SAT-Based Approaches to Reasoning in Choice Logics

Tuomo Lehtonen[1,*], Andreas Niskanen[2] and Matti Järvisalo[2]

[1]Aalto University, Finland
[2]University of Helsinki, Finland
ORCID (Tuomo Lehtonen): https://orcid.org/0000-0001-6117-4854, ORCID (Andreas Niskanen): https://orcid.org/0000-0003-3197-2075, ORCID (Matti Järvisalo): https://orcid.org/0000-0003-2572-063X

Abstract. Representing and reasoning about preferences is a fundamental task in artificial intelligence. Various logic-based languages for representing preferences have been proposed. However, developing practical algorithms for reasoning in such logic-based languages remains a challenge due to high computational complexity. In this work, we develop practical algorithms based on Boolean satisfiability (SAT) for computing preferred models and for deciding preferred model entailment in qualitative and conjunctive choice logics QCL and CCL under the so-called minmax, lexicographic, and inclusion-based preference semantics. For each of the problem variants, we detail an algorithm which adheres to the computational complexity of the reasoning task, based on either maximum satisfiability (MaxSAT) or SAT with preferences (PrefSAT) solvers. We empirically evaluate our implementation of the algorithms, and show that our approach scales significantly better than a recently proposed answer set programming approach to computing preferred models.

1 Introduction

Preferences are intrinsic to human endeavours and decision making in various real-world settings. As such, representing and reasoning about preferences is a fundamental task in artificial intelligence [20, 31]. Various logic-based languages for representing preferences have been proposed. These include qualitative [14] and conjunctive choice logics [13] QCL and CCL, motivated by various application settings [5, 28, 34]. Extending propositional logic with specific choice connectives, the semantics of QCL and CCL are based on preferred models, defined via the notion of the satisfaction degree of a formula, as opposed to treating all models of a propositional formula equally.

Integration of preferences often has a significant impact on the complexity of reasoning, as is the case for QCL and CCL [9]. While deciding satisfiability of a given propositional formula (i.e., the SAT problem [12]) is NP-complete, reasoning problems concerned with preferred models in QCL and CCL are even harder. Analogously, while deciding propositional entailment is "merely" coNP-complete, its natural generalization, preferred model entailment—defined over choice logic theories composed of multiple choice logic formulas—is significantly harder [9]. The exact complexity of preferred model entailment depends on the choice of preferred model semantics, prescribing what are preferred models. Examples studied in the literature include minmax semantics [9] (minimizing the maximum satisfaction degree over formulas in a choice logic theory), lexicographic semantics [14] (prescribing a notion of lexicographically preferred models based on the number of formulas satisfied to as low a degree as possible) and inclusion-based semantics [14] (based on set inclusion of formulas satisfied to as low a degree as possible). Preferred model entailment has been shown to be Θ_2^p-complete under minmax semantics, Δ_2^p-complete under lexicographic semantics, and Π_2^p-complete under inclusion-based semantics [9], in each case surpassing the complexity of classical propositional entailment.

While QCL and CCL have been studied from various more theoretical perspectives (see, e.g., [7, 21, 22, 23, 10, 16]), the complexity of computing preferred models and especially preferred model entailment [9] poses significant challenges to developing practical algorithms for reasoning in these logics. Recently an approach to computing preferred models in QCL and CCL based on encoding the problems to answer set programming (ASP) was developed [8]. However, the approach has not been extended to the problem of deciding preferred model entailment, nor has its scalability been thoroughly evaluated, to the best of our knowledge.

In this work, we develop novel approaches to computing preferred models in QCCL (the combination of QCL and CCL) [9] and for preferred model entailment in QCCL covering each of minmax, lexicographic, and inclusion-based preference semantics. We base our approach on Boolean satisfiability (SAT) solvers [12] and their extensions. A SAT-based approach is motivated both by the significant success of SAT solvers as a key technology for efficiently solving NP-hard problems, and by the fact that QCL and CCL extend propositional formulas and their semantics (through the introduction of non-classical choice connectives). For each of the reasoning problems we detail an algorithm which adheres to the computational complexity of the problem: (i) a direct maximum satisfiability (MaxSAT) [3] encoding which captures preferred models in QCCL, allowing for computing preferred models using state-of-the-art MaxSAT solvers, as well as e.g. equivalence checking of QCCL formulas; and approaches to preferred model entailment based on (ii) a combination of one MaxSAT solver and one SAT solver call for minmax semantics, (iii) an iterative MaxSAT-based lexicographic optimization approach [1, 29] for lexicographic semantics, and (iv) a counterexample-guided abstraction refinement [17, 18] style iterative approach based on a SAT with preferences (PrefSAT) [33, 32, 19] solver for inclusion-based semantics. We provide an open-source implementation of the algorithms and empirically evaluate our approach. The evaluation shows promising scalability, and our approach significantly outperforms the recent ASP-based approach where applicable, i.e., for computing preferred models.

* Corresponding Author. Email: tuomo.lehtonen@aalto.fi

2 Choice logics

We focus on QCL [14] and CCL [13] as prominent choice logic instantiations [9]. The algorithms presented in this work cover QCCL [9] as a combination of QCL and CCL, and hence are applicable individually to both QCL and CCL as well.

Let $\mathcal{U}$ be a (countably infinite) set of propositional variables. A choice logic $\mathcal{L}$ extends classical propositional formulas by including (in addition to classical connectives $\neg$, $\wedge$, $\vee$) a set $\mathcal{C}_\mathcal{L}$ of *choice connectives* which enable expressing preferences among a set of alternatives. Different choices of $\mathcal{C}_\mathcal{L}$ for choice connectives give rise to different choice logics. For a specific choice logic, the set of formulas $\mathcal{F}_\mathcal{L}$ is defined recursively by

- $a \in \mathcal{F}_\mathcal{L}$ for all $a \in \mathcal{U}$,
- if $F \in \mathcal{F}_\mathcal{L}$, then $\neg F \in \mathcal{F}_\mathcal{L}$,
- if $F, G \in \mathcal{F}_\mathcal{L}$, then $F \circ G \in \mathcal{F}_\mathcal{L}$ for any $\circ \in \{\wedge, \vee\} \cup \mathcal{C}_\mathcal{L}$.

We denote by $\mathrm{Vars}(F) \subseteq \mathcal{U}$ the set of variables occurring in $F \in \mathcal{F}_\mathcal{L}$. An *interpretation* for $F \in \mathcal{F}_\mathcal{L}$ is a function $\mathcal{I}\colon \mathrm{Vars}(F) \to \{0,1\}$. For convenience, we sometimes identify an interpretation with the set of variables that the interpretation maps to 1 (true).

Central to defining semantics for specific choice logics are the notions of *satisfaction degree* and *optionality*. Generally, for a given formula $F \in \mathcal{F}_\mathcal{L}$ and an interpretation $\mathcal{I}$ for F, the satisfaction degree of F under $\mathcal{I}$ is either a positive integer (if $\mathcal{I}$ satisfies F in the classical sense), or infinite (if $\mathcal{I}$ does not satisfy F). The lower the satisfaction degree, the more preferred $\mathcal{I}$ is.

Qualitative choice logic (QCL) [14] extends classical propositional formulas with the connective $\vec{\times}$ called *ordered disjunction*, i.e. $\mathcal{C}_{\text{QCL}} = \{\vec{\times}\}$. *Conjunctive choice logic* (CCL) [13] (we follow the definition of [9]) on the other hand extends propositional formulas with the connective $\vec{\odot}$ called *ordered conjunction*, i.e., $\mathcal{C}_{\text{CCL}} = \{\vec{\odot}\}$. Intuitively, $F\vec{\times}G$ states that at least F or G must hold and it is most preferred for F to hold, while $F\vec{\odot}G$ requires F to hold and prefers G to hold as well. QCCL [9] combines QCL and CCL directly by incorporating both connectives, i.e. $\mathcal{C}_{\text{QCCL}} = \{\vec{\times}, \vec{\odot}\}$. The semantics of QCCL are defined through the following definitions of optionality and satisfaction degrees. The optionality of a formula is the maximum finite satisfaction degree to which it can be satisfied by any interpretation.

Definition 1. *The optionality of a formula $F \in \mathcal{F}_{QCCL}$ is given by a function $opt(F) \mapsto \mathbb{Z}_+$ defined recursively as*

- *$opt(a) = 1$ for all $a \in \mathcal{U}$,*
- *$opt(\neg F) = 1$ for all $F \in \mathcal{F}_{QCCL}$,*
- *$opt(F \wedge G) = \max(opt(F), opt(G))$ for all $F, G \in \mathcal{F}_{QCCL}$,*
- *$opt(F \vee G) = \max(opt(F), opt(G))$ for all $F, G \in \mathcal{F}_{QCCL}$,*
- *$opt(F\vec{\times}G) = opt(F) + opt(G)$ for all $F, G \in \mathcal{F}_{QCCL}$,*
- *$opt(F\vec{\odot}G) = opt(F) + opt(G)$ for all $F, G \in \mathcal{F}_{QCCL}$.*

In other words, the optionality of a variable and a negated formula is exactly one. For the classical conjunction and disjunction, the optionality of a formula is the maximum of the optionalities of its subformulas. For ordered conjunction and disjunction, however, the optionality of a formula is the sum of the optionalities of its subformulas. This is because a satisfied disjunction or conjunction has a degree of at most the maximum of the degrees of its members, while ordered conjunction and disjunction can have a degree up to the sum of its optionalities, as is formalized next.

Definition 2. *The satisfaction degree of a formula $F \in \mathcal{F}_{QCCL}$ under interpretation $\mathcal{I}$ is given by a function $deg(\mathcal{I}, F) \mapsto \mathbb{Z}_+ \cup \{\infty\}$ defined recursively as follows. For subformulas $F, G \in \mathcal{F}_{QCCL}$ and an interpretation $\mathcal{I}$, we denote $o_F = opt(F)$, $o_G = opt(G)$, $d_F = deg(\mathcal{I}, F)$, and $d_G = deg(\mathcal{I}, G)$. Now, for $a \in \mathcal{U}$ and $F, G \in \mathcal{F}_{QCCL}$,*

- $deg(\mathcal{I}, a) = \begin{cases} 1 & \text{if } \mathcal{I}(a) = 1, \\ \infty & \text{otherwise;} \end{cases}$
- $deg(\mathcal{I}, \neg F) = \begin{cases} 1 & \text{if } d_F = \infty, \\ \infty & \text{otherwise;} \end{cases}$
- $deg(\mathcal{I}, F \wedge G) = \max(d_F, d_G)$;
- $deg(\mathcal{I}, F \vee G) = \min(d_F, d_G)$;
- $deg(\mathcal{I}, F\vec{\times}G) = \begin{cases} d_F & \text{if } d_F < \infty, \\ d_G + o_F & \text{if } d_F = \infty \text{ and } d_G < \infty, \\ \infty & \text{otherwise;} \end{cases}$
- $deg(\mathcal{I}, F\vec{\odot}G) = \begin{cases} d_G & \text{if } d_F = 1 \text{ and } d_G < \infty, \\ d_F + o_G & \text{if } d_F < \infty \text{ and} \\ & (d_F > 1 \text{ or } d_G = \infty), \\ \infty & \text{otherwise.} \end{cases}$

If $\deg(\mathcal{I}, F) = k$ with $k < \infty$, we say that $\mathcal{I}$ satisfies F to degree k, and $\mathcal{I}$ is a model of F. If $\deg(\mathcal{I}, F) = \infty$, then $\mathcal{I}$ does not satisfy F. We denote the set of all models of F by $\mathrm{Mod}(F)$. Note that if $\mathcal{I} \in \mathrm{Mod}(F)$, then $\deg(\mathcal{I}, F) \leq \mathrm{opt}(F)$.

For intuition on ordered disjunction, the satisfaction degree of $F\vec{\times}G$ under interpretation $\mathcal{I}$ is finite if at least one of F and G is satisfied by $\mathcal{I}$ (to a finite degree). In this case, the degree is $\deg(\mathcal{I}, F)$ if $\mathcal{I} \in \mathrm{Mod}(F)$, and otherwise (i.e., when $\mathcal{I} \in \mathrm{Mod}(G)$) the degree is $\deg(\mathcal{I}, G) + \mathrm{opt}(F)$. In other words, not satisfying F is penalized by the optionality of F in addition to the degree of G.

Considering ordered conjunction, the satisfaction degree of $F\vec{\odot}G$ under interpretation $\mathcal{I}$ is finite if F is satisfied by $\mathcal{I}$ (to a finite degree). In this case, the degree is $\deg(\mathcal{I}, G)$ if $\deg(\mathcal{I}, F) = 1$ and $\mathcal{I} \in \mathrm{Mod}(G)$, and otherwise (i.e., when $\deg(\mathcal{I}, F) > 1$ or $\mathcal{I} \notin \mathrm{Mod}(G)$) the degree is $\deg(\mathcal{I}, F) + \mathrm{opt}(G)$. In other words, both not satisfying F to the smallest degree and not satisfying G at all are penalized by an additive term given by the optionality of G.

Finally, a model $\mathcal{I} \in \mathrm{Mod}(F)$ is a *preferred model* of F if its satisfaction degree is the smallest possible over all interpretations, that is, for all interpretations $\mathcal{I}'$ of F it holds that $\deg(\mathcal{I}, F) \leq \deg(\mathcal{I}', F)$. We denote the set of all preferred models of F by $\mathrm{Prf}(F)$.

Example 1. *Consider the QCCL formulas $F_1 = (a\vec{\times}b)$, $F_2 = (a\vec{\odot}b)$, $F_1 \wedge F_2$ and $F_1 \vee F_2$. The satisfaction degree of F_1 under the interpretations $\{a, b\}$ and $\{a\}$ is 1, i.e., $deg(\{a, b\}, F_1) = deg(\{a\}, F_1) = 1$. Further, $deg(\{b\}, F_1) = 2$ and $deg(\emptyset, F_1) = \infty$. For F_2, $deg(\{a, b\}, F_2) = 1$, $deg(\{a\}, F_2) = 2$, and $deg(\{b\}, F_2) = deg(\emptyset, F_2) = \infty$. The degree of a conjunction is the maximum of the conjuncts, so $deg(\{a, b\}, F_1 \wedge F_2) = 1$ while $deg(\{a\}, F_1 \wedge F_2) = 2$ and $deg(\{b\}, F_1 \wedge F_2) = deg(\emptyset, F_1 \wedge F_2) = \infty$. For disjunction, the degree is the minimum of the disjuncts: $deg(\{a, b\}, F_1 \vee F_2) = deg(\{a\}, F_1 \vee F_2) = 1$, $deg(\{b\}, F_1 \vee F_2) = 2$, and $deg(\emptyset, F_1 \vee F_2) = \infty$. Thus the preferred models of these formulas are $\mathrm{Prf}(F_1) = \{\{a, b\}, \{a\}\}$, $\mathrm{Prf}(F_2) = \{\{a, b\}\}$, $\mathrm{Prf}(F_1 \wedge F_2) = \{\{a, b\}\}$, and $\mathrm{Prf}(F_1 \vee F_2) = \{\{a, b\}, \{a\}\}$.*

3 Preferred models as MaxSAT

We detail a MaxSAT encoding for QCCL such that preferred models of QCCL formulas correspond to optimal solutions to the MaxSAT encoding. The encoding also forms the basis for our algorithms for preferred model entailment.

For background on SAT and MaxSAT [12, 3], recall that for a Boolean variable x, there are two literals, x and $\neg x$. A clause C is a disjunction ($\vee$) of literals. A conjunctive normal form (CNF) formula F is a conjunction ($\wedge$) of clauses. We may view clauses as sets of literals and formulas as sets of clauses. We denote by $\text{Vars}(F)$ and $\text{Lits}(F)$ the set of variables and literals of F, respectively. A truth assignment $\tau\colon \text{Vars}(F) \to \{0,1\}$ maps each variable to 0 (false) or 1 (true), and extends to literals via $\tau(\neg x) = 1 - \tau(x)$, to clauses via $\tau(C) = \max\{\tau(l) \mid l \in C\}$, and to formulas via $\tau(F) = \min\{\tau(C) \mid C \in F\}$ (in words, a clause is true iff one of its literals is true and a formula is true iff all of its clauses are true). We denote by $\tau[V]$ the restriction of a truth assignment τ to variables V. Given a CNF formula F, the *Boolean satisfiability* problem (SAT) asks whether there is an assignment τ with $\tau(F) = 1$. If there is, F is satisfiable, and otherwise F is unsatisfiable. In the *(partial) maximum satisfiability* problem (MaxSAT in short) [3], the input consists of "hard" clauses F_{hard} and "soft" clauses F_{soft}. The task is to find a truth assignment τ which satisfies F_{hard} and minimizes the cost $c(\tau) = \sum_{C \in F_{\text{soft}}} (1 - \tau(C))$ incurred by not satisfying individual soft clauses.

Towards our MaxSAT encoding for QCCL, we recall the standard "Tseitin" encoding [35] of a given propositional formula F to CNF. Take for each non-atomic subformula S of F a Boolean variable x^S with $\tau(x^S) = 1$ iff S is satisfied by τ, and express locally as clauses the equivalence of x^S and its immediate subformula(s) S_1 (and S_2) depending on the connective of S: $x^S \leftrightarrow \neg x^{S_1}$ for $\neg$, $x^S \leftrightarrow x^{S_1} \vee x^{S_2}$ for $\vee$, and $x^S \leftrightarrow x^{S_1} \wedge x^{S_2}$ for $\wedge$, using the original variables for atomic formulas, to obtain the CNF formula $\text{TSEITIN}(F)$. The CNF formula $\text{TSEITIN}(F) \wedge x^F$ is satisfiable iff F is satisfiable.

Let F be an input QCCL formula. Our goal is to express preferred models of F as optimal solutions of a MaxSAT instance. To be able to reason about satisfaction degrees, we modify the Tseitin encoding to additionally consider the satisfaction degrees of each subformula. For each subformula S of F, in addition to variable x^S indicating whether S is satisfied, for each $k = 1, \dots, \text{opt}(S) + 1$, we declare a Boolean variable d_k^S which is true iff $\deg(\tau, S) \geq k$ (or equivalently, is false iff $\deg(\tau, S) < k$).[1] Intuitively, the set of variables $\{d_k^S \mid k = 1, \dots, \text{opt}(S) + 1\}$ is an order encoding of the satisfaction degree of S (with $d_k^S = 1$ for each $k = 1, \dots, \text{opt}(S) + 1$ interpreted as satisfaction degree ∞). We define $\text{ENCODE}(S)$ as follows.

- If $S = a$ for $a \in \mathcal{U}$, then $\text{ENCODE}(S)$ is
  $(x^S \leftrightarrow a) \wedge (d_1^S) \wedge (d_2^S \leftrightarrow \neg a)$,
  stating that S is satisfied iff a is true, the satisfaction degree is at least one, and greater than one iff a is not true.
- If $S = \neg F$, $\text{ENCODE}(S)$ is
  $(x^S \leftrightarrow \neg x^F) \wedge (d_1^S) \wedge (d_2^S \leftrightarrow x^F)$,
  stating that S is satisfied iff F is falsified, the satisfaction degree is at least one, and greater than one iff F is satisfied.
- If $S = F \wedge G$ (with $\text{opt}(F) \geq \text{opt}(G)$ wlog), $\text{ENCODE}(S)$ is

$$(x^S \leftrightarrow (x^F \wedge x^G)) \wedge \bigwedge_{i=1}^{\text{opt}(G)+1} (d_i^S \leftrightarrow (d_i^F \vee d_i^G)) \wedge \bigwedge_{i=\text{opt}(G)+2}^{\text{opt}(F)+1} (d_i^S \leftrightarrow (\neg x^G \vee d_i^F)).$$

  Here we state that S is satisfied iff F and G are both satisfied, and encode the satisfaction degree by taking the bitwise *maximum* with the following intuition. Up to $\text{opt}(G) + 1$, the satisfaction degree of S is at least i iff the satisfaction degree of either F or G is at least i. Starting from $\text{opt}(G) + 2$, the satisfaction degree of S is at least i iff either G is falsified (i.e., has infinite satisfaction degree) or the satisfaction degree of F is at least i.
- If $S = F \vee G$ (with $\text{opt}(F) \geq \text{opt}(G)$ wlog), $\text{ENCODE}(S)$ is

$$(x^S \leftrightarrow (x^F \vee x^G)) \wedge \bigwedge_{i=1}^{\text{opt}(G)+1} (d_i^S \leftrightarrow (d_i^F \wedge d_i^G)) \wedge \bigwedge_{i=\text{opt}(G)+2}^{\text{opt}(F)+1} (d_i^S \leftrightarrow (\neg x^G \wedge d_i^F)).$$

  Analogously to the above, we state that S is satisfied iff F or G is, and encode the satisfaction degree by taking the bitwise *minimum*: up to $\text{opt}(G) + 1$, the satisfaction degree of S is at least i iff the satisfaction degree of both F and G is at least i; starting from $\text{opt}(G)+2$, the satisfaction degree of S is at least i iff G is falsified and the satisfaction degree of F is at least i.
- If $S = (F \vec{\times} G)$, $\text{ENCODE}(S)$ is

$$(x^S \leftrightarrow (x^F \vee x^G)) \wedge \left(x^F \to \left(\bigwedge_{i=1}^{\text{opt}(F)} (d_i^S \leftrightarrow d_i^F) \wedge \bigwedge_{i=\text{opt}(F)+1}^{\text{opt}(S)+1} \neg d_i^S \right) \right) \wedge \left(\neg x^F \to \left(\bigwedge_{i=1}^{\text{opt}(F)} d_i^S \wedge \bigwedge_{i=1}^{\text{opt}(G)+1} (d_{\text{opt}(F)+i}^S \leftrightarrow d_i^G) \right) \right).$$

  Here we state that S is satisfied iff F or G is, similarly to the previous case. The satisfaction degree of S, following the definition of $\vec{\times}$, is encoded as follows. If F is satisfied, copy the satisfaction degree of F bit by bit. If F is falsified, encode *addition* by setting the first $\text{opt}(F)$ bits to true and the next bits according to the satisfaction degree of G.
- If $S = (F \vec{\odot} G)$, $\text{ENCODE}(S)$ is

$$(x^S \leftrightarrow x^F) \wedge (x^G \wedge \neg d_2^F) \to \left(\bigwedge_{i=1}^{\text{opt}(G)} (d_i^S \leftrightarrow d_i^G) \wedge \bigwedge_{i=\text{opt}(G)+i}^{\text{opt}(S)+1} \neg d_i^S \right) \wedge (\neg x^G \vee d_2^F) \to \left(\bigwedge_{i=1}^{\text{opt}(G)} d_i^S \wedge \bigwedge_{i=1}^{\text{opt}(F)+1} (d_{\text{opt}(G)+i}^S \leftrightarrow d_i^F) \right).$$

  Similarly as above, but now we state that S is satisfied iff F is. The satisfaction degree according to the definition of $\vec{\odot}$ is encoded as follows. If F is satisfied to the smallest degree (i.e., the degree is at most one) and G is satisfied, copy the satisfaction degree of G bit by bit. If F is satisfied to a degree of at least two or G is falsified, add $\text{opt}(G)$ to the satisfaction degree of F.

Now, for a given QCCL formula F, let $\mathcal{S}_F$ be the set of all subformulas of F. We define $\text{QCCLTOCNF}(F) = \bigwedge_{S \in \mathcal{S}_F} \left(\text{ENCODE}(S) \wedge \bigwedge_{k=1}^{\text{opt}(S)} (d_{k+1}^S \to d_k^S) \right)$. This encoding enforces that if the satisfaction degree of S is at least $k+1$, it is also at least k. Note that $\text{QCCLTOCNF}(F)$ can be locally converted to a CNF formula.

The following properties of QCCLTOCNF follow from Definitions 1–2 and the previous discussion.

[1] For each subformula S of F, the variable d_1^S is redundant, as it is always true. It also holds that $x^S \leftrightarrow \neg d_{\text{opt}(S)+1}^S$ and thus x^S and $d_{\text{opt}(S)+1}^S$ can be replaced with each other. We keep both for clarity of presentation.

Proposition 1. *Let F be a QCCL formula and $k = 1, \ldots, opt(F)$ a satisfaction degree. Each satisfying truth assignment of* $\text{QCCLtoCNF}(F) \wedge \neg d^F_{k+1}$*, when projected to Vars(F), corresponds to a model $\mathcal{I} \in Mod(F)$ with $deg(\mathcal{I}, F) \leq k$. Vice versa, each $\mathcal{I} \in Mod(F)$ with $deg(\mathcal{I}, F) \leq k$ uniquely extends to a satisfying truth assignment of* $\text{QCCLtoCNF}(F) \wedge \neg d^F_{k+1}$.

Corollary 2. *Let F be a QCCL formula. Consider the MaxSAT instance with $F^{\text{Prf}}_{hard} = \text{QCCLtoCNF}(F) \wedge x^F$ and $F^{\text{Prf}}_{soft} = \bigwedge_{k=1}^{opt(F)} \neg d^F_k$. Each optimal MaxSAT solution of $(F^{\text{Prf}}_{hard}, F^{\text{Prf}}_{soft})$, when projected to Vars(F), corresponds to a preferred model $\mathcal{I} \in \text{Prf}(F)$ and the optimal MaxSAT cost is exactly $deg(\mathcal{I}, F)$. Vice versa, each $\mathcal{I} \in \text{Prf}(F)$ uniquely extends to an optimal MaxSAT solution of $(F^{\text{Prf}}_{hard}, F^{\text{Prf}}_{soft})$.*

By Proposition 1, for a given QCCL formula F and for any $k = 1, \ldots, \text{opt}(F)$, we can obtain an interpretation $\mathcal{I} \in \text{Mod}(F)$ with $\text{deg}(\mathcal{I}, F) \leq k$ if one exists (solving the NP-complete DEGREESAT problem [9]) by calling a SAT solver on the formula $\text{QCCLtoCNF}(F) \wedge \neg d^F_{k+1}$. If the formula is satisfiable, a satisfying truth assignment τ directly corresponds to such an interpretation $\mathcal{I} = \tau[\text{Vars}(F)]$. Hence, to extract a preferred model $\mathcal{I} \in \text{Prf}(F)$, it suffices to search for the smallest k for which $\text{QCCLtoCNF}(F) \wedge \neg d^F_{k+1}$ is satisfiable. By Corollary 2, this is achieved by calling a MaxSAT solver on the formula $\text{QCCLtoCNF}(F) \wedge x^F$ (enforcing that F is satisfied in the classical sense) with soft clauses $\bigwedge_{k=1}^{\text{opt}(F)} \neg d^F_k$ (to minimize satisfaction degree). Any optimal solution τ corresponds to a preferred model $\mathcal{I} = \tau[\text{Vars}(F)]$.

We note that this approach extends naturally to solving the following central NP-hard decision problems in QCCL [9].

PMCHECKING (coNP-complete): Given a QCCL formula F and an interpretation $\mathcal{I}$, decide whether $\mathcal{I} \in \text{Prf}(F)$.

PMCONTAINMENT (Θ_2^p-complete): Given a QCCL formula F and a variable $a \in \mathcal{U}$, decide whether there is an interpretation $\mathcal{I} \in \text{Prf}(F)$ such that $\mathcal{I}(a) = 1$.

DEGREEEQUIV (coNP-complete): Given QCCL formulas A and B, decide whether $\text{deg}(\mathcal{I}, A) = \text{deg}(\mathcal{I}, B)$ holds for all interpretations $\mathcal{I}$.

To solve PMCHECKING, first determine $\text{deg}(\mathcal{I}, F)$. If $\text{deg}(\mathcal{I}, F) = \infty$, $\mathcal{I} \notin \text{Mod}(F)$ and hence $\mathcal{I} \notin \text{Prf}(F)$. If $\text{deg}(\mathcal{I}, F) = 1$, $\mathcal{I} \in \text{Prf}(F)$. Otherwise, call a SAT solver on the formula $\text{QCCLtoCNF}(F) \wedge (\neg d^F_{\text{deg}(\mathcal{I},F)})$. By Proposition 1, any satisfying truth assignment corresponds to an interpretation that satisfies F to degree less than $\text{deg}(\mathcal{I}, F)$. If the formula is unsatisfiable, we have $\mathcal{I} \in \text{Prf}(F)$. Otherwise $\mathcal{I} \notin \text{Prf}(F)$ and the obtained satisfying truth assignment is a counterexample to $\mathcal{I}$ being a preferred model.

For PMCONTAINMENT, by Corollary 2, a preferred model of $F \wedge a$ can be extracted by calling a MaxSAT solver on the hard clauses $\text{QCCLtoCNF}(F) \wedge x^F \wedge a$ and soft clauses $\bigwedge_{i=1}^{\text{opt}(F)} \neg d^F_i$. If there is no solution, there is no $\mathcal{I} \in \text{Mod}(F) \supseteq \text{Prf}(F)$ with $\mathcal{I}(a) = 1$. Otherwise any optimal MaxSAT solution τ is a preferred model of $F \wedge a$, and $c(\tau)$ is the optimal satisfaction degree of $F \wedge a$. By Proposition 1, this is a preferred model of F if and only if $\text{QCCLtoCNF}(F) \wedge (\neg d^F_{c(\tau)})$ is unsatisfiable.

To decide DEGREEEQUIV, assume wlog that $\text{opt}(A) \geq \text{opt}(B)$. Consider the formula

$$F_d(A, B) = \text{QCCLtoCNF}(A) \wedge x^A \wedge \text{QCCLtoCNF}(B) \wedge x^B \wedge \left(\bigvee_{k=1}^{\text{opt}(B)+1} (d^A_k \oplus d^B_k) \vee \bigvee_{k=\text{opt}(B)+2}^{\text{opt}(A)+1} (d^A_k \oplus \neg x^B) \right),$$

where $\oplus$ is the exclusive-or connective, encoding that both A and B are satisfied and that their satisfaction degrees are different, i.e., there exists k so that d^A_k and d^B_k take on different truth values (where d^B_k is replaced by $\neg x^B$ for $k > \text{opt}(B) + 1$). Now $F_d(A, B)$ is unsatisfiable if and only if A and B are degree-equivalent. A satisfying truth assignment to $F_d(A, B)$ is a counterexample to degree equivalence. This can be extended to deciding full equivalence by first checking whether $\text{opt}(A) = \text{opt}(B)$; if not A and B are not fully equivalent; otherwise A and B are fully equivalent iff they are degree-equivalent [9].

4 Algorithms for preferred model entailment in choice logic theories

A *choice logic theory* T is a finite set of choice logic formulas. In this work, we consider QCCL theories consisting of QCCL formulas. An interpretation $\mathcal{I}$ is a model of T, denoted by $\mathcal{I} \in \text{Mod}(T)$, if $\mathcal{I} \in \text{Mod}(F)$ for each $F \in T$. The notion of optionality extends to theories via $\text{opt}(T) = \max\{\text{opt}(F) \mid F \in T\}$. For $k = 1, \ldots, \text{opt}(T)$, we denote by $T^{opt}_k = \{F \in T \mid \text{opt}(F) \geq k\}$ the set of formulas with optionality at least k. We focus on the minmax [9] (mm), lexicographic [14] (lex) and inclusion-based [14] (inc) semantics for preferred models of choice logic theories [9]. For semantics $\sigma \in \{\text{mm}, \text{lex}, \text{inc}\}$, we denote the σ-preferred models of a QCCL theory T by $\text{Prf}^\sigma(T)$. Given a QCCL theory T and a classical formula Q, the *preferred model entailment* problem [9] is to decide whether $\mathcal{I} \in \text{Prf}^\sigma(T)$ implies $\mathcal{I} \models Q$ in the classical sense. In words, a theory T entails a formula Q under a semantics σ iff all σ-preferred models of T are (classical) models of Q.

4.1 *Minmax preferred model entailment*

Let T be a QCCL theory. Minmax preferred models minimize the maximum satisfaction degree over all formulas in T [9], viewing a theory as a conjunction of its formulas.

Definition 3. *A model $\mathcal{I} \in Mod(T)$ is a* minmax preferred model *of T, denoted by $\mathcal{I} \in \text{Prf}^{mm}(T)$, if for all interpretations $\mathcal{I}' \in Mod(T)$, $\mathcal{I}' \neq \mathcal{I}$ we have $\max\{deg(\mathcal{I}, F) \mid F \in T\} \leq \max\{deg(\mathcal{I}', F) \mid F \in T\}$.*

Deciding minmax preferred model entailment is Θ_2^p-complete [9].

Example 2. *Consider the QCCL theory $T = \{(a \vec{\times} c), (b \vec{\times} c), (\neg b \vec{\times} \neg a)\}$. The satisfaction degree of at least one formula of T must be greater than 1 in a model of T and thus any classical model of T is in this case a minmax preferred model: $\text{Prf}^{mm}(T) = \{\{a, c\}, \{b, c\}, \{c\}\}$. Clearly c, and no other literal, is minmax entailed by T.*

To express the maximum satisfaction degree over all formulas $F \in T$, we use auxiliary variables s_k for each $k = 1, \ldots, \text{opt}(T)$. We define the formula $F_{\text{mm}}(T)$ as

$$\bigwedge_{F \in T} \left(\text{QCCLtoCNF}(F) \wedge x^F \right) \wedge \left(\bigwedge_{k=1}^{\text{opt}(T)} \left(s_k \rightarrow \bigwedge_{F \in T^{opt}_k} \neg d^F_{k+1} \right) \right)$$

encoding that if s_k is true, the maximum satisfaction degree is at most k. The optimal MaxSAT solutions of $F_{\text{mm}}(T)$ with soft clauses $\bigwedge_{k=1}^{\text{opt}(T)} s_k$ are the minmax preferred models of T.

Proposition 3. *Let T be a QCCL theory. Consider the MaxSAT instance with $F^{mm}_{hard} = F_{mm}(T)$ and $F^{mm}_{soft} = \bigwedge_{k=1}^{opt(T)} s_k$. Each optimal MaxSAT solution of $(F^{mm}_{hard}, F^{mm}_{soft})$, when projected to Vars$(T)$, corresponds to a minmax preferred model $\mathcal{I} \in \text{Prf}^{mm}(T)$. Moreover, the optimal MaxSAT cost is $\min_{\mathcal{I}\in Mod(T)} \max_{F\in T} deg(\mathcal{I}, F) - 1$. Vice versa, each minmax preferred model $\mathcal{I} \in \text{Prf}^{mm}(T)$ uniquely extends to an optimal MaxSAT solution of $(F^{mm}_{hard}, F^{mm}_{soft})$.*

To decide entailment under minmax semantics, note that T trivially entails Q if $\bigwedge_{F\in T} F \wedge \neg Q$ is unsatisfiable. Otherwise, we compute a minmax preferred model of T and obtain its maximum satisfaction degree via a MaxSAT solver call by Proposition 3. Finally, we check whether there is a minmax preferred model of T (enforcing a satisfaction degree at most the degree obtained via MaxSAT) which falsifies Q with a SAT solver call: T entails Q under the minmax semantics iff this is not the case.

Proposition 4. *Let T be a QCCL theory and Q a classical formula. If $\bigwedge_{F\in T} F \wedge \neg Q$ is unsatisfiable, T entails Q under minmax semantics. Otherwise, let τ be an optimal solution to the MaxSAT instance with hard clauses $F_{mm}(T)$ and soft clauses $\bigwedge_{k=1}^{opt(T)} s_k$. The formula $F_{mm(T)} \wedge \text{TSEITIN}(Q) \wedge (\neg x^Q) \wedge s_{c(\tau)}$ is unsatisfiable if and only if T entails Q under minmax semantics.*

4.2 Lexicographically preferred model entailment

The lexicographic semantics [14] is concerned with the satisfaction degree of each individual formula in a theory. For an interpretation $\mathcal{I}$ and $k \in \mathbb{Z}_+$, let $T^{deg}_k(\mathcal{I}) = \{F \in T \mid \deg(\mathcal{I}, F) = k\}$ denote the set of formulas in T satisfied by $\mathcal{I}$ to degree k.

Definition 4. *Let T be a QCCL theory. A model $\mathcal{I} \in Mod(T)$ is a lexicographically preferred model of T, denoted by $\mathcal{I} \in \text{Prf}^{lex}(T)$, if there is no $\mathcal{I}' \in Mod(T)$, $\mathcal{I}' \neq \mathcal{I}$ such that $|T^{deg}_k(\mathcal{I})| < |T^{deg}_k(\mathcal{I}')|$ for some $k \in \mathbb{Z}_+$, and $|T^{deg}_m(\mathcal{I})| = |T^{deg}_m(\mathcal{I}')|$ for all $m < k$.*

In words, preferred models under lexicographic semantics are interpretations which satisfy a greatest number of formulas with smallest possible satisfaction degrees, in a lexicographic sense.

Example 3. *Consider the QCCL theory from Example 2: $T = \{(a \vec{\times} c), (b \vec{\times} c), (\neg b \vec{\times} \neg a)\}$. The only lexicographically preferred model is $\{a, c\}$: two formulas of T are satisfied with degree 1, and one with degree 2, while the other models satisfy one formula with degree 1 and two with degree 2. Thus, e.g. a, c, and $a \wedge c$ are lexicographically entailed by T.*

Deciding lexicographic preferred model entailment is Δ^p_2-complete [9]. In line with this, we detail an algorithm for MaxSAT-based lexicographic optimization [1, 29], as Algorithm 1. We begin by initializing a set of hard clauses F_{hard} as $\bigwedge_{F\in T}(\text{QCCLtoCNF}(F) \wedge x^F)$ (line 1), stating that each $F \in T$ must be satisfied (to a finite degree). We continue by iterating through every possible degree $k = 1, \ldots, \text{opt}(T)$ (lines 2–7), and define a set of soft clauses F_{soft} as $\bigwedge_{F\in T^{opt}_k}(\neg d^F_{k+1})$ (line 3). Via a MaxSAT solver call with instance $(F_{\text{hard}}, F_{\text{soft}})$ (line 4) we obtain an optimal solution τ with cost c which maximizes $|T^{deg}_k(\tau)|$, i.e., the number of formulas satisfied to degree k. If there is no satisfying truth assignment (line 5), all $F \in T$ cannot be satisfied, meaning that Q is trivially entailed as a consequence of an inconsistent theory. If, on the other hand, all formulas are satisfied to degree at most k (line 6), τ is a preferred model of T, so we exit the loop. Otherwise, we add a *cardinality constraint* $\sum_{F\in T^{opt}_k} d^F_{k+1} \leq c$ to F_{hard} (line 7), which

Algorithm 1 MaxSAT-based algorithm for deciding preferred model entailment under lexicographic semantics.

Input: QCCL theory T and classical formula Q
Output: YES if T entails Q, NO otherwise
1: $F_{\text{hard}} \leftarrow \bigwedge_{F\in T}(\text{QCCLtoCNF}(F) \wedge x^F)$
2: **for** $k = 1, \ldots, \text{opt}(T)$ **do**
3: $\quad F_{\text{soft}} \leftarrow \bigwedge_{F\in T^{opt}_k}(\neg d^F_{k+1})$
4: $\quad (c, \tau) \leftarrow \text{MaxSAT}(F_{\text{hard}}, F_{\text{soft}})$
5: $\quad$ **if** $c = \infty$ **return** YES
6: $\quad$ **if** $c = 0$ **break**
7: $\quad F_{\text{hard}} \leftarrow F_{\text{hard}} \wedge (\sum_{F\in T^{opt}_k} d^F_{k+1} \leq c)$
8: **if** $\text{SAT}(F_{\text{hard}} \wedge \text{TSEITIN}(Q) \wedge (\neg x^Q))$ **return** NO
9: **else return** YES

enforces that the number of formulas satisfied to degree at most k must be at least $|T^{deg}_k(\tau)| = |T| - c$. In other words, the cardinality constraints ensure that when considering degrees higher than k on successive iterations, the number of formulas satisfied to degree k does not decrease. Note that various CNF encodings of cardinality constraints have been developed (see e.g [4, 2]) and are readily available [25]. After exiting the loop, the hard clauses F_{hard} encode lexicographically preferred models of T.

Proposition 5. *For $k = 1, \ldots, opt(T)$, let c_k be the optimal cost of a MaxSAT instance with hard clauses $\text{QCCLtoCNF}(F) \wedge x^F \wedge \bigwedge_{m=1}^{k-1} \left(\sum_{F\in T_m} d^F_{m+1} \leq c_m\right)$ and soft clauses $\bigwedge_{F\in T^{opt}_k}(\neg d^F_{k+1})$, and $F_{lex}(T) = \text{QCCLtoCNF}(F) \wedge x^F \wedge \bigwedge_{k=1}^{opt(T)} \left(\sum_{F\in T^{opt}_k} d^F_{k+1} \leq c_k\right)$. Each satisfying truth assignment of $F_{lex}(T)$, when projected to Vars(T), corresponds to a lexicographically preferred model $\mathcal{I} \in \text{Prf}^{lex}(T)$. Vice versa, each $\mathcal{I} \in \text{Prf}^{lex}(T)$ uniquely extends to a satisfying truth assignment of $F_{lex}(T)$.*

It remains to check whether there is a preferred model which falsifies Q. This is achieved using a SAT solver call on input $F_{\text{hard}} \wedge \text{TSEITIN}(Q) \wedge (\neg x^Q)$ (line 9): Q is entailed by T in a lexicographic sense iff this formula is unsatisfiable.

Proposition 6. *Let T be a QCCL theory and Q a classical formula. With input T and Q, Algorithm 1 returns YES if and only if T entails Q under lexicographic semantics.*

4.3 Inclusion-based preferred model entailment

The inclusion-based semantics [14] refines lexicographic semantics.

Definition 5. *Let T be a QCCL theory. A model $\mathcal{I} \in Mod(T)$ is an inclusion-based preferred model of T, denoted by $\mathcal{I} \in \text{Prf}^{inc}(T)$, if there is no $\mathcal{I}' \in Mod(T)$, $\mathcal{I}' \neq \mathcal{I}$ such that $T^{deg}_k(\mathcal{I}) \subsetneq T^{deg}_k(\mathcal{I}')$ for some $k \in \mathbb{Z}_+$ and $T^{deg}_m(\mathcal{I}) = T^{deg}_m(\mathcal{I}')$ for all $m < k$.*

Deciding inclusion-based entailment is Π^p_2-complete [9].

Example 4. *Consider again the QCCL theory $T = \{(a \vec{\times} c), (b \vec{\times} c), (\neg b \vec{\times} \neg a)\}$. Both $\{a, c\}$ and $\{b, c\}$ are inclusion-based preferred models. The latter is preferred because it satisfies $(b \vec{\times} c)$ with degree 1, unlike other interpretations. The literal c is inclusion-based entailed.*

Inclusion-based entailment requires reasoning about *subsets* of formulas entailed to a specific degree. Due to this, we develop an

Algorithm 2 PrefSAT-based algorithm for deciding preferred model entailment under inclusion-based semantics.

Input: QCCL theory T and classical formula Q
Output: YES if T entails Q, NO otherwise

```
1: F ← ⋀_{F∈T}(QCCLtoCNF(F) ∧ x^F) ∧ Tseitin(Q)
2: w ← {¬d^F_{k+1} ↦ opt(T) − k + 1 | F ∈ T, k = 1, ..., opt(F)}
3: while true do
4:     (result, τ) ← PrefSAT(F ∧ (¬x^Q), w)
5:     if result = unsat then return YES
6:     I ← τ[Vars(T)]
7:     (result, _) ← PrefSAT(F ∧ MorePref(I, T), w)
8:     if result = unsat then return NO
9:     F ← F ∧ Refine(I, T)
```

approach based on *SAT with preferences* (PrefSAT) [33, 32, 19]. In PrefSAT, the input is a CNF formula F with a weight function $w\colon \text{Lits}(F) \to \mathbb{Z}_+$. For a satisfying truth assignment τ to F and an integer k, let $\Lambda_k(\tau) = \{l \in \text{Lits}(F) \mid \tau(l) = 1, w(l) = k\}$. A truth assignment τ is preferred to another truth assignment τ', if there is a $k > 0$ with $\Lambda_k(\tau) \supsetneq \Lambda_k(\tau')$ and $\Lambda_i(\tau) = \Lambda_i(\tau')$ for all $i > k$. The task is to find a most preferred satisfying assignment τ to F.

The Π_2^p-completeness of the entailment problem suggests a counterexample-guided abstraction refinement (CEGAR) [17, 18] approach. In CEGAR [17], an abstraction as an overapproximation of the set of solutions is iteratively solved to obtain candidate solutions. In each iteration it is checked whether the obtained candidate is an actual solution by searching for counterexamples. If there are no counterexamples, the candidate is an actual solution to the problem. Otherwise, the abstraction is refined to exclude at least the candidate.

We initialize the abstraction F as $\bigwedge_{F\in T}(\text{QCCLtoCNF}(F) \wedge x^F) \wedge \text{Tseitin}(Q)$ (line 1) enforcing that each formula in T must be satisfied to a finite degree, and introducing the auxiliary variable x^Q which is true iff the classical formula Q is satisfied. We assign the PrefSAT weights w to map each literal $\neg d^F_{k+1}$ to $\text{opt}(T) - k + 1$ (line 2), stating the higher preference to satisfy a formula the lower the satisfaction degree of the formula is. Now PrefSAT solutions to (F, w) correspond to inclusion-based preferred models.

Proposition 7. *Let T be a QCCL theory. Consider the formula $F_{inc}(T) = \bigwedge_{F\in T}(\text{QCCLtoCNF}(F) \wedge x^F)$, and the weight function $w_{inc}(T)$ which, for each $F \in T$ and $k = 1, \ldots, opt(F)$, maps literal $\neg d^F_{k+1}$ to $opt(T) - k + 1$. Each PrefSAT solution of $F_{inc}(T)$ with weights $w_{inc}(T)$, when projected to Vars(T), corresponds to an inclusion-based preferred model $\mathcal{I} \in \text{Prf}^{inc}(T)$. Vice versa, each $\mathcal{I} \in \text{Prf}^{inc}(T)$ uniquely extends to a PrefSAT solution of $F_{inc}(T)$ with weights $w_{inc}(T)$.*

For a satisfaction degree $k = 2, \ldots, \text{opt}(T)$ and $\mathcal{I} \in \text{Mod}(T)$ let

$$\text{NHD}_k(\mathcal{I}) = \bigwedge_{\substack{F\in T_k^{opt} \\ \deg(\mathcal{I},F)<k}} \neg d_k^F, \quad \text{ELD}_k(\mathcal{I}) = \bigvee_{\substack{F\in T_k^{opt} \\ \deg(\mathcal{I},F)\geq k}} \neg d_k^F.$$

$\text{NHD}_k(\mathcal{I})$ (short for *No Higher Degree*) encodes that all formulas $F \in T$ with satisfaction degree less than k under $\mathcal{I}$ still have satisfaction degree less than k. $\text{ELD}_k(\mathcal{I})$ (short for *Exists Lower Degree*) encodes that at least one formula $F \in T$ with satisfaction degree at least k under $\mathcal{I}$ has satisfaction degree less than k. Let $\text{Eq}_2(\mathcal{I}) = \text{NHD}_2(\mathcal{I}) \wedge \neg\text{ELD}_2(\mathcal{I})$, and $\text{Eq}_k(\mathcal{I}) = \text{Eq}_{k-1}(\mathcal{I}) \wedge \text{NHD}_k(\mathcal{I}) \wedge \neg\text{ELD}_k(\mathcal{I})$ for each $k = 3, \ldots, \text{opt}(T)$. $\text{Eq}_k(\mathcal{I})$ encodes that up to satisfaction degree k, formulas have the same satisfaction degrees as in $\mathcal{I}$.

In the CEGAR loop of Algorithm 2 (lines 3–9), we iteratively call a PrefSAT solver on formula $F \wedge (\neg x^Q)$ with weights w (line 4). If no solution exists (line 5), T entails Q. Otherwise, by Proposition 7, a PrefSAT solution τ corresponds to an inclusion-based preferred model $\mathcal{I}$ of T under the constraint that Q is falsified (line 6). We continue by checking for a counterexample, i.e., an inclusion-based preferred model of T which is more preferred than $\mathcal{I}$. A PrefSAT solver call on the formula $F \wedge \text{MorePref}(\mathcal{I}, T)$ and weights w (line 7), where $\text{MorePref}(\mathcal{I}, T) = (\text{NHD}_2(\mathcal{I}) \wedge \text{ELD}_2(\mathcal{I})) \vee \bigvee_{k=3}^{\text{opt}(T)} (\text{NHD}_k(\mathcal{I}) \wedge \text{ELD}_k(\mathcal{I}) \wedge \text{Eq}_{k-1}(\mathcal{I}))$ checks this. If no such models exist (line 8), $\mathcal{I}$ is a lexicographically preferred model of T and falsifies Q, and hence T does not entail Q. Otherwise, we refine the abstraction by conjoining to the abstraction $\text{Refine}(\mathcal{I}, T)$ (line 9), where $\text{Refine}(\mathcal{I}, T) = (\text{NHD}_2(\mathcal{I}) \vee \text{ELD}_2(\mathcal{I})) \wedge \bigwedge_{k=3}^{\text{opt}(T)-1} (\text{Eq}_{k-1}(\mathcal{I}) \to (\text{NHD}_k(\mathcal{I}) \vee \text{ELD}_k(\mathcal{I}))) \wedge \big(\text{Eq}_{\text{opt}(T)-1}(\mathcal{I}) \to \text{ELD}_{\text{opt}(T)}(\mathcal{I})\big)$ enforces that for all possible satisfaction degrees k, if formulas are satisfied with the same satisfaction degree up to degree $k-1$, then it is not possible to satisfy a strict subset of formulas satisfied on level k.

Proposition 8. *Let T be a QCCL theory and Q a classical formula. With input T and Q, Algorithm 2 returns YES if and only if T entails Q under inclusion-based semantics.*

5 Empirical Evaluation

We present results on the empirical runtime performance of the implementation, named choiceSAT, of our approach of the SAT-based approaches detailed in Sections 3 and 4. The implementation, available in open source [27], uses the PySAT Python library (1.8.dev12) [25], RC2 [26] as the MaxSAT solver, CaDiCaL (1.9.5) [11] as the SAT solver, and Minipref [19] as the PrefSAT solver. For computing a preferred model, we compare choiceSAT to the recent ASP-based approach [8] referred to here as QCCL-ASP, using Clingo (5.4.0) [24] as the ASP solver. To the best of our knowledge, QCCL-ASP is the only existing implementation for

	#solved (mean runtime over solved (s))					
	QCCL-ASP					
$\#F$	QCL		CCL		QCCL	
30	5	(294.7)	5	(139.0)	5	(570.9)
>30	0	—	0	—	0	—
	choiceSAT					
$\#F$	QCL		CCL		QCCL	
30	5	(0.1)	5	(0.1)	5	(0.1)
60	5	(0.2)	5	(0.1)	5	(0.2)
90	5	(0.3)	5	(0.3)	5	(0.2)
120	5	(0.7)	5	(0.6)	5	(0.3)
150	5	(5.4)	5	(4.0)	5	(0.4)
180	5	(15.2)	5	(12.3)	5	(0.4)
210	5	(112.2)	5	(25.0)	5	(0.5)
240	5	(261.5)	4	(178.4)	5	(0.7)
270	0	—	1	(189.7)	5	(0.6)
300	0	—	1	(425.3)	5	(0.9)
330	0	—	0	—	5	(1.0)
360	0	—	0	—	5	(1.2)
390	0	—	0	—	5	(1.2)
420	0	—	0	—	5	(1.3)

Table 1. choiceSAT vs QCCL-ASP on computing a preferred model.

#solved (mean runtime over solved (s))																		
	Minmax						Lexicographic						Inclusion-based					
$\#F$	QCL		CCL		QCCL		QCL		CCL		QCCL		QCL		CCL		QCCL	
30	5	(0.1)	5	(0.1)	5	(0.1)	5	(0.3)	5	(0.2)	5	(1.1)	5	(0.2)	5	(0.2)	5	(0.3)
60	5	(0.2)	5	(0.2)	5	(0.2)	5	(0.8)	5	(0.4)	5	(3.7)	5	(0.4)	5	(0.4)	5	(0.4)
90	5	(0.4)	5	(0.3)	5	(0.3)	5	(2.8)	5	(1.0)	5	(54.1)	5	(4.6)	5	(2.1)	5	(0.6)
120	5	(1.0)	5	(0.9)	5	(0.4)	5	(20.4)	5	(1.6)	2	(361.8)	5	(39.3)	5	(11.5)	5	(0.8)
150	5	(6.7)	5	(3.3)	5	(0.4)	5	(81.4)	5	(6.0)	0	—	0	—	5	(22.8)	5	(0.9)
180	5	(18.8)	5	(11.6)	5	(0.5)	0	—	5	(23.0)	0	—	1	(521.5)	5	(39.2)	5	(1.1)
210	5	(96.4)	2	(90.7)	5	(0.6)	0	—	5	(28.1)	0	—	0	—	5	(88.3)	5	(1.4)
240	5	(305.8)	5	(127.7)	5	(1.0)	0	—	5	(116.8)	0	—	0	—	4	(164.9)	5	(1.5)
270	0	—	2	(159.3)	5	(0.7)	0	—	2	(486.5)	0	—	0	—	2	(262.1)	5	(1.7)
300	0	—	1	(576.8)	5	(1.2)	0	—	2	(432.5)	0	—	0	—	1	(19.6)	5	(2.3)
330	0	—	0	—	5	(1.1)	0	—	0	—	0	—	0	—	0	—	5	(2.5)
360	0	—	0	—	5	(1.5)	0	—	0	—	0	—	0	—	0	—	5	(2.3)

Table 2. Runtime performance of CHOICESAT on preferred model entailment.

computing preferred models in QCCL and no previous implementations exist for preferred model entailment. The experiments were run on 2.40-GHz Intel Xeon Gold 6148 CPUs under a per-instance time limit of 600 seconds and memory limit of 32 GB.

Lacking a standard benchmark library for choice logics, we base our benchmarks on a model for non-clausal negation normal form propositional formulas viewed as balanced trees with alternative levels of $\vee$ and $\wedge$ nodes [30]. The model exhibits a phase transition behavior [30], where (depending on chosen number of levels/formula depth) at a specific ratio r^* of immediate subformulas to variables a sharp peak in empirical hardness occurs for deciding satisfiability. At r^*, there is a sharp change from almost all formulas being satisfiable to almost all being unsatisfiable, in analogy to phase transitions in random k-SAT. We used the model to generate depth-5 formulas with $\vee$ in the root and binary connectives. The empirical phase transition ratio for formulas that are the conjunction of a set of such so-called $\langle 2,2,2,2,2\rangle$-shape formulas is $r = 5.42$ [30]. To construct expectedly challenging satisfiable formulas, we set $r = 5$ and generated instances with a varying number of subformulas, namely $\#F = 30..450$ of $\langle 2,2,2,2,2\rangle$-shape subformulas in increments of 30. The number of variables is thus $|\text{Vars}| = \#F/5$. Hard but satisfiable instances are desirable for evaluating algorithms for choice logics, since unsatisfiable instances are uninteresting from the point of view of preferences. To obtain QCL, CCL, and QCCL formulas, we replaced in the obtained propositional formulas each $\vee$ (resp. $\wedge$) by $\vec{\times}$ (resp. $\vec{\odot}$) with a fixed probability, using combinations of probabilities $(0, 0.5)$ (resulting in QCL formulas), $(0.5, 0)$ (CCL) and $(0.5, 0.5)$ (QCCL). We generated five QCL, CCL and QCCL formulas for each value of $\#F$. For preferred model entailment, we interpreted each constructed QCCL formula as the choice logic theory T consisting of its immediate subformulas (i.e. the $\langle 2,2,2,2,2\rangle$-formulas), and selected a random literal l to obtain the instance for deciding if l is entailed from T under minmax, lexicographic or inclusion-based semantics.

A comparison between CHOICESAT and QCCL-ASP on the problem of computing a preferred model in terms of number of instances solved (out of 5) and mean runtime over solved instances for each value of $\#F$ is shown in Table 1. QCCL-ASP could not solve any instances beyond $\#F = 30$. In contrast, CHOICESAT scales significantly further, solving all instances up to $\#F = 120$ in less than one second on average. Scalability depends to an extent on the choice connective(s) in the formulas. On the QCL instances, which can be observed to be the hardest ones to solve, CHOICESAT scales up to $\#F = 240$. The better scalability on instances with ordered conjunction may be due to fact that the ordered conjunction is intuitively a relaxed variant of classical conjunction, and hence replacing conjunction with ordered conjunction moves the underlying formulas further away from the phase transition threshold towards more easy-to-solve satisfiable instances.

Turning to preferred model entailment, results for CHOICESAT under the three different semantics are shown in Table 2. CHOICESAT exhibits good scalability even for these computationally harder problems. Under minmax semantics the performance on entailment is similar to computing a preferred model, which is explained by the fact that our minmax algorithm performs only a single additional SAT solver call after computing a preferred model. The relative runtimes under the three semantics are in line with the relative complexity of the problems (Θ_2^p for minmax, Δ_2^p for lexicographic, and Π_2^p for inclusion-based), with the exception of large QCCL instances being easier to solve under inclusion-based than lexicographic semantics. We suspect that this due to the fact that adding more ordered connectives increases the optionality of formulas, increasing the iterations required to find a lexicographically preferred model. Additionally, at times the algorithm for inclusion-based entailment might be able to quickly find a counterexample, while the lexicographic algorithm always needs to perform multiple MaxSAT calls up to the level where each formula is satisfied to degree at most this level. Overall, the results for CHOICESAT are encouraging, and CHOICESAT considerably outperform the recent QCCL-ASP approach.

6 Conclusions

We developed SAT-based approaches for computing preferred models and for preferred model entailment for QCCL, the combination of QCL and CCL. We captured preferred models via MaxSAT, forming the basis of algorithms for preferred model entailment and related problems such as equivalence checking of QCCL formulas. For preferred model entailment, we detailed iterative procedures that pertain to the known complexity of the entailment problems, capturing minmax semantics via MaxSAT and SAT, lexicographic semantics via MaxSAT-based lexicographic optimization, and inclusion-based semantics by PrefSAT-based CEGAR. Empirically, our implementation of the approaches scales significantly better than a recent approach proposed for computing preferred models, and allows for deciding entailment for reasonably-sized choice logic theories with hundreds of formulas. This motivates studying the potential of developing SAT-based approaches to further related choice logics [6, 15].

Acknowledgements

Work financially supported by Academy of Finland (grants 347588 and 356046). The authors thank the Finnish Computing Competence Infrastructure (FCCI) for computational and data storage resources.

References

[1] J. Argelich, I. Lynce, and J. P. M. Silva. On solving boolean multilevel optimization problems. In *IJCAI*, pages 393–398. ijcai.org, 2009.

[2] R. Asín, R. Nieuwenhuis, A. Oliveras, and E. Rodríguez-Carbonell. Cardinality networks: a theoretical and empirical study. *Constraints An Int. J.*, 16(2):195–221, 2011.

[3] F. Bacchus, M. Järvisalo, and R. Martins. Maximum satisfiabiliy. In *Handbook of Satisfiability - Second Edition*, volume 336 of *FAIA*, pages 929–991. IOS Press, 2021.

[4] O. Bailleux and Y. Boufkhad. Efficient CNF encoding of boolean cardinality constraints. In *CP*, volume 2833 of *LNCS*, pages 108–122. Springer, 2003.

[5] S. Benferhat and K. Sedki. Alert correlation based on a logical handling of administrator preferences and knowledge. In *SECRYPT*, pages 50–56. INSTICC Press, 2008.

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

[7] S. Benferhat, D. L. Berre, and K. Sedki. Handling qualitative preferences using normal form functions. In *FLAIRS*, pages 38–43. AAAI Press, 2007.

[8] M. Bernreiter, J. Maly, and S. Woltran. Encoding choice logics in ASP. In *ICLP 2020 Workshop Proceedings*, volume 2678 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2020.

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

[10] M. Bernreiter, A. Lolic, J. Maly, and S. Woltran. Sequent calculi for choice logics. *J. Autom. Reason.*, 68(2):8, 2024.

[11] A. Biere, K. Fazekas, M. Fleury, and M. Heisinger. CaDiCaL, Kissat, Paracooba, Plingeling and Treengeling entering the SAT Competition 2020. In *Proc. of SAT Competition 2020 – Solver and Benchmark Descriptions*, volume B-2020-1 of *Department of Computer Science Report Series B*, pages 51–53. University of Helsinki, 2020.

[12] A. Biere, M. Heule, H. van Maaren, and T. Walsh, editors. *Handbook of Satisfiability - Second Edition*, volume 336 of *FAIA*. IOS Press, 2021. ISBN 978-1-64368-160-3.

[13] A. Boudjelida and S. Benferhat. Conjunctive choice logic. In *ISAIM*, 2016.

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

[15] A. Charalambidis, G. Papadimitriou, P. Rondogiannis, and A. Troumpoukis. A many-valued logic for lexicographic preference representation. In *KR*, pages 646–650, 2021.

[16] A. Charalambidis, P. Rondogiannis, and A. Troumpoukis. A logical characterization of the preferred models of logic programs with ordered disjunction. *Theory Pract. Log. Program.*, 21(5):629–645, 2021.

[17] E. M. Clarke, O. Grumberg, S. Jha, Y. Lu, and H. Veith. Counterexample-guided abstraction refinement for symbolic model checking. *J. ACM*, 50(5):752–794, 2003.

[18] E. M. Clarke, A. Gupta, and O. Strichman. Sat-based counterexample-guided abstraction refinement. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.*, 23(7):1113–1123, 2004.

[19] C. Dodaro and A. Previti. Minipref: A tool for preferences in SAT. In *RCRA*, volume 2538 of *CEUR Workshop Proc.* CEUR-WS.org, 2019.

[20] C. Domshlak, E. Hüllermeier, S. Kaci, and H. Prade. Preferences in AI: an overview. *Artif. Intell.*, 175(7-8):1037–1052, 2011.

[21] W. Faber, H. Tompits, and S. Woltran. Notions of strong equivalence for logic programs with ordered disjunction. In *KR*, pages 433–443. AAAI Press, 2008.

[22] R. Freiman and M. Bernreiter. Truth and preferences - A game approach for qualitative choice logic. In *JELIA*, volume 14281 of *Lecture Notes in Computer Science*, pages 547–560. Springer, 2023.

[23] R. Freiman and M. Bernreiter. Validity in choice logics - A game-theoretic investigation. In *WoLLIC*, volume 13923 of *Lecture Notes in Computer Science*, pages 211–226. Springer, 2023.

[24] M. Gebser, R. Kaminski, B. Kaufmann, M. Ostrowski, T. Schaub, and P. Wanko. Theory solving made easy with clingo 5. In *ICLP 2016 TCs*, volume 52 of *OASIcs*, pages 2:1–2:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2016.

[25] A. Ignatiev, A. Morgado, and J. Marques-Silva. PySAT: A python toolkit for prototyping with SAT oracles. In *SAT*, volume 10929 of *LNCS*, pages 428–437. Springer, 2018.

[26] A. Ignatiev, A. Morgado, and J. Marques-Silva. RC2: an efficient MaxSAT solver. *J. Satisf. Boolean Model. Comput.*, 11(1):53–64, 2019.

[27] T. Lehtonen, A. Niskanen, and M. Järvisalo. ChoiceSAT: SAT-based algorithms for choice logics, 2024. URL https://bitbucket.org/coreo-group/choicesat/.

[28] L. Liétard, A. Hadjali, and D. Rocacher. Towards a gradual QCL model for database querying. In *IPMU*, volume 444 of *Communications in Computer and Information Science*, pages 130–139. Springer, 2014.

[29] J. Marques-Silva, J. Argelich, A. Graça, and I. Lynce. Boolean lexicographic optimization: algorithms & applications. *Ann. Math. Artif. Intell.*, 62(3-4):317–343, 2011.

[30] J. A. N. Pérez and A. Voronkov. Generation of hard non-clausal random satisfiability problems. In *AAAI*, pages 436–442. AAAI Press / The MIT Press, 2005.

[31] G. Pigozzi, A. Tsoukiàs, and P. Viappiani. Preferences in artificial intelligence. *Ann. Math. Artif. Intell.*, 77(3-4):361–401, 2016.

[32] E. D. Rosa and E. Giunchiglia. Combining approaches for solving satisfiability problems with qualitative preferences. *AI Commun.*, 26(4): 395–408, 2013.

[33] E. D. Rosa, E. Giunchiglia, and M. Maratea. Solving satisfiability problems with preferences. *Constraints An Int. J.*, 15(4):485–515, 2010.

[34] K. Sedki, J. Lamy, and R. Tsopra. Qualitative choice logic for modeling experts recommendations of antibiotics. In *FLAIRS*, 2022.

[35] G. S. Tseitin. On the complexity of derivation in propositional calculus. In *Automation of Reasoning: 2: Classical Papers on Computational Logic 1967–1970*, pages 466–483. Springer, 1983.