

## SAT-INSPIRED HIGHER-ORDER ELIMINATIONS

JASMIN BLANCHETTE <sup>a,b</sup>  AND PETAR VUKMIROVIĆ <sup>c</sup> 

<sup>a</sup>Ludwig-Maximilians-Universität München, Institut für Informatik, Munich, Germany  
*e-mail address:* jasmin.blanchette@ifi.lmu.de

<sup>b</sup>Max-Planck-Institut für Informatik, Saarland Informatics Campus, Saarbrücken, Germany  
*e-mail address:* jasmin.blanchette@mpi-inf.mpg.de

<sup>c</sup>Vrije Universiteit Amsterdam, Department of Computer Science, Amsterdam, The Netherlands  
*e-mail address:* petar.vukmirovic2@gmail.com

---

**ABSTRACT.** We generalize several propositional preprocessing techniques to higher-order logic, building on existing first-order generalizations. These techniques eliminate literals, clauses, or predicate symbols from the problem, with the aim of making it more amenable to automatic proof search. We also introduce a new technique, which we call *quasipure literal elimination*, that strictly subsumes pure literal elimination. The new techniques are implemented in the Zipperposition theorem prover. Our evaluation shows that they sometimes help prove problems originating from Isabelle formalizations and the TPTP library.

### 1. INTRODUCTION

Processing techniques are an important optimization in SAT (Boolean satisfiability) solving. Following up on early work in the 1990s [GO92, Ohl96], there has recently been renewed interest in adapting propositional techniques to first-order logic [KK16, KSS<sup>+</sup>17, VBH23], resulting in a noticeable increase of the success rate of automatic theorem provers based on superposition [BG94].

In this article, we consider the extension of four main classes of SAT preprocessing techniques to classical higher-order logic. These extensions are called hidden-literal-based elimination (Section 3), predicate elimination (Section 4), blocked clause elimination (Section 5), and quasipure literal elimination (Section 6). Elimination techniques make the problem simpler and hence possibly more amenable to automatic proof search. The techniques can be used either to preprocess the problem or to transform the prover’s current clause set during proof search, a use that is sometimes called *inprocessing*.

An advantage of preprocessing is its greater generality: Preprocessing techniques can be used in tandem with any higher-order proof calculus, as long as the calculus is built around a notion of clause. We assume that a clausifier is run as a preprocessor and introduces some clausal structure. The more clausal structure it produces, the more effective the elimination techniques can be. Examples of provers compatible with the techniques are  $\lambda E$  [VBS23], Leo-III [SB18], Vampire [BR20], and Zipperposition [BBTV21].

Our setting is a clausal version of classical rank-1 polymorphic higher-order logic (Section 2). Since previous work focused on an untyped or monomorphic logic, our work also generalizes this aspect. Higher-order logic also distinguishes between standard and general (Henkin) semantics. Since calculi are proved complete with respect to general semantics, this is the semantics we adopt. Our techniques preserve the unsatisfiability of problems, and therefore their provability by a complete prover. In addition, they preserve the satisfiability of problems, and therefore their unprovability by a sound prover.

The main difficulty we face in higher-order logic concerns predicate elimination and blocked clause elimination, which are both based on resolution. In first-order logic, a literal  $\neg \mathbf{p}(\vec{s})$  can only be resolved against a literal  $\mathbf{p}(\vec{t})$ , with the same predicate symbol  $\mathbf{p}$ . By contrast, in higher-order logic,  $\neg \mathbf{p} \vec{s}$  can be resolved against any variable-headed literal  $y \vec{t}$ , for example by taking  $y := \lambda \vec{x}. \mathbf{p} \vec{s}$ , where the bound variables  $\vec{x}$  are fresh. We will see that this issue can be circumvented: A key finding of this article is that we can ignore variable-headed resolvents and focus on the  $\mathbf{p}$ -literals.

Another potential issue is that higher-order logic can have infinitely many resolvents. For example, resolving  $\neg \mathbf{p}(\mathbf{f}(y \mathbf{a}))$  and  $\mathbf{p}(y(\mathbf{f} \mathbf{a}))$  produces infinitely many conclusions of the form  $\mathbf{p}(\mathbf{f}(\dots(\mathbf{f} \mathbf{a})\dots))$ . Again, the issue is not as severe as it looks, because the variant of resolution we use—flat resolution—does not unify terms.

**Example 1.1.** To give a flavor of our elimination techniques, let us review an example involving blocked clause elimination. Let  $\mathbf{a} : \iota$ ,  $\mathbf{p} : \iota \rightarrow o$ , and  $\text{choice} : (\iota \rightarrow o) \rightarrow \iota$  be symbols, where  $o$  is the type of Booleans and  $\iota$  is a base type. Consider the clause set

$$N = \{\neg y z \vee y(\text{choice } y), \mathbf{q} \mathbf{a}, \neg \mathbf{q}(\text{choice } \mathbf{q}), \mathbf{p} \mathbf{a}, \neg \mathbf{p} z \vee z \approx \mathbf{a}\}$$

The set is clearly inconsistent because the  $\{y \mapsto \mathbf{q}, z \mapsto \mathbf{a}\}$  instance of the first clause is inconsistent with the second and third clauses.

Under some basic conditions, a clause  $C$  containing a literal  $L$  is said to be “blocked” if all of its so-called binary flat  $L$ -resolvents with clauses from  $N \setminus \{C\}$  are tautologies. (We will see in Section 5 what this means exactly.) The fourth clause is blocked by its literal  $\mathbf{p} \mathbf{a}$  because its only binary flat  $(\mathbf{p} \mathbf{a})$ -resolvent, with the fifth clause, is the tautology  $\mathbf{a} \not\approx z \vee z \approx \mathbf{a}$ . Similarly, the fifth clause is blocked by its literal  $\neg \mathbf{p} z$  because its only binary flat  $(\neg \mathbf{p} z)$ -resolvent, with the fourth clause, is the same tautology. Either or both clauses can be removed without from  $N$  losing unsatisfiability.

All the techniques are implemented in the higher-order prover Zipperposition (Section 7), allowing us to measure their effectiveness on benchmarks originating from Isabelle [NPW02] formalizations and the TPTP library [Sut17] (Section 8). The raw experimental data are available online.<sup>1</sup>

We will mention closely related research in the relevant sections. We point to Vukmirović et al. [VBH23, Section 8] for a more detailed discussion of related work.

## 2. CLAUSAL HIGHER-ORDER LOGIC

The logic we use as a basis of our work is a rank-1 polymorphic higher-order logic with general semantics and both functional and Boolean extensionality. It corresponds essentially to the logic embodied by the TPTP TH1 format [KSR16], including Hilbert choice. Our

<sup>1</sup><https://zenodo.org/record/7448169>

conventions largely follow those used by Bentkamp, Blanchette, Tourret, and Vukmirović to define  $\lambda$ -superposition [BBTV21]. Our presentation is based on theirs.

In higher-order logic, formulas are simply terms of Boolean type. Briefly, our version of the logic also has a clausal outer structure, as found in several higher-order provers. Clauses are then built around terms as an extra layer of structure. We write formula-level Boolean operators in bold (e.g.,  $\neg$ ,  $\mathbf{V}$ ,  $\mathbf{V}$ ) to distinguish them from clause-level operators.

**2.1. Syntax.** Let us define the syntax of our logic more precisely, starting with types. Throughout this article, we use the notation  $\vec{a}_n$ , or simply  $\vec{a}$ , to denote an  $n$ -tuple  $(a_1, \dots, a_n)$ . Sometimes we might also write  $\vec{a}_i$ , meaning  $(a_{i_1}, \dots, a_{i_n})$ , to be distinguished from  $\vec{a}_i = (a_1, \dots, a_i)$ .

We start by fixing an infinite set  $\mathcal{V}_{\text{ty}}$  of type variables. A set  $\Sigma_{\text{ty}}$  of type constructors with associated arities is a *type signature*. We require the presence of a nullary type constructor  $o$ , for Booleans, and a binary type constructor  $\rightarrow$  for functions. We let  $\alpha$  range over type variables and  $\kappa$  over type constructors. A *type*, ranged over by  $\tau$  and  $v$ , is defined inductively to be either a variable  $\alpha \in \mathcal{V}_{\text{ty}}$  or an expression  $\kappa(\vec{\tau}_n)$ , where  $\kappa$  is an  $n$ -ary type constructor and  $\vec{\tau}$  in an  $n$ -tuple of types. If  $n = 0$ , we write  $\kappa$  instead of  $\kappa()$ . In addition, expressions with  $\kappa = \rightarrow$  are written in infix notation, as  $\tau_1 \rightarrow \tau_2$ . A *type declaration* is an expression  $\Pi \vec{\alpha}_m. \tau$ , where  $\vec{\alpha}$  consists of distinct type variables and all the type variables occurring in  $\tau$  belong to  $\vec{\alpha}$ . If  $m = 0$ , we write  $\tau$  instead of  $\Pi. \tau$ .

Next, we fix a type signature  $\Sigma_{\text{ty}}$  and a set  $\mathcal{V}$  of term variables with associated types. We require that there are infinitely many variables of each type. A *term signature* is a set  $\Sigma$  of *symbols*  $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{f}, \mathbf{p}, \mathbf{q}, \dots$ , each associated with a type declaration. (Often, symbols are called “constants” in the higher-order logic literature.) We write  $\mathbf{f} : \Pi \vec{\alpha}. \tau$  to indicate that symbol  $\mathbf{f}$  has type signature  $\Pi \vec{\alpha}. \tau$ . We require the presence in  $\Sigma$  of the logical symbols  $\perp, \top, \neg, \wedge, \mathbf{V}, \rightarrow, \mathbf{V}, \exists, \approx, \not\approx$  with their usual type declarations (e.g.,  $\Pi \alpha. \alpha \rightarrow \alpha \rightarrow o$  for  $\approx$ ). We also assume the presence of the Hilbert choice operator  $\boldsymbol{\varepsilon} : \Pi \alpha. (\alpha \rightarrow o) \rightarrow \alpha$ . We will generally leave the signature implicit, assuming some fixed signature  $\Sigma$ .

Polymorphic higher-order logic does not distinguish between function and predicate symbols. Nonetheless, it will be convenient to refer to symbols that can yield Boolean values as predicate symbols. Specifically, we will call a symbol with type declaration  $\mathbf{p} : \Pi \vec{\alpha}. \tau$  that admits an instance of the form  $\mathbf{p}(\vec{\tau}) : v_1 \rightarrow \dots \rightarrow v_n \rightarrow o$  a *predicate symbol*.

We now introduce three notions of terms: raw  $\lambda$ -terms,  $\lambda$ -terms, and actual terms. The  $\lambda$ -terms are  $\alpha$ -equivalence classes of raw  $\lambda$ -terms, and the actual terms are  $\beta\eta$ -equivalence classes of  $\lambda$ -terms.

More precisely, the *raw  $\lambda$ -terms* are defined inductively as follows:

- Every variable  $x$  of type  $\tau$  is a raw  $\lambda$ -term of type  $\tau$ .
- If  $\mathbf{f}$  has type declaration  $\Pi \vec{\alpha}_m. \tau$  in  $\Sigma$  and  $\vec{v}_m$  is a tuple of types, called *type arguments*, then  $\mathbf{f}(\vec{v}_m)$  is a raw  $\lambda$ -term of type  $\tau\{\vec{\alpha}_m \mapsto \vec{v}_m\}$ . If  $m = 0$ , we simply write  $\mathbf{f}$  instead of  $\mathbf{f}(\langle \rangle)$ .
- If  $x$  is a variable of type  $\tau$  and  $t$  is a term of type  $v$ , then the  *$\lambda$ -abstraction*  $\lambda x. t$  is a raw  $\lambda$ -term of type  $\tau \rightarrow v$ .
- If  $s$  is a term of type  $\tau \rightarrow v$  and  $t$  is a term of type  $\tau$ , then the *application*  $s t$  is a raw  $\lambda$ -term of type  $v$ .

We abbreviate  $\lambda x_1. \dots \lambda x_n. t$  to  $\lambda x_1 \dots x_n. t$  or  $\lambda \vec{x}_n. t$ ,  $\mathbf{V}(\lambda x. t)$  to  $\mathbf{V}x. t$ , and similarly for  $\exists$ . Abusing notation, we also write  $t \vec{u}_n$  for  $t u_1 \dots u_n$ . We assume standard notions of

free and bound variables as well as subterms. To indicate that a term  $t$  has a type  $\tau$ , we write  $t : \tau$ .

The  $\alpha$ -renaming rule of the  $\lambda$ -calculus relates two raw  $\lambda$ -terms if the two are equal up to (capture-avoiding) renaming of their bound variables. For example,  $\lambda x. f x x$  and  $\lambda y. f y y$  are  $\alpha$ -renamings of each other. Two raw  $\lambda$ -terms are  $\alpha$ -equivalent if they can be made equal by  $\alpha$ -renaming their subterms. The  $\lambda$ -terms consist of the equivalence classes of raw  $\lambda$ -terms modulo  $\alpha$ -equivalence of subterms. We assume the standard notion of (capture-avoiding) substitution on  $\lambda$ -terms. We also define a notion of replacement:  $t[f \mapsto u]$  denotes the term obtained by replacing all occurrences of  $f$  in  $t$  with a term  $u$  of the same type.

The  $\beta$ -reduction rule relates two  $\lambda$ -terms if the first one has the form  $(\lambda x. s) t$  and the second one has the form  $s\{x \mapsto t\}$ , where bound variables in  $s$  are implicitly renamed to avoid capture. The  $\eta$ -reduction rule relates two  $\lambda$ -terms if the first one has the form  $\lambda x. t x$  and the second one has the form  $t$ , where  $t$  contains no free occurrences of  $x$ . For example,  $(\lambda x. f x x) \mathbf{b}$   $\beta$ -reduces to  $f \mathbf{b} \mathbf{b}$ , and  $\lambda x. f x$   $\eta$ -reduces to  $f$ . Two  $\lambda$ -terms are  $\beta\eta$ -equivalent if they can be made equal by  $\beta$ - and  $\eta$ -reducing their subterms. The *terms* consist of the equivalence classes of  $\lambda$ -terms modulo  $\beta\eta$ -equivalence of subterms. The *formulas* are the terms of type  $o$ . We let  $\varphi, \psi$  range over formulas.

**Convention 2.1.** When inspecting the structure of a term, we will consider a representative in  $\eta$ -short  $\beta$ -normal form, obtained by exhaustively applying  $\beta$ - and  $\eta$ -reduction on subterms. Such a representative is unique up to  $\alpha$ -equivalence.

An alternative to the  $\eta$ -short  $\beta$ -normal form is the  $\eta$ -long  $\beta$ -normal form, in which unapplied functions are  $\eta$ -expanded rather than  $\eta$ -reduced (i.e.,  $\eta$ -reduction is applied in reverse on these). The techniques presented in this article work unchanged in such a setting.

Two terms  $t, u$  are *unifiable* if there exists a substitution  $\sigma$  such that  $t\sigma = u\sigma$ . For example,  $\mathbf{a}$  and  $y \mathbf{a}$  are unifiable by taking  $y := \lambda x. x$ . (This works because terms are equal up to  $\beta$ -reduction.) Unification of higher-order terms is undecidable. Types, however, are isomorphic to first-order terms, and hence their unification problem is decidable. Moreover, if a unifier exists, then a most general unifier exists (up to the naming of variables). For example, the most general unifier for the unification problem  $\text{pair}(\alpha, \text{nat}) \stackrel{?}{=} \text{pair}(\text{int}, \beta)$  is  $\{\alpha \mapsto \text{int}, \beta \mapsto \text{nat}\}$ .

Finally, we define literals and clauses on top of terms. An atom is an equation  $s \approx t$  corresponding to an unordered pair  $\{s, t\}$ . (We reserve  $=$  for syntactic equality of terms.) A literal is an equation  $s \approx t$  or a disequation  $s \not\approx t$ . Given a predicate symbol  $\mathbf{p}$ , the literal  $\mathbf{p}(\vec{\alpha}) \vec{s} \approx \mathbf{T}$  is abbreviated to  $\mathbf{p}(\vec{\alpha}) \vec{s}$ , and its complement  $\mathbf{p}(\vec{\alpha}) \vec{s} \not\approx \mathbf{T}$  is abbreviated to  $\neg \mathbf{p}(\vec{\alpha}) \vec{s}$ . Moreover, a  *$\mathbf{p}$ -literal* is a literal of the form  $(\neg) \mathbf{p}(\vec{\tau}) \vec{t}$ . Note that it is possible in higher-order logic for a non- $\mathbf{p}$ -literal to contain  $\mathbf{p}$ , or even for the arguments  $\vec{t}$  of a  $\mathbf{p}$ -literal to contain  $\mathbf{p}$ . Given a literal  $L$ , we write  $\neg L$  to denote its complement, with  $\neg \neg L = L$ .

A clause  $C$  is a finite multiset of literals, written as  $L_1 \vee \cdots \vee L_n$  and interpreted disjunctively. Clauses are often defined as sets of literals, but multisets are better behaved with respect to substitution: If  $C$  has  $n$  literals, so has  $C\sigma$  regardless of whether  $\sigma$  unifies some of  $C$ 's literals. The type and term variables contained in a clause are implicitly quantified universally. (Within terms,  $\forall$  and  $\exists$  can be used to quantify over term variables.)

A type, term, or clause is *monomorphic* if it contains no type variables. A term or clause is *closed* if it contains no free term variables.

It is sometimes useful to encode a clause  $C$  as a formula. The formula  $[C]$  representing the clause  $C$  is defined by replacing the nonbold symbols  $\approx$ ,  $\not\approx$ , and  $\vee$  by their bold counterparts

$\approx$ ,  $\not\approx$ , and  $\mathbf{V}$ . This formula is uniquely defined up to the orientation of the equations and the order of the literals, neither of which affects the semantics.

**2.2. Semantics.** A *type interpretation*  $\mathcal{J}_{\text{ty}} = (\mathcal{U}, \mathcal{J}_{\text{ty}})$  consists of two components. The *universe*  $\mathcal{U}$  is a collection of nonempty sets, the *domains*. We require  $\mathcal{U}$  to contain the domain  $\{0, 1\}$ , where 0 represents falsehood and 1 represents truth. The function  $\mathcal{J}_{\text{ty}}$  associates with each  $n$ -ary type constructor  $\kappa$  a function  $\mathcal{J}_{\text{ty}}(\kappa) : \mathcal{U}^n \rightarrow \mathcal{U}$ , with the requirements that  $\mathcal{J}_{\text{ty}}(o) = \{0, 1\}$  and that the set  $\mathcal{J}_{\text{ty}}(\rightarrow)(\mathcal{D}_1, \mathcal{D}_2)$  is a subset of the (total) function space from  $\mathcal{D}_1$  to  $\mathcal{D}_2$  for all domains  $\mathcal{D}_1, \mathcal{D}_2 \in \mathcal{U}$ . The semantics is *standard* if  $\mathcal{J}_{\text{ty}}(\rightarrow)(\mathcal{D}_1, \mathcal{D}_2)$  is the entire function space for all  $\mathcal{D}_1, \mathcal{D}_2 \in \mathcal{U}$ . A *type valuation*  $\xi$  is a function that maps every type variable to a domain.

The *denotation* of a type in a type interpretation  $\mathcal{J}_{\text{ty}}$  under a type valuation  $\xi$  is defined by the recursive equations  $\llbracket \alpha \rrbracket_{\mathcal{J}_{\text{ty}}, \xi} = \xi(\alpha)$  and  $\llbracket \kappa(\vec{\tau}) \rrbracket_{\mathcal{J}_{\text{ty}}, \xi} = \mathcal{J}_{\text{ty}}(\kappa)(\llbracket \vec{\tau} \rrbracket_{\mathcal{J}_{\text{ty}}, \xi})$ . For monomorphic types  $\tau$ , the denotation does not depend on the valuation  $\xi$ , allowing us to write  $\llbracket \tau \rrbracket_{\mathcal{J}_{\text{ty}}}$  instead of  $\llbracket \tau \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}$ .

A type valuation  $\xi$  can be extended to be a *valuation* by additionally assigning an element  $\xi(x) \in \llbracket \tau \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}$  to each variable  $x : \tau$ . We will sometimes use partial functions as valuations if the values outside the function's domain are irrelevant. An *interpretation function*  $\mathcal{J}$  for a type interpretation  $\mathcal{J}_{\text{ty}}$  associates with each symbol  $f : \Pi \vec{\alpha}_m. \tau$  and domain tuple  $\vec{\mathcal{D}} \in \mathcal{U}^m$  a value  $\mathcal{J}(f, \vec{\mathcal{D}}) \in \llbracket \tau \rrbracket_{\mathcal{J}_{\text{ty}}, [\vec{\alpha} \rightarrow \vec{\mathcal{D}}]}$ . We require that the logical symbols are interpreted in the usual way in terms of 0 and 1. For example,  $\mathcal{J}(\perp) = 0$  and  $\mathcal{J}(\mathbf{V})(v, w) = \max\{v, w\}$ . Note that in the presence of  $\epsilon$  in the signature, every type  $\tau$  must be interpreted by a nonempty set, for  $\mathcal{J}(\epsilon(\tau)(\lambda x. \perp))$  to be defined.

We require the comprehension principle to hold: Every function designated by a  $\lambda$ -abstraction is contained in the domain associated with its type. We initially allow  $\lambda$ -abstractions to designate arbitrary elements of the domain. This enables us to define the denotation of a term. Then we impose restrictions to rule out undesirable  $\lambda$ -abstraction designations.

A  *$\lambda$ -designation function*  $\mathcal{L}$  for a type interpretation  $\mathcal{J}_{\text{ty}}$  is a function that maps a valuation  $\xi$  and a  $\lambda$ -abstraction of type  $\tau$  to an element of  $\llbracket \tau \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}$ . An *interpretation*  $\mathcal{J} = (\mathcal{J}_{\text{ty}}, \mathcal{J}, \mathcal{L})$  combines a type interpretation, an interpretation function, and a  $\lambda$ -designation function.

For an interpretation  $\mathcal{J}$  and a valuation  $\xi$ , the denotation of a term is defined recursively as  $\llbracket x \rrbracket_{\mathcal{J}, \xi} = \xi(x)$ ,  $\llbracket f(\vec{\tau}) \rrbracket_{\mathcal{J}, \xi} = \mathcal{J}(f, \llbracket \vec{\tau} \rrbracket_{\mathcal{J}_{\text{ty}}, \xi})$ ,  $\llbracket s t \rrbracket_{\mathcal{J}, \xi} = \llbracket s \rrbracket_{\mathcal{J}, \xi}(\llbracket t \rrbracket_{\mathcal{J}, \xi})$ , and  $\llbracket \lambda x. t \rrbracket_{\mathcal{J}, \xi} = \mathcal{L}(\xi, \lambda x. t)$ . For monomorphic closed terms  $t$ , the denotation does not depend on the valuation  $\xi$ , allowing us to write  $\llbracket t \rrbracket_{\mathcal{J}}$  instead of  $\llbracket t \rrbracket_{\mathcal{J}, \xi}$ .

An interpretation  $\mathcal{J}$  is *proper* if  $\llbracket \lambda x : v. t \rrbracket_{\mathcal{J}, \xi}(v) = \llbracket t \rrbracket_{\mathcal{J}, \xi[x \rightarrow v]}$  for every  $\lambda$ -abstraction  $\lambda x : v. t$ , every valuation  $\xi$ , and every value  $v \in \llbracket v \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}$ . We will assume throughout that all interpretations are proper and will construct only proper interpretations. If a type interpretation  $\mathcal{J}_{\text{ty}}$  and an interpretation function  $\mathcal{J}$  can be extended by a  $\lambda$ -designation function  $\mathcal{L}$  to an interpretation  $(\mathcal{J}_{\text{ty}}, \mathcal{J}, \mathcal{L})$ , then this  $\mathcal{L}$  is unique [Fit02, Proposition 2.18].

Given an interpretation  $\mathcal{J}$  and a valuation  $\xi$ , an equation  $s \approx t$  is true if  $\llbracket s \rrbracket_{\mathcal{J}, \xi}$  and  $\llbracket t \rrbracket_{\mathcal{J}, \xi}$  are equal and it is false otherwise. A disequation  $s \not\approx t$  is true if  $s \approx t$  is false. A clause is true if at least one of its literals is true. A clause set is true if all the clauses it contains are true. An interpretation  $\mathcal{J}$  is a *model* of a clause set  $N$ , written  $\mathcal{J} \models N$ , if  $N$  is true in  $\mathcal{J}$  for every valuation  $\xi$ .

A clause  $C$  is a *tautology* if  $J \models \{C\}$  for every interpretation  $J$ . It is *satisfiable* if there exists an interpretation  $J$  such that  $J \models \{C\}$ ; otherwise, it is *unsatisfiable*. Notice that these concepts are defined with respect to general, and not necessarily standard, interpretations.

It is sometimes convenient to assume that the interpretation of monomorphic types is injective on types—that is, for all monomorphic  $\tau, v$ , if  $\llbracket \tau \rrbracket_{\mathcal{J}_{\text{ty}}} = \llbracket v \rrbracket_{\mathcal{J}_{\text{ty}}}$ , then  $\tau = v$ . This assumption is reasonable because the elements of  $\mathcal{J}_{\text{ty}}(\kappa)$ , where  $\kappa \notin \{o, \rightarrow\}$ , are immaterial and can be renamed if desired and because the set-theoretic representation of functions, as nonempty sets of pairs, preserves this property. (The sets of pairs are nonempty thanks to the presence of  $\epsilon$  in the signature, as noted above.) We call this principle the *distinct domain assumption*.

### 3. HIDDEN-LITERAL-BASED ELIMINATION

In propositional logic [HJB11] and clausal first-order logic [VBH23], a hidden literal for a literal  $L$  and a clause set  $N$  is a literal that can be added or removed from any clause containing  $L$  without affecting its truth value in models of  $N$ . Several elimination techniques are based on hidden literals; in particular, hidden literal elimination removes hidden literals from clauses in which they occur.

**Example 3.1.** Consider the literal  $c$  and the clause set  $N = \{\neg a \vee b, \neg b \vee c\}$ . Then  $b$  is a hidden literal: Since  $b$  implies  $c$  according to  $N$ , we have that  $b \vee c$  and  $c$  have the same truth value in models of  $N$ . Similarly, since  $a$  implies  $c$  (by transitivity),  $a$  is also a hidden literal. Thus, hidden literal elimination would reduce the clause  $a \vee b \vee c$  to  $c$ .

The first-order definitions of hidden literals, hidden tautologies, hidden literal elimination, hidden tautology elimination, failed literal elimination, hidden tautology reduction, and failed literal reduction [VBH23] work verbatim in clausal higher-order logic, for both preprocessing and inprocessing. All the techniques preserve satisfiability and unsatisfiability. We call them collectively *hidden-literal-based elimination* (HLBE).

### 4. PREDICATE ELIMINATION

*Predicate elimination* (PE) [GO92, KK16, VBH23] is a set of techniques that remove all occurrences of some predicate symbol in a first-order problem by resolving clauses that contain it. Predicate elimination generalizes variable elimination in propositional logic [SP04, CS00]. In this section, we generalize two specific techniques to higher-order logic.

**4.1. Singular Predicate Elimination.** The first technique is called singular (or “non-self-referential”) predicate elimination. Definitions 4.1 to 4.5 below are adapted from monomorphic first-order logic.

**Definition 4.1.** A predicate symbol  $p$  *occurs deep* in a clause  $C$  if it occurs in a position other than as the head of the atom of a  $p$ -literal somewhere in  $C$ . The symbol  $p$  *occurs deep* in a clause set  $N$  if it occurs deep in one of its clauses  $C \in N$ .

**Definition 4.2.** A predicate symbol  $p$  is called *singular* for a clause  $C$  if these conditions are met:

- (1)  $C$  contains at most one  $p$ -literal;
- (2)  $p$  does not occur deep in  $C$ .

The symbol  $\mathbf{p}$  is *singular* for a clause set  $N$  if  $\mathbf{p}$  is singular for every clause contained in  $N$ .

**Definition 4.3.** The clause  $C = (\neg) \mathbf{p}(\vec{\tau}) \vec{t} \vee C'$  is *polymorphism-safe* for its literal  $(\neg) \mathbf{p}(\vec{\tau}) \vec{t}$  if all type variables occurring in  $C$  occur in  $\vec{\tau}$ . A clause  $C$  is *polymorphism-safe* for  $\mathbf{p}$  if it is polymorphism-safe for all its  $\mathbf{p}$ -literals. A clause set  $N$  is *polymorphism-safe* for  $\mathbf{p}$  if all the clauses it contains are polymorphism-safe for  $\mathbf{p}$ .

**Definition 4.4.** Let  $C = \mathbf{p}(\vec{\tau}) \vec{s}_n \vee C'$  and  $D = \neg \mathbf{p}(\vec{v}) \vec{t}_n \vee D'$ . The *flat resolvent* of  $C$  and  $D$  on  $\mathbf{p}(\vec{\tau}) \vec{s}_n$  and  $\neg \mathbf{p}(\vec{v}) \vec{t}_n$  is the clause  $(s_1 \not\approx t_1 \vee \dots \vee s_n \not\approx t_n \vee C' \vee D')\sigma$ , where  $\sigma$  is the most general unifier of  $\vec{\tau} \stackrel{?}{=} \vec{v}$ . The flat resolvent is not defined if  $\vec{\tau}$  and  $\vec{v}$  are not unifiable.

Flat resolvents were already present in the first-order setting [GO92, KK16, VBH23]. They are reminiscent of Huet's approach to unification in higher-order resolution [Hue73], which is also used by Benzmüller and colleagues [BK98, SB21].

**Definition 4.5.** Let  $M, N$  be clause sets and  $\mathbf{p}$  be a singular predicate for  $M$ . Let  $\rightsquigarrow$  be the following relation on clause set pairs and clause sets:

- (1)  $(M, \{(\neg) \mathbf{p}(\vec{\tau}) \vec{s} \vee C'\} \uplus N) \rightsquigarrow (M, N' \cup N)$  if  $N'$  is the set that consists of all clauses, up to variable renaming, that are flat resolvents on  $(\neg) \mathbf{p}(\vec{\tau}) \vec{s}$  and  $(\neg) \neg \mathbf{p}(\vec{v}) \vec{t}$  of  $(\neg) \mathbf{p}(\vec{\tau}) \vec{s} \vee C'$  and a clause  $(\neg) \neg \mathbf{p}(\vec{v}) \vec{t} \vee D'$  from  $M$  as premises. The premises' variables are renamed apart.
- (2)  $(M, N) \rightsquigarrow N$  if  $N$  contains no  $\mathbf{p}$ -literals.

The *resolved set*  $M \rtimes_{\mathbf{p}} N$  is the  $(\Sigma \setminus \{\mathbf{p}\})$ -clause set  $N'$  such that  $(M, N) \rightsquigarrow^* N'$ .

For finite sets  $M, N$ , the resolved set  $N'$  is reached in a finite number of steps, and it is unique up to variable renaming. The argument is as for first-order logic [VBH23, Lemma 4.4]. Note that the result may contain deep occurrences of  $\mathbf{p}$  if the initial set  $N$  contains such occurrences.

**Definition 4.6.** Let  $N$  be a clause set and  $\mathbf{p}$  be a singular predicate symbol for  $N$ . Let  $N_{\mathbf{p}}^+$  consist of all clauses belonging to  $N$  that contain a positive  $\mathbf{p}$ -literal, let  $N_{\mathbf{p}}^-$  consist of all clauses belonging to  $N$  that contain a negative  $\mathbf{p}$ -literal, let  $N_{\mathbf{p}} = N_{\mathbf{p}}^+ \cup N_{\mathbf{p}}^-$ , and let  $\bar{N}_{\mathbf{p}} = N \setminus N_{\mathbf{p}}$ .

**Definition 4.7.** Let  $N$  be a finite clause set that is polymorphism-safe for  $\mathbf{p}$  and  $\mathbf{p}$  be a singular predicate for  $N$ . *Singular predicate elimination* (SPE) of  $\mathbf{p}$  in  $N$  replaces  $N$  by the  $(\Sigma \setminus \{\mathbf{p}\})$ -clause set  $\bar{N}_{\mathbf{p}} \cup (N_{\mathbf{p}}^+ \rtimes_{\mathbf{p}} N_{\mathbf{p}}^-)$ .

SPE preserves satisfiability: The only clauses added are flat resolvents, and flat resolution is clearly sound. In first-order logic, SPE also preserves unsatisfiability [KK16, Theorem 1]. With a small restriction on polymorphism, this result extends to polymorphic higher-order logic. Also note that deep occurrences of  $\mathbf{p}$  are not possible in the result, because of the requirement that  $\mathbf{p}$  be a singular predicate for the input  $N$ .

**Example 4.8.** Thanks to the use of flat resolvents, the unification work is left to the proof calculus. This is convenient, because higher-order unification is undecidable and hence, in general, could not be done exhaustively in a preprocessing technique.

Consider the clause set  $N = \{\mathbf{p} z z \vee q z, \neg \mathbf{p} (f (y a)) (y (f a)), \neg q b\}$ . Applying SPE to  $\mathbf{p}$  transforms  $N$  into  $N' = \{f (y a) \not\approx z \vee y (f a) \not\approx z \vee q z, \neg q b\}$ . It is then the calculus's

task to enumerate values for  $z$  that solve the unification problem  $z \stackrel{?}{=} f(y \ a) \stackrel{?}{=} y(f \ a)$ . These values are

$$f \ a, f(f \ a), f(f(f \ a)), \dots$$

In  $\lambda$ -superposition [BBTV21], this enumeration would be the responsibility of the ERES inference rule. The first clause in  $N'$  could be simplified to  $y(f \ a) \not\approx f(y \ a) \vee q(f(y \ a))$ , by eliminating  $z$ . Then ERES would unify the two sides of the first literal, producing one conclusion per unifier. Since there are infinitely many unifiers, this would lead to infinitely many conclusions:

$$q(f \ a), q(f(f \ a)), q(f(f(f \ a))), \dots$$

Using dovetailing, this infinite enumeration can be interwoven with other inferences and other activities of the prover [VBB<sup>+</sup>22, Section 6].

**Example 4.9.** Consider the satisfiable clause set  $N = \{p(f \ z) \vee q \ z, \neg p(f \ a)\}$ . SPE of  $p$  transforms  $N$  into the equally satisfiable set  $N' = \{f \ a \not\approx f \ z \vee q \ z\}$ .

Note that although  $p$  is absent from  $N'$ , a predicate that meets its specification can be created based on the first clause of  $N$ , in which the  $p$ -literal is positive. This predicate is  $\lambda x. \exists y. x \approx f \ y \wedge \neg q \ y$ . In general, we would use all the clauses in which the  $p$ -literal is positive and ignore the other clauses. (Dually, we could have defined the predicate in terms of the clauses in which the  $p$ -literal is negative.) This predicate makes  $p$  true only when it must be true to satisfy the first clause—namely, when the clause’s non- $p$ -literal is false.

If we replace  $p$  with this  $\lambda$ -abstraction in  $N$  and  $\beta$ -reduce, we obtain the satisfiable set  $N'' = \{(\exists y. f \ z \approx f \ y \wedge \neg q \ y) \vee q \ z, \neg(\exists y. f \ a \approx f \ y \wedge \neg q \ y)\}$ . Nothing essential is lost by eliminating  $p$ —if we need  $p$ , we can use the  $\lambda$ -abstraction. This idea is the key to the proof of Theorem 4.11 below.

**Example 4.10.** Consider the clause set  $N = \{\neg y \ a, p \ a\}$ . It is easy to see that the set is unsatisfiable, by taking  $y := p$ . SPE of  $p$  transforms  $N$  into the set  $N' = \{\neg y \ a\}$ . Although  $y \ a$  is unifiable with the literal  $p \ a$ , the first clause is left unchanged by SPE. Like  $N$ ,  $N'$  is unsatisfiable. We cannot witness unsatisfiability by taking  $y := p$ . We can take  $y := \lambda x. x \approx a$ , simulating  $p$  on the input  $a$ . Alternatively, we can take  $y := \lambda x. \mathbf{T}$ . Indeed, we could have taken either instantiation for  $y$  to show that  $N$  is unsatisfiable, without exploiting the presence of  $p \ a$ .

Example 4.10 corroborates our choice of ignoring literals headed by a variable in the definition of SPE, focusing instead on  $p$ -literals. The intuition is that often  $p$  is unnecessary to have unsatisfiability, and when it is necessary it can be simulated by a  $\lambda$ -abstraction that does not contain it.

**Theorem 4.11.** *Let  $N$  be a finite clause set that is polymorphism-safe for  $p$  and  $p$  be a singular predicate symbol for  $N$ . Let  $N'$  be the result of applying SPE of  $p$  to  $N$ . Then  $N'$  is satisfiable if and only if  $N$  is satisfiable.*

*Proof.* The “if” direction follows immediately from the soundness of flat resolution. For the other direction, our strategy is inspired by Khasidashvili and Korovin [KK16, Theorem 1].

Let  $\mathcal{J} = (\mathcal{J}_{\text{ty}}, \mathcal{J}, \mathcal{L})$  be a model of  $N' = \bar{N}_p \cup (N_p^+ \times_p N_p^-)$ , with  $\mathcal{J}_{\text{ty}} = (\mathcal{U}, \mathcal{J}_{\text{ty}})$ . We assume without loss of generality that this model satisfies the distinct domain assumption (Section 2). We will define a new interpretation  $\mathcal{J}'$  that extends  $\mathcal{J}$  with a semantics for  $p$  and show that  $\mathcal{J}'$  is a model of  $N = \bar{N}_p \cup N_p^+ \cup N_p^-$ . To achieve this, we will separately show that (1)  $\mathcal{J}' \models \bar{N}_p$ ; (2)  $\mathcal{J}' \models N_p^+$ ; and (3)  $\mathcal{J}' \models N_p^-$ .



The interpretation  $\mathcal{J}' = (\mathcal{J}_{\text{ty}}, \mathcal{J}', \mathcal{L}')$  we construct is identical to  $\mathcal{J}$  except that  $\mathcal{J}'$  is extended so that  $\mathcal{J}'(\mathbf{p}, \vec{\mathcal{D}})$  and  $\mathcal{L}'$  are defined as follows, by mutual recursion.

We start with  $\mathcal{J}'(\mathbf{p}, \vec{\mathcal{D}})$ . We construct a suitable interpretation for  $\mathbf{p}$  as a curried  $n$ -ary predicate. Let  $\vec{\tau}'$  be monomorphic types such that  $\llbracket \vec{\tau}' \rrbracket_{\mathcal{J}_{\text{ty}}} = \vec{\mathcal{D}}$ . By the distinct domain assumption, these types are unique if they exist. If no such types exist, let  $\mathcal{J}'(\mathbf{p}, \vec{\mathcal{D}})$  be the predicate  $\llbracket \lambda \vec{x}. \perp \rrbracket_{\mathcal{J}}$ . This predicate is guaranteed to exist in the interpretation of the type of  $\mathbf{p}(\vec{\tau}')$  thanks to the comprehension principle (Section 2). The predicate plays the role of a placeholder for impossible interpretations of  $\mathbf{p}$ ; its value is irrelevant.

In the case where the types  $\vec{\tau}'$  exist, we will construct a right-hand side for  $\mathcal{J}'(\mathbf{p}, \vec{\mathcal{D}})$  using the same idea as in Example 4.9. To cope with polymorphism, we will filter out any clauses whose  $\mathbf{p}$ 's type arguments cannot be instantiated to  $\vec{\tau}'$  and instantiate the remaining clauses.

More precisely, let  $M$  be the smallest set such that for each clause  $C = \mathbf{p}(\vec{v}) \vec{t}_n \vee C'$  contained in  $N_{\mathbf{p}}$ , if there exists a substitution  $\sigma$  such that  $\vec{v}\sigma = \vec{\tau}'$ , then have  $M$  contain  $C\sigma$ . Notice that the polymorphism-safety hypothesis ensures that  $C\sigma$  is monomorphic and uniquely specified. This is desirable because we want to assign a unique right-hand side to  $\mathcal{J}'(\mathbf{p}, \vec{\mathcal{D}})$ .

We now define a term  $u$  whose interpretation  $\llbracket u \rrbracket_{\mathcal{J}} = \llbracket u \rrbracket_{\mathcal{J}'}$  will give us the right-hand side. Let  $\vec{x}_n$  be a tuple of fresh variables. With each clause  $\mathbf{p}(\vec{\tau}') \vec{t}_n \vee C'$  contained in  $M$  and whose free variables are  $\vec{y}$ , associate the formula

$$\exists \vec{y}. x_1 \approx t_1 \wedge \cdots \wedge x_n \approx t_n \wedge \neg [C']$$

(Recall from Section 2 that  $[C']$  denotes a formula representing the clause  $C'$ .) Let  $\varphi_1, \dots, \varphi_k$  be all such formulas, and let  $\varphi = \varphi_1 \mathbf{V} \cdots \mathbf{V} \varphi_k$ . Then we take  $u = \lambda \vec{x}. \varphi$ . This choice of  $u$  will ensure that  $\mathcal{J}'$  satisfies every clause in  $N_{\mathbf{p}}^+$ , and thanks to the comprehension principle, the predicate denoted by  $u$  is guaranteed to exist in the interpretation of  $\mathbf{p}(\vec{\tau}')$ 's type.

To finish the definition of  $\mathcal{J}'$ , we must specify  $\mathcal{L}'$ . Given a term  $t$ , let  $t[u/\mathbf{p}]$  denote the variant of the term  $t$  in which all occurrences of  $\mathbf{p}$  are replaced by the term  $u$  as defined above (for suitable types  $\vec{\tau}'$ ). For all valuations  $\xi$  and  $\lambda$ -abstractions  $\lambda x : v. t$ , we define  $\mathcal{L}'(\xi, \lambda x. t)$  as the function that maps each  $v \in \llbracket v \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}$  to  $\llbracket t[u/\mathbf{p}] \rrbracket_{\mathcal{J}, \xi[x \mapsto v]}$ . This function exists in the domain associated with the  $\lambda$ -abstraction's type because  $\mathcal{J}$  obeys the comprehension principle. Moreover, because  $t[u/\mathbf{p}]$  replaces  $\mathbf{p}$  by a term with the same semantics according to  $\mathcal{J}'$ , the interpretation  $\mathcal{J}'$  is proper.

We are now ready to tackle the three conditions we need to prove. To prove (1), we start from  $\mathcal{J} \models \bar{N}_{\mathbf{p}}$  and show  $\mathcal{J}' \models \bar{N}_{\mathbf{p}}$ . More precisely, we must show that  $\llbracket \bar{N}_{\mathbf{p}} \rrbracket_{\mathcal{J}', \xi}$  is true for every valuation  $\xi$ . This is obvious because  $\mathcal{J}$  and  $\mathcal{J}'$  only differ on  $\mathbf{p}$ , which does not occur in  $\bar{N}_{\mathbf{p}}$ .

To prove (2), we show that  $\mathcal{J}' \models N_{\mathbf{p}}^+$  holds by construction of  $\mathcal{J}'$ . More precisely, we must show that  $\llbracket N_{\mathbf{p}}^+ \rrbracket_{\mathcal{J}', \xi}$  is true for every valuation  $\xi$ . Let  $C = \mathbf{p}(\vec{v}) \vec{t} \vee C'$  be a clause in  $N_{\mathbf{p}}^+$ . By definition,

$$\mathcal{J}'(\mathbf{p}, \llbracket \vec{v} \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}) = \llbracket \lambda \vec{x}. \underbrace{\cdots \mathbf{V} (\exists \vec{y}. x_1 \approx t_1 \wedge \cdots \wedge x_n \approx t_n \wedge \neg [C']) \mathbf{V} \cdots}_{\psi} \rrbracket_{\mathcal{J}, \xi}$$

If  $\llbracket C' \rrbracket_{\mathcal{J}', \xi}$  is true, then clearly  $\llbracket C \rrbracket_{\mathcal{J}', \xi}$  is true. Otherwise, the literal  $\mathbf{p}(\vec{v}) \vec{t}$  is true, because the disjunct  $\psi$  is true. This can be seen by taking the values of  $\vec{y}$  under  $\xi$  as the existential

witnesses. The arguments  $\vec{t}$  in  $\mathbf{p}\langle\vec{v}\rangle \vec{t}$  passed for  $\vec{x}$  match those expected by the equalities  $x_i \approx t_i$ , and since  $\llbracket C' \rrbracket_{\mathcal{J}', \xi}$  is false,  $\llbracket \neg[C'] \rrbracket_{\mathcal{J}', \xi}$  is true.

For the proof of (3), the argument is similar to Khasidashvili and Korovin's [KK16, Theorem 1]. It relies on the presence of the flat resolvents  $N_{\mathbf{p}}^+ \times_{\mathbf{p}} N_{\mathbf{p}}^-$  in the result set. Let  $D = \neg \mathbf{p}\langle\vec{v}\rangle \vec{u} \vee D'$  be a clause in  $N_{\mathbf{p}}^-$ . If  $\llbracket D' \rrbracket_{\mathcal{J}, \xi}$  is true, then clearly  $\llbracket D \rrbracket_{\mathcal{J}', \xi}$  is true. Otherwise,  $\llbracket D' \rrbracket_{\mathcal{J}, \xi}$  is false, and  $\llbracket \neg \mathbf{p}\langle\vec{v}\rangle \vec{u} \rrbracket_{\mathcal{J}', \xi}$  is either true or false. In the true case,  $\llbracket D \rrbracket_{\mathcal{J}', \xi}$  is true, as desired. As for the remaining case, it is impossible for the following reason. If  $\llbracket \neg \mathbf{p}\langle\vec{v}\rangle \vec{u} \rrbracket_{\mathcal{J}', \xi}$  were false, this would mean  $\llbracket \mathbf{p}\langle\vec{v}\rangle \vec{u} \rrbracket_{\mathcal{J}', \xi}$  is true. By definition of  $\mathcal{J}'$ , this would then mean  $\llbracket \cdot \cdot \cdot \vee (\exists \vec{y}. u_1 \approx t_1 \wedge \cdots \wedge u_n \approx t_n \wedge \neg[C']) \vee \cdots \rrbracket_{\mathcal{J}, \xi}$  is true.

Suppose the displayed disjunct is one of those that makes the whole big disjunction true. This means that there exists a clause  $C = \mathbf{p}\langle\vec{\tau}\rangle \vec{t} \vee C'$  in  $N_{\mathbf{p}}^+$  such that  $\llbracket \vec{t} \rrbracket_{\mathcal{J}', \xi} = \llbracket \vec{u} \rrbracket_{\mathcal{J}', \xi}$ , and  $\llbracket C' \rrbracket_{\mathcal{J}', \xi}$  is false, where  $\vec{\tau}$  is unifiable with  $\vec{v}$ . Let  $\sigma$  be the most general unifier of  $\vec{\tau} \doteq \vec{v}$ . If  $C$  exists, the flat resolvent  $(t_1 \not\approx u_1 \vee \cdots \vee t_n \not\approx u_n \vee C' \vee D')\sigma$  of  $C$  and  $D$  must be false in  $\mathcal{J}'$  under  $\xi$ , and since it does not contain  $\mathbf{p}$ , it would be false in  $\mathcal{J}$  under  $\xi$  as well, contradicting the hypothesis that  $\mathcal{J}$  satisfies it.  $\square$

As Khasidashvili and Korovin observed, eliminating all singular predicates indiscriminately can dramatically increase the number of clauses in the problem. To prevent this explosion, Vukmirović et al. proposed the following criterion. Let  $K_{\text{tol}} \in \mathbb{N}$  be a tolerance parameter. The application of SPE from  $N$  to  $N'$  is allowed if  $\lambda(N') < \lambda(N) + K_{\text{tol}}$  or  $\mu(N') < \mu(N)$  or  $|N'| < |N| + K_{\text{tol}}$ , where  $\lambda(N)$  is the number of literals in  $N$  and  $\mu(N)$  is the sum, for all clauses  $C \in N$ , of the square of the number of unique variables in  $C$ .

**4.2. Defined Predicate Elimination.** The next technique we generalize from first-order logic to higher-order logic is called defined predicate elimination. It generalizes, in turn, the propositional technique of elimination by substitution [EB05].

Given a clause set  $N$ , the basic idea is that the set  $N_{\mathbf{p}}$  of clauses containing  $\mathbf{p}$  is partitioned between a definition set (or “gate”)  $G$  and the remaining clauses  $R$ . The definition set fully characterizes  $\mathbf{p}$  for all input in a unique way and can be seen as constituting a definition of the form  $\mathbf{p} \vec{x} \leftrightarrow \varphi$ , where the variables  $\vec{x}$  are distinct,  $\mathbf{p}$  does not occur in  $\varphi$ , and the variables in  $\varphi$  are all among  $\vec{x}$ . Because of clasification,  $G$  will usually consist of multiple clauses that together are equivalent to  $\mathbf{p} \vec{x} \leftrightarrow \varphi$  for some  $\varphi$ .

We define definition sets largely as in monomorphic first-order logic, but with additional requirements on the type arguments and type variables.

**Definition 4.12.** Let  $G$  be a clause set and  $\mathbf{p}$  be a predicate symbol. The set  $G$  is a *definition set* for  $\mathbf{p}$  if

- (1)  $\mathbf{p}$  is singular for  $G$ ;
- (2)  $G$  consists of clauses of the form  $(\neg) \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C'$  up to variable renaming, where  $\vec{\alpha}$  are distinct type variables and  $\vec{x}$  are distinct term variables;
- (3) the type variables in  $C'$  are all among  $\vec{\alpha}$ ;
- (4) the term variables in  $C'$  are all among  $\vec{x}$ ;
- (5) all clauses in  $G_{\mathbf{p}}^+ \times_{\mathbf{p}} G_{\mathbf{p}}^-$  are tautologies; and
- (6)  $E(\vec{\tau}, \vec{\tau})$  is unsatisfiable, where the *environment*  $E(\vec{\alpha}, \vec{x})$  consists of all subclauses  $C'$  of any  $(\neg) \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C' \in G$ ,  $\vec{\tau}$  is a tuple of distinct nullary type constructors substituted in for  $\vec{\alpha}$ , and  $\vec{\tau}$  is a tuple of distinct fresh symbols substituted in for  $\vec{x}$ .

Intuitively, conditions 1 and 2 check that the definition set looks like the clausification of a definition; conditions 3, 4, and 5 check that the definition is not overconstrained (e.g.,  $\mathbf{p} a$  is not required to be both true and false); and condition 6 checks that it is not underconstrained (e.g.,  $\mathbf{p} a$  is not unspecified).

**Definition 4.13.** Given a definition set  $G$  for  $\mathbf{p}$ , its *associated definition* is the formula  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$ , up to variable renaming, where  $\varphi$  is the disjunction  $\varphi_1 \vee \dots \vee \varphi_n$  of all formulas  $\varphi_j$  of the form  $\neg[C']$  such that  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C'$  is contained in  $G$  up to variable renaming.

Note that our notion of definition set ensures that in Definition 4.13 the type variables  $\vec{\alpha}$  are distinct, the term variables  $\vec{x}$  are distinct,  $\mathbf{p}$  does not occur in  $\varphi$ , the type variables in  $\varphi$  are all among  $\vec{\alpha}$ , and the variables in  $\varphi$  are all among  $\vec{x}$ .

**Example 4.14.** For the formula  $\mathbf{p} x y \leftrightarrow \mathbf{q} x \vee r y$ , clausification would produce the clause set  $\{\neg \mathbf{p} x y \vee \mathbf{q} x \vee r y, \mathbf{p} x y \vee \neg \mathbf{q} x, \mathbf{p} x y \vee \neg r y\}$ , which qualifies as a definition set for  $\mathbf{p}$ . The associated definition is  $\mathbf{p} x y \leftrightarrow \neg \neg \mathbf{q} x \vee \neg \neg r y$ .

**Lemma 4.15.** *Let  $G$  be a definition set for  $\mathbf{p}$ . Then  $G$  is equivalent to the definition  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$  associated with  $G$ .*

*Proof.* We will show that under every valuation  $\xi$ , any model of  $G$  is a model of  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$  and vice versa.

Let  $J \models G$ . We will show that  $\llbracket \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \rrbracket_{J,\xi} = \llbracket \varphi \rrbracket_{J,\xi}$ . If  $\llbracket \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \rrbracket_{J,\xi}$  is false, then for each clause  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C' \in G$ , we have that  $\llbracket C' \rrbracket_{J,\xi}$  must be true. This in turn makes the right-hand side  $\varphi$  false. Otherwise,  $\llbracket \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \rrbracket_{J,\xi}$  is true. Then for each clause  $\neg \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C' \in G$ , we have that  $\llbracket C' \rrbracket_{J,\xi}$  must be true. By condition 6, there must exist a clause  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee D' \in G$  such that  $\llbracket D' \rrbracket_{J,\xi}$  is false. This means that  $\llbracket \neg D' \rrbracket_{J,\xi}$  is true and hence the entire disjunction  $\varphi$  is true, as desired.

Now let  $J \models \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$ . We need to show that  $J \models G$ . We will first prove the case of clauses in  $G$  where the  $\mathbf{p}$ -literal is positive; then we will consider the negative case. Let  $C = \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C' \in G$ . If  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x}$  is true, then  $C$  is true, as desired. Otherwise, from  $J \models \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$  we have that  $\varphi$  is false. This means that all of its disjuncts are false and hence that  $C'$  is true, meaning that  $C$  is true. For the remaining case, let  $C = \neg \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee C' \in G$ . If  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x}$  is false, then  $C$  is true, as desired. Otherwise, from  $J \models \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$  we have that  $\varphi$  is true. This means that there exists a clause  $D = \mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \vee D' \in G$  such that  $\llbracket D' \rrbracket_{J,\xi}$  is false. By condition 5, the resolvent  $C' \vee D'$  of  $C$  and  $D$  must be a tautology. Hence  $C'$  is true and thus  $C$  is true, as desired.  $\square$

Once a definition is identified, it is expanded in the remaining clauses  $R$ . For  $\mathbf{p}$ -literals in  $R$ , this is achieved as in first-order logic using flat resolution. For deeper occurrences of  $\mathbf{p}\langle\vec{\tau}\rangle$  in  $R$ , which may arise in higher-order logic, this is achieved by replacing them by the  $\lambda$ -abstraction  $\lambda \vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}$ . An alternative would be to replace *all* occurrences of  $\mathbf{p}\langle\vec{\tau}\rangle$  and not only deep occurrences, but this would leave more work for the clausifier.

**Definition 4.16.** Let  $N$  be a clause set and  $\mathbf{p}$  be a predicate symbol. Let  $G \subseteq N$  be a definition set for  $\mathbf{p}$  with associated definition  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$ . Let  $R = N_{\mathbf{p}} \setminus G$ . *Defined predicate elimination* (DPE) of  $\mathbf{p}$  in  $N$  replaces  $N$  by  $\bar{N}_{\mathbf{p}} \cup (G \times_{\mathbf{p}} R)[\mathbf{p}\langle\vec{\tau}\rangle \mapsto \lambda \vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}]$ .

The key result is that DPE preserves satisfiability and unsatisfiability. The proof builds on three lemmas.

**Lemma 4.17.** *Let  $G$  be a definition set for  $\mathbf{p}$  and  $R$  be an arbitrary clause set. If  $(G, R) \rightsquigarrow (G, R')$ , then  $G \cup R$  and  $G \cup R'$  are equivalent.*

*Proof.* The proof is essentially as in the first-order case [VBH23, Lemma 4.11].  $\square$

**Lemma 4.18.** *Let  $G$  be a definition set for  $\mathbf{p}$  with associated definition  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$ , and let  $R$  be a clause set. Then  $G \cup R$  and  $G \cup R[\mathbf{p}\langle\vec{\tau}\rangle \mapsto \lambda\vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}]$  are equivalent.*

*Proof.* By Lemma 4.15,  $G$  entails the characterization  $\mathbf{p}\langle\vec{\alpha}\rangle \vec{x} \leftrightarrow \varphi$ . Hence, by functional extensionality,  $G$  entails  $\mathbf{p}\langle\vec{\alpha}\rangle \approx \lambda\vec{x}. \varphi$ . Thus, in any model of  $G$ ,  $\mathbf{p}\langle\vec{\alpha}\rangle$  has the same interpretation as  $\lambda\vec{x}. \varphi$ . In particular, this applies to their instances:  $\mathbf{p}\langle\vec{\tau}\rangle$  has the same interpretation as  $\lambda\vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}$ .  $\square$

**Lemma 4.19.** *Let  $G$  be a definition set for  $\mathbf{p}$  and  $R$  be a clause set with no occurrences of  $\mathbf{p}$ . Then  $G \cup R$  is satisfiable if and only if  $R$  is satisfiable.*

*Proof.* The proof is essentially as in the first-order case [VBH23, Lemma 4.12].  $\square$

**Theorem 4.20.** *The result of applying DPE to a finite clause set  $N$  is satisfiable if and only if  $N$  is satisfiable.*

*Proof.* Let  $\mathbf{p}$  be a predicate symbol and  $G \subseteq N$  be the definition set used by DPE. Let  $R = N_{\mathbf{p}} \setminus G$ . The core of DPE is the computation of  $G \times_{\mathbf{p}} R$ , via a derivation  $(G, R) \rightsquigarrow^n (G, R') \rightsquigarrow R'$ . Applying Lemma 4.17  $n$  times, we get that  $G \cup R$  is equivalent to  $G \cup R'$ . Moreover, by Lemma 4.18,  $G \cup R'$  is equivalent to  $G \cup R'[\mathbf{p}\langle\vec{\tau}\rangle \mapsto \lambda\vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}]$ . Finally, by Lemma 4.19,  $G \cup R'[\mathbf{p}\langle\vec{\tau}\rangle \mapsto \lambda\vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}]$  is equivalent to  $R'[\mathbf{p}\langle\vec{\tau}\rangle \mapsto \lambda\vec{x}. \varphi\{\vec{\alpha} \mapsto \vec{\tau}\}]$ .  $\square$

**4.3. Portfolio Predicate Elimination.** A reasonable strategy for applying predicate elimination is to use a portfolio of DPE and SPE, first trying to apply DPE and, if this fails, trying SPE as a fallback.

**Definition 4.21.** Let  $N$  be a clause set and  $\mathbf{p}$  be a predicate symbol. If there exists a definition set  $G \subseteq N$  for  $\mathbf{p}$ , *portfolio predicate elimination* (PPE) on  $\mathbf{p}$  applies DPE on  $\mathbf{p}$ . Otherwise, if  $\mathbf{p}$  is singular in  $N$ , PPE applies SPE on  $\mathbf{p}$ . In all other cases, PPE is not applicable.

Like SPE and DPE (Theorems 4.11 and 4.20), PPE can be used as a preprocessor without affecting satisfiability. As for inprocessing, Vukmirović et al. [VBH23] explained that under a reasonable condition, the first-order version of PPE can be used at any point during proof search in a superposition prover without compromising refutational completeness. Inspection of the proofs reveals that the same applies to higher-order PPE and  $\lambda$ -superposition.

## 5. BLOCKED CLAUSE ELIMINATION

In propositional logic, a powerful technique for simplifying a clause set is to identify and remove so-called blocked clauses. These are clauses whose resolvents with other clauses in the set are all tautologies. Removing such clauses preserves unsatisfiability. Blocked clause elimination has been extended to first-order logic with equality by Kiesl et al. [KSS<sup>+</sup>17]. They call their key notion “equality-blocked clauses,” but since we consider only a logic with equality, we simply call these clauses “blocked.”

Blocked clause elimination has been shown by Vukmirović et al. [VBH23, Section 5] to be incompatible with the saturation loop of a superposition prover. Nevertheless, the technique can still be used effectively as a preprocessor, or even as an inprocessing technique within the prover’s saturation loop at the cost of potential divergence on some unsatisfiable problems.

Our extension to polymorphic higher-order logic is based on a slightly weaker definition of blocked clause than Kiesl et al. We were unsuccessful at showing that a generalization of blocked clause elimination based on their concept preserves unsatisfiability with respect to general interpretations. The notion we propose allows the generalization.

**Definition 5.1.** Let  $C = L \vee C'$  and  $D = L' \vee D'$  be clauses such that

- (1) the atom of  $L$  is  $\mathbf{p}\langle\vec{\tau}\rangle \vec{s}_n$ ;
- (2) the atom of  $L'$  is  $\mathbf{p}\langle\vec{v}\rangle \vec{t}_n$ ;
- (3) the literal  $L'$  is of opposite polarity to  $L$ ;
- (4)  $C$  and  $D$  have no (type or term) variables in common; and
- (5)  $\sigma$  is the most general unifier of  $\vec{\tau} \stackrel{?}{=} \vec{v}$ .

The clause  $((\bigvee_{j=1}^n s_j \not\approx t_j) \vee C' \vee D')\sigma$  is a *binary flat  $L$ -resolvent* of  $C$  and  $D$ .

We already see a first key difference with Keisl et al.: They consider  $n$ -ary flat resolvents, whereas we need to consider only binary resolvents, for reasons illustrated below (Example 5.4). Another, more superficial difference is that our definition is generalized to polymorphic higher-order logic.

**Definition 5.2.** Let  $L = (\neg)\mathbf{p}\langle\vec{\tau}\rangle \vec{s}$  be a predicate literal,  $C = L \vee C'$  be a clause, and  $N$  be a clause set. Let  $N'$  consist of all clauses from  $N \setminus \{C\}$  with their type and term variables renamed so that  $N'$  shares no variables with  $C$ . The clause  $C$  is *blocked* by  $L$  in the set  $N$  if the following conditions are met:

- (1)  $C$  is polymorphism-safe for  $L$ ;
- (2)  $N$  contains no deep occurrences of  $\mathbf{p}$ ;
- (3)  $C'$  contains no  $\mathbf{p}$ -literals with the same polarity as  $L$ ; and
- (4) all binary flat  $L$ -resolvents between  $C$  and clauses in  $N'$  are tautologies.

We now see another key difference with Keisl et al.: They have no restriction corresponding to condition 3 of Definition 5.2. In this respect, our notion is less powerful than theirs. (They also have no restriction corresponding to conditions 1 and 2, but these conditions are trivially satisfied in a monomorphic first-order setting.) As a result, their notion and our notion of blocked clause are incomparable in strength.

**Example 5.3.** This example is based on Keisl et al. [KSS<sup>+</sup>17, Example 1]. Let  $C = \neg\mathbf{p} \vee \mathbf{q}$ , and take  $N = \{C, \mathbf{p} \vee \neg\mathbf{q}, \neg\mathbf{q} \vee \mathbf{r}\}$  as the clause set. The clause  $C$  is blocked by  $\neg\mathbf{p}$  in  $N$  according to Definition 5.2 because the only resolvent of  $C$  on  $\neg\mathbf{p}$  is the tautology  $\mathbf{q} \vee \neg\mathbf{q}$  resulting from resolution against  $\mathbf{p} \vee \neg\mathbf{q}$ . The clause  $C$  is also blocked according to the definition in Kiesl et al.

**Example 5.4.** The next example is also based on Keisl et al. [KSS<sup>+</sup>17, Example 4]:  $C = \mathbf{p} x y \vee \mathbf{p} y x$ ,  $D = \neg\mathbf{p} x y \vee \neg\mathbf{p} y x$ , and  $N = \{C, D\}$ . The set  $N$  is unsatisfiable, because  $C$  entails  $\mathbf{p} x x$  and  $D$  entails  $\neg\mathbf{p} x x$ . On the other hand,  $D$  alone is satisfiable. Hence, removing  $C$  from  $N$  does *not* preserve unsatisfiability, and therefore  $C$  should *not* be considered blocked. With Definition 5.2,  $C$  correctly cannot be blocked on a  $\mathbf{p}$ -literal by condition 3, because of the presence of another  $\mathbf{p}$ -literal in the clause. With Kiesl et al.,

this condition is missing, but since they consider all  $n$ -ary resolvents, the nontautological resolvent  $\mathbf{p} x x$  is computed. In both cases,  $C$  is correctly considered not blocked.

**Example 5.5.** Let  $C = \mathbf{p}a \vee \mathbf{p}b \vee \neg \mathbf{q}$ ,  $D = \neg \mathbf{p}x \vee \mathbf{q}$ , and  $N = \{C, D\}$ . With Definition 5.2, condition 3 prevents  $C$  from being considered blocked. In contrast, the clause is considered blocked by Kiesl et al.

We will now show that removing a blocked clause from a clause set preserves the set's unsatisfiability. Our strategy is loosely inspired by Kiesl et al. [KSS<sup>+</sup>17, Section 4].

**Definition 5.6.** Let  $\mathcal{J} = (\mathcal{J}_{\text{ty}}, \mathcal{J}, \mathcal{L})$  be an interpretation and  $L \vee C'$  be a clause that is polymorphism-safe for  $L$  and where  $L = (\neg) \mathbf{p}(\vec{\tau}) \vec{s}_n$ . Let  $\vec{y}$  be the tuple of all free variables in  $L \vee C'$  and  $\vec{x}_n$  be a tuple of fresh variables. The interpretation  $\mathcal{J}^* = (\mathcal{J}_{\text{ty}}, \mathcal{J}^*, \mathcal{L}^*)$  obtained by *flipping* the truth value of  $L$  in  $L \vee C'$  is defined as follows by mutual recursion. We let  $\mathcal{J}^*$  be the function defined as follows:

$$\mathcal{J}^*(\mathbf{f}, \vec{\mathcal{D}}) = \begin{cases} \llbracket \lambda \vec{x}_n. \varphi \rrbracket_{\mathcal{J}, \xi} & \text{if } \mathbf{f} = \mathbf{p} \text{ and } \llbracket \vec{\tau} \rrbracket_{\mathcal{J}_{\text{ty}}, \xi} = \vec{\mathcal{D}} \text{ for some type valuation } \xi \\ \mathcal{J}(\mathbf{f}, \vec{\mathcal{D}}) & \text{otherwise} \end{cases}$$

where  $\varphi$  is defined by

$$\varphi = \begin{cases} \mathbf{p}(\vec{\tau}) \vec{x} \mathbf{V} (\exists \vec{y}. x_1 \approx s_1 \wedge \cdots \wedge x_n \approx s_n \wedge \neg [C']) & \text{if } L \text{ is positive} \\ \mathbf{p}(\vec{\tau}) \vec{x} \wedge (\forall \vec{y}. x_1 \approx s_1 \wedge \cdots \wedge x_n \approx s_n \rightarrow [C']) & \text{if } L \text{ is negative} \end{cases}$$

Moreover, for all valuations  $\xi$  and  $\lambda$ -abstractions  $\lambda x : v. t$ , we let  $\mathcal{L}^*(\xi, \lambda x. t)$  be the function that maps each  $v \in \llbracket v \rrbracket_{\mathcal{J}_{\text{ty}}, \xi}$  to  $\llbracket t \rrbracket_{\mathcal{J}^*, \xi[x \mapsto v]}$ .

This definition introduces a well-formed interpretation. Because  $L \vee C'$  is polymorphism-safe for  $L$ , the right-hand side  $\llbracket \lambda \vec{x}_n. \varphi \rrbracket_{\mathcal{J}, \xi}$  of  $\mathcal{J}^*(\mathbf{f}, \vec{\mathcal{D}})$  is uniquely defined. Moreover, the comprehension principle guarantees that the corresponding predicate exists in the interpretation of  $\mathbf{p}$ 's type. Similarly, the function that provides the interpretation for a  $\lambda$ -abstraction  $\lambda x. t$  exists in the domain associated with the  $\lambda$ -abstraction's type. This is because the semantics in  $\mathcal{J}^*$  of any occurrences of  $\mathbf{p}$  in  $t$  corresponds to the semantics in  $\mathcal{J}$  of  $\lambda \vec{x}_n. \varphi$ , and  $\mathcal{J}$  is a well-formed interpretation.

The intuition behind  $\mathcal{J}^*$  is that whenever the clause  $L \vee C'$  is blocked and  $\mathcal{J} \models N \setminus \{L \vee C'\}$ , we have  $\mathcal{J}^* \models N$ . Since adding the blocked clause preserves satisfiability, removing it preserves unsatisfiability.

**Example 5.7.** We will try to justify the definition of  $\varphi$  above with an example. Consider the clause set  $\{\mathbf{p} a, \neg \mathbf{p} z \vee z \approx a \vee z \approx b\}$  and an interpretation  $\mathcal{J}$  that maps  $\mathbf{p}$  to the uniformly false predicate. Clearly,  $\mathcal{J}$  is not a model of the first clause,  $\mathbf{p} a$ . The interpretation  $\mathcal{J}^*$  obtained by flipping the truth value of the literal  $\mathbf{p} a$  in the first clause interprets  $\mathbf{p}$  in the same way as  $\mathcal{J}$  interprets

$$\lambda x. \mathbf{p} x \mathbf{V} (x \approx a \wedge \neg \perp)$$

In other words,  $\mathcal{J}^*$  makes  $\mathbf{p}$  true for arguments interpreted as equal to  $a$  and false otherwise. Intuitively, the interpretation  $\mathcal{J}$  is flipped to satisfy the clause  $\mathbf{p} a$ .

Next, assume instead that  $\mathcal{J}$  interprets  $\mathbf{p}$  as the uniformly true predicate, and consider the interpretation  $\mathcal{J}^*$  obtained by flipping the truth value of  $\neg \mathbf{p} z$  in the second clause,  $\neg \mathbf{p} z \vee z \approx a \vee z \approx b$ . Then  $\mathcal{J}^*$  interprets  $\mathbf{p}$  in the same way as  $\mathcal{J}$  interprets

$$\lambda x. \mathbf{p} x \wedge (\forall z. x \approx z \rightarrow z \approx a \vee z \approx b)$$

This means that  $\mathcal{J}^*$  makes  $\mathbf{p}$  true for arguments interpreted as equal to  $\mathbf{a}$  or  $\mathbf{b}$  and false otherwise. Intuitively, the interpretation  $\mathcal{J}$  is flipped to satisfy the clause  $\neg \mathbf{p}z \vee z \approx \mathbf{a} \vee z \approx \mathbf{b}$ ; whenever  $z \not\approx \mathbf{a}$  and  $z \not\approx \mathbf{b}$ , we have  $\neg \mathbf{p}z$ .

**Lemma 5.8.** *Let  $N$  be a clause set and  $C$  be a clause contained in  $N$  such that  $C$  has no variables in common with  $N \setminus \{C\}$ . Assume  $C$  is blocked by  $L$  in  $N$ . Let  $\mathcal{J}$  be an interpretation and  $\mathcal{J}^*$  be the interpretation obtained by flipping the truth value of  $L$ . For every  $D \in N \setminus \{C\}$ , if  $\mathcal{J} \models D$ , then  $\mathcal{J}^* \models D$ .*

*Proof.* Let  $C = L \vee C'$  where  $L = (\neg) \mathbf{p}(\vec{\tau}) \vec{s}_n$  (as per condition 1 of Definition 5.1). Assuming that  $\mathcal{J} \models D$ , we will show that  $\llbracket D \rrbracket_{\mathcal{J}^*, \xi}$  is true for every valuation  $\xi$ .

Let  $L'$  be a literal of  $D$  of opposite polarity to  $L$ , whose atom is  $\mathbf{p}(\vec{v}) \vec{t}$ , and such that, for each  $j$ ,  $\llbracket L' \rrbracket_{\mathcal{J}, \xi}$  is true and  $\llbracket L' \rrbracket_{\mathcal{J}^*, \xi}$  is false. Intuitively,  $L'$  is a literal whose truth value goes from true in  $\mathcal{J}$  to false in  $\mathcal{J}^*$ . We distinguish two cases: The case where such a literal  $L'$  exists and the case where it does not.

**CASE WHERE  $L'$  DOES NOT EXIST:** If  $D$  contains literals  $L'$  of opposite polarity to  $L$ , then either  $\llbracket L' \rrbracket_{\mathcal{J}, \xi}$  is false and then the truth value of  $\llbracket L' \rrbracket_{\mathcal{J}^*, \xi}$  is irrelevant or  $\llbracket L' \rrbracket_{\mathcal{J}, \xi}$  is true and then  $\llbracket L' \rrbracket_{\mathcal{J}^*, \xi}$  is true. As for the occurrences of  $\mathbf{p}$  in  $D$  that have the same polarity as  $L$ , if these were true in  $\mathcal{J}$ , they are by construction also true in  $\mathcal{J}^*$ . Finally,  $D$  contains no deep occurrences of  $\mathbf{p}$  (by condition 2 of Definition 5.2). Since all the literals either keep their truth value or go from false to true when taking the step from  $\mathcal{J}$  to  $\mathcal{J}^*$ , we have that  $\llbracket D \rrbracket_{\mathcal{J}^*, \xi}$  is true.

**CASE WHERE  $L'$  EXISTS:** We will show that  $\llbracket D \rrbracket_{\mathcal{J}^*, \xi} = \llbracket D \rrbracket_{\mathcal{J}, \xi}$  even if  $L'$  has gone from true in  $\mathcal{J}$  to false in  $\mathcal{J}^*$  under  $\xi$ . Since  $\mathcal{J} \models D$ ,  $\llbracket D \rrbracket_{\mathcal{J}^*, \xi}$  will then be true.

Since  $\mathcal{J}^*$  flips the truth value of  $L'$ , by construction of  $\mathcal{J}^*$ , this flipping must be triggered by  $C$ . Hence, there must exist a valuation  $\xi'$  such that  $\llbracket \tau_i \rrbracket_{\mathcal{J}, \xi'} = \llbracket v_i \rrbracket_{\mathcal{J}, \xi}$  and  $\llbracket s_j \rrbracket_{\mathcal{J}, \xi'} = \llbracket t_j \rrbracket_{\mathcal{J}, \xi}$  for every  $i, j$ . Let  $\xi''$  be the valuation that coincides with  $\xi'$  on  $C$ 's free variables and with  $\xi$  on  $D$ 's free variables. Let  $D = L' \vee D'$ .

Since  $C$  is blocked by  $L$  in  $N$ , all binary flat  $L$ -resolvents of  $C$  with clauses from  $N$  are tautologies (by condition 4 of Definition 5.2). In particular, consider the binary flat  $L$ -resolvent of  $C$  and  $D$  of the form  $((\bigvee_{j=1}^n s_j \not\approx t_j) \vee C' \vee D')\sigma$ , where  $\sigma$  is the most general unifier of  $\vec{\tau} \stackrel{?}{=} \vec{v}$ . This binary flat  $L$ -resolvent, which must exist by the five conditions of Definition 5.1, is a tautology, and it must be satisfied by  $\mathcal{J}^*$  under  $\xi$ .

By definition of  $\xi''$ ,  $\llbracket s_j \not\approx t_j \rrbracket_{\mathcal{J}, \xi''}$  must be false, and since the terms  $\vec{s}, \vec{t}$  do not contain  $\mathbf{p}$  (by condition 2 of Definition 5.2),  $\llbracket s_j \not\approx t_{ij} \rrbracket_{\mathcal{J}^*, \xi''}$  must be false as well. Moreover, by construction of  $\mathcal{J}^*$ , the only way for the interpretation of  $\mathbf{p}(\vec{\tau}) \vec{s}_n$  in  $\mathcal{J}^*$  under  $\xi''$  to differ from that in  $\mathcal{J}$  under  $\xi''$  is if  $\llbracket C' \rrbracket_{\mathcal{J}, \xi''}$  is false, and since  $C'$  contains only occurrences of  $\mathbf{p}$  of opposite polarity to  $L$  (by conditions 2 and 3 of Definition 5.2),  $\llbracket C' \rrbracket_{\mathcal{J}^*, \xi''}$  must still be false after we flipped  $L$  to make it true.

Finally, since both  $\llbracket s_j \not\approx t_j \rrbracket_{\mathcal{J}, \xi''}$  and  $\llbracket C' \rrbracket_{\mathcal{J}^*, \xi''}$  are false and the binary flat  $L$ -resolvent of  $C$  and  $D$  is a tautology,  $D'\sigma$  must be true in  $\mathcal{J}^*$  under every valuation. Since  $\sigma$  is a most general unifier and  $\xi''$  assigns the same semantics to  $\vec{\tau}$  and  $\vec{v}$ , effectively “unifying” them, we also have that  $\llbracket D' \rrbracket_{\mathcal{J}^*, \xi''}$  is true and hence  $\llbracket D' \rrbracket_{\mathcal{J}^*, \xi}$  is true. Thus  $\llbracket D \rrbracket_{\mathcal{J}^*, \xi}$  is true, as desired.  $\square$

**Lemma 5.9.** *Let  $N$  be a clause set and  $C$  be a clause contained in  $N$ . If  $C$  is blocked by a literal in  $N$ , then  $N \setminus \{C\}$  is satisfiable if and only if  $N$  is satisfiable.*

*Proof.* The “if” direction is trivial. For the other direction, let  $N'$  consist of all clauses from  $N \setminus \{C\}$  with their type and term variables renamed so that they share no variables with  $C$ . Let  $\mathcal{J}$  be a model of  $N'$ . By Lemma 5.8,  $\mathcal{J}^*$  is a model of each  $D \in N'$ .

We also need to show that  $\mathcal{J}^*$  is a model of  $C$ . Specifically, we must show that  $\llbracket C \rrbracket_{\mathcal{J}^*, \xi}$  is true for any valuation  $\xi$ . Let  $C = L \vee C'$ . If  $\llbracket C' \rrbracket_{\mathcal{J}^*, \xi}$  is true, we are done. Otherwise, first suppose  $L$  is positive. Then  $\xi$  provides the necessary witnesses for the  $\exists$  quantifier in the definition of  $\varphi$  in Definition 5.6, making the interpretation of  $\mathbf{p}$  by  $\mathcal{J}^*$  true in that case, as in the first part of Example 5.7. Hence  $\llbracket L \rrbracket_{\mathcal{J}^*, \xi}$  is true, and thus  $\llbracket C \rrbracket_{\mathcal{J}^*, \xi}$  is true. Next, suppose  $L$  is negative. Then  $\xi$  provides a counterexample to the  $\forall$  quantifier in the definition of  $\varphi$ , making the interpretation of  $\mathbf{p}$  by  $\mathcal{J}^*$  false in that case, as in the second part of Example 5.7. Hence  $\llbracket L \rrbracket_{\mathcal{J}^*, \xi}$  is true, and thus  $\llbracket C \rrbracket_{\mathcal{J}^*, \xi}$  is true.

Since  $\mathcal{J}^* \models N' \cup \{C\}$ , we have that  $N$  is satisfiable.  $\square$

**Definition 5.10.** Given a finite clause set, *blocked clause elimination* (BCE) repeatedly removes blocked clauses until no such clauses remain.

The procedure is confluent and hence yields a unique result. This is easy to see because removing a blocked clause will only make more clauses blocked; it can never “unblock” a clause.

**Theorem 5.11.** *The result of applying BCE to a clause set  $N$  is satisfiable if and only if  $N$  is satisfiable.*

*Proof.* This follows by iteration of Lemma 5.9.  $\square$

## 6. QUASIPURE LITERAL ELIMINATION

*Pure literal elimination* (PLE) is one of the simplest optimizations implemented in SAT solvers. It is a special case of variable elimination [SP04, CS00]: If a given variable always occurs with the same polarity in a problem, the solver can assign it that polarity without loss of generality, making all the clauses that contain it tautologies. PLE consists of recursively deleting all such clauses. PLE’s generalization to first-order logic considers literals  $(\neg) \mathbf{p}(\vec{s})$ , where the arguments  $\vec{s}$  are ignored by the analysis; only the polarity is considered. The same idea carries over to higher-order logic.

**Example 6.1.** Consider the clause set  $N = \{\mathbf{p}x \vee \mathbf{q}ax, \mathbf{p}(fx), \neg \mathbf{q}aa\}$ . Since  $\mathbf{p}$  occurs only positively in  $N$ , PLE removes the two first clauses. At that point,  $\mathbf{q}$  occurs only negatively in the remaining singleton clause set and can be removed as well. The result is the empty set, which is obviously satisfiable, indicating that  $N$  is satisfiable. As model of  $N$ , we can take an interpretation  $\mathcal{J}$  that makes all  $\mathbf{p}$ -literals true and all  $\mathbf{q}$ -literals false.

**Example 6.2.** Consider the clause set  $N' = \{\mathbf{p}a, \neg \mathbf{p}x \vee \mathbf{p}(fx)\}$ . PLE does not apply because  $\mathbf{p}$  occurs with both polarities. Yet we notice that  $\mathbf{p}$  occurs positively in both clauses, and hence that the same reasoning as in Example 6.1 applies: We can satisfy both clauses by making  $\mathbf{p}$ -literals true. Using  $\mathcal{J}$  from Example 6.1, we have  $\mathcal{J} \models N'$ .

**Example 6.3.** Consider the clause set  $N'' = \{\mathbf{p}a, \mathbf{q}x \vee \mathbf{p}(fx), \neg \mathbf{q}(fa), \neg \mathbf{p}x \vee \neg \mathbf{q}(h\mathbf{p}(\mathbf{p}b))\}$ . PLE does not apply because  $\mathbf{p}$  and  $\mathbf{q}$  occur with both polarities. In addition,  $\mathbf{p}$  also occurs unapplied and deep within a term. Yet each clause contains either a positive  $\mathbf{p}$ -literal or a negative  $\mathbf{q}$ -literal. Thus  $\mathcal{J} \models N''$ , where  $\mathcal{J}$  is as in Example 6.1. The additional literals are harmless.



Examples 6.2 and 6.3 suggest that pure literals are a needlessly restrictive criterion in first- and higher-order logic. We propose a generalization to “quasipure literals.” Although we present the criterion in a higher-order setting, it is equally applicable for first-order logic. (In contrast, it is uninteresting for propositional logic, because the only clauses with the same predicates, or rather variables, with opposite polarities are tautologies, and these would be deleted before pure literal elimination is attempted.)

**Definition 6.4.** A *polarity map* is a function that maps each predicate symbol in  $\Sigma$  to a polarity (+ or  $-$ ). A set  $P$  of predicate symbols is *quasipure* in a clause set  $N$  with a polarity map  $m$  if for every clause in  $N$  that contains an element of  $P$ , there exists a predicate symbol  $\mathbf{p} \in P$  such that the clause contains a  $\mathbf{p}$ -literal with polarity  $m_{\mathbf{p}}$ . The set  $P$  is *quasipure* in  $N$  if there exists a polarity map  $m$  such that  $P$  is quasipure in  $N$  with  $m$ .

In Example 6.1,  $\{\mathbf{p}, \mathbf{q}\}$  is quasipure in  $N$  with  $m_{\mathbf{p}} = +$  and  $m_{\mathbf{q}} = -$ . In Example 6.2,  $\{\mathbf{p}\}$  is quasipure in  $N'$  with  $m_{\mathbf{p}} = +$ . In Example 6.3,  $\{\mathbf{p}, \mathbf{q}\}$  is quasipure in  $N''$  with  $m_{\mathbf{p}} = +$  and  $m_{\mathbf{q}} = -$ . For this last example, it is crucial to consider  $\mathbf{p}$  and  $\mathbf{q}$  together; neither of the singletons  $\{\mathbf{p}\}$  and  $\{\mathbf{q}\}$  is quasipure in  $N''$ .

**Definition 6.5.** A predicate symbol  $\mathbf{p}$  is *quasipure* in a clause set  $N$  with polarity  $s \in \{+, -\}$  if there exists a set  $P$  of predicate symbols with  $\mathbf{p} \in P$  and a polarity map  $m$  such that  $m_{\mathbf{p}} = s$  and  $P$  is quasipure in  $N$  with  $m$ . The symbol  $\mathbf{p}$  is *quasipure* in  $N$  if there exists a polarity  $m_{\mathbf{p}} \in \{+, -\}$  such that  $\mathbf{p}$  is quasipure in  $N$  with  $m_{\mathbf{p}}$ . A literal  $L = (-) \mathbf{p} \dots$  is *quasipure* in  $N$  if  $\mathbf{p}$  is quasipure in  $N$  with  $L$ 's polarity.

Notice that a predicate symbol that does not occur in a clause set is trivially quasipure in that clause set.

Deleting a clause containing a quasipure literal might create new opportunities for quasipure literal elimination, but it never ruins existing ones. Therefore, the following nondeterministic procedure is confluent and hence yields a unique result:

**Definition 6.6.** Given a finite clause set, *quasipure literal elimination* (QLE) repeatedly removes clauses containing quasipure literals until no such literals remain.

Although QLE is defined by iteration, it is always possible to remove all clauses at the same time:

**Lemma 6.7.** *Let  $N$  be a finite clause set and let  $N'$  be the result of QLE. Then there exists a predicate symbol set  $P$  and a polarity map  $m$  such that  $P$  is quasipure in  $N$  and for every clause in  $N \setminus N'$  there exists a predicate symbol  $\mathbf{q} \in P$  such that the clause contains a  $\mathbf{q}$ -literal with polarity  $m_{\mathbf{q}}$ .*

*Proof.* The iterative process defining QLE gives rise to a finite sequence  $(P_1, m_1), \dots, (P_n, m_n)$  of predicate symbol sets and polarity maps. Without loss of generality, we assume that each  $P_j$  does not contain predicate symbols that do not occur in the clause set at iteration  $j$ . Then the sets  $P_j$  are clearly mutually disjoint and we can take  $P = P_1 \cup \dots \cup P_n$  as the desired witness. As for the polarity map  $m$ , we associate each  $\mathbf{p} \in P_j$  with  $m_j(\mathbf{p})$ .  $\square$

The key property of QLE is that it preserves unsatisfiability:

**Lemma 6.8.** *Let  $C \in N$  be a clause containing a quasipure literal. If  $N \setminus \{C\}$  is satisfiable, then  $N$  is satisfiable.*

*Proof.* Let  $\mathcal{J}$  be a model of  $N \setminus \{C\}$ . Let  $P$  be the set of predicate symbols and  $m$  the polarity map whose existence is guaranteed by Lemma 6.7. Let  $N_0 \subseteq N \setminus \{C\}$  be the result of applying QLE on  $N$ . Clearly,  $N_0$  contains no occurrences of the symbols in  $P$ . Define  $\mathcal{J}'$  based on  $\mathcal{J}$  by redefining the semantics of each monomorphic instance  $\mathbf{p}\langle\vec{\tau}\rangle : v_1 \rightarrow \dots \rightarrow v_n \rightarrow o$  of symbol  $\mathbf{p}$  such that  $m_{\mathbf{p}} = +$  or  $m_{\mathbf{p}} = -$ : If  $m_{\mathbf{p}} = +$ , interpret  $\mathbf{p}\langle\vec{\tau}\rangle$  as the predicate that is uniformly true; otherwise, interpret  $\mathbf{p}\langle\vec{\tau}\rangle$  as the predicate that is uniformly false. By the comprehension principle, both of these predicates are guaranteed to exist in the interpretation of the type  $v_1 \rightarrow \dots \rightarrow v_n \rightarrow o$ . Now  $\mathcal{J}'$  coincides with  $\mathcal{J}$ , since  $N_0$  contains no  $P$  symbols, and thus  $\mathcal{J}'$  is a model of  $N_0$ . In addition,  $\mathcal{J}'$  is a model of  $N$ , because each clause in  $N \setminus N_0$  contains a quasipure literal, which is satisfied by  $\mathcal{J}'$ .  $\square$

Definition 6.6 suggests a naive, nondeterministic procedure for discovering and eliminating quasipure predicate symbols: Choose a predicate symbol  $\mathbf{p}$  and a polarity  $m_{\mathbf{p}}$ , and take  $P = \{\mathbf{p}\}$ . If the predicate symbol occurs with the wrong polarity in a clause  $(\neg) \mathbf{p} \dots \vee (\neg) \mathbf{q}_1 \dots \vee \dots \vee (\neg) \mathbf{q}_n \dots \vee C$ , try to extend the set  $P$  with one of the  $(\neg) \mathbf{q}_i$ 's and the polarity map  $m$  accordingly, and continue recursively with  $\mathbf{q}_i$ . In Section 7, we will see a more efficient approach based on a SAT encoding.

## 7. IMPLEMENTATION

We implemented the elimination techniques described in Sections 3 to 6 in the Zipperposition prover. For HLBE, PE, and BCE, which had been studied by Vukmirović et al., we could directly adapt their code [VBH23, Section 6]. The data structure and algorithms they described and implemented could be generalized to handle polymorphic higher-order logic. For QLE, we developed our own code.

Zipperposition is a higher-order prover that implements the  $\lambda$ -superposition calculus [BBTV21], a generalization of standard superposition to classical rank-1 polymorphic higher-order logic—the logic described in Section 2. The prover is highly competitive: It won in the higher-order theorem division of the CADE ATP Systems Competition (CASC) [SD22] in 2020, 2021, and 2022.

By default, Zipperposition *immediately clausifies* the initial problem as much as possible; then it optionally invokes preprocessing elimination techniques. If inprocessing is enabled, the prover also invokes the elimination techniques at regular intervals from within the saturation loop. Immediate clausification performs well in practice, but an even more successful strategy is to *delay clausification*, interleaving clausification with superposition-style calculus rules [VBB<sup>+</sup>22, Section 4]. Then it makes little sense to apply preprocessing elimination techniques; inprocessing seems more appropriate.

**7.1. Hidden-Literal-Based Elimination.** HLBE relies on matching. In our setting, it needs to consider type variables and higher-order terms. Our implementation uses an efficient approximation of higher-order matching, which recognizes  $\lambda y. y a$  as an instance of  $\lambda y. y x$ , but not  $a$  as an instance of  $y a$  (with  $y := \lambda x. x$ ). The same approximated matching algorithm is used in Zipperposition to efficiently recognize subsumed clauses. This weaker matching reduces the applicability of HLBE, but it does not compromise its soundness.

**7.2. Predicate Elimination.** Compared with HLBE, more work was needed to make predicate elimination cope with polymorphism and higher-order logic. For polymorphism, the previous code simply did not consider polymorphic predicates as predicates. Predicates needed to have a type declaration of the form  $\tau_1 \times \dots \times \tau_n \rightarrow o$ . Now, predicate symbols can have type declarations of the form  $\Pi \vec{\alpha}. \tau_1 \rightarrow \dots \rightarrow \tau_n \rightarrow o$  as well as  $\Pi \vec{\alpha}. \tau_1 \rightarrow \dots \rightarrow \tau_n \rightarrow \alpha_i$ ; via instantiation,  $\alpha_i$  can become  $o$  or  $\dots \rightarrow o$ .

For SPE, the type arguments of eliminated predicate symbols must be unified, as per Definition 4.4. For DPE, the type arguments must be distinct type variables. In addition, for DPE, we must check that the predicate is polymorphism-safe.

Adding support for higher-order logic required changing the definition of singular predicates, which is used by both SPE and DPE, to check that the predicate symbol to eliminate does not occur deep in the clauses that define the symbol. In addition, for DPE, we need to synthesize a  $\lambda$ -abstraction to replace any deep occurrences of the predicate outside the definition set.

**7.3. Blocked Clause Elimination.** Adding support for polymorphism to BCE was straightforward: We simply ensured that type arguments are unified when computing flat resolvents and added a polymorphism-safety check.

To support higher-order logic, we added a check that the  $\mathbf{p}$ -literal on which the clause is resolved in the only  $\mathbf{p}$ -literal of that polarity (corresponding to condition 3 of Definition 5.2). We also disabled the code that computed  $n$ -ary resolvents for  $n > 2$ . Finally, we added code to compute the list of all deep predicate symbols  $\mathbf{q}$  and made sure clauses are never blocked on a  $\mathbf{q}$ -literal. These mechanisms come into play only if some higher-order construct is detected in the input problem; otherwise, the first-order formulation of BCE is used.

In the presence of equality in the logic, BCE relies on a congruence closure algorithm to detect valid clauses [KSS<sup>+</sup>17, Section 6.1]. In our implementation, we rely on a first-order congruence closure algorithm, which can handle higher-order constructs but does not take advantage of them. For example, resolving  $\mathbf{p} \mathbf{a} \vee \neg \mathbf{q} \mathbf{a}$  against  $\neg \mathbf{p} \mathbf{b} \vee \mathbf{q} \mathbf{b}$  on the  $\mathbf{p}$ -literal yields the clause  $\mathbf{a} \approx \mathbf{b} \vee \neg \mathbf{q} \mathbf{a} \vee \mathbf{q} \mathbf{b}$ . Our congruence closure algorithm can detect the validity of such a clause. On the other hand, because the algorithm views  $\lambda$ -abstractions as black boxes, it fails to recognize  $\mathbf{a} \approx \mathbf{b} \vee \neg \mathbf{q} (\lambda x. x \mathbf{a}) \vee \mathbf{q} (\lambda x. x \mathbf{b})$  as valid.

**7.4. Quasipure Literal Elimination.** A simple and efficient implementation of quasipure literal elimination uses a SAT solver. Let  $N$  be a finite clause set. Without loss of generality, we can assume that  $\Sigma$  is finite. The signature of the generated SAT problem consists of the variables  $\mathbf{p}^+, \mathbf{p}^-$  for each predicate symbol  $\mathbf{p}$  in  $\Sigma$ , where  $\mathbf{p}^s$  means “ $\mathbf{p}$  (possibly together with other predicate symbols) is quasipure with polarity  $s$ .” The SAT problem consists of the following clauses:

- (1) For each clause in  $N$  containing  $n$  predicate literals, headed by  $\mathbf{q}_1, \dots, \mathbf{q}_n$  and with respective polarities  $s_1, \dots, s_n$ , generate  $n$  clauses of the form

$$\mathbf{q}_1^{s_1} \vee \dots \vee \mathbf{q}_{j-1}^{s_{j-1}} \vee \neg \mathbf{q}_i^{-s_j} \vee \mathbf{q}_{j+1}^{s_{j+1}} \vee \dots \vee \mathbf{q}_n^{s_n}$$

where  $-s$  flips the polarity  $s$ . Such clauses ensure that whenever a literal  $(\neg) \mathbf{q}_i \dots$  has the wrong polarity according to the current variable assignment, one of the other predicate literals must be quasipure.

- (2) For each clause in  $N$  containing  $n$  predicate literals, headed by  $\mathbf{q}_1, \dots, \mathbf{q}_n$  and with respective polarities  $s_1, \dots, s_n$  and containing a deep occurrence of  $\mathbf{p}$  (in an argument to a  $\mathbf{q}_j$  or in a functional literal), generate the two clauses

$$\neg \mathbf{p}^+ \vee \mathbf{q}_1^{s_1} \vee \dots \vee \mathbf{q}_n^{s_n} \qquad \neg \mathbf{p}^- \vee \mathbf{q}_1^{s_1} \vee \dots \vee \mathbf{q}_n^{s_n}$$

These clauses ensures that whenever  $\mathbf{p}$  occurs deep and is nonetheless considered quasipure, one of the predicate literals must be quasipure.

- (3) For each predicate symbol  $\mathbf{p}$  in  $\Sigma$ , generate the clause  $\neg \mathbf{p}^+ \vee \neg \mathbf{p}^-$ . It ensures that a single polarity is assigned to a quasipure predicate symbol.
- (4) Generate the clause  $\mathbf{p}_1^+ \vee \mathbf{p}_1^- \vee \dots \vee \mathbf{p}_n^+ \vee \mathbf{p}_n^-$ , where  $\{\mathbf{p}_1, \dots, \mathbf{p}_n\}$  are the predicate symbols in  $\Sigma$ . It tells the SAT solver to look for a nontrivial solution, in which at least one predicate symbol is quasipure.

From a satisfying assignment, we can easily read off a predicate symbol set and a polarity map. The process can be iterated until we reach a maximal solution, at which point the SAT solver returns a verdict of “unsatisfiable.”

It is unclear whether this problem is **NP**-complete. Given the nondeterministic nature of the naive procedure, we suspect that it is, but we have not found a reduction from 3-SAT. Thus, it is unclear whether our use of a SAT solver is fully satisfactory from a theoretical point of view, even though it works well in practice.

## 8. EVALUATION

We evaluate the techniques presented above by running Zipperposition [BBTV21] on various benchmarks in various configurations. We consider seven benchmark sets:

- *S0*: a randomly selected subset of 1000 higher-order monomorphic (TH0) problems from the Sledgehammer-generated Seventeen benchmark suite [DVBW22] in the base configuration, in the language fragment called TH0<sup>−</sup> in the Seventeen paper;
- *S1*: a randomly selected subset of 1000 higher-order polymorphic (TH1) problems from the Seventeen benchmark suite in the base configuration, in the language fragment called TH1<sup>−</sup> in the Seventeen paper;
- *TH0*: a randomly selected subset of 1000 higher-order monomorphic (TH0) problems from the TPTP [Sut17] version 8.0.0;
- *TH1*: a randomly selected subset of 1000 higher-order polymorphic (TH1) problems from the TPTP version 8.0.0;
- *CF*: a predefined set of 1000 first-order untyped (CNF and FOF) problems from the TPTP;
- *TF0*: all 389 first-order monomorphic (TF0) problems without arithmetic from the TPTP;
- *TF1*: all 678 first-order polymorphic (TF1) problems without arithmetic from the TPTP.

The first four benchmark sets are used to determine how much our techniques can help on higher-order problems. Among these, S1 and TH1 exercise polymorphism. As for the remaining three, they are included for comparison; they show how well the techniques work on first-order problems. The benchmark sets contain some problems known to be unprovable, which we use to check soundness of the techniques.

We consider 12 Zipperposition configurations, all derived from the portfolio of time slices that was used at the 2022 edition of CASC. This portfolio does not use any of our

techniques. Even on first-order problem, it applies the higher-order  $\lambda$ -superposition calculus. The 12 configurations are as follows:

- *None*: the baseline, corresponding to the portfolio used at CASC;
- *X-pre*: the baseline modified to use technique  $X \in \{\text{PE, BCE, PLE, QLE}\}$  as a preprocessor in all the time slices;
- *X-in*: the baseline modified to use technique  $X \in \{\text{HLBE, PE, BCE, PLE, QLE}\}$  as an inprocessor in all the time slices;
- *All-pre*: the baseline modified to use all of PE, BCE, and QLE as preprocessors in all the time slices;
- *All-in*: the baseline modified to use all of HLBE, PE, BCE, and QLE as inprocessors in all the time slices.

In addition, we define the virtual configuration *Union*, consisting of the virtual portfolio of all other 12 configurations. All problems that are solved in at least one of the 12 configurations are considered solved by Union, and only those.

The experiments were carried out on StarExec Miami [SST14] servers equipped with Intel Xeon E5-2620 v4 CPUs clocked at 2.10 GHz. We used CPU and wallclock time limits of 120 s. The raw results are available online.<sup>2</sup>

Figure 1 reports how many problems were solved for each combination of benchmark set and configuration. The last row of the table presents the total of the seven rows above it. Bold singles out the best configuration (other than Union) for each benchmark set.

The results are sobering. We see substantial gains on the untyped first-order benchmarks, but the gains are much more modest, if actually present, on the typed first-order and the higher-order benchmarks. A possible explanation is that there is less clausal structure in a higher-order problem. Most of Zipperposition’s time slices clausify the problem lazily, meaning that little information is visible to our techniques, especially when used as preprocessors. Another possible explanation might be that the higher-order Seventeen and TPTP benchmarks look quite different from the untyped first-order TPTP benchmarks, and our techniques are less applicable. For example, the higher-order TPTP problems tend to be much smaller than their first-order counterparts. Moreover, we selected only subsets of the TPTP benchmarks—an evaluation on entire benchmark suites might yield different results.

The picture is more positive if we look at the Union column of the table. Clearly, in a portfolio setting, with enough time, the new techniques can make a useful contribution.

We also notice that inprocessing often performs worse than preprocessing. This corroborates the findings of Vukmirović et al. [VBH23]. An explanation might be the heavy cost of running the techniques multiple time, during proof search. In addition, PE, BCE, PLE, and QLE rely on a global analysis of the clause set and tend to become less applicable as the clause set grows.

Finally, we see that PLE and QLE help, especially on first-order problems. Unexpectedly, PLE generally outperforms the more general QLE.

## 9. CONCLUSION

We presented four SAT-inspired techniques for transforming higher-order problems with the aim of making them more amenable to automatic proof search. Three of the techniques

<sup>2</sup><https://zenodo.org/record/6997515>

	None	HLBE	PE		BCE		PLE		QLE		All		Union
		in	pre	in	pre	in	pre	in	pre	in	pre	in	
S0	<b>575</b>	574	574	572	<b>575</b>	574	574	<b>575</b>	<b>575</b>	574	572	<b>575</b>	589
S1	349	343	351	350	347	348	352	352	352	353	<b>354</b>	344	366
TH0	711	713	713	709	713	<b>714</b>	710	713	711	705	711	706	722
TH1	338	334	339	340	<b>341</b>	339	337	<b>341</b>	337	336	338	336	354
CF	504	513	512	510	509	507	513	505	508	505	517	<b>519</b>	545
TF0	138	<b>142</b>	138	137	<b>142</b>	139	141	141	139	141	139	137	148
TF1	227	227	229	<b>231</b>	228	230	229	230	229	230	228	224	234
Total	2842	2846	2856	2849	2855	2851	2856	2857	2851	2844	<b>2859</b>	2841	2958

Figure 1: Number of solved problems per benchmark set and configuration

(HLBE, PE, and BCE) had been previously generalized to first-order logic; we now generalized them further to higher-order logic. The fourth technique (QLE) is new.

On the theoretical side, we showed that the techniques preserve satisfiability and unsatisfiability of problems with respect to Henkin semantics. On the practical side, we implemented the techniques in the higher-order prover Zipperposition. Regrettably, the techniques did not perform as well on higher-order problems as they do on first-order problems. This could be due to the nature of the benchmark sets.

**Acknowledgment.** Alexander Bentkamp gave us some advice about general interpretations. The anonymous reviewers made many useful suggestions.

This research has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 713999, Matryoshka). The research has also received funding from the Netherlands Organization for Scientific Research (NWO) under the Vidi program (project No. 016.Vidi.189.037, Lean Forward).

## REFERENCES

- [BBTV21] Alexander Bentkamp, Jasmin Blanchette, Sophie Tourret, and Petar Vukmirović. Superposition for full higher-order logic. In André Platzer and Geoff Sutcliffe, editors, *CADE-28*, volume 12699 of *LNCS*, pages 396–412. Springer, 2021.
- [BG94] Leo Bachmair and Harald Ganzinger. Rewrite-based equational theorem proving with selection and simplification. *J. Log. Comput.*, 4(3):217–247, 1994.
- [BK98] Christoph Benzmüller and Michael Kohlhase. Extensional higher-order resolution. In Claude Kirchner and Hélène Kirchner, editors, *CADE-15*, volume 1421 of *LNCS*, pages 56–71. Springer, 1998.
- [BR20] Ahmed Bhayat and Giles Reger. A combinator-based superposition calculus for higher-order logic. In Nicolas Peltier and Viorica Sofronie-Stokkermans, editors, *IJCAR 2020, Part I*, volume 12166 of *LNCS*, pages 278–296. Springer, 2020.
- [CS00] Philippe Chatalic and Laurent Simon. ZRES: The old Davis–Putnam procedure meets ZBDD. In David A. McAllester, editor, *CADE-18*, volume 1831 of *LNCS*, pages 449–454. Springer, 2000.
- [DVBW22] Martin Desharnais, Petar Vukmirovic, Jasmin Blanchette, and Makarius Wenzel. Seventeen provers under the hammer. In June Andronick and Leonardo de Moura, editors, *ITP 2022*, volume 237 of *LIPICs*, pages 8:1–8:18. Schloss Dagstuhl—Leibniz-Zentrum für Informatik, 2022.
- [EB05] Niklas Eén and Armin Biere. Effective preprocessing in SAT through variable and clause elimination. In Fahiem Bacchus and Toby Walsh, editors, *SAT 2005*, volume 3569 of *LNCS*, pages 61–75. Springer, 2005.

- [Fit02] Melvin Fitting. *Types, Tableaux, and Gödel's God*. Kluwer, 2002.
- [GO92] Dov M. Gabbay and Hans Jürgen Ohlbach. Quantifier elimination in second-order predicate logic. In Bernhard Nebel, Charles Rich, and William R. Swartout, editors, *KR '92*, pages 425–435. Morgan Kaufmann, 1992.
- [HJB11] Marijn J. H. Heule, Matti Järvisalo, and Armin Biere. Efficient CNF simplification based on binary implication graphs. In Karem A. Sakallah and Laurent Simon, editors, *SAT 2011*, volume 6695 of *LNCS*, pages 201–215. Springer, 2011.
- [Hue73] Gérard P. Huet. A mechanization of type theory. In Nils J. Nilsson, editor, *IJCAI 1973*, pages 139–146. William Kaufmann, 1973.
- [KK16] Zurab Khasidashvili and Konstantin Korovin. Predicate elimination for preprocessing in first-order theorem proving. In Nadia Creignou and Daniel Le Berre, editors, *SAT 2016*, volume 9710 of *LNCS*, pages 361–372. Springer, 2016.
- [KSR16] Cezary Kaliszyk, Geoff Sutcliffe, and Florian Rabe. TH1: The TPTP typed higher-order form with rank-1 polymorphism. In Pascal Fontaine, Stephan Schulz, and Josef Urban, editors, *PAAR-2016*, volume 1635 of *CEUR Workshop Proceedings*, pages 41–55. CEUR-WS.org, 2016.
- [KSS<sup>+</sup>17] Benjamin Kiesl, Martin Suda, Martina Seidl, Hans Tompits, and Armin Biere. Blocked clauses in first-order logic. In Thomas Eiter and David Sands, editors, *LPAR-21*, volume 46 of *EPiC Series in Computing*, pages 31–48. EasyChair, 2017.
- [NPW02] Tobias Nipkow, Lawrence C. Paulson, and Markus Wenzel. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer, 2002.
- [Ohl96] Hans Jürgen Ohlbach. SCAN—elimination of predicate quantifiers. In Michael A. McRobbie and John K. Slaney, editors, *CADE-13*, volume 1104 of *LNCS*, pages 161–165. Springer, 1996.
- [SB18] Alexander Steen and Christoph Benzmüller. The higher-order prover Leo-III. In Didier Galmiche, Stephan Schulz, and Roberto Sebastiani, editors, *IJCAR 2018*, volume 10900 of *LNCS*, pages 108–116. Springer, 2018.
- [SB21] Alexander Steen and Christoph Benzmüller. Extensional higher-order paramodulation in Leo-III. *J. Autom. Reason.*, 65(6):775–807, 2021.
- [SD22] Geoff Sutcliffe and Martin Desharnais. The CADE-28 Automated Theorem Proving System Competition—CASC-28. *AI Commun.*, 34(4):259–276, 2022.
- [SP04] Sathiamoorthy Subbarayan and Dhiraj K. Pradhan. NiVER: Non-increasing variable elimination resolution for preprocessing SAT instances. In Holger H. Hoos and David G. Mitchell, editors, *SAT 2004*, volume 3542 of *LNCS*, pages 276–291. Springer, 2004.
- [SST14] Aaron Stump, Geoff Sutcliffe, and Cesare Tinelli. StarExec: A cross-community infrastructure for logic solving. In Stéphane Demri, Deepak Kapur, and Christoph Weidenbach, editors, *IJCAR 2014*, volume 8562 of *LNCS*, pages 367–373. Springer, 2014.
- [Sut17] Geoff Sutcliffe. The TPTP problem library and associated infrastructure—from CNF to TH0, TPTP v6.4.0. *J. Autom. Reason.*, 59(4):483–502, 2017.
- [VBB<sup>+</sup>22] Petar Vukmirovic, Alexander Bentkamp, Jasmin Blanchette, Simon Cruanes, Visa Nummelin, and Sophie Tourret. Making higher-order superposition work. *J. Autom. Reason.*, 66(4):541–564, 2022.
- [VBH23] Petar Vukmirovic, Jasmin Blanchette, and Marijn J. H. Heule. SAT-inspired eliminations for superposition. *ACM Trans. Comput. Log.*, 24(1):7:1–7:25, 2023.
- [VBS23] Petar Vukmirović, Jasmin Blanchette, and Stephan Schulz. Extending a high-performance prover to higher-order logic. In Natasha Sharygina and Sriram Sankaranarayanan, editors, *TACAS 2023*, LNCS. Springer, 2023.