \beta \succ_{\pi} \gamma$  and  $\alpha \equiv_{\pi'} \gamma \succ_{\pi'} \beta$ .

**Composition of lexicographic models:** We define an important composition operation on lex models, which can be shown to be associative. Let  $\pi = (Y_1, \geq_{Y_1}), \dots, (Y_k, \geq_{Y_k})$ , and  $\pi' = (Z_1, \geq_{Z_1}), \dots, (Z_l, \geq_{Z_l})$  be two lexicographic models. Let  $\pi''$  be the sequence  $\pi'$  but where pairs  $(Z_i, \geq_{Z_i})$  are omitted if  $Z_i \in V_{\pi}$ . Define lex model  $\pi \circ \pi'$  to be  $\pi$  followed by  $\pi''$ . Note that  $V_{\pi \circ \pi'} = V_{\pi} \cup V_{\pi'}$ .

**Definition 1** For lex models  $\pi$  and  $\pi'$  in  $\mathcal{G}$  we say that  $\pi'$  extends  $\pi$  if  $\pi' \neq \pi$  and the sequence  $\pi'$  begins with  $\pi$ . We then write  $\pi' \sqsupseteq \pi$ , and write  $\pi' \sqsupset \pi$  to mean that  $\pi'$  extends or equals  $\pi$ . This holds if and only if there exists a lex model  $\pi''$  such that  $\pi' = \pi \circ \pi''$ .

The composition of the two previously defined lex models is  $\pi \circ \pi' = (\text{airline}, \text{KLM} > \text{LAN}), (\text{time}, \text{day} > \text{night}), (\text{class}, \text{economy} > \text{business})$  and  $\pi \circ \pi' \sqsupseteq \pi$ .

## 3 Lexicographic Inference and Strongly Compositional Preference Statements

In this section we define *lexicographic inference*, which can also be expressed in terms of (lexicographic) consistency. We define a notion of *strong compositionality*, and show that this property enables a greedy algorithm for checking consistency, and hence for preference inference. More specifically, a preference statement  $\varphi$  is strongly compositional if the composition of two lex models satisfies  $\varphi$  whenever the second and some extension of the first satisfy  $\varphi$ .

Throughout this section, we assume some language  $\mathcal{L}$ , and satisfaction relation  $\models \subseteq \mathcal{G} \times \mathcal{L}$ . Here, the language  $\mathcal{L}$  is a set of preference statements and the relation  $\models$  describes when a lex model satisfies a statement. For example, we can consider the simple language  $\mathcal{L}^O = \{\alpha \geq \beta \mid \alpha, \beta \in \underline{V}\}$  with the satisfaction relation  $\pi \models \alpha \geq \beta \Leftrightarrow \alpha \succ_{\pi} \beta$  as described in the previous section. We extend the satisfaction relation to subsets  $\Gamma$  of  $\mathcal{L}$  in the usual way: for  $\Gamma \subseteq \mathcal{L}$ , define  $\pi \models \Gamma \Leftrightarrow \pi \models \varphi$  for each  $\varphi \in \Gamma$ . If  $\pi \models \Gamma$  then we say that  $\pi$  *satisfies*  $\Gamma$ , or  $\pi$  is a *[lexicographic] model* of  $\Gamma$  (and similarly, for a single preference statement  $\varphi$ ). We say that  $\Gamma$  is *consistent*, if there exists some  $\pi \in \mathcal{G}$  satisfying  $\Gamma$ ; otherwise,  $\Gamma$  is *inconsistent*. We will sometimes use negations of preference statements: when necessary we can always extend the language  $\mathcal{L}$  and the relation  $\models$ , by defining  $\pi \models \neg\varphi \Leftrightarrow \pi \not\models \varphi$ , when  $\varphi \in \mathcal{L}$ .

A key problem is to determine if a set  $\Gamma \subseteq \mathcal{L}$  of preference statements is consistent. We will derive methods for determining this in Section 3.3.

### Some Basic Preference Statements

For outcomes  $\alpha, \beta \in \underline{V}$ , we can define preference statements  $\alpha \geq \beta$  and  $\alpha > \beta$ , where we define  $\pi \models \alpha \geq \beta \Leftrightarrow$

$\alpha \succ_{\pi} \beta$ , and  $\pi \models \alpha > \beta \iff \alpha \succ_{\pi} \beta$ . We also define preference statement  $\alpha \equiv \beta$ , with  $\pi \models \alpha \equiv \beta \iff \alpha \equiv_{\pi} \beta$ , which is if and only if  $\pi \models \alpha \geq \beta$  and  $\pi \models \beta \geq \alpha$ .

We have that  $\pi \models \alpha \geq \beta \iff \pi \not\models \beta > \alpha$ , i.e.,  $\pi \models \neg(\beta > \alpha)$ . So  $\alpha \geq \beta$  and  $\neg(\beta > \alpha)$  are equivalent preference statements, in that they are satisfied by exactly the same set of lex models. Similarly,  $\alpha > \beta$  and  $\neg(\beta \geq \alpha)$  are equivalent preference statements.

Consider the example of lex models  $\pi, \pi'$  and outcomes  $\alpha, \beta, \gamma$  as before. Then  $\pi \models \alpha \geq \beta$  and thus  $\pi \not\models \beta > \alpha$ . Also,  $\pi' \models \alpha \equiv \gamma$ , i.e.,  $\pi' \models \alpha \geq \gamma$  and  $\pi' \models \gamma \geq \alpha$ .

### 3.1 Lexicographic Inference

We define the lexicographic inference relation  $\Gamma \models \varphi \iff \pi \models \varphi$  for all  $\pi \in \mathcal{G}$  such that  $\pi \models \Gamma$ . For instance, we have, for  $\Gamma \subseteq \mathcal{L}$  and  $\alpha, \beta \in \mathcal{V}$ ,  $\Gamma \models \alpha \geq \beta \iff \alpha \succ_{\pi} \beta$  holds for all  $\pi \in \mathcal{G}$  such that  $\pi \models \Gamma$ . Similarly,  $\Gamma \models \alpha > \beta \iff \alpha \succ_{\pi} \beta$  holds for all  $\pi \in \mathcal{G}$  such that  $\pi \models \Gamma$ .

Lexicographic inference can be reduced to checking consistency, since we have  $\Gamma \models \varphi \iff \Gamma \cup \{\neg\varphi\}$  is inconsistent; and  $\Gamma \models \neg\varphi \iff \Gamma \cup \{\varphi\}$  is inconsistent. Thus,  $\Gamma \models \alpha \geq \beta \iff \Gamma \cup \{\beta > \alpha\}$  is inconsistent; and  $\Gamma \models \alpha > \beta \iff \Gamma \cup \{\beta \geq \alpha\}$  is inconsistent.

Consider flight connections as before. Let  $\Gamma$  be the set of statements  $\{\alpha > \beta, \beta \geq \gamma\}$  with  $\alpha, \beta, \gamma$  defined as in the previous section. Then the lex model  $\pi$  satisfies  $\Gamma$ ,  $\pi \models \Gamma$ , and thus  $\Gamma$  is consistent. In fact, the only preference models that satisfy  $\Gamma$  are:  $\pi = (\text{airline}, KLM > LAN), (\text{time}, \text{day} > \text{night})$ ,  $\pi'' = (\text{airline}, KLM > LAN), (\text{class}, \text{economy} > \text{business})$ , and any extension of either of these two models. Thus for  $\delta = (LAN, \text{night}, \text{business})$ ,  $\Gamma \models \gamma > \delta$ .

#### The Induced Relation $\models^*$ and Maximal Models

From the relation  $\models$  we also define the derived relation  $\models^*$  as follows, where  $\pi \in \mathcal{G}$  and  $\varphi \in \mathcal{L}$ :  $\pi \models^* \varphi$  if and only if there exists  $\pi' \in \mathcal{G}$  either extending or equalling  $\pi$  such that  $\pi' \models \varphi$ . Thus,  $\pi \models^* \varphi$  holds either if  $\pi$  satisfies  $\varphi$  or some extension of  $\pi$  satisfies  $\varphi$ . We extend the relation to sets of statements in the usual way: for  $\Gamma \subseteq \mathcal{L}$ , define  $\pi \models^* \Gamma$  if and only if  $\pi \models^* \varphi$  holds for every  $\varphi \in \Gamma$ . The following lemma follows easily from the definitions.

**Lemma 1** Let  $\pi, \pi' \in \mathcal{G}$  and  $\Gamma \subseteq \mathcal{L}$ .

- (i)  $\pi \models \Gamma \Rightarrow \pi \models^* \Gamma$ .
- (ii) Suppose that  $\pi'$  extends  $\pi$ . Then  $\pi' \models^* \Gamma \Rightarrow \pi \models^* \Gamma$ .

**Definition 2** We say that  $\pi \in \mathcal{G}$  is a maximal model of  $\Gamma$  if  $\pi \models \Gamma$  and for all lex models  $\pi'$  extending  $\pi$  we have  $\pi' \not\models \Gamma$ .

Clearly, for  $\pi \in \mathcal{G}$  such that  $\pi \models \Gamma$ , either  $\pi$  is a maximal model of  $\Gamma$  or there exists a maximal model of  $\Gamma$  that extends  $\pi$ . Thus, if  $\Gamma$  is consistent there exists a maximal model of  $\Gamma$ . Analogously, we define a maximal  $\models^*$ -model of  $\Gamma$  to be an element  $\pi \in \mathcal{G}$  such that (i)  $\pi \models^* \Gamma$  and (ii) there does not exist  $\pi'$  extending  $\pi$  with  $\pi' \models^* \Gamma$ .

Consider preference statements  $\Gamma$  relating to flight connections  $\alpha, \beta, \gamma$  as before. The lex model  $(\text{airline}, KLM > LAN)$  satisfies  $(\text{airline}, KLM > LAN) \models^* \Gamma$ . The model  $(\text{airline}, KLM > LAN), (\text{time}, \text{day} > \text{night}), (\text{class}, \text{business} > \text{economy})$  is a maximal model of  $\Gamma$  (and a maximal  $\models^*$ -model of  $\Gamma$ ).

### 3.2 Compositional and Strongly Compositional Preference Statements

We formulate a pair of properties of preference statements that have strong implications regarding lexicographic inference and optimality.

**Definition 3 ((Strongly) Compositional)** Let  $\varphi \in \mathcal{L}$ . We say that  $\varphi$  is compositional if for all  $\pi, \pi' \in \mathcal{G}$ ,

$$\pi \models \varphi \text{ and } \pi' \models \varphi \text{ implies } \pi \circ \pi' \models \varphi.$$

We say that  $\varphi$  is strongly compositional if for all  $\pi, \pi' \in \mathcal{G}$ ,

$$\pi \models^* \varphi \text{ and } \pi' \models \varphi \text{ implies } \pi \circ \pi' \models \varphi.$$

For  $\Gamma \subseteq \mathcal{L}$ , we define  $\Gamma$  to be compositional if every element of  $\Gamma$  is compositional. Similarly, we say that  $\Gamma$  is strongly compositional if every element of  $\Gamma$  is strongly compositional.

#### Instances of Strongly Compositional Statements

Although being strongly compositional might appear to be quite a restrictive assumption, it turns out that it is satisfied by many natural preference statements, as illustrated by the next proposition. First, we give a lemma which, roughly speaking, states that the property of being [strongly] compositional is closed under conjunction.

**Lemma 2** Let  $\Gamma \subseteq \mathcal{L}$  and let  $\psi \in \mathcal{L}$ . Suppose that  $\psi$  is such that for all  $\pi \in \mathcal{G}$ ,  $\pi \models \psi \iff \pi \models \Gamma$ . If  $\Gamma$  is compositional then  $\psi$  is compositional. If  $\Gamma$  is strongly compositional then  $\psi$  is strongly compositional and, for all  $\pi \in \mathcal{G}$ ,  $[\pi \models^* \psi \iff \pi \models^* \Gamma \text{ and } \Gamma \text{ is consistent}]$ .

**Proposition 1** For any outcomes,  $\alpha, \beta \in \mathcal{V}$ , statements  $\alpha \geq \beta$  and  $\alpha > \beta$  are strongly compositional, and, for  $\pi \in \mathcal{G}$ ,  $\pi \models^* \alpha \geq \beta \iff \pi \models \alpha \geq \beta$ . Also, if  $\alpha \neq \beta$ ,  $\pi \models^* \alpha > \beta \iff \pi \models \alpha > \beta$ .

In addition, given  $\mathfrak{R} \subseteq \mathcal{V} \times \mathcal{V}$ , let  $\varphi^{\mathfrak{R}}$  be some statement satisfying:  $\pi \models \varphi^{\mathfrak{R}}$  if and only if  $\pi \succ_{\pi} \mathfrak{R}$ . Then  $\varphi^{\mathfrak{R}}$  is strongly compositional, and, for any  $\pi \in \mathcal{G}$ , we have  $\pi \models^* \varphi^{\mathfrak{R}} \iff \pi \models \varphi^{\mathfrak{R}}$ .

Here, the second half of Proposition 1 follows from the first using Lemma 2. Further important examples of strongly compositional statements will be discussed in Section 4.

#### A Characterisation of Consistency for Strongly Compositional $\Gamma$

The definitions immediately imply the following.

**Lemma 3** Let  $\Gamma \subseteq \mathcal{L}$ , and let  $\pi, \pi' \in \mathcal{G}$ .

- If  $\Gamma$  is strongly compositional then it is compositional.
- If  $\Gamma$  is compositional then  $\pi \models \Gamma$  and  $\pi' \models \Gamma$  imply  $\pi \circ \pi' \models \Gamma$ .
- If  $\Gamma$  is strongly compositional then  $\pi \models^* \Gamma$  and  $\pi' \models \Gamma$  imply  $\pi \circ \pi' \models \Gamma$ .

The last point implies that, for strongly compositional and consistent  $\Gamma$ , if  $\pi \models^* \Gamma$  then there exists a model of  $\Gamma$  either equalling or extending  $\pi$ . In fact we have:

**Lemma 4** Suppose that  $\Gamma$  is strongly compositional, and let  $\pi$  be an element of  $\mathcal{G}$ . Then [there exists  $\pi' \in \mathcal{G}$  with  $\pi' \supseteq \pi$  and  $\pi' \models \Gamma$ ] if and only if [ $\Gamma$  is consistent and  $\pi \models^* \Gamma$ ].

Suppose that  $\pi'$  satisfies strongly compositional  $\Gamma$  and that  $\pi$  is a maximal  $\models^*$ -model of  $\Gamma$ . Since we have  $\pi \circ \pi' \models^* \Gamma$  (because we have  $\pi \circ \pi' \models \Gamma$ ) and  $\pi \circ \pi' \supseteq \pi$ , then the maximality of  $\pi$  implies that  $\pi = \pi \circ \pi'$ , and thus,  $\pi \models \Gamma$ . We also have that  $V_\pi \supseteq V_{\pi'}$ . This implies the following theorem and corollary:

**Theorem 1** *If  $\Gamma$  is consistent and strongly compositional then every maximal  $\models^*$ -model of  $\Gamma$  satisfies  $\Gamma$ . Also, if  $\pi$  and  $\pi'$  are two maximal  $\models^*$ -models of  $\Gamma$  then  $V_\pi = V_{\pi'}$ .*

**Corollary 1** *Let  $\pi$  be any maximal  $\models^*$ -model of strongly compositional  $\Gamma$ . Then  $\Gamma$  is consistent if and only if  $\pi \models \Gamma$ .*

### 3.3 Checking Consistency of a Set of Strongly Compositional Preferences

Corollary 1 shows that we can test consistency of strongly compositional  $\Gamma$ , by finding any maximal  $\models^*$ -model  $\pi$  of it, and checking if  $\pi$  satisfies  $\Gamma$ . In this section we show how a maximal  $\models^*$ -model of  $\Gamma$  can be constructed iteratively, starting from the empty model  $\{\}$ . We first check if  $\{\} \models^* \Gamma$ . The following lemma shows that this is equivalent to  $\Gamma$  being  $\models^*$ -consistent, which is a very weak property, i.e., it just requires that each element of  $\Gamma$  is (individually) consistent.

**Lemma 5** *Let  $\Gamma \subseteq \mathcal{L}$ .  $\Gamma$  is  $\models^*$ -consistent (i.e., there exists some  $\pi \in \mathcal{G}$  with  $\pi \models^* \Gamma$ ) if and only if, for each  $\varphi \in \Gamma$ ,  $\varphi$  is consistent. This also holds if and only if  $\{\} \models^* \Gamma$ .*

We say that  $\pi'$  *minimally extends*  $\pi$  if  $\pi'$  extends  $\pi$  and  $|V_{\pi'}| = |V_\pi| + 1$ , i.e.,  $\pi'$  involves one more variable than  $\pi$ . The next lemma implies that in building a maximal  $\models^*$ -model, we only need consider adding one variable at a time.

**Lemma 6**  *$\pi$  is a maximal  $\models^*$ -model of  $\Gamma$  if and only if  $\pi \models^* \Gamma$  and there exists no  $\pi'$  minimally extending  $\pi$  such that  $\pi' \models^* \Gamma$ .*

Starting with the empty model we grow a maximal  $\models^*$ -model of  $\Gamma$ , by (iteratively) replacing the model with one minimally extending it and still  $\models^*$ -satisfying  $\Gamma$ , if such a model exists. Otherwise, we have a maximal  $\models^*$ -model  $\pi$  of  $\Gamma$ . By Corollary 1, we can test if  $\Gamma$  is consistent by checking  $\pi \models \Gamma$ .

For bounded domain size, testing consistency of strongly compositional  $\Gamma$  can be performed with polynomial number of tests of the form  $\pi \models \varphi$  or  $\pi \models^* \varphi$ . For this algorithm to be efficient we need that each preference statement allows efficient tests of  $\pi \models \varphi$  and  $\pi \models^* \varphi$ . In Section 4 we define languages that satisfy these properties.

## 4 Preference Languages $\mathcal{L}_{pqT}$ and $\mathcal{L}'_{pqT}$

Here we show that certain relatively expressive compact preference languages are strongly compositional. This includes forms of the statements  $\varphi^{\mathfrak{R}}$  from Proposition 1, where  $\mathfrak{R}$  is a set of pairs of outcomes. In many natural situations,  $\mathfrak{R}$  can be exponentially large; in the languages discussed here, we are able to express certain exponentially large sets  $\mathfrak{R}$  compactly.

### 4.1 The Language $\mathcal{L}_{pqT}$

We will consider preference statements of the form  $p \triangleright q \parallel T$ , where  $\triangleright$  is either  $\geq$ , or  $\gg$  or  $>$ , and  $P, Q$  and  $T$  are subsets

of  $V$ , with  $(P \cup Q) \cap T = \emptyset$ , and  $p \in P$  is an assignment to  $P$ , and  $q \in Q$  is an assignment to  $Q$ . The statement  $p \triangleright q \parallel T$  represents that  $p$  is preferred to  $q$  if  $T$  is held constant, i.e., any outcome  $\alpha$  extending  $p$  is preferred to any outcome  $\beta$  that extends  $q$  and agrees with  $\alpha$  on variables  $T$ . Formally, the semantics of this statement relates to the set  $\varphi^*$  which is defined to be the set of pairs  $(\alpha, \beta)$  of outcomes such that  $\alpha$  extends  $p$ , and  $\beta$  extends  $q$ , and  $\alpha$  and  $\beta$  agree on  $T$ , i.e.,  $\alpha(T) = \beta(T)$ .

Statements of the form  $p \geq q \parallel T$  are called *non-strict*; statements of the form  $p \gg q \parallel T$ , are called *fully strict*, and statements of the form  $p > q \parallel T$  are called *weakly strict*.

Let  $\mathcal{L}_{pqT}$  be the set of all preference statements  $\varphi$  of the form  $p \triangleright q \parallel T$ , as defined above. For any statement  $\varphi \in \mathcal{L}_{pqT}$  equalling  $p \triangleright q \parallel T$ , we define  $\varphi^{(\geq)}$  to be  $p \geq q \parallel T$ , the *non-strict version of  $\varphi$* . For lex model  $\pi$ , we define:

- $\pi$  satisfies  $\varphi^{(\geq)}$  if  $\alpha \succ_{\pi} \beta$  for all  $(\alpha, \beta) \in \varphi^*$ .
- $\pi$  satisfies fully strict  $\varphi$  if  $\alpha \succ_{\pi} \beta$  for all  $(\alpha, \beta) \in \varphi^*$ .
- $\pi$  satisfies weakly strict  $\varphi$  if  $\pi$  satisfies  $\varphi^{(\geq)}$  and if  $\alpha \succ_{\pi} \beta$  for some  $(\alpha, \beta) \in \varphi^*$ .

For outcomes  $\alpha$  and  $\beta$ , a non-strict preference of  $\alpha$  over  $\beta$  can be represented as  $\alpha \geq \beta \parallel \emptyset$ , which is equivalent to the preference statement  $\alpha \geq \beta$  introduced in Section 3, so we abbreviate it to that. Similarly, we abbreviate  $\alpha > \beta \parallel \emptyset$  to  $\alpha > \beta$  (which is also equivalent to  $\alpha \gg \beta \parallel \emptyset$ ).

We can write a statement  $\varphi \in \mathcal{L}_{pqT}$  as  $ur \triangleright us \parallel T$ , where  $u \in \underline{U}$ ,  $r \in \underline{R}$ ,  $s \in \underline{S}$ , and  $U, T$  and  $R \cup S$  are (possibly empty) mutually disjoint subsets of  $V$ , and for all  $X \in R \cap S$ ,  $r(X) \neq s(X)$ . For such a representation, we write  $u_{\varphi} = u$ ,  $r_{\varphi} = r$ ,  $s_{\varphi} = s$ ,  $U_{\varphi} = U$ ,  $R_{\varphi} = R$ ,  $S_{\varphi} = S$  and  $T_{\varphi} = T$ . We assume, without loss of generality, that if  $|\underline{X}| = 1$  then  $X \in T_{\varphi}$ . This ensures that such a representation is unique. We also define  $W_{\varphi} = V - (R_{\varphi} \cup S_{\varphi} \cup T_{\varphi} \cup U_{\varphi})$ .

As discussed in [Wilson, 2009] (which, however, just considers non-strict statements), this is a relatively expressive preference language. As well as preferences between outcomes, preferences between partial tuples can be expressed. *Ceteris paribus* statements can be expressed by using  $\varphi$  with  $W_{\varphi} = \emptyset$ . More generally, any variables in  $W_{\varphi}$  are forced to be less important than variables in  $R_{\varphi}$  and  $S_{\varphi}$ .

The following result characterises when a lex model satisfies a non-strict preference statement in  $\mathcal{L}_{pqT}$ .

**Proposition 2** *Let  $\varphi$  be a non-strict element of  $\mathcal{L}_{pqT}$  (so that  $\varphi = \varphi^{(\geq)}$ ), which we write as  $u_{\varphi} r_{\varphi} \geq u_{\varphi} s_{\varphi} \parallel T_{\varphi}$ . Let  $\pi \in \mathcal{G}$  be the model  $(Y_1, \geq_1), \dots, (Y_k, \geq_k)$ . Let  $i$  be the smallest index such that  $Y_i \in R_{\varphi} \cap S_{\varphi}$ , or let  $i = k + 1$  if  $R_{\varphi} \cap S_{\varphi} \cap V_{\pi} = \emptyset$ . Then,  $\pi \models \varphi$  if and only if (i)  $r_{\varphi}(Y_i) >_i s_{\varphi}(Y_i)$  if  $R_{\varphi} \cap S_{\varphi} \cap V_{\pi} \neq \emptyset$ , and (ii) for all  $j < i$ ,*

- $Y_j \notin W_{\varphi}$ ;
- if  $Y_j \in R_{\varphi} \setminus S_{\varphi}$  then for all  $y \in \underline{Y}_j$ ,  $r_{\varphi}(Y_j) \geq_j y$ , i.e.,  $r_{\varphi}(Y_j)$  is the best value of  $Y_j$ ; and
- if  $Y_j \in S_{\varphi} \setminus R_{\varphi}$  then for all  $y \in \underline{Y}_j$ ,  $y \geq_j s_{\varphi}(Y_j)$ , i.e.,  $s_{\varphi}(Y_j)$  is the worst value of  $Y_j$ .

*In particular, if  $V_{\pi} \subseteq T_{\varphi} \cup U_{\varphi}$  then  $\pi \models \varphi$ .*

The next result gives the extra conditions required for satisfying strict statements.

**Proposition 3** *Let  $\varphi \in \mathcal{L}_{pqT}$  and  $\pi \in \mathcal{G}$ . Then:*

- if  $\varphi$  is a fully strict statement then  $\pi \models \varphi$  if and only if  $\pi \models \varphi^{(\geq)}$  and  $R_\varphi \cap S_\varphi \cap V_\pi \neq \emptyset$ ;
- if  $\varphi$  is a weakly strict statement then  $\pi \models \varphi$  if and only if  $\pi \models \varphi^{(\geq)}$  and  $(R_\varphi \cup S_\varphi) \cap V_\pi \neq \emptyset$ .

Proposition 1 can be seen to imply that the non-strict elements of the language  $\mathcal{L}_{pqT}$  are strongly compositional. In fact this also holds for both kinds of strict statements and certain negations.

**Theorem 2** *Consider any  $\varphi \in \mathcal{L}_{pqT}$ . Then  $\varphi$  is strongly compositional and  $\pi \models^* \varphi$  if and only if  $\pi \models \varphi^{(\geq)}$ . If  $\varphi$  is non-strict then  $\neg\varphi$  is compositional; if also,  $R_\varphi = S_\varphi$  then  $\neg\varphi$  is strongly compositional, and  $\pi \models^* \neg\varphi$  if and only if either  $\pi \models \neg\varphi$  or  $V_\pi \cap S_\varphi = \emptyset$ .*

## 4.2 Checking Consistency for Subsets of $\mathcal{L}'_{pqT}$

Theorem 2 suggests the feasibility of checking consistency of subsets of the language  $\mathcal{L}'_{pqT}$ , which is  $\mathcal{L}_{pqT}$  with certain negated statements also included. Formally, define  $\mathcal{L}'_{pqT}$  to be the union of  $\mathcal{L}_{pqT}$  with  $\{\neg\varphi : \varphi \in \mathcal{L}_{pqT}, \varphi \text{ non-strict, and } R_\varphi = S_\varphi\}$ .

We use the method of Section 3.3 to determine the consistency of a set of preference statements  $\Gamma \subseteq \mathcal{L}'_{pqT}$ , by incrementally extending a maximal  $\models^*$ -model  $\pi$  of  $\Gamma$ , and then checking whether or not  $\pi \models \Gamma$  holds; this makes use of Propositions 2 and 3.

Let  $\Gamma \subseteq \mathcal{L}'_{pqT}$ , let  $X \in V$  and let  $\pi \in \mathcal{G}$ . We make the following definitions, where  $\bar{\Gamma}$  is the set of all  $\varphi \in \Gamma \cap \mathcal{L}_{pqT}$  such that  $R_\varphi \cap S_\varphi \cap V_\pi = \emptyset$ .

- $\text{Best}_\Gamma^\pi(X) = \{r_\varphi(X) : \varphi \in \bar{\Gamma} \ \& \ X \in R_\varphi \setminus S_\varphi\}$ .
- $\text{Worst}_\Gamma^\pi(X) = \{s_\varphi(X) : \varphi \in \bar{\Gamma} \ \& \ X \in S_\varphi \setminus R_\varphi\}$ .
- $\text{Pairs}_\Gamma^\pi(X) = \text{Pos}_\Gamma^\pi(X) \cup \text{Neg}_\Gamma^\pi(X)$ , where

$\text{Pos}_\Gamma^\pi(X) = \{(r_\varphi(X), s_\varphi(X)) : \varphi \in \bar{\Gamma} \ \& \ X \in R_\varphi \cap S_\varphi\}$ ; and  $\text{Neg}_\Gamma^\pi(X)$  is the set of all pairs  $(s_\varphi(X), r_\varphi(X))$  such that  $\neg\varphi \in \Gamma$  and  $T_\varphi \cup U_\varphi \supseteq V_\pi$ , and  $X \in R_\varphi (= S_\varphi)$ .

Given  $\Gamma$  and  $\pi \in \mathcal{G}$  with  $\pi \models^* \Gamma$ , we say that  $X$  can be chosen next if:  $X \in V - V_\pi$  and

- if  $\varphi \in \Gamma \cap \mathcal{L}_{pqT}$  and  $R_\varphi \cap S_\varphi \cap V_\pi = \emptyset$  then  $X \notin W_\varphi$ ;
- $\text{Pairs}_\Gamma^\pi(X)$  is acyclic;
- $|\text{Best}_\Gamma^\pi(X)| \leq 1$  and  $|\text{Worst}_\Gamma^\pi(X)| \leq 1$ ;
- if  $x \in \text{Best}_\Gamma^\pi(X)$  then  $x$  is undominated in  $\text{Pairs}_\Gamma^\pi(X)$ , i.e., there exists no element of the form  $(x', x)$  in  $\text{Pairs}_\Gamma^\pi(X)$ ;
- if  $x \in \text{Worst}_\Gamma^\pi(X)$  then  $x$  is not dominating in  $\text{Pairs}_\Gamma^\pi(X)$ , i.e., there exists no element of the form  $(x, x')$  in  $\text{Pairs}_\Gamma^\pi(X)$ .

We also say that  $(X, \geq_X)$  is a valid extension of  $\pi$  if (i)  $X$  can be chosen next; and (ii)  $\geq_X \supseteq \text{Pairs}_\Gamma^\pi(X)$ ; (iii) if  $x \in \text{Best}_\Gamma^\pi(X)$  then  $x$  is the best element in  $\underline{X}$  with respect to  $\geq_X$

(so that  $x \geq y$  for all  $y \in \underline{X}$ ); and (iv) if  $x' \in \text{Worst}_\Gamma^\pi(X)$  then  $x'$  is the worst element in  $\underline{X}$  with respect to  $\geq_X$ .

The following result states the conditions needed for minimally extending  $\pi$  to maintain the  $\models^*$ -satisfaction of  $\Gamma$ .

**Proposition 4** *Suppose that  $\Gamma \subseteq \mathcal{L}'_{pqT}$ , and that  $\pi \models^* \Gamma$ . Let  $X$  be a variable in  $V - V_\pi$  and let  $\pi' = \pi \circ (X, \geq_X)$ , where  $\geq_X$  is a total ordering on  $\underline{X}$ . Then  $\pi' \models^* \Gamma$  if and only if  $(X, \geq_X)$  is a valid extension of  $\pi$ .*

In summary, at each stage of the incremental algorithm we see if there is a variable  $X$  that can be chosen next, and if so, generate a valid extension; if not, we then have generated a maximal  $\models^*$ -model  $\pi$  of  $\Gamma$ . We check consistency of  $\Gamma$  by determining if  $\pi$  satisfies  $\Gamma$ . Using the fact that  $|\text{Pairs}_\Gamma^\pi(X)| \leq |\Gamma|$ , it can be shown that the overall complexity of checking that  $\Gamma \subseteq \mathcal{L}'_{pqT}$  is consistent is  $\mathcal{O}(|V|^2|\Gamma|)$ .

## 5 Optimality with Respect to Sets of Compositional Preference Statements

We consider a finite set of alternatives  $\mathcal{A} \subseteq \underline{V}$ , and we assume that we have elicited a set  $\Gamma$  of preference statements; we would like to find the optimal alternatives among  $\mathcal{A}$ . As we will see, there are several natural definitions of *optimal* [Gelain *et al.*, 2010; Wilson and O'Mahony, 2011]. We compare some of these and analyse their computational cost in the context of lex models and compositional statements.

We define pre-order relation  $\succsim_\Gamma$  on outcomes by  $\alpha \succsim_\Gamma \beta \iff \Gamma \models \alpha \geq \beta$ , and we define equivalence relation  $\equiv_\Gamma$  by  $\alpha \equiv_\Gamma \beta \iff \Gamma \models \alpha \equiv \beta$ , i.e., if  $\alpha$  and  $\beta$  are equivalent in all models of  $\Gamma$ . Using a similar argument to that for Theorem 1, it can be shown that, for compositional  $\Gamma$ , we have  $\Gamma \models \alpha \equiv \beta$  holds if and only if  $\alpha \equiv_\pi \beta$ , where  $\pi$  is an arbitrary maximal model of  $\Gamma$ . We also define  $\succ_\Gamma$  to be the strict part of  $\succsim_\Gamma$ , so that  $\alpha \succ_\Gamma \beta$  if and only if  $\alpha \succsim_\Gamma \beta$  and  $\alpha \not\equiv_\Gamma \beta$ . We then say that  $\alpha$  strictly dominates  $\beta$ .

We define  $\text{CSD}_\Gamma(\mathcal{A})$  ('Can Strictly Dominate') to be the set of maximal, i.e., undominated, elements of  $\mathcal{A}$  w.r.t.  $\succ_\Gamma$ .  $\alpha \in \text{CSD}_\Gamma(\mathcal{A})$  if and only if for all  $\beta \in \mathcal{A}$  which are not  $\equiv_\Gamma$ -equivalent to  $\alpha$  there exists some  $\pi \models \Gamma$  with  $\alpha \succ_\pi \beta$ .

We define  $\text{O}_\pi(\mathcal{A})$  to be the subset of the alternatives that are optimal in model  $\pi$ , i.e.,  $\{\alpha \in \mathcal{A} : \forall \beta \in \mathcal{A}, \alpha \not\succ_\pi \beta\}$ . We say that  $\alpha \in \mathcal{A}$  is necessarily optimal in  $\mathcal{A}$ , written  $\alpha \in \text{NO}_\Gamma(\mathcal{A})$ , if  $\alpha$  is optimal in every model, i.e., if for all  $\pi \in \mathcal{G}$  with  $\pi \models \Gamma$  we have  $\alpha \in \text{O}_\pi(\mathcal{A})$ .

We say that  $\alpha$  is possibly optimal, written  $\alpha \in \text{PO}_\Gamma(\mathcal{A})$ , if  $\alpha$  is optimal in some model of  $\Gamma$ , so that  $\text{PO}_\Gamma(\mathcal{A}) = \bigcup_{\pi \models \Gamma} \text{O}_\pi(\mathcal{A})$ . Similarly, we say that  $\alpha \in \text{POM}_\Gamma(\mathcal{A})$  if  $\alpha$  is optimal in some maximal model of  $\Gamma$ .  $\alpha$  is possibly strictly optimal in  $\mathcal{A}$ , written  $\alpha \in \text{PSO}_\Gamma(\mathcal{A})$ , if there exists some  $\pi \models \Gamma$  with  $\text{O}_\pi(\mathcal{A}) \ni \alpha$  and  $\Gamma \models \alpha \equiv \beta$  for all  $\beta \in \text{O}_\pi(\mathcal{A})$ . Thus  $\alpha$  is in  $\text{PSO}_\Gamma(\mathcal{A})$  if there is a model of  $\Gamma$  in which  $\alpha$  is optimal, and all other optimal elements are equivalent to  $\alpha$ .

Let  $\text{Opt}_\Gamma^A(\alpha)$  be the set of models  $\pi$  of  $\Gamma$  that make  $\alpha$  optimal in  $\mathcal{A}$ , i.e.,  $\{\pi \models \Gamma : \text{O}_\pi(\mathcal{A}) \ni \alpha\}$ . We define  $\alpha \in \text{MPO}_\Gamma(\mathcal{A})$  if  $\text{Opt}_\Gamma^A(\alpha)$  is maximal, in the sense that there exists no  $\beta \in \mathcal{A}$  with  $\text{Opt}_\Gamma^A(\beta)$  a strict superset of  $\text{Opt}_\Gamma^A(\alpha)$ . We say that  $\alpha$  is maximally possibly optimal in  $\mathcal{A}$  given  $\Gamma$ ; this

holds if and only if there is no alternative that is optimal in the same set of lex models and more.

Let  $\pi_1, \dots, \pi_k$  be a finite sequence of models. Define  $\mathcal{A}_{\pi_1}$  to be  $O_{\pi_1}(\mathcal{A})$ . For  $i = 1, \dots, k$  we iteratively define  $\mathcal{A}_{\pi_1, \dots, \pi_i}$  to be  $O_{\pi_i}(\mathcal{A}_{\pi_1, \dots, \pi_{i-1}})$ . We define the *extreme elements*  $\text{EXT}_{\Gamma}(\mathcal{A})$  as follows.  $\alpha \in \text{EXT}_{\Gamma}(\mathcal{A})$  if and only if there exists a sequence  $\pi_1, \dots, \pi_k$  of models of  $\Gamma$  such that  $\mathcal{A}_{\pi_1, \dots, \pi_k} \ni \alpha$  and for all  $\beta \in \mathcal{A}_{\pi_1, \dots, \pi_k}$ ,  $\Gamma \models \alpha \equiv \beta$ . Therefore,  $\alpha \in \text{EXT}_{\Gamma}(\mathcal{A})$  if there is a sequence of lex models such that iteratively maximising with respect to each lex model in turn leads to a set containing  $\alpha$  and only other alternatives that are equivalent to  $\alpha$ .

**Proposition 5** *For any  $\mathcal{A} \subseteq \underline{V}$  and  $\Gamma \subseteq \mathcal{L}$  we have  $\text{PSO}_{\Gamma}(\mathcal{A}) \subseteq \text{POM}_{\Gamma}(\mathcal{A}) \cap \text{MPO}_{\Gamma}(\mathcal{A}) \cap \text{EXT}_{\Gamma}(\mathcal{A})$ . If  $\Gamma$  is compositional then  $\text{PSO}_{\Gamma}(\mathcal{A}) = \text{POM}_{\Gamma}(\mathcal{A}) = \text{MPO}_{\Gamma}(\mathcal{A}) = \text{EXT}_{\Gamma}(\mathcal{A})$ .*

Without making assumptions about  $\Gamma$  we have the following properties, which follow from basic arguments, that apply in a very general context [Wilson and O'Mahony, 2011].

**Proposition 6** *Consider any  $\mathcal{A} \subseteq \underline{V}$  and  $\Gamma \subseteq \mathcal{L}$ . Then, the following all hold. (i)  $\text{NO}_{\Gamma}(\mathcal{A}) \cup \text{PSO}_{\Gamma}(\mathcal{A}) \subseteq \text{MPO}_{\Gamma}(\mathcal{A}) \cap \text{EXT}_{\Gamma}(\mathcal{A})$ ; (ii)  $\text{EXT}_{\Gamma}(\mathcal{A}) \subseteq \text{CSD}_{\Gamma}(\mathcal{A}) \cap \text{PO}_{\Gamma}(\mathcal{A})$ ; (iii)  $\text{MPO}_{\Gamma}(\mathcal{A}) \subseteq \text{PO}_{\Gamma}(\mathcal{A})$ ; (iv)  $\text{MPO}_{\Gamma}(\mathcal{A}) \cap \text{EXT}_{\Gamma}(\mathcal{A})$  is always non-empty. (v) If  $\text{NO}_{\Gamma}(\mathcal{A})$  is non-empty then  $\text{NO}_{\Gamma}(\mathcal{A}) = \text{MPO}_{\Gamma}(\mathcal{A}) = \text{EXT}_{\Gamma}(\mathcal{A}) = \text{CSD}_{\Gamma}(\mathcal{A})$ .*

Propositions 5 and 6 imply the following result, showing that there are substantial simplifications of the optimality classes when  $\Gamma$  is compositional.

**Theorem 3** *Consider any  $\mathcal{A} \subseteq \underline{V}$  and compositional  $\Gamma \subseteq \mathcal{L}$ . Then  $\text{NO}_{\Gamma}(\mathcal{A}) \subseteq \text{PSO}_{\Gamma}(\mathcal{A}) = \text{EXT}_{\Gamma}(\mathcal{A}) = \text{MPO}_{\Gamma}(\mathcal{A}) = \text{POM}_{\Gamma}(\mathcal{A}) \subseteq \text{CSD}_{\Gamma}(\mathcal{A}) \cap \text{PO}_{\Gamma}(\mathcal{A})$ .*

**Computational Results:** One approach to compute  $\Omega(\mathcal{A})$  for  $\Omega \in \{\text{PO}_{\Gamma}, \text{PSO}_{\Gamma}, \text{CSD}_{\Gamma}, \text{NO}_{\Gamma}\}$  is to test membership for every alternative. By the results in Section 4.2, we can decide consistency for  $g$  statements in  $\mathcal{L}'_{pg\Gamma}$  and  $n$  variables in  $\mathcal{O}(n^2g)$ . To test if  $\alpha \in \text{PO}_{\Gamma}(\mathcal{A})$ , we test whether  $\Gamma \cup \{\alpha \geq \beta \mid \beta \in \mathcal{A} - \{\alpha\}\}$  is consistent in  $\mathcal{O}(n^2(g+m))$ , where  $|\mathcal{A}| = m$ . Similarly, we test if  $\alpha \in \text{PSO}_{\Gamma}(\mathcal{A})$  in  $\mathcal{O}(n^2(g+m))$ , by checking if  $\Gamma \cup \{\alpha > \beta \mid \beta \in \mathcal{A}, \beta \not\equiv_{\Gamma} \alpha\}$  is consistent. To test if  $\alpha \in \text{CSD}_{\Gamma}(\mathcal{A})$ , we check for all  $\beta \in \mathcal{A}$  with  $\beta \not\equiv_{\Gamma} \alpha$  if  $\Gamma \cup \{\alpha > \beta\}$  is consistent in  $\mathcal{O}(mn^2g)$ . To test if  $\alpha \in \text{NO}_{\Gamma}(\mathcal{A})$ , we test for all  $\beta \in \mathcal{A} - \{\alpha\}$  if  $\Gamma \cup \{\alpha < \beta\}$  is consistent in  $\mathcal{O}(mn^2g)$ . Optimality sets can also be computed incrementally [Wilson *et al.*, 2015b].

We implemented these methods for computing  $\text{PO}_{\Gamma}, \text{PSO}_{\Gamma}, \text{CSD}_{\Gamma}$  and  $\text{NO}_{\Gamma}$  and randomly generated 10 consistent instances for each configuration of  $g \in \{100, \dots, 1000\}$  (equally divided into fully strict, weakly strict, non-strict and negated non-strict) and  $n \in \{100, \dots, 1000\}$  variables with random domain sizes 2 or 3. We tested the methods on every instance for  $m = 100, \dots, 500$  alternatives. All experiments result in  $\text{NO}_{\Gamma}(\mathcal{A}) \subseteq \text{PO}_{\Gamma}(\mathcal{A}) = \text{PSO}_{\Gamma}(\mathcal{A}) \subseteq \text{CSD}_{\Gamma}(\mathcal{A})$  which extends the relations from Theorem 3. In our experiments, CSD is computed much slower than PO, PSO and NO. Also, often PO is computed faster than NO. In practice, it

is often desired to compute relatively small sets of optimal solutions, e.g., not to overwhelm a user with too many choices. Since  $\text{NO}_{\Gamma}$  is usually empty in our experiments, computing  $\text{PSO}_{\Gamma}$  or  $\text{PO}_{\Gamma}$  may be the best choice because of its size and computational cost.

## 6 Discussion

We have developed a new approach for preference inference based on lexicographic models, using a notion of strong compositionality that allows a greedy algorithm. This is shown to be polynomial for relatively expressive preference languages. We also examined different notions of optimality, and proved relationships between them. Our experimental results confirm that the preference inference/consistency algorithm is fast, and examined the problem of generating optimal solutions, where it was found that generating the sets of possibly optimal and possibly strictly optimal solutions were significantly faster than generating the undominated solutions.

There are other common forms of preference statement that are strongly compositional, and for which the greedy algorithm will enable checking consistency. For instance, a restriction on the allowed value orderings of each variable is always strongly compositional. This can include, for instance, highly disjunctive statements, for instance, structural properties of the value orderings, such as being single-peaked [Conitzer, 2009]. Certain kinds of restrictions on variable orderings are also strongly compositional.

A lexicographic order can also be viewed as a static variable and value search ordering, for a constraint satisfaction or optimisation problem. In a configuration problem, for instance, we may want to generate solutions in an ordering which fits in with what one knows about user preferences and the structure of the problem [Junker and Mailharro, 2003]. We can thus take, as input, past user preferences between solutions, or even partial tuples; or that all solutions of one constraint are preferred to all those satisfying another constraint. There may be natural restrictions on the set of value orderings, e.g., that there is a single peak or trough with a numerical domain. Given such inputs, which are strongly compositional, one can apply our approach to see if there is a compatible lexicographic order, which can then be used to generate solutions in a best first order.

Our approach in Section 3 was rather abstract, in that no explicit structure on the language  $\mathcal{L}$  was assumed. In fact, a closer examination of the proofs of some of the main results suggests that these will hold more generally than for lexicographic orders: essentially they just depend on basic properties of composition and extension, but not otherwise on the structure of models. In particular, it is natural to look at the generalisation to conditional lexicographic orders, which are closely related to conditional preference languages [Wilson, 2009]. It would also be interesting to explore the potential connection with efficient classes of planning problems.

## Acknowledgments

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number SFI/12/RC/2289.

## References

- [Agrawal and Wimmers, 2000] Rakesh Agrawal and Edward L. Wimmers. A framework for expressing and combining preferences. In *Proc. ACM SIGMOD 2000*, pages 297–306, New York, NY, USA, 2000. ACM.
- [Arrow and Raynaud, 1986] Kenneth J. Arrow and Hervé Raynaud. *Social Choice and Multicriterion Decision-Making*, volume 1. The MIT Press, 1 edition, 1986.
- [Booth *et al.*, 2010] Richard Booth, Yann Chevaleyre, Jérôme Lang, Jérôme Mengin, and Chattrakul Sombatheera. Learning conditionally lexicographic preference relations. In *Proc. ECAI 2010*, pages 269–274, 2010.
- [Boutilier *et al.*, 2004] Craig Boutilier, Ronen I. Brafman, Carmel Domshlak, Holger H. Hoos, and David Poole. CP-nets: A tool for reasoning with conditional *ceteris paribus* preference statements. *Journal of Artificial Intelligence Research*, 21:135–191, 2004.
- [Bräuning and Hüllermeier, 2012] Michael Bräuning and Eyke Hüllermeier. Learning conditional lexicographic preference trees. In *Preference Learning (PL-12), ECAI 2012 workshop*, pages 11–15, 2012.
- [Chen and Pu, 2007] Li Chen and Pearl Pu. *Preference-Based Organization Interfaces: Aiding User Critiques in Recommender Systems*, pages 77–86. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.
- [Conitzer, 2009] Vincent Conitzer. Eliciting single-peaked preferences using comparison queries. *Journal of Artificial Intelligence Research*, 35:161–191, 2009.
- [Dombi *et al.*, 2007] József Dombi, Csanád Imreh, and Nándor Vincze. *European Journal of Operational Research*, 183(2):748–756, 2007.
- [Gelain *et al.*, 2010] Mirco Gelain, Maria Silvia Pini, Francesca Rossi, Kristen Brent Venable, and Nic Wilson. Interval-valued soft constraint problems. *Ann. Math. Artif. Intell.*, 58(3-4):261–298, 2010.
- [George and Wilson, 2016] Anne-Marie George and Nic Wilson. Preference inference based on Pareto models. In *Proc. SUM 2016*, pages 170–183, 2016.
- [Junker and Mailharro, 2003] Ulrich Junker and Daniel Mailharro. Preference programming: Advanced problem solving for configuration. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 17(1):13–29, 2003.
- [Kießling, 2002] Werner Kießling. Foundations of preferences in database systems. In *Proc. VLDB 2002*, pages 311–322. VLDB Endowment, 2002.
- [Kohli and Jedidi, 2007] Rajeev Kohli and Kamel Jedidi. Representation and inference of lexicographic preference models and their variants. *Marketing Science*, 26(3):380–399, 2007.
- [Lang, 2002] Jérôme Lang. From preference representation to combinatorial vote. In *Proc. KR 2002*, pages 277–288, 2002.
- [Montazery and Wilson, 2016] Mojtaba Montazery and Nic Wilson. Learning user preferences in matching for ridesharing. In *Proc. ICAART 2016*, pages 63–73, 2016.
- [Sandholm and Boutilier, 2006] Tuomas Sandholm and Craig Boutilier. Preference elicitation in combinatorial auctions. In *Combinatorial Auctions*, ed. P. Cramton, Y. Shoham, and R. Steinberg, Chapter 10, pages 233–264. Cambridge, MA: MIT Press, 2006.
- [Trabelsi *et al.*, 2013] Walid Trabelsi, Nic Wilson, and Derek G. Bridge. Comparative preferences induction methods for conversational recommenders. In *Proc. ADT 2013*, pages 363–374, 2013.
- [Wilson and George, 2017] Nic Wilson and Anne-Marie George. *Efficient Inference and Computation of Optimal Alternatives for Preference Languages Based On Lexicographic Models (extended version of IJCAI’17 paper including proofs)*. <http://ucc.insight-centre.org/nwilson/LexInfOptLonger.pdf>, 2017.
- [Wilson and O’Mahony, 2011] Nic Wilson and Conor O’Mahony. The relationships between qualitative notions of optimality for decision making under logical uncertainty. In *Proc. AICS 2011*, pages 66–75, 2011.
- [Wilson *et al.*, 2015a] Nic Wilson, Anne-Marie George, and Barry O’Sullivan. Computation and complexity of preference inference based on hierarchical models. In *Proc. IJCAI 2015*, pages 3271–3277, 2015.
- [Wilson *et al.*, 2015b] Nic Wilson, Abdul Razak, and Radu Marinescu. Computing possibly optimal solutions for multi-objective constraint optimisation with tradeoffs. In *Proc. IJCAI 2015*, pages 815–821. AAAI Press, 2015.
- [Wilson, 2009] Nic Wilson. Efficient inference for expressive comparative preference languages. In *Proc. IJCAI-09*, pages 961–966, 2009.
- [Wilson, 2014] Nic Wilson. Preference inference based on lexicographic models. In *Proc. ECAI 2014*, pages 921–926, 2014.
- [Yaman *et al.*, 2010] Fusun Yaman, Thomas J. Walsh, Michael L. Littman, and Marie desJardins. Learning lexicographic preference models. In Johannes Fürnkranz and Eyke Hüllermeier, editors, *Preference Learning*, pages 251–272. Springer, 2010.