

The Alethe Proof Format

An Evolving Specification and Reference

Haniel Barbosa¹ Mathias Fleury² Pascal Fontaine³
Hans-Jörg Schurr⁴

¹ Universidade Federal de Minas Gerais, Brazil

² Albert-Ludwig-Universität Freiburg, Germany

³ Université de Liège, Belgium

⁴ The University of Iowa, Iowa City, USA

Contents

1	Introduction	2
1.1	Notations	3
2	The Alethe Language	4
2.1	The Syntax	7
3	Checking Alethe Proofs	14
3.1	Contexts and Metaterms	17
3.2	Soundness	19
4	Core Concepts of the Alethe Rules	22
4.1	Bitvector Reasoning with Bitblasting	23
4.2	RARE rules	25
5	The Alethe Rules	25
5.1	Classifications of the Rules	25
5.2	Rule List	30
5.3	Index of Rules	62
	Changelog	64
	Index	68
	References	68

Foreword

This document is a speculative specification and reference of a proof format for SMT solvers. The format consists of a language to express proofs and a set of proof rules. On the one side, the language is inspired by natural-deduction and is based on the widely used SMT-LIB format. The language also includes a flexible mechanism to reason about bound variables which allows fine-grained preprocessing proofs. On the other side, the rules are structured around resolution and the introduction of theory lemmas, in the same way as CDCL(T)-based SMT solvers.

The specification is not yet cast in stone, but it will evolve over time. It emerged from a list of proof rules used by the SMT solver veriT collected in a document called “Proofonomicon”. Following the fate presupposed by its name, it informally circulated among researchers interested in the proofs produced by veriT after a few months. We now polished this document and gave it a respectable name.

Instead of aiming for theoretical purity, our approach is pragmatic: the specification describes the format as it is in use right now. It will develop in parallel with practical support for the format within SMT solvers, proof checkers, and other tools. We believe it is not a perfect specification that fosters the adaption of a format, but great tooling. This document will be a guide to develop such tools.

Nevertheless, it not only serves as a norm to ensure compatibility between tools, it also allows us to uncover the unsatisfactory aspects that would otherwise be hidden deep within the nooks and crannies of solver and checker implementations. Every uncovered problem presents an opportunity to improve the format. The authors of this document overlap with the authors of those tools and we are committed to improve the tools, the format, and ultimately the specification together. This document is also an invitation to other researchers to join these efforts. To read the reference and provide feedback, or to even implement support for Alethe into their own tools. Please get in touch!

The authors.

1 Introduction

This document is a reference of the Alethe¹ proof format. Alethe is designed to be a flexible format to represent unsatisfiability proofs generated by SMT solvers. Alethe proofs can be consumed by other systems, such as interactive theorem provers or proof checkers. The design is based on natural-deduction style structure and rules generating

¹Alethe is a genus of small birds that occur in West Africa [9]. The name was chosen because it resembles the Greek word ἀλήθεια (alítheia) – truth.

and operating on first-order clauses. The Alethe proof format consists of two parts: the proof language based on SMT-LIB and a collection of proof rules. Section 2 introduces the language. First as an abstract language, then as a concrete syntax. Section 3 then discusses an abstract procedure to check Alethe proofs. This abstract checking procedure specifies the semantics of Alethe proofs. The Alethe proof rules are discussed in two sections. First, Section 4 discusses the core concepts behind the rules. Second, Section 5 presents a list of all proof rules currently used by `veriT`.

Alethe follows a few core design principles. First, proofs should be easy to understand by humans to ensure working with Alethe proofs is easy. Second, the language of the format should directly correspond to the language used by the solver. Since many solvers use the SMT-LIB language, Alethe also uses this language. Therefore, Alethe’s base logic is the many-sorted first-order logic of SMT-LIB. Third, the format should be uniform for all theories used by SMT solvers. With the exception of clauses for propositional reasoning, there is no dedicated syntax for any theory.

The Alethe format was originally developed for the SMT solver `veriT`. If requested by the user, `veriT` outputs a proof if it can deduce that the input problem is unsatisfiable. In proof production mode, `veriT` supports the theory of uninterpreted functions, the theory of linear integer and real arithmetic, and quantifiers. The SMT solver `cvc5` [1] (the successor of `CVC4`) supports Alethe experimentally as one of its multiple proof output formats. Alethe proofs can be reconstructed by the `smt` tactic of the proof assistant `Isabelle/HOL` [7, 8]. The `SMTCoq` tool can reconstruct an older version of the format in the proof assistant `Coq` [6]. An effort to update the tool to the latest version of Alethe is ongoing. Furthermore, `Carcara` is an experimental high-performance proof checker written in Rust.²

In addition to this reference, the proof format has been discussed in past publications, which provide valuable background information. The core of the format goes back to 2011 when two publications at the PxTP workshop outlined the fundamental ideas behind the format [4] and proposed rules for quantifier instantiation [5]. More recently, the format has gained support for reasoning typically used for processing, such as skolemization, substitutions, and other manipulations of bound variables [2].

1.1 Notations

The notation used in this document is similar to the notation used by the SMT-LIB standard. The Alethe proof format uses the SMT-LIB logic. Since the SMT-LIB language is based on S-expressions, SMT-LIB formulas are written using a λ -calculus style. That is, instead of $f(1, 2)$, we write $(f\ 1\ 2)$. However, connectives that are usually written using infix notation, also use infix notation here. That is, we write $t_1 \vee t_2$, not $(\vee\ t_1\ t_2)$.

We use x, y, z to indicate variables, f, g for functions, and P, Q for predicates (functions with co-domain sort **Bool**). To indicate terms we use t, u and to indicate formulas (terms of sort **Bool**) we use φ, ψ . To distinguish syntactic equality and the SMT-LIB equality predicate, we write $=$ for the former, and $\approx$ for the latter. We will write pre-defined

²Available at <https://github.com/ufmg-smite/carcara>.

SMT-LIB symbols, such as sorts and functions, in bold (e.g., **Bool**, **ite**).

We will use θ to denote a substitution. The notation $[x_1 \mapsto t_1, \dots, x_n \mapsto t_n]$, provided variables $x_1, \dots, x_n$ are pairwise distinct, denotes the substitution that maps x_i to t_i for $1 \leq i \leq n$ and corresponds to the identity function for all other variables. If θ and η are two substitutions, then $\theta\eta$ denotes the result of first applying θ and then η (i.e., $\eta(\theta(\cdot))$). A substitution can naturally be extended to a function that maps terms to terms by replacing the occurrences of free variables. The application of a substitution θ to a term t (i.e., $\theta(t)$) is capture-avoiding; bound variables in t are renamed as necessary.

We write $t[u]$ for a term that contains the term u as a subterm. If u is subsequently replaced by a term v , we write $t[v]$ for the new term. We also use this notation with multiple terms. The notation $t[u_1, \dots, u_n]$ stands for a term that may contain the pairwise distinct terms $u_1, \dots, u_n$. Then, $t[s_1, \dots, s_n]$ is the respective term where the variables $u_1, \dots, u_n$ are simultaneously replaced by $s_1, \dots, s_n$. Usually, $u_1, \dots, u_n$ will be variables.

Note that we will introduce the Alethe specific notation to write proof steps in the following sections.

2 The Alethe Language

This section provides an overview of the core concepts of the Alethe language and also introduces some notation used throughout this chapter. The section first introduces an abstract notation to write Alethe proofs. Then, it introduces the concrete, SMT-LIB-based syntax. Finally, we show how a concrete Alethe proof can be checked.

Example 1. The following example shows a simple Alethe proof expressed in the abstract notation used in this document. It uses quantifier instantiation and resolution to show a contradiction. The paragraphs below describe the concepts necessary to understand the proof step by step.

1. $\triangleright$	$\forall x. (P x)$	assume
2. $\triangleright$	$\neg(P a)$	assume
3. $\triangleright$	$\neg(\forall x. (P x)) \vee (P a)$	forall_inst $[a]$
4. $\triangleright$	$\neg(\forall x. (P x)), (P a)$	(or 3)
5. $\triangleright$	$\perp$	(resolution 1, 2, 4)

Many-Sorted First-Order Logic. Alethe builds on the SMT-LIB language. This includes its many-sorted first-order logic. The available sorts depend on the selected SMT-LIB theory/logic as well as on those defined by the user. However, some sorts are always available. This includes a distinguished **Bool** sort, and the two sorts **Int** and **Real**. The reason for this is that values of these sorts are used in the arguments of certain rules. Alethe also extends this logic with Hilbert's choice operator ε . The term $\varepsilon x. \varphi[x]$ stands for a value v such that $\varphi[v]$ is true if such a value exists. Any value is possible otherwise. Alethe requires that ε is functional with respect to logical equivalence: if for two formulas φ, ψ that contain the free variable x , it holds that $(\forall x. \varphi \approx \psi)$, then

$(\varepsilon x. \varphi) \approx (\varepsilon x. \psi)$ must also hold. Note that choice terms can only appear in Alethe proofs, not in SMT-LIB problems.

A subtle difference between Alethe and SMT-LIB is the handling of variadic operators. SMT-LIB handles variadic function application purely as syntactic sugar for chained binary applications of annotated associative functions, while the description of Alethe rules treats them as first-class syntax. This is described in detail on Page 11.

Steps. A proof in the Alethe language is an indexed list of steps. To mimic the concrete syntax of Alethe proofs, proof steps in the abstract notation have the form

$$i. \ c_1, \dots, c_j \triangleright \qquad l_1, \dots, l_k \qquad (\text{rule } p_1, \dots, p_n) [a_1, \dots, a_m]$$

Each step has a unique index $i \in \mathbb{I}$, where $\mathbb{I}$ is a countable infinite set of valid indices. In the concrete syntax all SMT-LIB symbols are valid indices, but for examples we will use natural numbers. Furthermore, $l_1, \dots, l_k$ is a clause with the literals l_i . It is the conclusion of the step. If a step has the empty clause as its conclusion (i.e., $k = 0$) we write $\perp$. While this muddles the water a bit with regard to steps which have the unit clause with the unit literal $\perp$ as their conclusion, it simplifies the notation. We will remark on the difference if it is relevant. The rule name *rule* is taken from a set of possible proof rules (see Section 5). Furthermore, each step has a possibly empty set of premises $\{p_1, \dots, p_n\} \subseteq \mathbb{I}$, and a rule-dependent and possibly empty list of arguments $[a_1, \dots, a_m]$. The list of premises only references earlier steps, such that the proof forms a directed acyclic graph. If the list of premises is empty, we will drop the parentheses around the proof rule. The arguments a_i are either terms or tuples (x_i, t_i) where x_i is a variable and t_i is a term. The interpretation of the arguments is rule specific. The list $c_1, \dots, c_j$ is the *context* of the step. Contexts are discussed below. Every proof ends with a step that has the empty clause as the conclusion and an empty context. The list of proof rules in Section 5 also uses this notation to define the proof rules.

The example above consists of five steps. Step 4 and 5 use premises. Since step 3 introduces a tautology, it uses no premises. However, it uses arguments to express the substitution $[x \mapsto a]$ used to instantiate the quantifier. Step 4 translates the disjunction into a clause. In the example above, the contexts are all empty.

Assumptions. An *assume* step introduces a term as an assumption. The proof starts with a number of *assume* steps. Each such step corresponds to an input assertion. Within a subproof, additional assumptions can be introduced too. In this case, each assumption must be discharged with an appropriate step. The rule *subproof* can be used to do so. In the concrete syntax, *assume* steps have a dedicated command **assume** to clearly distinguish them from normal steps that use the **step** command (see Section 2.1).

The example above uses two assumptions which are introduced in the first two steps.

Subproofs and Lemmas. Alethe uses subproofs to prove lemmas and to create and manipulate the context. To prove lemmas, a subproof can introduce local assumptions. The *subproof rule* discharges the local assumptions. From an assumption φ and a formula

ψ proved from φ , the **subproof** rule deduces the clause $\neg\varphi, \psi$ that discharges the local assumption φ . A **subproof** step cannot use a premise from a subproof nested within the current subproof.

Subproofs are also used to manipulate the context. As the example below shows, the abstract notation denotes subproofs by a frame around the steps in the subproof. In this case the subproof concludes with a step that does not use the **subproof** rule, but another rule, such as the **bind** rule.

Example 2. This example shows a refutation of the formula $(2 + 2) \approx 5$. The proof uses a subproof to prove the lemma $((2 + 2) \approx 5) \Rightarrow 4 \approx 5$.

1. $\triangleright$	$(2 + 2) \approx 5$	assume
2. $\left \triangleright$	$(2 + 2) \approx 5$	assume
3. $\left \triangleright$	$(2 + 2) \approx 4$	sum_simplify
4. $\left \triangleright$	$4 \approx 5$	(trans 2, 3)
5. $\triangleright$	$\neg((2 + 2) \approx 5), 4 \approx 5$	subproof
6. $\triangleright$	$(4 \approx 5) \approx \perp$	eq_simplify
7. $\triangleright$	$\neg((4 \approx 5) \approx \perp), \neg(4 \approx 5), \perp$	equiv_pos2
8. $\triangleright$	$\perp$	(resolution 1, 5, 6, 7)

Contexts. A specificity of the Alethe proofs is the step context. Alethe contexts are a general mechanism to write substitutions and to change them by attaching new elements. A context is a possibly empty list $c_1, \dots, c_l$, where each element is either a variable or a variable-term tuple denoted $x_i \mapsto t_i$. In the first case, we say that c_i *fixes* the variable. The second case is a mapping. Throughout this chapter, Γ denotes an arbitrary context.

Every context Γ induces a capture-avoiding substitution $\text{subst}(\Gamma)$. If Γ is the empty list, $\text{subst}(\Gamma)$ is the empty substitution, i.e., the identity function. Otherwise, if Γ ends with variable x_n , this fixes the variable and allows the context to shadow a previously defined substitution for x_n :

$$\text{subst}([c_1, \dots, c_{n-1}, x_n]) \text{ is } \text{subst}([c_1, \dots, c_{n-1}]), \text{ but } x_n \text{ maps to } x_n.$$

When Γ ends in a mapping, the substitution is extended with this mapping:

$$\text{subst}([c_1, \dots, c_{n-1}, x_n \mapsto t_n]) = \text{subst}([c_1, \dots, c_{n-1}]) \circ \{x_n \mapsto t_n\}.$$

The following example illustrates this idea.

$$\begin{aligned} \text{subst}([x \mapsto 7, x \mapsto g(x)]) &= \{x \mapsto g(7)\} \\ \text{subst}([x \mapsto 7, x, x \mapsto g(x)]) &= \{x \mapsto g(x)\} \end{aligned}$$

Contexts are used to express proofs about the preprocessing of terms. The conclusions of proof rules that use contexts always have the form

i. $\Gamma \triangleright$ $t \approx u$ (rule, ...)

where the term t is the original term, and u is the term after preprocessing. The symbol $\approx$ is the ordinary SMT-LIB equality (usually denoted $\approx$), but we highlight because of its special role in rules that operate modulo a context. Tautologies with contexts correspond to judgments $\models_T \text{subst}(\Gamma)(t) \approx u$. Note that, some proof rules require an empty context. In the list of proof rules in Section 5 this is indicated by omitting the Γ .

The substitution induced by Γ is capture-avoiding. Hence, some bound variables could be renamed in $\text{subst}(\Gamma)(t)$ with respect to the original term t . A consequence of this is that steps that use a context must be checked under α -equivalence. The bind rule can be used to express renaming of bound variables explicitly. The refl rule, on the other hand, can be exploited to directly rename bound variables without an explicit proof.

Formally, the context can be translated to λ -abstractions and applications. This is discussed in Section 3.

Example 3. This example shows a proof that uses a subproof with a context to rename a bound variable.

1.	$\triangleright$	$\forall x. (P x)$	assume
2.	$\triangleright$	$\neg(\forall y. (P y))$	assume
3.	$y, x \mapsto y \triangleright$	$x \approx y$	refl
4.	$y, x \mapsto y \triangleright$	$(P x) \approx (P y)$	(cong 3)
5.	$\triangleright$	$\forall x. (P x) \approx \forall y. (P y)$	bind
6.	$\triangleright$	$\neg(\forall x. (P x) \approx \forall y. (P y)), \neg(\forall x. (P x)), (\forall y. (P y))$	equiv_pos2
7.	$\triangleright$	$\perp$	(resolution 1, 2, 5, 6)

Implicit Reordering of Equalities. In addition to the explicit steps, solvers might reorder equalities, i.e., apply symmetry of the equality predicate, without generating steps. The sole exception is the topmost equality in the conclusion of steps with non-empty context. The order of the arguments of this equality can never change. As described above, all rules that accept a non-empty context have a conclusion of the form $t \approx u$. Since the context represents a substitution applied to the left-hand side, this equality symbol is not symmetric.

The SMT solver `veriT` currently applies this freedom in a restricted form: equalities are reordered only when the term below the equality changes during proof search. One such case is the instantiation of universally quantified variables. If an instantiated variable appears below an equality, then the equality might have an arbitrary order after instantiation. Nevertheless, consumers of Alethe must consider the possible implicit reordering of equalities everywhere.

The SMT solver `cvc5` is more strict. It never applies reordering of equalities implicitly.

2.1 The Syntax

The concrete text representation of the Alethe proofs is based on the SMT-LIB standard. Figure 1 shows an example proof as printed by `veriT` with light edits for readability.

```

(assume h1 (not (p a)))
(assume h2 (forall ((z1 U)) (forall ((z2 U)) (p z2))))
...
(anchor :step t9 :args ((vr4 U) (:= (z2 U) vr4)))
(step t9.t1 (cl (= z2 vr4)) :rule refl)
(step t9.t2 (cl (= (p z2) (p vr4)))
  :rule cong :premises (t9.t1))
(step t9 (cl (= (forall ((z2 U)) (p z2))
  (forall ((vr4 U)) (p vr4))))
  :rule bind)
...
(step t14 (cl (forall ((vr5 U)) (p vr5)))
  :rule th_resolution :premises (t11 t12 t13))
(step t15 (cl (or (not (forall ((vr5 U)) (p vr5)))
  (p a)))
  :rule forall_inst :args (a))
(step t16 (cl (not (forall ((vr5 U)) (p vr5))) (p a))
  :rule or :premises (t15))
(step t17 (cl) :rule resolution :premises (t16 h1 t14))

```

Figure 1: Example proof output. Assumptions are introduced; a subproof renames bound variables; the proof finishes with instantiation and resolution steps.

The format broadly follows the surface syntax specified by SMT-LIB. Input problems in the SMT-LIB format are scripts. An SMT-LIB script is a list of commands that manipulate the SMT solver. For example, **assert** introduces an assertion, **check-sat** starts solving, and **get-proof** instructs the SMT solver to print the proof. Alethe mirrors this structure. Therefore, beside the SMT-LIB logic and term language, it also uses commands to structure the proof. An Alethe proof is a list of commands.

Every Alethe proof is associated with an SMT-LIB problem that is proved by the Alethe proof. This can either be a concrete problem file or, if the incremental solving commands of SMT-LIB are used, the implicit problem constructed at the invocation of a **get-proof** command. In this document, we will call this SMT-LIB problem the *input problem*. An Alethe proof inherits the namespace of its SMT-LIB problem. All symbols declared or defined in the input problem remain declared or defined, respectively. Furthermore, the symbolic names introduced by the **:named** annotation also stay valid and can be used in the script. For the purpose of the proof rules, terms are treated as if proxy names introduced by **:named** annotations have been expanded and annotations have been removed. For example, the term `(or (! (P a) :named baz) (! baz :foo))` and `(or (P a) (P a))` are considered to be syntactically equal. Here **:foo** is an arbitrary SMT-LIB annotation.

Figure 2 shows the grammar of the proof text. It is based on the SMT-LIB grammar, as defined in the SMT-LIB standard [3, Appendix B]. The non-terminals `<attribute>`,

A $\langle \text{symbol} \rangle$ is an SMT-LIB $\langle \text{symbol} \rangle$ that is not a
 $-\langle \text{numeral} \rangle / \langle \text{positive_numeral} \rangle$,
 $-\langle \text{numeral} \rangle$, or $-\langle \text{decimal} \rangle$.

$\langle \text{proof} \rangle ::= \langle \text{proof_command} \rangle^*$
 $\langle \text{proof_command} \rangle ::= (\text{assume } \langle \text{symbol} \rangle \langle \text{proof_term} \rangle \langle \text{attribute} \rangle^*)$
 $\quad | (\text{step } \langle \text{symbol} \rangle \langle \text{clause} \rangle : \text{rule } \langle \text{symbol} \rangle$
 $\quad \quad \langle \text{premises_annotation} \rangle^?$
 $\quad \quad \langle \text{args_annotation} \rangle^? \langle \text{attribute} \rangle^*)$
 $\quad | (\text{anchor :step } \langle \text{symbol} \rangle$
 $\quad \quad \langle \text{context_annotation} \rangle^? \langle \text{attribute} \rangle^*)$
 $\quad | (\text{define-fun } \langle \text{function_def} \rangle)$
 $\langle \text{clause} \rangle ::= (\text{cl } \langle \text{proof_term} \rangle^*)$
 $\langle \text{proof_term} \rangle ::= \langle \text{term} \rangle \text{ extended with}$
 $\quad (\text{choice } (\langle \text{sorted_var} \rangle) \langle \text{proof_term} \rangle)$
 $\quad | \langle \text{rational} \rangle$
 $\quad | \langle \text{nonpositive_numeral} \rangle$
 $\quad | \langle \text{nonpositive_decimal} \rangle$
 $\langle \text{premises_annotation} \rangle ::= \text{:premises } (\langle \text{symbol} \rangle^+)$
 $\langle \text{args_annotation} \rangle ::= \text{:args } (\langle \text{step_arg} \rangle^+)$
 $\langle \text{step_arg} \rangle ::= \langle \text{proof_term} \rangle | (@\text{list } \langle \text{proof_term} \rangle^+) | @\text{list}$
 $\langle \text{context_annotation} \rangle ::= \text{:args } (\langle \text{context_assignment} \rangle^+)$
 $\langle \text{context_assignment} \rangle ::= \langle \text{sorted_var} \rangle$
 $\quad | (:= \langle \text{sorted_var} \rangle \langle \text{proof_term} \rangle)$
 $\langle \text{positive_numeral} \rangle ::= \text{any } \langle \text{numeral} \rangle \text{ except } 0$
 $\langle \text{rational} \rangle ::= -^? \langle \text{numeral} \rangle / \langle \text{positive_numeral} \rangle$
 $\langle \text{nonpositive_numeral} \rangle ::= -\langle \text{numeral} \rangle$
 $\langle \text{nonpositive_decimal} \rangle ::= -\langle \text{decimal} \rangle$

Figure 2: The proof grammar.

$\langle \text{function_def} \rangle$, $\langle \text{sorted_var} \rangle$, and $\langle \text{term} \rangle$ are as defined in the standard. A special restriction applies to the $\langle \text{symbol} \rangle$ non-terminal. Alethe has an extended set of number literals. Since these can start with a negation sign, they overlap SMT-LIB's $\langle \text{symbol} \rangle$ non-terminal. For example, -1 is a valid $\langle \text{symbol} \rangle$ in SMT-LIB. These sequences cannot be used as symbols when using Alethe. Note that symbols are also used to name user defined constants and functions in the input problem. Hence, Alethe cannot express proofs about problems that use such symbols.

Alethe proofs are a list of commands. The **assume** command introduces a new assumption. While this command could also be expressed using the **step** command with a special rule, the special semantics of an assumption justifies the presence of a dedicated command: assumptions are neither tautological nor derived from premises. The **step** command, on the other hand, introduces a derived or tautological clause. Both com-

mands **assume** and **step** require an index as the first argument to later refer back to it. This index is an arbitrary SMT-LIB symbol. It must be unique for each **assume** and **step** command. A special restriction applies to the **assume** commands not within a subproof, which reference assertions in the input SMT-LIB problem. To simplify proof checking, the **assume** command must use the name assigned to the asserted formula if there is any. For example, if the input problem contains `(assert (! (P c) :named foo))`, then the proof must refer to this assertion (if it is needed in the proof) as `(assume foo (P c))`. Note that an SMT-LIB problem can assign a name to a term at any point, not only at its first occurrence. If a term has more than one name, any can be picked.

The second argument of **step** and **assume** is the conclusion of the command. For a **step**, this term is always a clause. To express disjunctions in SMT-LIB the **or** operator is used. This operator, however, needs at least two arguments and cannot represent unary or empty clauses. To circumvent this, we introduce a new **cl** operator. It corresponds to the standard **or** function extended to one argument, where it is equal to the identity, and zero arguments, where it is equal to **false**. Every step must use the **cl** operator, even if its conclusion is a unit clause. The **anchor** and **define-fun** commands are used for subproofs and sharing, respectively. The **define-fun** command corresponds exactly to the **define-fun** command of the SMT-LIB language.

Furthermore, the syntax uses annotations as used by SMT-LIB. The original SMT-LIB syntax uses the non-terminal `<attribute>`. The Alethe syntax uses some predefined annotation. To simplify parsing, the order in which those must be printed is strict. The **:premises** annotation denotes the premises and is skipped if the rule does not require premises. If the rule carries arguments, the **:args** annotation is used to denote them. Anchors have two annotations: **:step** provides the name of the step that concludes the subproof and **:args** provides the context as sorted variables and assignments. Note that in this annotation, the `<symbol>` non-terminal is intended to be a variable. After those pre-defined annotations, the solver can use additional annotations. This can be used for debugging, or other purposes. A consumer of Alethe proofs *must* be able to parse proofs that contain such annotations.

Terms The non-terminal `<proof_term>` is an extended version of the SMT-LIB non-terminal `<term>`. First, it has an additional production for the **choice** binder. Second, it has productions to express rationals and negative integers concisely. A difficulty when parsing SMT-LIB terms is that numerical constants are not easy to distinguish from general terms. For example, $-\frac{1}{2}$ is written as `(/ (- 1) 2)`. The `<rational>` non-terminal makes it possible to write this constant as a single literal: `-1/2`. Furthermore, the non-terminals `<nonpositive_numeral>` and `<nonpositive_decimal>` achieve the same for unary negation.

The sorting rules for `<proof_term>` are as for SMT-LIB terms with one key difference. The sort of terms produced by `<rational>`, `<decimal>`, and `<nonpositive_decimal>` is always **Real**. The sort of terms produced by `<integer>` and `<nonpositive_numeral>` is always **Int**. For example, in standard SMT-LIB, the term `(+ 5 3)` has the sort **Int** in the logic QF_LIA, but has the sort **Real** in QF_LRA. In Alethe, this term always has the

sort `Int`.

Variadic Operators SMT-LIB allows certain operators, like `and`, to be used with a variable number of arguments. SMT-LIB achieves this by handling applications of more than two arguments as syntactic sugar. For example, the operator `and` is annotated in SMT-LIB with the `:left-assoc` keyword. This has the effect that a term like `(and a b c)` can be desugared to `(and (and a b) c)`. In other words, after parsing an SMT-LIB S-expression into an abstract syntax tree, the variadic application of these applications is broken up into binary applications before type checking.

However, most SMT solvers do not handle these operators this way. Instead, their internal abstract syntax tree maintains the variadic application. To take this into account, all rules of Alethe operate on SMT-LIB terms before variadic applications are broken into binary applications. Consider, for example, the rule `and`. This rule implements `and`-elimination, and takes the index of the conjunct to extract as an argument. The following is a correct application of the `and` rule:

```
(assume h1 (and a b c))
(step t1 (cl b) :rule and :premises h1 :args (2))
```

If Alethe rules would operate on terms after desugaring, this step would instead conclude with `c`.

In the abstract notation of this document, we write these operators using infix notation, and we add parentheses to distinguish nested applications. For example, the following are two correct applications of the `and` rule.

$i.$	$\triangleright$	$a \wedge b \wedge c$	($\dots$)
$j.$	$\triangleright$	b	($\text{and } i$) 2
$i.$	$\triangleright$	$(a \wedge b) \wedge c$	($\dots$)
$j.$	$\triangleright$	c	($\text{and } i$) 2

A consequence of this behavior is that Alethe proofs cannot be handled by consumers that transform terms as suggested by SMT-LIB, because the two input terms `(and a b c)` and `(and (and a b) c)` cannot be distinguished after desugaring. One way to work around this is to desugar these terms like lists in functional programming languages. The operator takes the role of the list-constructor, and the neutral element of the operator takes the role of the empty list. For example, the term `(and a b c)` would desugar to `(and (and (and a b) c) true)` and the term `(and (and a b) c)` would desugar to `(and (and (and (and a b) true) c) true)`.

Subproofs The abstract notation denotes subproofs by marking them with a vertical line. To map this notation to a list of commands, Alethe uses the `anchor` command. This command indicates the start of a subproof. A subproof is concluded by a matching `step` command. This step must use a *concluding rule* (such as `subproof`, `bind`, and so forth).

After the `anchor` command, the proof uses `assume` commands to list the assumptions of the subproof. Subsequently, the subproof is a list of `step` commands that can use

prior steps in the subproofs as premises. It is not allowed to issue **assume** commands after the first **step** command of a subproof has been issued.

In the abstract notation, every step is associated with a context. The concrete syntax uses anchors to optimize this. The context is manipulated in a nested way: if a step pops $c_1, \dots, c_n$ from a context Γ , there is an earlier step which pushes precisely $c_1, \dots, c_n$ onto the context. Since contexts can only be manipulated by push and pop, context manipulations are nested. The **anchor** commands push onto the context and the corresponding **step** commands pop from the context. To indicate these changes to the context, every anchor is associated with a list of fixed variables and mappings. The list is provided by the **:args** annotation. If the list is empty, the **:args** annotation is omitted³. Note that, when an **anchor** command extends a context Γ with some mappings $x_1 \mapsto t_1, \dots, x_n \mapsto t_n$, then the terms t_i are normalized by applying the substitution $\text{subst}(\Gamma)$ to t_i . This is because the definition on page 6 extends the context by composing the substitutions.

The **:step** annotation of the anchor command is used to indicate the **step** command that will end the subproof. The clause of this **step** command is the conclusion of the subproof. While it is possible to infer the step that concludes a subproof from the structure of the proof, the explicit annotation simplifies proof checking and makes the proof easier to read. If the subproof uses a context, the **:args** annotation of the **anchor** command indicates the arguments added to the context for this subproof. The annotation provides the sort of fixed variables.

In the example proof (Figure 1) a subproof starts at the **anchor** command. It ends with the bind steps that finishes the proof for the renaming of the bound variable **z2** to **vr4**.

A further restriction applies: only the conclusion of a subproof can be used as a premise outside the subproof. Hence, a proof checking tool can remove the steps of the subproof from memory after checking it.

Example 4. This example shows the proof from Example 3 expressed as a concrete proof.

```
(assume h1 (forall ((x S)) (P x)))
(assume h2 (not (forall ((y S)) (P y))))
(anchor :step t5 :args ((y S) (:= (x S) y)))
(step t3 (cl (= x y)) :rule refl)
(step t4 (cl (= (P x) (P y))) :rule cong :premises (t3))
(step t5 (cl (= (forall ((x S)) (P x)) (forall ((y S)) (P y))))
         :rule bind)
(step t6 (cl (not (= (forall ((x S)) (P x)) (forall ((y S)) (P y))))
         (not (forall ((x S)) (P x)))
         (forall ((y S)) (P y))) :rule equiv_pos2)
(step t7 (cl) :rule resolution :premises (h1 h2 t5 t6))
```

³The only rule that allows an empty list is the subproof rule. Since this rule corresponds to implication introduction, it does not interact with binders.

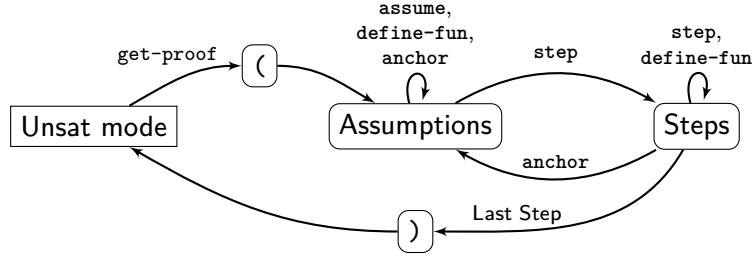


Figure 3: Abstract view of the transitions in an Alethe proof.

Alethe Proof Printing States Figure 2.1 shows the states of an Alethe proof abstractly. To generate a proof, the SMT solver must be in the *Unsat mode*, i.e., the solver determined that the problem is unsatisfiable. The SMT-LIB problem script then requests the proof by invoking the **get-proof** command. It is possible that this command fails. For example, if proof production was not activated. If there is no error, the proof is printed. Since SMT-LIB requires all response to be S-expressions, the solver first prints an opening parenthesis '`(`'. The solver then follows the behavior specified by the state machine to print the proof as a list of Alethe commands. After printing is done, the solver prints an additional closing parenthesis '`)`'. Note that these extra parentheses are printed to ensure that the proof output is SMT-LIB compliant. They are not part of the proof, and should only be printed when interacting with the solver through the SMT-LIB interface. For example, if a solver can write proofs directly into a file, it should not write these extra parentheses.

The printing state machine ensures that a constraint is maintained: assumptions can only appear at either the beginning of a proof or right after a subproof is started by the **anchor** command. They cannot be mixed with ordinary proof steps. This simplifies reconstruction. Each assumption printed at the beginning of the proof corresponds to assertions in the input problem, up to symmetry of equality. Proof printing concludes after the last step is printed and the solver returns to the Unsat mode and the user can issue further commands. Usually the last step is an outermost step (i.e., not within a subproof) that concludes the proof by deriving the empty clause, but this is not necessary. The solver is allowed to print some additional, useless, steps after the proof is concluded.

Sharing and Skolem Terms Usually, SMT solvers store terms internally in an efficient manner. A term data structure with perfect sharing ensures that every term is stored in memory precisely once. When printing the proof, this compact storage is unfolded. This leads to a blowup of the proof size.

Alethe can optionally use sharing⁴ to print common subterms only once. This is realized using the standard naming mechanism of SMT-LIB. A term t is annotated with a name n by writing `(! t :named n)` where n is a symbol. After a term is annotated with a name, the name can be used in place of the term. This is a purely syntactical

⁴For veriT this can be activated by the command-line option `--proof-with-sharing`.

replacement. Alethe continues to use the names already used in the input problem. Hence, terms that already have a name in the input problem can be replaced by that name and new names introduced in the proof must not use names already used in the input problem.

To limit the number of names, an SMT solver can use the following simple approach used by veriT. Before printing the proof, it iterates over all terms of the proof and recursively descend into the terms. It marks every unmarked subterm it visits. If a visited term is already marked, the solver assigns a new name to this term. If a term already has a name, it does not descend further into the term. By doing so, it ensures that only terms that appear as child of two different parent terms get a name. Since a named term is replaced with its name after its first appearance, a term that only appears as a child of one single term does not need a distinct name. Thanks to the perfect sharing representation, testing if a term is marked takes constant time and the overall traversal takes linear time in the proof size.

To simplify reconstruction, Alethe can optionally⁵ define Skolem constants as functions. In this case, the proof contains a list of **define-fun** commands that define shorthand 0-ary functions for the (**choice...**) terms needed. Without this option, no **define-fun** commands are issued, and the constants are expanded.

Implicit Transformations Overall, the following aspects are treated implicitly by Alethe.

- Symmetry of equalities that are not top-most equalities in steps with non-empty context.
- The order of literals in the clauses.
- The unfolding of names introduced by (**! t : named s**) in the original SMT-LIB problem or in the proof.
- The removal of other SMT-LIB annotations of the form (**! t ...**).
- The unfolding of function symbols introduced by **define-fun**.⁶
- If the input problem is in a logic without integers, then constants from **<numeral>** in the input problem will be printed using **<decimal>** or **<rational>** in the proof.

Alethe proofs contain steps for other aspects that are commonly left implicit, such as renaming of bound variables, and the application of substitutions.

3 Checking Alethe Proofs

In this section we present an abstract procedure to check if an Alethe proof is well-formed and valid. An Alethe proof is well-formed only if its anchors and steps are balanced.

⁵For veriT by using the command-line option `--proof-define-skolems`.

⁶For veriT this is only used when the user introduces veriT to print Skolem terms as defined functions. User defined functions in the input problem are not supported by veriT in proof production mode.

To check that this is the case, we replace innermost subproofs by holes until there is no subproof left. If the resulting reduced proof is free of anchors and steps that use concluding rules, then the overall proof is well-formed. To check if a proof is valid we have to check if all steps of a subproof adhere to the conditions of their rules before replacing the subproof by a hole. If all subproofs are valid and all steps in the reduced proof adhere to the conditions of their rule, then the entire proof is valid.

Formally, an Alethe proof P is a list $[C_1, \dots, C_n]$ of steps and anchors. Since every step uses an unique index, we assume that each step C_i in P uses i as its index. The context only changes at anchors and subproof-concluding steps. Therefore, the elements of $C_1, \dots, C_n$ that are steps are not associated with a context. Instead, the context can be computed from the prior anchors. The anchors only ever extend the context.

To check an Alethe proof we can iteratively eliminate the first-innermost subproof, i.e., the innermost subproof that does not come after a complete subproof. The restriction to the first subproofs simplifies the calculation of the context of the steps in the subproof.

Definition 4.1 (First-Innermost Subproof). Let P be the proof $[C_1, \dots, C_n]$ and $1 \leq start < end \leq n$ be two indices such that

- C_{start} is an anchor,
- C_{end} is a step that uses a concluding rule,
- no C_k with $k < start$ uses a concluding rule,
- no C_k with $start < k < end$ is an anchor or a step that uses a concluding rule.

Then $[C_{start}, \dots, C_{end}]$ is the first-innermost subproof of P .

Example 5. The proof in Example 4 has only one subproof and this subproof is also a first-innermost subproof. It is the subproof

```
(anchor :step t5 :args ((y S) (:= (x S) y)))
(step t3 (cl (= x y)) :rule refl)
(step t4 (cl (= (P x) (P y))) :rule cong :premises (t3))
(step t5 (cl (= (forall ((x S)) (P x)) (forall ((y S)) (P y))))
         :rule bind)
```

It is easy to calculate the context of the first-innermost subproof.

Definition 5.1 (Calculated Context). Let $[C_{start}, \dots, C_{end}]$ be the first-innermost subproof of P . Let $A_1, \dots, A_m$ be the anchors among $C_1, \dots, C_{start-1}$.

The calculated context of C_i is the context

$$c_{1,1}, \dots, c_{1,n_1}, \dots, c_{m,1}, \dots, c_{m,n_m}$$

where $c_{k,1}, \dots, c_{k,n_k}$ is the list of fixed variables and mappings associated with A_k .

Note that if C_i is an anchor, its calculated context does not contain the elements associated with C_i . Therefore, the context of C_{start} is the context of the steps before the subproof. Furthermore, the step C_{end} is the concluding step of the subproof and must have the same context as the steps surrounding the subproof. Hence, the context of C_{end} is the calculated context of C_{start} .

Example 6. The calculated context of the steps **t3** and **t5** in Example 4 is the context $x \mapsto y$. The calculated context of the concluding step **t5** and the anchor is empty.

A first-innermost subproof is valid if all its steps adhere to the conditions of their rule and only use premises that occur before them in the subproof. The conditions of each rule are listed in Section 5.

Definition 6.1 (Valid First-Innermost Subproof). Let $[C_{start}, \dots, C_{end}]$ be the first-innermost subproof of P . The subproof is *valid* if

- all steps C_i with $start < i < end$ only use premises C_j with $start < j < i$,
- all C_i that are steps adhere to the conditions of their rule under the calculated context of C_i ,
- the step C_{end} adheres to the conditions of its rule under the calculated context of C_{start} .

The only rule that can discharge assumptions in a subproof is the **subproof** rule. Therefore, an admissible subproof can only contain **assume** step if C_{end} is the **subproof** rule.

To eliminate a subproof we can replace the subproof with a hole that has at its conclusion the conclusion of the subproof. This is safe as long as the subproof that is eliminated is valid (see Section 3.2).

Definition 6.2. The function E eliminates the first-innermost subproof from a proof if there is one. Let P be a proof $[C_1, \dots, C_n]$. Then $E(P) = P$ if P has no first-innermost subproof. Otherwise, P has the first-innermost subproof $[C_{start}, \dots, C_{end}]$, and $E(P) = [C_1, \dots, C_{start-1}, C', C_{end+1}, \dots, C_n]$ where C' is a new step that uses the **hole** rule and has the index, conclusion, and premises of C_{end} .

It is important to add the premises of C_{end} to C' . The **let** rule can use additional premises and omitting those premises results in an unsound step. We can apply E iteratively to a proof P until we reach the least fixed point. Since P is finite we will always reach a fixed point in finitely many steps. The result is a list $[P_0, P_1, P_2, \dots, P_{last}]$ where $P_0 = P$, $P_1 = E(P)$, $P_2 = E(E(P))$ and $P_{last} = E(P_{last})$.

Example 7. Applying E to the proof in Example 4 gives us the proof

```
(assume h1 (forall ((x S)) (P x)))
(assume h2 (not (forall ((y S)) (P y))))
(step t5 (cl (= (forall ((x S)) (P x)) (forall ((y S)) (P y))))
```



```

                                :rule hole)
(step t6 (cl (= (forall ((x S)) (P x)) (forall ((y S)) (P y)))
          (not (forall ((x S)) (P x)))
          (forall ((y S)) (P y))) :rule equiv_pos2)
(step t7 (cl) :rule resolution :premises (h1 h2 t5 t6))

```

Since this proof contains no subproof, it is also P_{last} .

Definition 7.1 (Well-Formed Proof). The Alethe proof P is well-formed if every step uses a unique index and P_{last} contains no anchor or step that uses a concluding rule.

Definition 7.2 (Valid Alethe Proof). The proof P is a *valid Alethe proof* if

- P is well-formed,
- P does not contain any step that uses the hole rule,
- P_{last} contains a step that has the empty clause as its conclusion,
- the first-innermost subproof of every P_i , $i < last$ is valid,
- all steps C_i in P_{last} only use premises C_j in P_{last} with $1 \leq j < i$,
- all steps C_i in P_{last} adhere to the conditions of their rule under the empty context.

The condition that P contains no hole ensures that the original proof is complete and holes are only introduced by eliminating valid subproofs.

Example 8. The proof in Example 4 is valid. The only subproof is valid, the proof contains no hole, and P_{last} contains the step **t7** that concludes with the empty clause.

It is sometimes useful to speak about the steps that are not within a subproof. We call such a step an *outermost step*. In a well-formed proof those are the steps of P_{last} .

3.1 Contexts and Metaterms

We now direct our attention to subproofs with contexts. It is useful to give precise semantics to contexts to have the tools to check that rules that use contexts are sound. Contexts are local in the sense that they affect only the proof step they are applied to. For the full details on contexts see [2]. The presentation here is adapted from this publication, but omits some details.

To handle subproofs with contexts, we translate the contexts into λ -terms. This allows us to leverage the λ -calculus as an existing well-understood theory of binders. These λ -terms are called *metaterms*.

Definition 8.1 (Metaterm). Metaterms are expressions constructed by the grammar

$$M ::= \boxed{t} \mid \lambda x. M \mid (\lambda \bar{x}_n. M) \bar{t}_n$$

where t is an ordinary term and t_i and x_i have matching sorts for all $0 \leq i \leq 1$.

According to this definition, a metaterm is either a boxed term, a λ -abstraction, or an application to an uncurried λ -term. The annotation $\boxed{t}$ delimits terms from the context, a simple λ -abstraction is used to express fixed variables, and the application expresses simulations substitution of n terms.⁷

We use $=_{\alpha\beta}$ to denote syntactic equivalence modulo α -equivalence and β -reduction.

Proof steps with contexts can be encoded into proof steps with empty contexts, but with metaterms. A proof step

$$\text{i. } \Gamma \triangleright \quad t \approx u \quad (\text{rule } \bar{p}_n) [\bar{a}_m]$$

is encoded into

$$\text{i. } \triangleright \quad L(\Gamma)[t] \approx R(\Gamma)[u] \quad (\text{rule } \bar{p}_n) [\bar{a}_m]$$

where

$$\begin{aligned} L(\emptyset)[t] &= \boxed{t} & R(\emptyset)[u] &= \boxed{u} \\ L(x, \bar{c}_m)[t] &= \lambda x. L(\bar{c}_m)[t] & R(x, \bar{c}_m)[u] &= \lambda x. R(\bar{c}_m)[u] \\ L(\bar{x}_n \mapsto \bar{s}_n, \bar{c}_m)[t] &= (\lambda \bar{x}_n. L(\bar{c}_m)[t]) \bar{s}_n & R(\bar{x}_n \mapsto \bar{s}_n, \bar{c}_m)[u] &= R(\bar{c}_m)[u] \end{aligned}$$

To achieve the same effect as using the `subst()` function described above, we can translate the terms into metaterms, perform β -reduction, and rename bound variables if necessary [2, Lemma 11].

Example 9. The example on page 6 becomes

$$\begin{aligned} L(x \mapsto 7, x \mapsto g(x))[x] &= (\lambda x. (\lambda x. \boxed{x}) (g(x))) 7 =_{\alpha\beta} \boxed{g(7)} \\ L(x \mapsto 7, x, x \mapsto g(x))[x] &= (\lambda x. \lambda x. (\lambda x. \boxed{x}) (g(x))) 7 =_{\alpha\beta} \lambda x. \boxed{g(x)} \end{aligned}$$

Most proof rules that operate with contexts can easily be translated into proof rules using metaterms. The exception are the tautologous rules, such as `refl` and the `..._simplify` rules.

Steps that use such rules always encode a judgment $\models \Gamma \triangleright t \approx u$. With the encoding described above we get $L(\Gamma)[t] \approx R(\Gamma)[u] =_{\alpha\beta} \lambda \bar{x}_n. \boxed{t'} \approx \lambda \bar{x}_n. \boxed{u'}$ with some terms t', u' . To handle those terms, we use the `reify()` function. This function is defined as

$$\text{reify}(\lambda \bar{x}_n. \boxed{t} \approx \lambda \bar{x}_n. \boxed{u}) = \forall \bar{x}_n. (t \approx u).$$

Therefore, all tautologous rules with contexts represent a judgment $\models \text{reify}(T \approx U)$ where $T =_{\alpha\beta} L(\Gamma)[t]$ and $U =_{\alpha\beta} R(\Gamma)[u]$.

⁷The box annotation used here is unrelated to the boxes within the SMT solver discussed in the introduction.

Example 10. Consider the step

$$\text{i. } y, x \mapsto y \triangleright \quad x + 0 \approx y \quad \text{sum_simplify}$$

Translating the context into metaterms leads to

$$\text{i. } \triangleright \quad (\lambda y. (\lambda x. \boxed{x + 0}) y) \approx (\lambda y. \boxed{y}) \quad \text{sum_simplify}$$

Applying β -reduction leads to

$$\text{i. } \triangleright \quad (\lambda y. \boxed{y + 0}) \approx (\lambda y. \boxed{y}) \quad \text{sum_simplify}$$

Finally, using `reify()` leads to

$$\text{i. } \triangleright \quad \forall y. (y + 0 \approx y) \quad \text{sum_simplify}$$

This obviously holds in the theory of linear arithmetic.

3.2 Soundness

Any proof calculus should be sound. In the case of Alethe, most proof rules are standard rules, or simple tautologies. The rules that use context are unusual, but those proof rules were previously shown to be sound [2]. Alethe does not use any rules that are merely satisfiability preserving. The skolemization rules replace the bound variables with choice terms instead of fresh symbols.⁸ All Alethe rules express semantic implications. Overall, we assume in this document that the proof rules and proofs written in the abstract notation are sound.

Nevertheless, a modest gap remains. The concrete, command-based syntax does not precisely correspond to the abstract notation. In this section we will address the soundness of concrete Alethe proofs.

Theorem 10.1 (Soundness of Concrete Alethe Proofs). If there is a valid Alethe proof $P = [C_1, \dots, C_n]$ that has the formulas $\varphi_1, \dots, \varphi_m$ as the conclusions of the outermost assume steps, then

$$\varphi_1, \dots, \varphi_m \models \perp.$$

Here, $\models$ represents semantic consequence in the many-sorted first order logic of SMT-LIB with the theories of uninterpreted functions and linear arithmetic extended with the choice operator and clauses.

To show the soundness of a valid Alethe proof $P = [C_1, \dots, C_n]$, we can use the same approach as for the definition of validity: consider first-innermost subproof first and then replace them by holes. Since valid proofs do not contain holes, we have to generalize the induction to allow holes that were introduced by the elimination of subproofs. We start with simple subproofs with empty contexts and without nested subproofs.

⁸The `define-fun` function can introduce fresh symbols, but we will assume here that those commands have been eliminated by unfolding the definition.

Lemma 10.1. Let P be a proof that contains a valid first-innermost subproof where C_{end} is a subproof step. Let ψ be the conclusion of C_{end} . Then $\models \psi$ holds.

Proof. First, we use induction on the number of steps n after C_{start} . Let ψ_n be the conclusion of $C_{start+n}$ and V_n the conclusions of the **assume** steps in $[C_{start}, \dots, C_{start+n}]$. Assumptions always introduce unit clauses. We will identify unit clauses with their single literal. We will show $V_n \models \psi_n$ if $start + n < end$.

If $n = 1$, then $C_{start+n} = C_{start+1}$ must either be a tautology, or an **assume** step. In the first case, $\models \psi_{start+1}$ holds, and in the second case $\psi_{start+1} \models \psi_{start+1}$ holds.

For subsequent n , $C_{start+n}$ is either an ordinary step, or an **assume** step. In the second case, $\psi_{start+n} \models \psi_{start+n}$ which can be weakened to $V_n \models \psi_{start+n}$. In the first case, the step $C_{start+n}$ has a set of premises S . For each step $C_{start+i} \in S$ we have $i < n$ and $V_i \models \psi_{start+i}$ due to the induction hypothesis. Since $i < n$, the premises V_i are a subset of V_n and we can weaken $V_i \models \psi_{start+i}$ to $V_n \models \psi_{start+i}$. Since all premises of $C_{start+n}$ are the consequence of V_n we get $V_n \models \psi_n$.

The step C_{end-1} is the last step of the subproof that does not use a concluding rule. At this step we have $V_{end-1} \models \psi_{end-1}$. Since C_{end} is not an **assume** step, the set $V_{end-1} = \{\varphi_1, \dots, \varphi_m\}$ contains all assumptions in the subproof. By the deduction theorem we get

$$\models \varphi_1 \wedge \dots \wedge \varphi_m \rightarrow \psi_{end-1}.$$

This can be transformed into the clause

$$\models \neg\varphi_1, \dots, \neg\varphi_m, l_1, \dots, l_o.$$

where $l_1, \dots, l_o$ are the literals of ψ_{end-1} . This clause is exactly the conclusion of the step C_{end} according to the definition of the **subproof** rule. $\square$

We can do the same reasoning as for Lemma 10.1 for subproofs with contexts. This is slightly complicated by the **let** rule that can use extra premises.

Lemma 10.2. Let P be a proof that contains a valid first-innermost subproof where C_{end} is a step using one of: **bind**, **sco_ex**, **sco_forall**, **onepoint**, **let**.

Then $V \models \Gamma \triangleright \psi$ where Γ is the calculated context of C_{start} and ψ is the conclusion of C_{end} . The set V is empty if C_{end} does not use the **let** rule. Otherwise, it contains all conclusions of the **assume** steps among $[C_\delta, \dots, C_{start}]$ where δ is either the largest index $\delta < start$ such that s_δ is an anchor, or 1 if no such index exist. Hence, there is no anchor between C_δ and C_{start} .

Proof. The step C_{start} is an anchor due to the definition of innermost-first subproof. Let $c_1, \dots, c_n$ be the context introduced by the anchor C_{start} , and let Γ be the calculated context of C_{start} . $\Gamma' = \Gamma, c_1, \dots, c_n$ is the calculated context of the steps in the subproof after C_{start} .

The step C_{end} is a step

$$\begin{array}{c} | \qquad \qquad \qquad \dots \end{array}$$

$$\frac{\text{end} - 1. \Gamma' \triangleright \psi'}{\text{end. } \Gamma \triangleright \psi} \quad (\text{rule } i_1, \dots, i_n)$$

Since we assume the step C_{end} is correctly employed, $\models \Gamma \triangleright \psi$ holds, as long as $\models \Gamma' \triangleright \psi'$ holds.

We perform the same induction as for Lemma 10.1 over the steps in $[C_{\text{start}}, \dots, C_{\text{end}}]$. Since C_{end} does not use the **subproof** rule, the subproof does not contain any assumptions and V_i stays empty. Again, we are interested in the step $C_{\text{end}-1}$. At this step we get $\models \Gamma' \triangleright \psi'$.

Only the **let** rule uses additional premises $C_{i_1}, \dots, C_{i_n}$. Hence, for all other rules, the conclusion cannot depend on any step outside the subproof and V is empty. Due to the definition of first-innermost subproof, all steps $C_{i_1}, \dots, C_{i_n}$ are in the same subproof that starts at C_δ .

The steps $C_{i_1}, \dots, C_{i_n}$ might depend on some **assume** steps that appear before them in their subproof. This is the case if the steps are outermost steps, or if the subproof that starts at C_δ concludes with a **subproof** step. In this case we can, as we saw in the proof of Lemma 10.1, weaken their judgments to include all assumptions in $[C_\delta, \dots, C_{\text{start}}]$.

If the subproof that starts at C_δ concludes with any other rule, then there cannot be any assumptions and V is empty. $\square$

By using Lemma 10.1 and Lemma 10.2 we can now show that a valid, concrete Alethe proof is sound. That is, we can show Theorem 10.1.

Proof. Since $P = [C_1, \dots, C_n]$ is valid, all steps that do not use the **hole** rule adhere to their rule. Since we assume that the abstract notation and the rules are sound, we only have to worry about the steps using the **hole** rule. Those should be sound, i.e., for a **hole** step with the conclusion ψ , premises V , and context Γ the judgment $V \models \Gamma \triangleright \psi$ must hold.

Since P is a valid proof there is a sequence $[P_0, \dots, P_{\text{last}}]$ as discussed in Section 3. For $i < \text{last}$, $E(P_i) = P_{i+1}$ replaces the first-innermost subproof in P_i by a **hole** with the conclusion ψ . Furthermore, the context of the introduced **hole** corresponds to the context Γ of the start of the subproof. Since P is a valid proof, the first-innermost subproof eliminated by E is always valid. Therefore, we can apply Lemma 10.1 or Lemma 10.2 to conclude that the **hole** introduced by E is sound.

Since P_0 does not contain any **holes**, the **holes** in each proof P_i are all introduced by innermost-first subproof elimination. Therefore, they are sound. In consequence, all **holes** in P_{last} are sound and we can perform the same argument as in the proof of Lemma 10.1 to the proof P_{last} .

Let j be the index of the step in P_{last} that concludes with the empty clause. Let $\text{start} = 1$ and $\text{end} = j$ in the argument of Lemma 10.1. This shows that $V \models \perp$, where V is the conclusion of the **assume** steps in the sublist $[C_1, \dots, C_j]$ of P_{last} . We can weaken this by adding the conclusions of the **assume** steps in $[C_j, \dots, C_n]$ of P_{last} to get $\varphi_1, \dots, \varphi_m \models \perp$. $\square$

4 Core Concepts of the Alethe Rules

Together with the language, the Alethe format also includes a set of proof rule. Section 5 gives a full list of all proof rules. Currently, the proof rules correspond to the rules that the solver `veriT` can emit. For the rest of this section, we will discuss some general concepts related to the rules.

Tautologous Rules and Simple Deduction Most rules introduce tautologies. One example is the `and_pos` rule: $\neg(\varphi_1 \wedge \varphi_2 \wedge \dots \wedge \varphi_n), \varphi_i$. Other rules derive their conclusion from a single premise. Those rules are primarily used to simplify Boolean connectives during preprocessing. For example, the `implies` rule eliminates an implication: From $\varphi_1 \rightarrow \varphi_2$, it deduces $\neg\varphi_1, \varphi_2$.

Resolution. CDCL(T)-based SMT solvers, and especially their SAT solvers, are fundamentally based on resolution of clauses. Hence, Alethe also has dedicated clauses and a resolution proof rule. However, since SMT solvers do not enforce a strict clausal normal form, ordinary disjunction is also used. Keeping clauses and disjunctions distinct simplifies rule checking. For example, in the case of resolution there is a clear distinction between unit clauses where the sole literal starts with a disjunction and non-unit clauses. The syntax for clauses uses the `cl` operator, while disjunctions use the standard SMT-LIB `or` operator. The `or` rule is responsible for converting disjunctions into clauses.

The Alethe proofs use a generalized propositional resolution rule with the name `resolution` or `th_resolution`. Both names denote the same rule. The difference only serves to distinguish if the rule was introduced by the SAT solver or by a theory solver. The resolution step is purely propositional; there is no notion of a unifier. The resolution rules also implicitly simplify repeated negations at the head of literals.

The premises of a resolution step are clauses, and the conclusion is a clause that has been derived from the premises by some binary resolution steps.

Quantifier Instantiation The rule `forall_inst` expresses quantifier instantiation. It produces a formula of the form $(\neg\forall\bar{x}_n.\varphi) \vee \varphi[\bar{t}_n]$, where φ is a term containing the free variables $\bar{x}_n$, and for each i the ground term t_i is a new term with the same sort as x_i .⁹

The argument of a `forall_inst` step, for a quantification of the form $\forall x_1, \dots, x_n.\varphi$, is the list $t_1, \dots, t_n$, where t_i denotes the substitute of variable x_i . While this information can be recovered from the term, providing it explicitly helps reconstruction because the implicit reordering of equalities obscures which terms have been used as instances. Existential quantifiers are handled by skolemization.

Linear Arithmetic Proofs for linear arithmetic use a number of straightforward rules, such as `la_totality` ($t_1 \leq t_2 \vee t_2 \leq t_1$)¹⁰ and the main rule `la_generic`. The conclusion of

⁹For historical reasons, `forall_inst` has a unit clause with a disjunction as its conclusion and not the clause $(\neg\forall\bar{x}_n.\varphi), \varphi[\bar{t}_n]$.

¹⁰This rule also has a unit clause with a disjunction as its conclusion.

an `la_generic` step is a tautology $\neg\varphi_1, \neg\varphi_2, \dots, \neg\varphi_n$ where the φ_i are linear (in)equalities. Checking the validity of this clause amounts to checking the unsatisfiability of the system of linear equations $\varphi_1, \varphi_2, \dots, \varphi_n$. The annotation of an `la_generic` step contains a coefficient for each (in)equality. The result of forming the linear combination of the literals with the coefficients is a trivial inequality between constants.

Example 11. The following example is the proof for the unsatisfiability of $(x + y < 1) \vee (3 < x)$, $x \approx 2$, and $0 \approx y$.

1. $\triangleright$	$(3 < x) \vee (x + y < 1)$	assume
2. $\triangleright$	$x \approx 2$	assume
3. $\triangleright$	$0 \approx y$	assume
4. $\triangleright$	$(3 < x), (x + y < 1)$	(or 1)
5. $\triangleright$	$\neg(3 < x), \neg(x \approx 2)$	<code>la_generic</code> [1.0, 1.0]
6. $\triangleright$	$\neg(3 < x)$	(resolution 2, 5)
7. $\triangleright$	$x + y < 1$	(resolution 4, 6)
8. $\triangleright$	$\neg(x + y < 1), \neg(x \approx 2) \vee \neg(0 \approx y)$	<code>la_generic</code> [1.0, -1.0, 1.0]
9. $\triangleright$	$\perp$	(resolution 8, 7, 2, 3)

Skolemization and Other Preprocessing Steps One typical example for a rule with context is the `sko_ex` rule that is used to express skolemization of an existentially quantified variable. The conclusion of a step that uses this rule is an equality. The left-hand side is a formula starting with an existential quantifier over some variable x . In the formula on the right-hand side, the variable is replaced by the appropriate Skolem term. To provide a proof for the replacement, the `sko_ex` step uses one premise. The premise has a context that maps the existentially quantified variable to the appropriate Skolem term.

i. $\Gamma, x \mapsto (\varepsilon x. \varphi) \triangleright$	$\varphi \approx \psi$	(...)
j. $\Gamma \triangleright$	$(\exists x. \varphi) \approx \psi$	(<code>sko_ex</code>)

Example 12. To illustrate how such a rule is applied, consider the following example taken from [2]. Here the term $\neg p(\varepsilon x. \neg p(x))$ is skolemized. The `refl` rule expresses a simple tautology on the equality (reflexivity in this case), `cong` is functional congruence, and `sko_forall` works like `sko_ex`, except that the choice term is $\varepsilon x. \neg \varphi$.

1. $x \mapsto (\varepsilon x. \neg(p x)) \triangleright$	$x \approx \varepsilon x. \neg(p x)$	<code>refl</code>
2. $x \mapsto (\varepsilon x. \neg(p x)) \triangleright$	$(p x) \approx p(\varepsilon x. \neg(p x))$	(<code>cong</code> 1)
3. $\triangleright$	$(\forall x. (p x)) \approx (p(\varepsilon x. \neg(p x)))$	(<code>sko_forall</code> 2)
4. $\triangleright$	$(\neg \forall x. (p x)) \approx \neg(p(\varepsilon x. \neg(p x)))$	(<code>cong</code> 3)

4.1 Bitvector Reasoning with Bitblasting

A standard approach to handle bitvector reasoning in SMT solvers is bitblasting. Bitblasting works by translating bitvector functions to propositional formulas that model the logical circuit of the bitvector function.

To express bitblasting in Alethe proof rules, the Alethe calculus uses multiple families of helper functions: **bbterm**, **bit_of_m**, **bvsize**, and **bv_nⁱ**. Functions in the families are distinguished either by overloading (**bbterm** and **bvsize**) or by explicit indexing (**bit_of_m** and **bv_nⁱ**). In concrete syntax, to avoid name clashes with user defined functions, **bbterm** is written as **@bbterm**, **bit_of** as **@bit_of**, **bvsize** as **@bvsize**, and **bv** as **@bv**. The SMT-LIB standard specifies that simple symbols starting with “@” are reserved for solver generated functions.

The family **bbterm** contains one function for each bitvector sort **BitVec n**.

$$\mathbf{bbterm} : \underbrace{\mathbf{Bool} \dots \mathbf{Bool}}_n \rightarrow \mathbf{BitVec} \, n.$$

Intuitively, the function **bbterm** takes a list of boolean arguments and packs them into a bitvector. Let $\langle u_0, \dots, u_{n-1} \rangle$ denote a bitvector of sort **BitVec n**, where $u_i = \top$ if the bit at position i is 1, and $u_i = \perp$ otherwise. The bit u_{n-1} is the least significant bit. This notation, which allows us to describe bitvectors in terms of sequences of booleans, will be useful when defining the several bitvector related functions. Note that:

$$\mathbf{bbterm} \, v_0 \dots v_{n-1} = \langle v_0, \dots, v_{n-1} \rangle.$$

Also, note that the **bbterm** functions could be defined in terms of standard SMT-LIB functions.

$$\begin{aligned} \mathbf{bbterm} \, v_0 \dots v_{n-1} := & \mathbf{concat} \left(\mathbf{concat} \left(\dots \right. \right. \\ & \left. \left. \mathbf{concat} \left(\mathbf{ite} \, v_0 \, \langle \top \rangle \, \langle \perp \rangle \right) \left(\mathbf{ite} \, v_1 \, \langle \top \rangle \, \langle \perp \rangle \right) \right) \right. \\ & \dots \\ & \left. \mathbf{ite} \, v_{n-2} \, \langle \top \rangle \, \langle \perp \rangle \right) \\ & \left. \mathbf{ite} \, v_{n-1} \, \langle \top \rangle \, \langle \perp \rangle \right) \end{aligned}$$

The function **bit_of_m** yields the Boolean corresponding to the m -th bit from a given bitvector. It is used as a symbol indexed by the desired m -th position, similarly to the built-in function **extract**. Hence, for $0 \leq m < n$, the concrete syntax expression (**@bit_of m**) will denote functions:

$$\mathbf{bit_of}_m : \mathbf{BitVec} \, n \rightarrow \mathbf{Bool}.$$

These functions are defined as:

$$\mathbf{bit_of}_m \langle u_0, \dots, u_{n-1} \rangle := u_m.$$

The functions **bvsize** return the size of a bitvector. Formally, there is one **bvsize** for each bitvector sort **BitVec n**. Each **bvsize** is a constant function that returns the size of the corresponding bitvector sort. Using the previously introduced notation:

$$\begin{aligned} \text{bvsiz}e &: \text{BitVec } n \rightarrow \mathbb{N} \\ \text{bvsiz}e \langle u_0, \dots, u_{n-1} \rangle &:= n \end{aligned}$$

Finally, bv_n^i is a family of constants indexed by two parameters: a bitvector length n , and a natural number i . We write $(_ @\text{bv}_n i)$ for bv_n^i , where the space before n (in the concrete-syntax term) is omitted for historical reasons. Each bv_n^i is the bitvector constant that represents the bitvector of length n that encodes the integer i . Formally, it corresponds to `nat2bv[n](i)`, where `nat2bv` is defined as in the SMT-LIB standard.¹¹

4.2 RARE rules

RARE is a domain-specific language for defining rewrite rules. There are two versions of the language, RARE 1.0 and RARE 2.0, and RARE 2.0 is defined as a fragment of the *Eunoia* logical framework. For details about the RARE 2.0 rewriting rules specification language and Eunoia, see <https://github.com/alethe-proofs/rare2-rules/>. The Alethe standard allows for steps in proof certificates to refer to rewrite rules specified in terms of the RARE 2.0 language. Generic interfaces are provided by the rules `rare_rewrite`, for a single rewrite, and `multi_rare_rewrite`, for multiple rewrites in the same step.

5 The Alethe Rules

This section provides a list of all proof rules supported by Alethe. To make this long list more accessible, the section first lists multiple classes of proof rules. The classes are not mutually exclusive: for example, the `la_generic` rule is both a linear arithmetic rule and introduces a tautology. The number in brackets is the position of the rule in the overall list of proof rules. Table 16 lists rules that serve a special purpose. Table 18 lists all rules that introduce tautologies. That is, regular rules that do not use premises.

The subsequent section, starting at 5.2, defines all rules and provides example proofs for complicated rules. The index of proof rules on page 62 can be used to quickly find the definition of rules.

5.1 Classifications of the Rules

Table 16: Special rules.

Rule	Description
<code>assume</code> (1)	Introduction of an assumption.
<code>hole</code> (2)	Placeholder for rules not defined here.
<code>subproof</code> (10)	Concludes a subproof and discharges local assumptions.

¹¹See <https://smt-lib.github.io/theories-FixedSizeBitVectors.shtml>.

Table 17: Resolution and related rules.

Rule	Description
resolution (7)	Chain resolution of two or more clauses.
th_resolution (6)	As resolution, but used by theory solvers.
tautology (8)	Simplification of tautological clauses to $\top$.
contraction (9)	Removal of duplicated literals.

Table 18: Rules introducing tautologies.

Rule	Description
true (3)	$\top$
false (4)	$\neg \perp$
not_not (5)	$\neg(\neg\neg\varphi), \varphi$
la_generic (11)	Tautologous disjunction of linear inequalities.
lia_generic (12)	Tautologous disjunction of linear integer inequalities.
la_disequality (13)	$t_1 \approx t_2 \vee \neg(t_1 \leq t_2) \vee \neg(t_2 \leq t_1)$
la_totality (14)	$t_1 \leq t_2 \vee t_2 \leq t_1$
la_tautology (15)	A trivial linear tautology.
la_mult_pos (16)	Multiplication with a positive factor.
la_mult_neg (17)	Multiplication with a negative factor.
forall_inst (21)	Quantifier instantiation.
refl (22)	Reflexivity after applying the context.
eq_reflexive (25)	$t \approx t$ without context.
eq_transitive (26)	$\neg(t_1 \approx t_2), \dots, \neg(t_{n-1} \approx t_n), t_1 \approx t_n$
eq_congruent (27)	$\neg(t_1 \approx u_1), \dots, \neg(t_n \approx u_n), f(t_1, \dots, t_n) \approx f(u_1, \dots, u_n)$
eq_congruent_pred (28)	$\neg(t_1 \approx u_1), \dots, \neg(t_n \approx u_n), P(t_1, \dots, t_n) \approx P(u_1, \dots, u_n)$
qnt_cnf (29)	Clausification of a quantified formula.
and_pos (48)	$\neg(\varphi_1 \wedge \dots \wedge \varphi_n), \varphi_k$
and_neg (49)	$(\varphi_1 \wedge \dots \wedge \varphi_n), \neg\varphi_1, \dots, \neg\varphi_n$
or_pos (50)	$\neg(\varphi_1 \vee \dots \vee \varphi_n), \varphi_1, \dots, \varphi_n$
or_neg (51)	$(\varphi_1 \vee \dots \vee \varphi_n), \neg\varphi_k; \text{ with } 1 \leq k \leq n$
xor_pos1 (52)	$\neg(\varphi_1 \mathbf{xor} \varphi_2), \varphi_1, \varphi_2$
xor_pos2 (53)	$\neg(\varphi_1 \mathbf{xor} \varphi_2), \neg\varphi_1, \neg\varphi_2$
xor_neg1 (54)	$\varphi_1 \mathbf{xor} \varphi_2, \varphi_1, \neg\varphi_2$
xor_neg2 (55)	$\varphi_1 \mathbf{xor} \varphi_2, \neg\varphi_1, \varphi_2$
implies_pos (56)	$\neg(\varphi_1 \rightarrow \varphi_2), \neg\varphi_1, \varphi_2$
implies_neg1 (57)	$\varphi_1 \rightarrow \varphi_2, \varphi_1$
implies_neg2 (58)	$\varphi_1 \rightarrow \varphi_2, \neg\varphi_2$
equiv_pos1 (59)	$\neg(\varphi_1 \approx \varphi_2), \varphi_1, \neg\varphi_2$
equiv_pos2 (60)	$\neg(\varphi_1 \approx \varphi_2), \neg\varphi_1, \varphi_2$
equiv_neg1 (61)	$\varphi_1 \approx \varphi_2, \neg\varphi_1, \neg\varphi_2$
equiv_neg2 (62)	$\varphi_1 \approx \varphi_2, \varphi_1, \varphi_2$

ite_pos1 (65)	$\neg(\text{ite } \varphi_1 \varphi_2 \varphi_3), \varphi_1, \varphi_3$
ite_pos2 (66)	$\neg(\text{ite } \varphi_1 \varphi_2 \varphi_3), \neg\varphi_1, \varphi_2$
ite_neg1 (67)	$(\text{ite } \varphi_1 \varphi_2 \varphi_3), \varphi_1, \neg\varphi_3$
ite_neg2 (68)	$(\text{ite } \varphi_1 \varphi_2 \varphi_3), \neg\varphi_1, \neg\varphi_2$
connective_def (71)	Definition of some connectives.
and_simplify (72)	Simplification of a conjunction.
or_simplify (73)	Simplification of a disjunction.
not_simplify (74)	Simplification of a Boolean negation.
implies_simplify (75)	Simplification of an implication.
equiv_simplify (76)	Simplification of an equivalence.
bool_simplify (77)	Simplification of combined Boolean connectives.
ac_simp (78)	Flattening and removal of duplicates for $\vee$ or $\wedge$.
ite_simplify (79)	Simplification of if-then-else.
qnt_simplify (80)	Simplification of constant quantified formulas.
qnt_join (82)	Joining of consecutive quantifiers.
qnt_rm_unused (83)	Removal of unused quantified variables.
eq_simplify (84)	Simplification of equality.
div_simplify (85)	Simplification of division.
prod_simplify (86)	Simplification of products.
unary_minus_simplify (87)	Simplification of the unary minus.
minus_simplify (88)	Simplification of subtractions.
sum_simplify (89)	Simplification of sums.
comp_simplify (90)	Simplification of arithmetic comparisons.
distinct_elim (93)	Elimination of the distinct operator.
la_rw_eq (94)	$(t \approx u) \approx (t \leq u \wedge u \leq t)$
nary_elim (95)	Eliminate n -ary application of operators via binary applications.
eq_symmetric (114)	Symmetry of equality as equivalence.
miniscope_distribute (115)	Distribution of quantified variables over $\vee, \wedge$.
miniscope_split (116)	Splitting of quantified variables over $\vee, \wedge$.
miniscope_ite (117)	Miniscoping of if-then-else.

Table 19: Linear arithmetic rules.

Rule	Description
la_generic (11)	Tautologous disjunction of linear inequalities.
lia_generic (12)	Tautologous disjunction of linear integer inequalities.
la_disequality (13)	$t_1 \approx t_2 \vee \neg(t_1 \leq t_2) \vee \neg(t_2 \leq t_1)$
la_totality (14)	$t_1 \leq t_2 \vee t_2 \leq t_1$
la_tautology (15)	A trivial linear tautology.
la_mult_pos (16)	Multiplication with a positive factor.
la_mult_neg (17)	Multiplication with a negative factor.
la_rw_eq (94)	$(t \approx u) \approx (t \leq u \wedge u \leq t)$

div_simplify (85)	Simplification of division.
prod_simplify (86)	Simplification of products.
unary_minus_simplify (87)	Simplification of the unary minus.
minus_simplify (88)	Simplification of subtractions.
sum_simplify (89)	Simplification of sums.
comp_simplify (90)	Simplification of arithmetic comparisons.

Table 20: Quantifier handling.

Rule	Description
forall_inst (21)	Instantiation of a universal quantifier.
bind (18)	Renaming of bound variables.
sko_ex (19)	Skolemization of an existential quantifier.
sko_forall (20)	Skolemization of an universal quantifier.
qnt_cnf (29)	Clausification of quantified formulas.
qnt_simplify (80)	Simplification of constant quantified formulas.
onepoint (81)	The one-point rule.
qnt_join (82)	Joining of consecutive quantifiers.
qnt_rm_unused (83)	Removal of unused quantified variables.
miniscope_distribute (115)	Distribution of quantified variables over $\vee$, $\wedge$.
miniscope_split (116)	Splitting of quantified variables over $\vee$, $\wedge$.
miniscope_ite (117)	Miniscoping of if-then-else.

Table 21: Skolemization rules.

Rule	Description
sko_ex (19)	Skolemization of existential variables.
sko_forall (20)	Skolemization of universal variables.

Table 22: Clausification rules. These rules can be used to perform propositional clausification.

Rule	Description
and (30)	And elimination.
not_or (31)	Elimination of a negated disjunction.
or (32)	Disjunction to clause.
not_and (36)	Distribution of negation over a conjunction.
xor1 (37)	From $(\mathbf{xor} \varphi_1 \varphi_2)$ to φ_1, φ_2 .
xor2 (38)	From $(\mathbf{xor} \varphi_1 \varphi_2)$ to $\neg\varphi_1, \neg\varphi_2$.
not_xor1 (39)	From $\neg(\mathbf{xor} \varphi_1 \varphi_2)$ to $\varphi_1, \neg\varphi_2$.
not_xor2 (40)	From $\neg(\mathbf{xor} \varphi_1 \varphi_2)$ to $\neg\varphi_1, \varphi_2$.
implies (41)	From $\varphi_1 \rightarrow \varphi_2$ to $\neg\varphi_1, \varphi_2$.

not_implies1 (42)	From $\neg(\varphi_1 \rightarrow \varphi_2)$ to φ_1 .
not_implies2 (43)	From $\neg(\varphi_1 \rightarrow \varphi_2)$ to $\neg\varphi_2$.
equiv1 (44)	From $\varphi_1 \approx \varphi_2$ to $\neg\varphi_1, \varphi_2$.
equiv2 (45)	From $\varphi_1 \approx \varphi_2$ to $\varphi_1, \neg\varphi_2$.
not_equiv1 (46)	From $\neg(\varphi_1 \approx \varphi_2)$ to φ_1, φ_2 .
not_equiv2 (47)	From $\neg(\varphi_1 \approx \varphi_2)$ to $\neg\varphi_1, \neg\varphi_2$.
and_pos (48)	$\neg(\varphi_1 \wedge \dots \wedge \varphi_n), \varphi_k$
and_neg (49)	$(\varphi_1 \wedge \dots \wedge \varphi_n), \neg\varphi_1, \dots, \neg\varphi_n$
or_pos (50)	$\neg(\varphi_1 \vee \dots \vee \varphi_n), \varphi_1, \dots, \varphi_n$
or_neg (51)	$(\varphi_1 \vee \dots \vee \varphi_n), \neg\varphi_k$
xor_pos1 (52)	$\neg(\varphi_1 \mathbf{xor} \varphi_2), \varphi_1, \varphi_2$
xor_pos2 (53)	$\neg(\varphi_1 \mathbf{xor} \varphi_2), \neg\varphi_1, \neg\varphi_2$
xor_neg1 (54)	$\varphi_1 \mathbf{xor} \varphi_2, \varphi_1, \neg\varphi_2$
xor_neg2 (55)	$\varphi_1 \mathbf{xor} \varphi_2, \neg\varphi_1, \varphi_2$
implies_pos (56)	$\neg(\varphi_1 \rightarrow \varphi_2), \neg\varphi_1, \varphi_2$
implies_neg1 (57)	$\varphi_1 \rightarrow \varphi_2, \varphi_1$
implies_neg2 (58)	$\varphi_1 \rightarrow \varphi_2, \neg\varphi_2$
equiv_pos1 (59)	$\neg(\varphi_1 \approx \varphi_2), \varphi_1, \neg\varphi_2$
equiv_pos2 (60)	$\neg(\varphi_1 \approx \varphi_2), \neg\varphi_1, \varphi_2$
equiv_neg1 (61)	$\varphi_1 \approx \varphi_2, \neg\varphi_1, \neg\varphi_2$
equiv_neg2 (62)	$\varphi_1 \approx \varphi_2, \varphi_1, \varphi_2$
let (91)	Elimination of the let operator.
distinct_elim (93)	Elimination of the distinct operator.
nary_elim (95)	Elimination of n-ary application of operators.

Table 23: Simplification rules. These rules represent typical operator-level simplifications.

Rule	Description
connective_def (71)	Definition of the Boolean connectives.
and_simplify (72)	Simplification of a conjunction.
or_simplify (73)	Simplification of a disjunction.
not_simplify (74)	Simplification of a Boolean negation.
implies_simplify (75)	Simplification of an implication.
equiv_simplify (76)	Simplification of an equivalence.
bool_simplify (77)	Simplification of combined Boolean connectives.
ac_simp (78)	Flattening and removal of duplicates for $\vee$ or $\wedge$.
aci_simp (118)	Simplification of terms through ACI reasoning.
ite_simplify (79)	Simplification of if-then-else.
qnt_simplify (80)	Simplification of constant quantified formulas.
onepoint (81)	The one-point rule.
qnt_join (82)	Joining of consecutive quantifiers.
qnt_rm_unused (83)	Removal of unused quantified variables.
eq_simplify (84)	Simplification of equalities.

<code>div_simplify</code> (85)	Simplification of division.
<code>prod_simplify</code> (86)	Simplification of products.
<code>unary_minus_simplify</code> (87)	Simplification of the unary minus.
<code>minus_simplify</code> (88)	Simplification of subtractions.
<code>sum_simplify</code> (89)	Simplification of sums.
<code>comp_simplify</code> (90)	Simplification of arithmetic comparisons.
<code>qnt_simplify</code> (80)	Simplification of constant quantified formulas.

Table 24: Bitvector rules.

Rule	Description
<code>bitblast_extract</code> (98)	Bitblasting of extract .
<code>bitblast_ult</code> (102)	Bitblasting of ult .
<code>bitblast_add</code> (104)	Bitblasting of add .
<code>bitblast_neg</code> (105)	Bitblasting of add .

Table 25: Rules used by `cvc5`, but not by `veriT`.

Rule	Description
<code>bitblast_extract</code> (98)	Bitblasting of extract .
<code>bitblast_ult</code> (102)	Bitblasting of ult .
<code>bitblast_add</code> (104)	Bitblasting of add .
<code>la_mult_pos</code> (16)	Multiplication with a positive factor.
<code>la_mult_neg</code> (17)	Multiplication with a negative factor.
<code>symm</code> (112)	Symmetry of equality.
<code>not_symm</code> (113)	Symmetry of not-equal.
<code>reordering</code> (34)	Permutation of the literals in a clause.
<code>shuffle</code> (35)	Permutation of the arguments of a commutative operator.

Table 26: Rules used by the Carcara elaborator.

Rule	Description
<code>weakening</code> (33)	Weakening of a clause.
<code>eq_symmetric</code> (114)	Symmetry of equality as equivalence.

5.2 Rule List

Rule 1: **assume**

i. $\triangleright$ φ **assume**
 where φ corresponds up to the orientation of equalities to a formula asserted in the input problem, or φ is a local assumption in a subproof.

Remark. This rule can not be used by the (**step...**) command. Instead it corresponds to the dedicated (**assume...**) command.

Rule 2: hole

$i. \triangleright \quad \varphi \quad (\text{hole } p_2, \dots, p_n) [a_1, \dots, a_m]$

where φ is any well-formed formula.

This rule can be used to express holes in the proof. It can be used by solvers as a placeholder for proof steps that are not yet expressed by the proof rules in this document. A proof checker *must not* accept a proof as valid that contains this rule even if the checker can somehow check this rule. However, it is possible for checkers to have a dedicated status for proofs that contain this rule and are otherwise valid. Any other tool can accept or reject proofs that contain this rule.

The premises and arguments are arbitrary, but must follow the syntax for premises and arguments.

Rule 3: true

$i. \triangleright \quad \top \quad \text{true}$

Rule 4: false

$i. \triangleright \quad \neg \perp \quad \text{false}$

Rule 5: not_not

$i. \triangleright \quad \neg(\neg\neg\varphi), \varphi \quad \text{not_not}$

Remark. This rule is useful to remove double negations from a clause by resolving a clause with the double negation on φ .

Rule 6: th_resolution

This rule is the chain resolution of two or more clauses.

$i_1. \triangleright \quad l_1^1, \dots, l_{k^1}^1 \quad (\dots)$
 $\vdots$
 $i_n. \triangleright \quad l_1^n, \dots, l_{k^n}^n \quad (\dots)$
 $j. \triangleright \quad l_{s_1}^{r_1}, \dots, l_{s_m}^{r_m} \quad (\text{th_resolution } i_1, \dots, i_n)$

where $l_{s_1}^{r_1}, \dots, l_{s_m}^{r_m}$ are disjuncts l_j^i (for $1 \leq i \leq n$, $1 \leq j \leq \max\{k^1, \dots, k^n\}$) from the clauses in premises $i_1, \dots, i_n$, which result from a chain of binary resolution steps on the clauses of these premises. It is possible that $m = 0$, i.e. that the result is the empty clause. When performing resolution steps, the rule implicitly merges repeated negations at the start of the formulas l_j^i . For example, the formulas $\neg\neg\neg P$ and $\neg\neg P$ can serve as pivots during resolution. The first formula is interpreted as $\neg P$ and the second as just P for the purpose of performing resolution steps.

This rule is only used when the resolution step is not emitted by the SAT solver. See the equivalent **resolution** rule for the rule emitted by the SAT solver.

Remark. The definition given here is very general. The motivation for this is to ensure the definition covers all possible resolution steps generated by the existing proof genera-

tion code in veriT. It will be restricted after a full review of the code. As a consequence of this checking this rule is theoretically NP-complete. In practice, however, the `th_resolution`-steps produced by veriT are simple. Experience with reconstructing the step in Isabelle/HOL shows that checking can be done by naive decision procedures. The vast majority of `th_resolution`-steps are binary resolution steps.

Rule 7: resolution

This rule is equivalent to the `th_resolution` rule, but it is emitted by the SAT solver instead of theory reasoners. The differentiation serves only informational purpose.

Rule 8: tautology

$i. \triangleright$ $l_1, \dots, l_k, \dots, l_m, \dots, l_n$ (...)
 $j. \triangleright$ $\top$ (tautology i)
 and l_k, l_m are such that

$$l_k = \underbrace{\neg \dots \neg}_o \varphi$$

$$l_m = \underbrace{\neg \dots \neg}_p \varphi$$

and one of o, p is odd and the other even. Either can be 0.

Rule 9: contraction

$i. \triangleright$ $l_1, \dots, l_n$ (...)
 $j. \triangleright$ $l_{k_1}, \dots, l_{k_m}$ (contraction i)
 where $m \leq n$ and $k_1 \dots k_m$ is a monotonic map to $1 \dots n$ such that $l_{k_1} \dots l_{k_m}$ are pairwise distinct and $\{l_1, \dots, l_n\} = \{l_{k_1}, \dots, l_{k_m}\}$. Hence, this rule removes duplicated literals.

Rule 10: subproof

The `subproof` rule completes a subproof and discharges local assumptions. Every subproof starts with some `assume` steps. The last step of the subproof is the conclusion.

$i_1.$	$\triangleright$	φ_1	assume
		$\vdots$	
$i_n.$	$\triangleright$	φ_n	assume
		$\vdots$	
$j.$	$\triangleright$	ψ	(...)
$k.$	$\triangleright$	$\neg\varphi_1, \dots, \neg\varphi_n, \psi$	subproof

Rule 11: la_generic

A step of the `la_generic` rule represents a tautological clause of linear disequalities. It can be checked by showing that the conjunction of the negated disequalities is unsatisfiable. After the application of some strengthening rules, the resulting conjunction is unsatisfiable, even if integer variables are assumed to be real variables.

A linear inequality is a term of the form:

$$\sum_{i=0}^n c_i \times t_i + d_1 \bowtie \sum_{i=n+1}^m c_i \times t_i + d_2$$

where $\bowtie \in \{\approx, <, >, \leq, \geq\}$, where $m \geq n$, c_i, d_1, d_2 are either integer or real constants, c_i and t_i have the same sort for all i ,

Let $l_1, \dots, l_n$ be linear inequalities and $a_1, \dots, a_n$ rational numbers, then a `la_generic` step has the form

$i. \triangleright \quad \varphi_1, \dots, \varphi_n \quad \text{la_generic } [a_1, \dots, a_n]$
 where φ_i is either $\neg l_i$ or l_i , but it never is an equality of the form $s_1 \approx s_2$, for terms s_1 and s_2 . However, the rules allow negated equalities: $\neg(s_1 \approx s_2)$.

The constants a_i are of sort `Real` and must be written following the production rules for the non-terminals `<rational>`, `<decimal>` or `<nonpositive_decimal>` (see Figure 2 in Section 2.1).

To check the unsatisfiability of the negation of $\varphi_1 \vee \dots \vee \varphi_o$ one performs the following steps for each literal φ_j and corresponding constant a_j :

1.
 - If $\varphi_j = s_1 > s_2$, then replace it by $s_2 \geq s_1$.
 - If $\varphi_j = s_1 \geq s_2$, then replace it by $s_2 > s_1$.
 - If $\varphi_j = s_1 < s_2$, then replace it by $s_1 \geq s_2$.
 - If $\varphi_j = s_1 \leq s_2$, then replace it by $s_1 > s_2$.

This negates the literal. Note that, in these cases, we always end up with inequalities expressed using only $>$ or $\geq$. This simplifies the description of the following steps.

2. If $\varphi_j = \neg(s_1 \bowtie s_2)$, then replace it by $s_1 \bowtie s_2$. If $\bowtie$ is $<$ or $\leq$, redefine the inequality $s_1 \bowtie s_2$ in terms of the relations $>$ or $\geq$, accordingly:
 - If $s_1 \bowtie s_2 = s_1 < s_2$, then replace it by $s_2 > s_1$.
 - If $s_1 \bowtie s_2 = s_1 \leq s_2$, then replace it by $s_2 \geq s_1$.
3. The resulting inequality has the form:

$$\sum_{i=0}^n c_i \times t_i + d_1 \bowtie \sum_{i=n+1}^m c_i \times t_i + d_2$$

Replace it by $s \bowtie d$, where $s = \sum_{i=0}^n c_i \times t_i - \sum_{i=n+1}^m c_i \times t_i$ and $d := d_2 - d_1$.

4. If all variables in s are integer sorted, replace $\bowtie$ and d by $\bowtie'$ and d' as defined in the following table:

$\bowtie$	If d is an integer	Otherwise
$>$	$\bowtie' = \geq, d' = d + 1$	$\bowtie' = \geq, d' = \lfloor d \rfloor + 1$
$\geq$	$\bowtie' = \geq, d' = d$	$\bowtie' = \geq, d' = \lfloor d \rfloor + 1$
$\approx$	$\bowtie' = \approx, d' = d$	$\bowtie' = \approx, d' = d$

If some variable in s is not integer sorted, consider $\bowtie' = \bowtie$ and $d' = d$ in the following steps.

5. Multiply both sides of the inequality by its corresponding coefficient a_j , if $\bowtie'$ is $\approx$, or by $|a_j|$, otherwise. We obtain:

$$\sum_{i=0}^m a_j \times c_i \times t_i \approx a_j \times d'$$

or

$$\sum_{i=0}^m |a_j| \times c_i \times t_i \bowtie' |a_j| \times d'$$

respectively.

Finally, the sum of the obtained n inequalities should result in a trivial contradiction. The sum results in an inequality of the form:

$$\sum_{k=1}^n \left(\sum_{i=0}^{m^k} a'_k * c_i^k * t_i^k \right) \bowtie_{sum} \sum_{k=1}^n d'^k$$

where $\sum_{i=0}^{m^k} a'_k * c_i^k * t_i^k$ is the left-hand side of the k -th inequality added (being a'_k , either, a_k or $|a_k|$, according to step 5), and d'^k is its right-hand side. Also, $\bowtie_{sum}$ is the “most strict” order among the relations that appear in the involved inequalities added, or just $\approx$. That is:

- $\bowtie_{sum}$ is some relation $\bowtie$, if all the relations involved are exactly $\bowtie$.
- $\bowtie_{sum}$ is $>$ if some of the relations involved are $>$.
- $\bowtie_{sum}$ is $\geq$ in any other case.

The coefficients $a_1, \dots, a_n$ provided to the rule must be such that the sum on the left-hand side is 0, and the right-hand side is > 0 (or ≥ 0 if $\bowtie_{sum}$ is $\geq$).

Example 11.1. A simple `la_generic` step in the logic LRA might look like this:

```
(step t10 (cl (not (> (f a) (f b))) (not (= (f a) (f b))))
:rule la_generic :args (1.0 -1.0))
```

To verify this we have to check the unsatisfiability of $(f a) > (f b) \wedge (f a) \approx (f b)$ (step 2). After step 3 we get $(f a) - (f b) > 0 \wedge (f a) - (f b) \approx 0$. Since we don't have an integer sort in this logic step 4 does not apply. Finally, after step 5 the conjunction is $(f a) - (f b) > 0 \wedge -(f a) + (f b) \approx 0$. This sums to $0 > 0$, which is a contradiction.

Example 11.2. The following `la_generic` step is from a QF_UFLIA problem:

```
(step t11 (cl (not (<= f3 0)) (<= (+ 1 (* 4 f3)) 1))
:rule la_generic :args (1.0 1/4))
```

After normalization we get $-f_3 \geq 0 \wedge 4 \times f_3 > 0$. This time step 4 applies and we can strengthen this to $-f_3 \geq 0 \wedge 4 \times f_3 \geq 1$ and after multiplication we get $-f_3 \geq 0 \wedge f_3 \geq \frac{1}{4}$. Which sums to the contradiction $0 \geq \frac{1}{4}$.

Rule 12: lia_generic

This rule is a placeholder rule for integer arithmetic solving. It takes the same form as `la_generic`, without the additional arguments.

$i. \triangleright \quad \varphi_1, \dots, \varphi_n \quad (\text{lia_generic})$
 with φ_i being linear inequalities. The disjunction $\varphi_1 \vee \dots \vee \varphi_n$ is a tautology in the theory of linear integer arithmetic.

Remark. Since this rule can introduce a disjunction of arbitrary linear integer inequalities without any additional hints, proof checking can be NP-hard. Hence, this rule should be avoided when possible.

Rule 13: la_disequality

$i. \triangleright \quad t_1 \approx t_2 \vee \neg(t_1 \leq t_2) \vee \neg(t_2 \leq t_1) \quad (\text{la_disequality})$

Rule 14: la_totality

$i. \triangleright \quad t_1 \leq t_2 \vee t_2 \leq t_1 \quad (\text{la_totality})$

Rule 15: la_tautology

This rule introduces a tautology of linear arithmetic. The conclusion is always a unit clause.

$i. \triangleright \quad \varphi \quad (\text{la_tautology})$

If the literal φ is either a linear inequality $s_1 \bowtie s_2$ or $\neg(s_1 \bowtie s_2)$, it is enough to perform steps 1 to 3 of the process for checking the `la_generic`. The result is trivially unsatisfiable. For example, if φ is $\neg(n \leq n - 1)$ we get $0 \leq -1$ after step 3.

If the literal is not a linear inequality, it is a binary clause $\varphi_1 \vee \varphi_2$. In this case the step is equivalent to

$i. \triangleright \quad \varphi_1, \varphi_2 \quad \text{la_tautology } [1.0, 1.0]$

Remark. The second form is intended to express bounds on linear combinations. In the case of `veriT`, the conclusion has one of the following forms, where d_1 and d_2 are decimals or integers, and s is a constant.

- $\neg(s \leq d_1) \vee s \leq d_2$ where $d_1 \leq d_2$
- $s \leq d_1 \vee \neg(s \leq d_2)$ where $d_1 = d_2$
- $\neg(s \geq d_1) \vee s \geq d_2$ where $d_1 \geq d_2$
- $s \geq d_1 \vee \neg(s \geq d_2)$ where $d_1 = d_2$
- $\neg(s \leq d_1) \vee \neg(s \geq d_2)$ where $d_1 < d_2$

Rule 16: la_mult_pos

Either of the form:

$$i. \triangleright \quad (t_1 > 0 \wedge t_2 \bowtie t_3) \rightarrow t_1 * t_2 \bowtie t_1 * t_3 \quad (\text{la_mult_pos})$$

with $\bowtie \in \{<, >, \leq, \geq, \approx\}$.

Or of the form:

$$i. \triangleright \quad (t_1 > 0 \wedge \neg(t_2 \approx t_3)) \rightarrow \neg(t_1 * t_2 \approx t_1 * t_3) \quad (\text{la_mult_pos})$$

Rule 17: la_mult_neg

Either of the form:

$$i. \triangleright \quad (t_1 < 0 \wedge t_2 \bowtie t_3) \rightarrow t_1 * t_2 \bowtie_{inv} t_1 * t_3 \quad (\text{la_mult_neg})$$

with $\bowtie \in \{<, >, \leq, \geq, \approx\}$ and $\bowtie_{inv}$ being defined according to the following table.

$\bowtie$	$\bowtie_{inv}$ is defined as:
$<$	$>$
$\leq$	$\geq$
$\approx$	$\approx$
$>$	$<$
$\geq$	$\leq$

Or of the form:

$$i. \triangleright \quad (t_1 < 0 \wedge \neg(t_2 \approx t_3)) \rightarrow \neg(t_1 * t_2 \approx t_1 * t_3) \quad (\text{la_mult_neg})$$

Rule 18: bind

The bind rule is used to rename bound variables.

$$\frac{j. \left[\begin{array}{c} \vdots \\ \Gamma, y_1, \dots, y_n, x_1 \mapsto y_1, \dots, x_n \mapsto y_n \triangleright \end{array} \right] \quad \varphi \approx \varphi' \quad (\dots)}{k. \Gamma \triangleright \quad Qx_1, \dots, x_n. \varphi \approx Qy_1, \dots, y_n. \varphi' \quad \text{bind}}$$

where $Q \in \{\forall, \exists\}$, and the variables $y_1, \dots, y_n$ are neither free in $Qx_1, \dots, x_n. \varphi$ nor occur in Γ .

Rule 19: sko_ex

The sko_ex rule skolemizes existential quantifiers.

$$\frac{j. \left[\begin{array}{c} \vdots \\ \Gamma, x_1 \mapsto \varepsilon_1, \dots, x_n \mapsto \varepsilon_n \triangleright \end{array} \right] \quad \varphi \approx \psi \quad (\dots)}{k. \Gamma \triangleright \quad \exists x_1, \dots, x_n. \varphi \approx \psi \quad \text{sko_ex}}$$

where ε_i stands for $\varepsilon x_i. (\exists x_{i+1}, \dots, x_n. \varphi)$.

Rule 20: sko_forall

The sko_forall rule skolemizes universal quantifiers.

$$\frac{j. \left[\begin{array}{c} \vdots \\ \Gamma, x_1 \mapsto (\varepsilon x_1. \neg \varphi), \dots, x_n \mapsto (\varepsilon x_n. \neg \varphi) \triangleright \end{array} \right] \quad \varphi \approx \psi \quad (\dots)}{k. \Gamma \triangleright \quad \forall x_1, \dots, x_n. \varphi \approx \psi \quad \text{sko_forall}}$$

Rule 21: forall_inst

$$i. \triangleright \quad \neg(\forall x_1, \dots, x_n. P) \vee P[x_1 \mapsto t_1, \dots, x_n \mapsto t_n] \quad \text{forall_inst } [t_1, \dots, t_n]$$

where x_i and t_i have the same sort, and $x_1, \dots, x_n$ are pairwise distinct.

Example 21.1. An application of the forall_inst rule.

```
(step t16 (cl (or (not (forall ((x S) (y T)) (P y x)))
                  (P b (f a)))
  :rule forall_inst :args ((f a) b))
```

Rule 22: refl

$$j. \Gamma \triangleright \quad t_1 \approx t_2 \quad \text{refl}$$

where, if $\sigma = \text{subst}(\Gamma)$, the terms $t_1\sigma$ and t_2 are syntactically equal up to renaming of bound variables and the orientation of equalities.

Remark. This is the only rule that requires the application of the context.

Rule 23: trans

$$\begin{array}{lll} i_1. \Gamma \triangleright & t_1 \approx t_2 & (\dots) \\ i_2. \Gamma \triangleright & t_2 \approx t_3 & (\dots) \\ & \vdots & \\ i_n. \Gamma \triangleright & t_n \approx t_{n+1} & (\dots) \\ j. \Gamma \triangleright & t_1 \approx t_{n+1} & (\text{trans } i_1, \dots, i_n) \end{array}$$

Rule 24: cong

$$\begin{array}{lll} i_1. \Gamma \triangleright & t_1 \approx u_1 & (\dots) \\ i_2. \Gamma \triangleright & t_2 \approx u_2 & (\dots) \\ & \vdots & \\ i_m. \Gamma \triangleright & t_m \approx u_m & (\dots) \\ j. \Gamma \triangleright & (f s_1 \dots s_n t_1 \dots t_m) \approx (f s_1 \dots s_n u_1 \dots u_m) & (\text{cong } i_1, \dots, i_m) \end{array}$$

where f is any $n + m$ -ary function symbol of appropriate sort.

Remark. Note that n can be 0. If $n > 0$ the omitted premises correspond to simple reflexivity steps.

Rule 25: eq_reflexive

$$i. \triangleright \quad t \approx t \quad \text{eq_reflexive}$$

Rule 26: eq_transitive

$$i. \triangleright \quad \neg(t_1 \approx t_2), \dots, \neg(t_{n-1} \approx t_n), t_1 \approx t_n \quad \text{eq_transitive}$$

Rule 27: eq_congruent

$$i. \triangleright \quad \neg(t_1 \approx u_1), \dots, \neg(t_n \approx u_n), (f t_1 \dots t_n) \approx (f u_1 \dots u_n) \quad \text{eq_congruent}$$

Rule 28: eq_congruent_pred

$$i. \triangleright \quad \neg(t_1 \approx u_1), \dots, \neg(t_n \approx u_n), (P t_1 \dots t_n) \approx (P u_1 \dots u_n) \quad \text{eq_congruent_pred}$$

where P is a function symbol with co-domain sort **Bool**.

Rule 29: qnt_cnf

$$i. \triangleright \quad \neg(\forall x_1, \dots, x_n. \varphi) \vee \forall x_{k_1}, \dots, x_{k_m}. \varphi' \quad \text{qnt_cnf}$$

This rule expresses clausification of a term under a universal quantifier. This is used by conflicting instantiation. φ' is one of the clause of the clause normal form of φ . The variables $x_{k_1}, \dots, x_{k_m}$ are a permutation of $x_1, \dots, x_n$ plus additional variables added by prenexing φ . Normalization is performed in two phases. First, the negative normal form is formed, then the result is prenexed. The result of the first step is $\Phi(\varphi, 1)$ where:

$$\begin{aligned}
\Phi(\neg\varphi, 1) &:= \Phi(\varphi, 0) \\
\Phi(\neg\varphi, 0) &:= \Phi(\varphi, 1) \\
\Phi(\varphi_1 \vee \dots \vee \varphi_n, 1) &:= \Phi(\varphi_1, 1) \vee \dots \vee \Phi(\varphi_n, 1) \\
\Phi(\varphi_1 \wedge \dots \wedge \varphi_n, 1) &:= \Phi(\varphi_1, 1) \wedge \dots \wedge \Phi(\varphi_n, 1) \\
\Phi(\varphi_1 \vee \dots \vee \varphi_n, 0) &:= \Phi(\varphi_1, 0) \wedge \dots \wedge \Phi(\varphi_n, 0) \\
\Phi(\varphi_1 \wedge \dots \wedge \varphi_n, 0) &:= \Phi(\varphi_1, 0) \vee \dots \vee \Phi(\varphi_n, 0) \\
\Phi(\varphi_1 \rightarrow \varphi_2, 1) &:= (\Phi(\varphi_1, 0) \vee \Phi(\varphi_2, 1)) \wedge (\Phi(\varphi_2, 0) \vee \Phi(\varphi_1, 1)) \\
\Phi(\varphi_1 \rightarrow \varphi_2, 0) &:= (\Phi(\varphi_1, 1) \wedge \Phi(\varphi_2, 0)) \vee (\Phi(\varphi_2, 1) \wedge \Phi(\varphi_1, 0)) \\
\Phi(\text{ite } \varphi_1 \varphi_2 \varphi_3, 1) &:= (\Phi(\varphi_1, 0) \vee \Phi(\varphi_2, 1)) \wedge (\Phi(\varphi_1, 1) \vee \Phi(\varphi_3, 1)) \\
\Phi(\text{ite } \varphi_1 \varphi_2 \varphi_3, 0) &:= (\Phi(\varphi_1, 1) \wedge \Phi(\varphi_2, 0)) \vee (\Phi(\varphi_1, 0) \wedge \Phi(\varphi_3, 0)) \\
\Phi(\forall x_1, \dots, x_n. \varphi, 1) &:= \forall x_1, \dots, x_n. \Phi(\varphi, 1) \\
\Phi(\exists x_1, \dots, x_n. \varphi, 1) &:= \exists x_1, \dots, x_n. \Phi(\varphi, 1) \\
\Phi(\forall x_1, \dots, x_n. \varphi, 0) &:= \exists x_1, \dots, x_n. \Phi(\varphi, 0) \\
\Phi(\exists x_1, \dots, x_n. \varphi, 0) &:= \forall x_1, \dots, x_n. \Phi(\varphi, 0) \\
\Phi(\varphi, 1) &:= \varphi \\
\Phi(\varphi, 0) &:= \neg\varphi
\end{aligned}$$

Remark. This is a placeholder rule that combines the many steps done during clausification.

Rule 30: and

$$\begin{aligned}
i. \triangleright & \quad \varphi_0 \wedge \dots \wedge \varphi_n & (\dots) \\
j. \triangleright & \quad \varphi_k & (\text{and } i) \ k \\
\text{and } 0 \leq k \leq n.
\end{aligned}$$

Rule 31: not_or

$$\begin{aligned}
i. \triangleright & \quad \neg(\varphi_0 \vee \dots \vee \varphi_n) & (\dots) \\
j. \triangleright & \quad \neg\varphi_k & (\text{not_or } i) \ k \\
\text{and } 0 \leq k \leq n.
\end{aligned}$$

Rule 32: or

$$\begin{aligned}
i. \triangleright & \quad \varphi_1 \vee \dots \vee \varphi_n & (\dots) \\
j. \triangleright & \quad \varphi_1, \dots, \varphi_n & (\text{or } i)
\end{aligned}$$

Remark. This rule deconstructs the `or` operator into a clause denoted by `cl`.

Example 32.1. An application of the `or` rule.

```
(step t15 (cl (or (= a b) (not (<= a b)) (not (<= b a))))
:rule la_disequality)
(step t16 (cl (= a b) (not (<= a b)) (not (<= b a)))
:rule or :premises (t15))
```

Rule 33: weakening

$i.$ $\triangleright \varphi_1, \dots, \varphi_n \quad (\dots)$
 $j.$ $\triangleright \varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m \quad (\text{weakening } i)$
 where $m \geq 1$.

Rule 34: reordering

$i.$ $\triangleright \varphi_1, \dots, \varphi_n \quad (\dots)$
 $j.$ $\triangleright \psi_1, \dots, \psi_n \quad (\text{reordering } i)$
 where the multisets $\{\varphi_1, \dots, \varphi_n\}$ and $\{\psi_1, \dots, \psi_n\}$ are the same. That is, the conclusion of the rule is a reordering of the literals in the premise.

Rule 35: shuffle

$i.$ $\Gamma \triangleright \circ t_1 \dots t_n \approx \circ u_1 \dots u_n \quad \text{shuffle}$

where $\circ$ is a commutative operator, and the multisets $\{t_1, \dots, t_n\}$ and $\{u_1, \dots, u_n\}$ are the same.

In the logics currently supported by Alethe $\circ \in \{+, *, \wedge, \vee\}$.

Remark. To permute the literals in a clause the `reordering` rule can be used.

Rule 36: not_and

$i.$ $\triangleright \neg(\varphi_1 \wedge \dots \wedge \varphi_n) \quad (\dots)$
 $j.$ $\triangleright \neg\varphi_1, \dots, \neg\varphi_n \quad (\text{not_and } i)$

Rule 37: xor1

$i.$ $\triangleright (\text{xor } \varphi_1 \varphi_2) \quad (\dots)$
 $j.$ $\triangleright \varphi_1, \varphi_2 \quad (\text{xor1 } i)$

Rule 38: xor2

$i.$ $\triangleright (\text{xor } \varphi_1 \varphi_2) \quad (\dots)$
 $j.$ $\triangleright \neg\varphi_1, \neg\varphi_2 \quad (\text{xor2 } i)$

Rule 39: not_xor1

$i.$ $\triangleright \neg(\text{xor } \varphi_1 \varphi_2) \quad (\dots)$
 $j.$ $\triangleright \varphi_1, \neg\varphi_2 \quad (\text{not_xor1 } i)$

Rule 40: not_xor2

$i.$ $\triangleright \neg(\text{xor } \varphi_1 \varphi_2) \quad (\dots)$
 $j.$ $\triangleright \neg\varphi_1, \varphi_2 \quad (\text{not_xor2 } i)$

Rule 41: implies

$i. \triangleright \varphi_1 \rightarrow \varphi_2 \quad (\dots)$
 $j. \triangleright \neg\varphi_1, \varphi_2 \quad (\text{implies } i)$

Rule 42: not_implies1

$i. \triangleright \neg(\varphi_1 \rightarrow \varphi_2) \quad (\dots)$
 $j. \triangleright \varphi_1 \quad (\text{not_implies1 } i)$

Rule 43: not_implies2

$i. \triangleright \neg(\varphi_1 \rightarrow \varphi_2) \quad (\dots)$
 $j. \triangleright \neg\varphi_2 \quad (\text{not_implies2 } i)$

Rule 44: equiv1

$i. \triangleright \varphi_1 \approx \varphi_2 \quad (\dots)$
 $j. \triangleright \neg\varphi_1, \varphi_2 \quad (\text{equiv1 } i)$

Rule 45: equiv2

$i. \triangleright \varphi_1 \approx \varphi_2 \quad (\dots)$
 $j. \triangleright \varphi_1, \neg\varphi_2 \quad (\text{equiv2 } i)$

Rule 46: not_equiv1

$i. \triangleright \neg(\varphi_1 \approx \varphi_2) \quad (\dots)$
 $j. \triangleright \varphi_1, \varphi_2 \quad (\text{not_equiv1 } i)$

Rule 47: not_equiv2

$i. \triangleright \neg(\varphi_1 \approx \varphi_2) \quad (\dots)$
 $j. \triangleright \neg\varphi_1, \neg\varphi_2 \quad (\text{not_equiv2 } i)$

Rule 48: and_pos

$i. \triangleright \neg(\varphi_0 \wedge \dots \wedge \varphi_n), \varphi_k \quad \text{and_pos } k$
 with $0 \leq k \leq n$.

Rule 49: and_neg

$i. \triangleright (\varphi_1 \wedge \dots \wedge \varphi_n), \neg\varphi_1, \dots, \neg\varphi_n \quad \text{and_neg}$

Rule 50: or_pos

$i. \triangleright \neg(\varphi_1 \vee \dots \vee \varphi_n), \varphi_1, \dots, \varphi_n \quad \text{or_pos}$

Rule 51: or_neg

$i. \triangleright (\varphi_0 \vee \dots \vee \varphi_n), \neg\varphi_k \quad \text{or_neg } k$
 with $0 \leq k \leq n$.

Rule 52: xor_pos1

$i. \triangleright \neg(\mathbf{xor} \varphi_1 \varphi_2), \varphi_1, \varphi_2 \quad \text{xor_pos1}$

Rule 53: xor_pos2

$i. \triangleright \neg(\mathbf{xor} \varphi_1 \varphi_2), \neg\varphi_1, \neg\varphi_2 \quad \text{xor_pos2}$

Rule 54: xor_neg1

$i. \triangleright (\mathbf{xor} \varphi_1 \varphi_2), \varphi_1, \neg\varphi_2 \quad \text{xor_neg1}$

Rule 55: xor_neg2

<i>i.</i> ▷	$(\text{xor } \varphi_1 \varphi_2), \neg\varphi_1, \varphi_2$	xor_neg2
Rule 56: implies_pos		
<i>i.</i> ▷	$\neg(\varphi_1 \rightarrow \varphi_2), \neg\varphi_1, \varphi_2$	implies_pos
Rule 57: implies_neg1		
<i>i.</i> ▷	$\varphi_1 \rightarrow \varphi_2, \varphi_1$	implies_neg1
Rule 58: implies_neg2		
<i>i.</i> ▷	$\varphi_1 \rightarrow \varphi_2, \neg\varphi_2$	implies_neg2
Rule 59: equiv_pos1		
<i>i.</i> ▷	$\neg(\varphi_1 \approx \varphi_2), \varphi_1, \neg\varphi_2$	equiv_pos1
Rule 60: equiv_pos2		
<i>i.</i> ▷	$\neg(\varphi_1 \approx \varphi_2), \neg\varphi_1, \varphi_2$	equiv_pos2
Rule 61: equiv_neg1		
<i>i.</i> ▷	$\varphi_1 \approx \varphi_2, \neg\varphi_1, \neg\varphi_2$	equiv_neg1
Rule 62: equiv_neg2		
<i>i.</i> ▷	$\varphi_1 \approx \varphi_2, \varphi_1, \varphi_2$	equiv_neg2
Rule 63: ite1		
<i>i.</i> ▷	$(\text{ite } \varphi_1 \varphi_2 \varphi_3)$	(...)
<i>j.</i> ▷	φ_1, φ_3	(ite1 <i>i</i>)
Rule 64: ite2		
<i>i.</i> ▷	$(\text{ite } \varphi_1 \varphi_2 \varphi_3)$	(...)
<i>j.</i> ▷	$\neg\varphi_1, \varphi_2$	(ite2 <i>i</i>)
Rule 65: ite_pos1		
<i>i.</i> ▷	$\neg(\text{ite } \varphi_1 \varphi_2 \varphi_3), \varphi_1, \varphi_3$	(ite_pos1)
Rule 66: ite_pos2		
<i>i.</i> ▷	$\neg(\text{ite } \varphi_1 \varphi_2 \varphi_3), \neg\varphi_1, \varphi_2$	(ite_pos2)
Rule 67: ite_neg1		
<i>i.</i> ▷	$(\text{ite } \varphi_1 \varphi_2 \varphi_3, \varphi_1, \neg\varphi_3)$	(ite_neg1)
Rule 68: ite_neg2		
<i>i.</i> ▷	$(\text{ite } \varphi_1 \varphi_2 \varphi_3, \neg\varphi_1, \neg\varphi_2)$	(ite_neg2)
Rule 69: not_ite1		
<i>i.</i> ▷	$\neg(\text{ite } \varphi_1 \varphi_2 \varphi_3)$	(...)
<i>j.</i> ▷	$\varphi_1, \neg\varphi_3$	(not_ite1 <i>i</i>)
Rule 70: not_ite2		
<i>i.</i> ▷	$\neg(\text{ite } \varphi_1 \varphi_2 \varphi_3)$	(...)
<i>j.</i> ▷	$\neg\varphi_1, \neg\varphi_2$	(not_ite2 <i>i</i>)

Rule 71: connective_def

This rule is used to replace connectives by their definition. It can be one of the following:

$i. \Gamma \triangleright$	$(\mathbf{xor} \varphi_1 \varphi_2) \approx ((\neg \varphi_1 \wedge \varphi_2) \vee (\varphi_1 \wedge \neg \varphi_2))$	connective_def
$i. \Gamma \triangleright$	$(\varphi_1 \approx \varphi_2) \approx ((\varphi_1 \rightarrow \varphi_2) \wedge (\varphi_2 \rightarrow \varphi_1))$	connective_def
$i. \Gamma \triangleright$	$(\mathbf{ite} \varphi_1 \varphi_2 \varphi_3) \approx ((\varphi_1 \rightarrow \varphi_2) \wedge (\neg \varphi_1 \rightarrow \varphi_3))$	connective_def
$i. \Gamma \triangleright$	$(\forall x_1, \dots, x_n. \varphi) \approx \neg(\exists x_1, \dots, x_n. \neg \varphi)$	connective_def
$i. \Gamma \triangleright$	$(\exists x_1, \dots, x_n. \varphi) \approx \neg(\forall x_1, \dots, x_n. \neg \varphi)$	connective_def

Rule 72: and_simplify

This rule simplifies an $\wedge$ term by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$$i. \Gamma \triangleright \quad \varphi_1 \wedge \dots \wedge \varphi_n \approx \psi \quad \text{and_simplify}$$

where ψ is the transformed term.

The possible transformations are:

- $\top \wedge \dots \wedge \top \Rightarrow \top$
- $\varphi_1 \wedge \dots \wedge \varphi_n \Rightarrow \varphi_1 \wedge \dots \wedge \varphi_{n'}$ where the right-hand side has all $\top$ literals removed.
- $\varphi_1 \wedge \dots \wedge \varphi_n \Rightarrow \varphi_1 \wedge \dots \wedge \varphi_{n'}$ where the right-hand side has all repeated literals removed.
- $\varphi_1 \wedge \dots \wedge \perp \wedge \dots \wedge \varphi_n \Rightarrow \perp$
- $\varphi_1 \wedge \dots \wedge \varphi_i \wedge \dots \wedge \varphi_j \wedge \dots \wedge \varphi_n \Rightarrow \perp$ and φ_i, φ_j are such that

$$\varphi_i = \underbrace{\neg \dots \neg}_n \psi$$

$$\varphi_j = \underbrace{\neg \dots \neg}_m \psi$$

and one of n, m is odd and the other even. Either can be 0.

Rule 73: or_simplify

This rule simplifies an $\vee$ term by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$$i. \Gamma \triangleright \quad (\varphi_1 \vee \dots \vee \varphi_n) \approx \psi \quad \text{or_simplify}$$

where ψ is the transformed term.

The possible transformations are:

- $\perp \vee \dots \vee \perp \Rightarrow \perp$
- $\varphi_1 \vee \dots \vee \varphi_n \Rightarrow \varphi_1 \vee \dots \vee \varphi_{n'}$ where the right-hand side has all $\perp$ literals removed.
- $\varphi_1 \vee \dots \vee \varphi_n \Rightarrow \varphi_1 \vee \dots \vee \varphi_{n'}$ where the right-hand side has all repeated literals removed.

- $\varphi_1 \vee \dots \vee \top \vee \dots \vee \varphi_n \Rightarrow \top$
- $\varphi_1 \vee \dots \vee \varphi_i \vee \dots \vee \varphi_j \vee \dots \vee \varphi_n \Rightarrow \top$ and φ_i, φ_j are such that

$$\varphi_i = \underbrace{\neg \dots \neg}_n \psi$$

$$\varphi_j = \underbrace{\neg \dots \neg}_m \psi$$

and one of n, m is odd and the other even. Either can be 0.

Rule 74: not_simplify

This rule simplifies an $\neg$ term by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$$i. \Gamma \triangleright \neg \varphi \approx \psi \quad \text{not_simplify}$$

where ψ is the transformed term.

The possible transformations are:

- $\neg(\neg \varphi) \Rightarrow \varphi$
- $\neg \perp \Rightarrow \top$
- $\neg \top \Rightarrow \perp$

Rule 75: implies_simplify

This rule simplifies an $\rightarrow$ term by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$$i. \Gamma \triangleright \varphi_1 \rightarrow \varphi_2 \approx \psi \quad \text{implies_simplify}$$

where ψ is the transformed term.

The possible transformations are:

- $\neg \varphi_1 \rightarrow \neg \varphi_2 \Rightarrow \varphi_2 \rightarrow \varphi_1$
- $\perp \rightarrow \varphi \Rightarrow \top$
- $\varphi \rightarrow \top \Rightarrow \top$
- $\top \rightarrow \varphi \Rightarrow \varphi$
- $\varphi \rightarrow \perp \Rightarrow \neg \varphi$
- $\varphi \rightarrow \varphi \Rightarrow \top$
- $\neg \varphi \rightarrow \varphi \Rightarrow \varphi$
- $\varphi \rightarrow \neg \varphi \Rightarrow \neg \varphi$

Rule 76: equiv_simplify

This rule simplifies a formula with the head symbol $\approx$: **Bool Bool Bool** by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$$i. \Gamma \triangleright \quad (\varphi_1 \approx \varphi_2) \approx \psi \quad \text{equiv_simplify}$$

where ψ is the transformed term.

The possible transformations are:

- $(\neg\varphi_1 \approx \neg\varphi_2) \Rightarrow (\varphi_1 \approx \varphi_2)$
- $(\varphi \approx \varphi) \Rightarrow \top$
- $(\varphi \approx \neg\varphi) \Rightarrow \perp$
- $(\neg\varphi \approx \varphi) \Rightarrow \perp$
- $(\top \approx \varphi) \Rightarrow \varphi$
- $(\varphi \approx \top) \Rightarrow \varphi$
- $(\perp \approx \varphi) \Rightarrow \neg\varphi$
- $(\varphi \approx \perp) \Rightarrow \neg\varphi$

Rule 77: bool_simplify

This rule simplifies a boolean term by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$$i. \Gamma \triangleright \quad \varphi \approx \psi \quad \text{bool_simplify}$$

where ψ is the transformed term.

The possible transformations are:

- $\neg(\varphi_1 \rightarrow \varphi_2) \Rightarrow (\varphi_1 \wedge \neg\varphi_2)$
- $\neg(\varphi_1 \vee \varphi_2) \Rightarrow (\neg\varphi_1 \wedge \neg\varphi_2)$
- $\neg(\varphi_1 \wedge \varphi_2) \Rightarrow (\neg\varphi_1 \vee \neg\varphi_2)$
- $(\varphi_1 \rightarrow (\varphi_2 \rightarrow \varphi_3)) \Rightarrow (\varphi_1 \wedge \varphi_2) \rightarrow \varphi_3$
- $((\varphi_1 \rightarrow \varphi_2) \rightarrow \varphi_2) \Rightarrow (\varphi_1 \vee \varphi_2)$
- $(\varphi_1 \wedge (\varphi_1 \rightarrow \varphi_2)) \Rightarrow (\varphi_1 \wedge \varphi_2)$
- $((\varphi_1 \rightarrow \varphi_2) \wedge \varphi_1) \Rightarrow (\varphi_1 \wedge \varphi_2)$

Rule 78: ac_simp

This rule simplifies nested occurrences of $\vee$ or $\wedge$:

$$i. \Gamma \triangleright \quad \psi \approx \varphi_1 \circ \dots \circ \varphi_n \quad \text{ac_simp}$$

where $\circ \in \{\vee, \wedge\}$ and ψ is a nested application of $\circ$. The literals φ_i are literals of the flattening of ψ with duplicates removed.

Rule 79: ite_simplify

This rule simplifies an if-then-else term by applying equivalence-preserving transformations until fixed point¹² It has the form

$$i. \Gamma \triangleright \quad (\text{ite } \varphi \, t_1 \, t_2) \approx u \quad \text{ite_simplify}$$

where u is the transformed term.

The possible transformations are:

- $(\text{ite } \top \, t_1 \, t_2) \Rightarrow t_1$
- $(\text{ite } \perp \, t_1 \, t_2) \Rightarrow t_2$
- $(\text{ite } \psi \, t \, t) \Rightarrow t$
- $(\text{ite } \neg\varphi \, t_1 \, t_2) \Rightarrow (\text{ite } \varphi \, t_2 \, t_1)$
- $(\text{ite } \psi \, (\text{ite } \psi \, t_1 \, t_2) \, t_3) \Rightarrow (\text{ite } \psi \, t_1 \, t_3)$
- $(\text{ite } \psi \, t_1 \, (\text{ite } \psi \, t_2 \, t_3)) \Rightarrow (\text{ite } \psi \, t_1 \, t_3)$
- $(\text{ite } \psi \, \top \, \perp) \Rightarrow \psi$
- $(\text{ite } \psi \, \perp \, \top) \Rightarrow \neg\psi$
- $(\text{ite } \psi \, \top \, \varphi) \Rightarrow \psi \vee \varphi$
- $(\text{ite } \psi \, \varphi \, \perp) \Rightarrow \psi \wedge \varphi$
- $(\text{ite } \psi \, \perp \, \varphi) \Rightarrow \neg\psi \wedge \varphi$
- $(\text{ite } \psi \, \varphi \, \top) \Rightarrow \neg\psi \vee \varphi$

Rule 80: qnt_simplify

This rule simplifies a $\forall$ -formula with a constant predicate.

$$i. \Gamma \triangleright \quad (\forall x_1, \dots, x_n. \varphi) \approx \varphi \quad \text{qnt_simplify}$$

where φ is either $\top$ or $\perp$.

Rule 81: onepoint

The onepoint rule is the “one-point-rule”. That is: it eliminates quantified variables that can only have one value.

$$\frac{j. \left[\Gamma, x_{k_1}, \dots, x_{k_m}, x_{j_1} \mapsto t_{j_1}, \dots, x_{j_o} \mapsto t_{j_o} \triangleright \quad \begin{array}{c} \vdots \\ \varphi \approx \varphi' \end{array} \quad (\dots) \right]}{k. \Gamma \triangleright Qx_1, \dots, x_n. \varphi \approx Qx_{k_1}, \dots, x_{k_m}. \varphi' \quad \text{onepoint}}$$

where $Q \in \{\forall, \exists\}$, $n = m + o$, $k_1, \dots, k_m$ and $j_1, \dots, j_o$ are monotone mappings to $1, \dots, n$, and no x_{k_i} appears in $x_{j_1}, \dots, x_{j_o}$.

The terms $t_{j_1}, \dots, t_{j_o}$ are the points of the variables $x_{j_1}, \dots, x_{j_o}$. Points are defined by equalities $x_i \approx t_i$ with positive polarity in the term φ .

¹²Note however that the order of the application is important, since the set of rules is not confluent. For example, the term $(\text{ite } \top \, t_1 \, t_2 \approx t_1)$ can be simplified into both p and $(\neg(\neg p))$ depending on the order of applications.

Remark. Since an eliminated variable x_i might appear free in a term t_j , it is necessary to replace x_i with t_i inside t_j . While this substitution is performed correctly, the proof for it is currently missing.

Example 81.1. An application of the `onepoint` rule on the term $(\forall x, y. x \approx y \rightarrow (f x) \wedge (f y))$ look like this:

```
(anchor :step t3 :args ((x S) (:= (y S) x)))
(step t3.t1 (cl (= x y)) :rule refl)
(step t3.t2 (cl (= (x y) (x x)))
  :rule cong :premises (t3.t1))
(step t3.t3 (cl (= x y)) :rule refl)
(step t3.t4 (cl (= (f y) (f x)))
  :rule cong :premises (t3.t3))
(step t3.t5 (cl (= (and (f x) (f y)) (and (f x) (f x))))
  :rule cong :premises (t3.t4))
(step t3.t6 (cl (= (=> (= x y) (and (f x) (f y)))
  (=> (= x x) (and (f x) (f x)))))
  :rule cong :premises (t3.t2 t3.t5))
(step t3 (cl (=
  (forall ((x S) (y S)) (=> (= x y) (and (f x) (f y))))
  (forall ((x S)) (=> (= x x) (and (f x) (f x))))))
  :rule onepoint)
```

Rule 82: `qnt_join`

$i. \Gamma \triangleright Qx_1, \dots, x_n. (Qx_{n+1}, \dots, x_m. \varphi) \approx Qx_{k_1}, \dots, x_{k_o}. \varphi$ qnt_join
 where $m > n$ and $Q \in \{\forall, \exists\}$. Furthermore, $k_1, \dots, k_o$ is a monotonic map to $1, \dots, m$ such that $x_{k_1}, \dots, x_{k_o}$ are pairwise distinct, and $\{x_1, \dots, x_m\} = \{x_{k_1}, \dots, x_{k_o}\}$.

Rule 83: `qnt_rm_unused`

$i. \Gamma \triangleright Qx_1, \dots, x_n. \varphi \approx Qx_{k_1}, \dots, x_{k_m}. \varphi$ qnt_rm_unused
 where $m \leq n$ and $Q \in \{\forall, \exists\}$. Furthermore, $k_1, \dots, k_m$ is a monotonic map to $1, \dots, n$ and if $x \in \{x_j \mid j \in \{1, \dots, n\} \wedge j \notin \{k_1, \dots, k_m\}\}$ then either x is not free in φ or $x \in \{x_j \mid j \in \{k_1, \dots, k_m\}\}$ (i.e., duplicates can be removed).

Rule 84: `eq_simplify`

This rule simplifies an $\approx$ term by applying equivalence-preserving transformations until a fixed point is reached. Hence, the general form is

$i. \Gamma \triangleright (t_1 \approx t_2) \approx \varphi$ eq_simplify

where φ is the transformed term.

The possible transformations are:

- $t \approx t \Rightarrow \top$
- $(t_1 \approx t_2) \Rightarrow \perp$ if t_1 and t_2 are different numeric constants.
- $\neg(t \approx t) \Rightarrow \perp$ if t is a numeric constant.

Rule 85: div_simplify

This rule simplifies a division by applying equivalence-preserving transformations. The general form is

$$i. \Gamma \triangleright \quad (t_1 / t_2) \approx t_3 \quad \text{div_simplify}$$

The possible transformations are:

- $t / t \Rightarrow 1$
- $t / 1 \Rightarrow t$
- $t_1 / t_2 \Rightarrow t_3$ if t_1 and t_2 are constants and t_3 is t_1 divided by t_2 according to the semantics of the current theory.

Rule 86: prod_simplify

This rule simplifies a product by applying equivalence-preserving transformations until a fixed point is reached. The general form is

$$i. \Gamma \triangleright \quad t_1 \times \cdots \times t_n \approx u \quad \text{prod_simplify}$$

where u is either a constant or a product.

The possible transformations are:

- $t_1 \times \cdots \times t_n \Rightarrow u$ where all t_i are constants and u is their product.
- $t_1 \times \cdots \times t_n \Rightarrow 0$ if any t_i is 0.
- $t_1 \times \cdots \times t_n \Rightarrow c \times t_{k_1} \times \cdots \times t_{k_n}$ where c is the product of the constants of $t_1, \dots, t_n$ and $t_{k_1}, \dots, t_{k_n}$ is $t_1, \dots, t_n$ with the constants removed.
- $t_1 \times \cdots \times t_n \Rightarrow t_{k_1} \times \cdots \times t_{k_n}$: same as above if c is 1.

Rule 87: unary_minus_simplify

This rule is either

$$i. \Gamma \triangleright \quad -(-t) \approx t \quad \text{unary_minus_simplify}$$

or

$$i. \Gamma \triangleright \quad -t \approx u \quad \text{unary_minus_simplify}$$

where u is the negated numerical constant t .

Rule 88: minus_simplify

This rule simplifies a subtraction by applying equivalence-preserving transformations. The general form is

$$i. \Gamma \triangleright \quad t_1 - t_2 \approx u \quad \text{minus_simplify}$$

The possible transformations are:

- $t - t \Rightarrow 0$
- $t_1 - t_2 \Rightarrow t_3$ where t_1 and t_2 are numerical constants and t_3 is t_2 subtracted from t_1 .
- $t - 0 \Rightarrow t$
- $0 - t \Rightarrow -t$

Rule 89: sum_simplify

This rule simplifies a sum by applying equivalence-preserving transformations until a fixed point is reached. The general form is

$$i. \Gamma \triangleright \quad t_1 + \dots + t_n \approx u \quad \text{sum_simplify}$$

where u is either a constant or a product.

The possible transformations are:

- $t_1 + \dots + t_n \Rightarrow c$ where all t_i are constants and c is their sum.
- $t_1 + \dots + t_n \Rightarrow c + t_{k_1} + \dots + t_{k_n}$ where c is the sum of the constants of $t_1, \dots, t_n$ and $t_{k_1}, \dots, t_{k_n}$ is $t_1, \dots, t_n$ with the constants removed.
- $t_1 + \dots + t_n \Rightarrow t_{k_1} + \dots + t_{k_n}$: same as above if c is 0.

Rule 90: comp_simplify

This rule simplifies a comparison by applying equivalence-preserving transformations until a fixed point is reached. The general form is

$$i. \Gamma \triangleright \quad t_1 \bowtie t_2 \approx \psi \quad \text{comp_simplify}$$

where $\bowtie \in \{<, >, \leq, \geq\}$.

The following transformation rules are used to simplify $t_1 \bowtie t_2$:

- $s_1 < s_2 \Rightarrow \top$ if s_1, s_2 are numerical constants and s_1 is strictly less than s_2
- $s_1 < s_2 \Rightarrow \perp$ if s_1, s_2 are numerical constants and s_1 is greater or equal than s_2
- $s_1 \leq s_2 \Rightarrow \top$ if s_1, s_2 are numerical constants and s_1 is less or equal than s_2
- $s_1 \leq s_2 \Rightarrow \perp$ if s_1, s_2 are numerical constants and s_1 is strictly greater than s_2
- $s < s \Rightarrow \perp$
- $s \leq s \Rightarrow \top$
- $s_1 \geq s_2 \Rightarrow s_2 \leq s_1$
- $s_1 < s_2 \Rightarrow \neg(s_2 \leq s_1)$
- $s_1 > s_2 \Rightarrow \neg(s_1 \leq s_2)$

Rule 91: let

This rule eliminates **let**. It has the form

$$\begin{array}{c} i_1. \Gamma \quad \triangleright \quad t_1 \approx s_1 \quad (\dots) \\ \vdots \\ i_n. \Gamma \quad \triangleright \quad t_n \approx s_n \quad (\dots) \\ \vdots \\ j. \frac{\Gamma, x_1 \mapsto s_1, \dots, x_n \mapsto s_n \triangleright \quad u \approx u' \quad (\dots)}{k. \Gamma \quad \triangleright \quad (\text{let } x_1 = t_1, \dots, x_n = t_n \text{ in } u) \approx u' \quad (\text{let } i_1, \dots, i_n)} \end{array}$$

The premise $i_1, \dots, i_n$ must be in the same subproof as the **let** step. If for $t_i \approx s_i$ the t_i and s_i are syntactically equal, the premise is omitted.

Rule 92: bind_let

This rule corresponds to the bind rule for **let**. It allows the renaming of the variables bound by the **let** step, the rewriting of the substituted terms, and the rewriting of the body of the **let**, resulting in a new **let** term. It has the form

$$\begin{array}{c}
 i_1. \Gamma \quad \triangleright \quad t_1 \approx s_1 \quad (\dots) \\
 \vdots \\
 i_n. \Gamma \quad \triangleright \quad t_n \approx s_n \quad (\dots) \\
 \vdots \\
 j. \left[\Gamma, y_1, \dots, y_n, x_1 \mapsto y_1, \dots, x_n \mapsto y_n \triangleright \quad u \approx u' \quad (\dots) \right] \\
 k. \frac{}{\Gamma \triangleright \text{let } x_1 = t_1, \dots, x_n = t_n \text{ in } u \approx \text{let } y_1 = s_1, \dots, y_n = s_n \text{ in } u' \text{ (bind_let } i_1, \dots, i_n)}
 \end{array}$$

The variables $y_1, \dots, y_n$ are neither free in $(\text{let } x_1 = t_1, \dots, x_n = t_n \text{ in } u)$ nor, for each y_i different from x_i , occur in Γ .

The premise $i_1, \dots, i_n$ must be in the same subproof as the **bind_let** step. If for $t_i \approx s_i$ the t_i and s_i are syntactically equal, the premise is omitted.

Example 92.1. The following example shows how this rule is used in a proof generated by Carcara's elaborator. It elaborates an implicit application of symmetry of equality.

```

(step t1 (cl (= (= 0 1) (= 1 0))) :rule eq_symmetric)
(anchor :step t2 :args ((p Bool) (:= (p Bool) p)))
(step t2.t1 (cl (= (= p false) (= false p))) :rule eq_symmetric)
(step t2 (cl (= (let ((p (= 0 1))) (= p false))
               (let ((p (= 1 0))) (= false p))))
  :rule bind_let :premises (t1))

```

Rule 93: distinct_elim

This rule eliminates the **distinct** predicate.

This rule behaves differently when the arguments are terms of type **Bool**. Then there are two cases:

$$\begin{array}{ll}
 i. \Gamma \triangleright & (\text{distinct } \varphi \psi) \approx \neg(\varphi \approx \psi) & \text{distinct_elim} \\
 \text{and} & \\
 i. \Gamma \triangleright & (\text{distinct } \varphi_1 \varphi_2 \varphi_3 \dots) \approx \perp & \text{distinct_elim}
 \end{array}$$

For all other sorts the rule is

$$i. \Gamma \triangleright \quad (\text{distinct } t_1 \dots t_n) \approx \bigwedge_{i=1}^n \bigwedge_{j=i+1}^n t_i \not\approx t_j \quad \text{distinct_elim}$$

Remark. SMT-LIB does not allow **distinct** to be used with only one argument. However, multiple solvers support this as a non-standard extension. While Alethe does not prescribe a behavior for this case, most solvers produce the following step:

$$i. \Gamma \triangleright \quad (\text{distinct } t) \approx \top \quad \text{distinct_elim}$$

Rule 94: la_rw_eq

$$i. \triangleright \quad (t \approx u) \approx (t \leq u \wedge u \leq t) \quad \text{la_rw_eq}$$

Rule 95: nary_elim

This rule replaces n -ary operators with their equivalent application of the binary operator. It is never applied to $\wedge$ or $\vee$.

Three cases are possible. If the operator $\circ$ is left associative, then the rule has the form

$$i. \Gamma \triangleright \quad \bigcirc_{i=1}^n t_i \approx (\dots ((t_1 \circ t_2) \circ t_3) \circ \dots t_n) \quad \text{nary_elim}$$

If the operator $\circ$ is right associative, then the rule has the form

$$i. \Gamma \triangleright \quad \bigcirc_{i=1}^n t_i \approx (t_1 \circ \dots \circ (t_{n-2} \circ (t_{n-1} \circ t_n)) \dots) \quad \text{nary_elim}$$

If the operator is *chainable*, then it has the form

$$i. \Gamma \triangleright \quad \bigcirc_{i=1}^n t_i \approx (t_1 \circ t_2) \oplus (t_2 \circ t_3) \oplus \dots \oplus (t_{n-1} \circ t_n) \quad \text{nary_elim}$$

where $\oplus$ is the binary operator indicated to perform the chaining.

Rule 96: bfun_elim

$$\begin{array}{ll} i. \triangleright & \psi \quad (\dots) \\ j. \triangleright & \varphi \quad (\text{bfun_elim } i) \end{array}$$

The formula φ is ψ after boolean functions have been simplified. This happens in a two step process. Both steps recursively iterate over ψ . The first step expands quantified variable of type **Bool**. Hence, $(\exists x. t)$ becomes $t[x \mapsto \perp] \vee t[x \mapsto \top]$ and $(\forall x. t)$ becomes $t[x \mapsto \perp] \wedge t[x \mapsto \top]$. If n variables of sort **Bool** appear in a quantifier, the disjunction (conjunction) has 2^n terms. Each term replaces the variables in t according to the bits of a number which is increased by one for each subsequent term starting from zero. The left-most variable corresponds to the least significant bit.

The second step expands function argument of boolean types by introducing appropriate if-then-else terms. For example, consider $(f x P y)$ where P is some formula. Then we replace this term by $(\text{ite } P (f x \top y) (f x \perp y))$. If the argument is already the constant $\top$ or $\perp$, it is ignored.

Rule 97: ite_intro

$$i. \triangleright \quad t \approx (t' \wedge u_1 \wedge \dots \wedge u_n) \quad (\text{ite_intro})$$

The term t (the formula φ) contains the **ite** operator. Let $s_1, \dots, s_n$ be the terms starting with **ite**, i.e. $s_i := \text{ite } \psi_i r_i r'_i$, then u_i has the form

$$\text{ite } \psi_i (s_i \approx r_i) (s_i \approx r'_i)$$

The term t' is equal to the term t up to the reordering of equalities where one argument is an **ite** term.

Remark. This rule stems from the introduction of fresh constants for if-then-else terms inside veriT. Internally s_i is a new constant symbol and the φ on the right side of the equality is φ with the if-then-else terms replaced by the constants. Those constants are unfolded during proof printing. Hence, the slightly strange form and the reordering of equalities.

Rule 98: bitblast_extract

$$i. \triangleright \quad (\text{extract } j \ i) \ x \approx \text{bbterm } \varphi_i \ \dots \ \varphi_j \quad \text{bitblast_extract}$$

where $\varphi_k = \text{bit_of}_k \ x$, for $i \leq k \leq j$.

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x is a **bbterm** application:

$$i. \triangleright \quad (\text{extract } j \ i) \ (\text{bbterm } x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n) \approx \text{bbterm } x_i \ \dots \ x_j \quad \text{bitblast_extract}$$

This alternative is based on the validity of the equality

$$\text{bit_of}_k \ (\text{bbterm } x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n) \approx x_k$$

for any bit-vector x of size $n + 1$, where $0 \leq k \leq n$.

Rule 99: bitblast_concat

$$i. \triangleright \quad \text{concat } x \ y \approx A_1 \quad \text{bitblast_concat}$$

The term “ A_1 ” is

$$\begin{aligned} & \text{bbterm } (\text{bit_of}_0 \ x) \ \dots \ (\text{bit_of}_{n-1} \ x) \\ & \quad (\text{bit_of}_0 \ y) \ \dots \ (\text{bit_of}_{m-1} \ y) \end{aligned}$$

Note that the vectors involved in the operation do not need to have the same width. Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that:

$$\begin{aligned} x &= \text{bbterm } x_0 \ \dots \ x_{n-1} \\ y &= \text{bbterm } y_0 \ \dots \ y_{m-1} \end{aligned}$$

then, the rule can be alternatively phrased as:

$$i. \triangleright \quad \text{concat } x \ y \approx A_2 \quad \text{bitblast_concat}$$

with $A_2 := \text{bbterm } x_0 \ \dots \ x_{n-1} \ y_0 \ \dots \ y_{m-1}$

Rule 100: bitblast_sext

$$i. \triangleright \quad \text{bvsextend } k \ x \approx A_1 \quad \text{bitblast_sext}$$

where x has the type **BitVec** n . The term “ A_1 ” is:

$$\begin{aligned} & \text{bbterm } (\text{bit_of}_0 \ x) \ \dots \ (\text{bit_of}_{n-1} \ x) \\ & \quad (\text{bit_of}_{n-1} \ x) \ \dots \ (\text{bit_of}_{n-1} \ x) \end{aligned}$$

That is, the original vector is extended just by repeating its most significant bit k times, such that the resulting vector has width $n + k$, and $\text{bit_of}_n \ (\text{bvsextend } k \ x) = \text{bit_of}_{n-1} \ x, \dots, \text{bit_of}_{n+k} \ (\text{bvsextend } k \ x) = \text{bit_of}_{n-1} \ x$. Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x is a “**bbterm**” application. So given that:

$$x = \text{bbterm } x_0 \ \dots \ x_{n-1}$$

then, the rule can be alternatively phrased as:

$i. \triangleright$ bitblast_sext
 $\text{bvsextend } k \ x \approx A_2$
 with $A_2 := \text{bbterm } x_0 \dots x_{n-1} x_{n-1} \dots x_{n-1}$, where, again, the most significant bit of x , that is, x_{n-1} , is repeated k times.

Rule 101: bitblast_eq

$i. \triangleright$ bitblast_eq
 $\text{bveq } x \ y \approx A_1$
 in which both x and y must have the same type **BitVec** n . The term “ A_1 ” is

$$\begin{aligned} & (\text{bit_of}_0 \ x \leftrightarrow \text{bit_of}_0 \ y) \wedge \\ & (\text{bit_of}_1 \ x \leftrightarrow \text{bit_of}_1 \ y) \wedge \\ & \dots \\ & (\text{bit_of}_{n-1} \ x \leftrightarrow \text{bit_of}_{n-1} \ y) \end{aligned}$$

Rule 102: bitblast_ult

$i. \triangleright$ bitblast_ult
 $\text{bvult } x \ y \approx \text{res}_{n-1}$
 in which both x and y must have the same type **BitVec** n and, for $i \geq 0$:

$$\begin{aligned} \text{res}_0 &= \neg(\text{bit_of}_0 \ x) \wedge \text{bit_of}_0 \ y \\ \text{res}_{i+1} &= ((\text{bit_of}_{i+1} \ x \approx \text{bit_of}_{i+1} \ y) \wedge \text{res}_i) \vee (\neg(\text{bit_of}_{i+1} \ x) \wedge \text{bit_of}_{i+1} \ y) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that

$$\begin{aligned} x &= \text{bbterm } x_0 \dots x_i \dots x_j \dots x_n \\ y &= \text{bbterm } y_0 \dots y_i \dots y_j \dots y_n \end{aligned}$$

then “ res_i ” can be defined, for $i \geq 0$, as

$$\begin{aligned} \text{res}_0 &= \neg x_0 \wedge y_0 \\ \text{res}_{i+1} &= ((x_{i+1} \approx y_{i+1}) \wedge \text{res}_i) \vee (\neg x_{i+1} \wedge y_{i+1}) \end{aligned}$$

Rule 103: bitblast_slt

$i. \triangleright$ bitblast_slt
 $\text{bvslt } x \ y \approx \text{res}$
 in which both x and y must have the same type **BitVec** n , where:

$$\text{res} = ((\text{bit_of}_{n-1} \ x \approx \text{bit_of}_{n-1} \ y) \wedge \text{res}_{n-2}) \vee (\text{bit_of}_{n-1} \ x \wedge \neg(\text{bit_of}_{n-1} \ y))$$

and, for $i \geq 0$:

$$\begin{aligned} \text{res}_0 &= \neg \text{bit_of}_0 \ x \wedge \text{bit_of}_0 \ y \\ \text{res}_{i+1} &= ((\text{bit_of}_{i+1} \ x \approx \text{bit_of}_{i+1} \ y) \wedge \text{res}_i) \vee (\neg(\text{bit_of}_{i+1} \ x) \wedge \text{bit_of}_{i+1} \ y) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that

$$\begin{aligned} x &= \mathbf{bbterm} \ x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n \\ y &= \mathbf{bbterm} \ y_0 \ \dots \ y_i \ \dots \ y_j \ \dots \ y_n \end{aligned}$$

then “res” can be defined as:

$$\text{res} = ((x_{n-1} \approx y_{n-1}) \wedge \text{res}_{n-2}) \vee (x_{n-1} \wedge \neg y_{n-1})$$

where, for $i \geq 0$:

$$\begin{aligned} \text{res}_0 &= x_0 \wedge \neg y_0 \\ \text{res}_{i+1} &= ((x_{i+1} \approx y_{i+1}) \wedge \text{res}_i) \vee (\neg x_{i+1} \wedge y_{i+1}) \end{aligned}$$

Rule 104: bitblast_add

$i. \triangleright$ bitblast_add
 $\mathbf{bvadd} \ x \ y \approx A_1$
in which both x and y must have the same type **BitVec** n . The term “ A_1 ” is

$$\begin{aligned} &\mathbf{bbterm} \ (((\mathbf{bit_of}_0 \ x) \mathbf{xor} (\mathbf{bit_of}_0 \ y)) \mathbf{xor} \text{carry}_0) \\ &\quad (((\mathbf{bit_of}_1 \ x) \mathbf{xor} (\mathbf{bit_of}_1 \ y)) \mathbf{xor} \text{carry}_1) \\ &\quad \dots \\ &\quad (((\mathbf{bit_of}_{n-1} \ x) \mathbf{xor} (\mathbf{bit_of}_{n-1} \ y)) \mathbf{xor} \text{carry}_{n-1}) \end{aligned}$$

and, for $i \geq 0$:

$$\begin{aligned} \text{carry}_0 &= \perp \\ \text{carry}_{i+1} &= ((\mathbf{bit_of}_i \ x) \wedge \mathbf{bit_of}_i \ y) \vee (((\mathbf{bit_of}_i \ x) \mathbf{xor} (\mathbf{bit_of}_i \ y)) \wedge \text{carry}_i) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that:

$$\begin{aligned} x &= \mathbf{bbterm} \ x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n \\ y &= \mathbf{bbterm} \ y_0 \ \dots \ y_i \ \dots \ y_j \ \dots \ y_n \end{aligned}$$

then, the rule can be alternatively phrased as:

$i. \triangleright$ bitblast_add
 $\mathbf{bvadd} \ x \ y \approx A_2$
with $A_2 := \mathbf{bbterm} \ ((x_0 \mathbf{xor} y_0) \mathbf{xor} \text{carry}_0) \ \dots \ ((x_{n-1} \mathbf{xor} y_{n-1}) \mathbf{xor} \text{carry}_{n-1})$, and
“ carry_k ” being defined, for $k \geq 0$, as:

$$\begin{aligned} \text{carry}_0 &= \perp \\ \text{carry}_{i+1} &= (x_i \wedge y_i) \vee ((x_i \mathbf{xor} y_i) \wedge \text{carry}_i) \end{aligned}$$

Rule 105: bitblast_neg

$i. \triangleright$ bitblast_neg
 $\text{bvneg } x \approx A_1$
 The term “ A_1 ” is

$$\begin{aligned} & \text{bbterm } (((\neg \text{bit_of}_0 \ x) \text{ xor } \perp) \text{ xor } \text{carry}_0) \\ & \quad (((\neg \text{bit_of}_1 \ x) \text{ xor } \perp) \text{ xor } \text{carry}_1) \\ & \quad \dots \\ & \quad (((\neg \text{bit_of}_{n-1} \ x) \text{ xor } \perp) \text{ xor } \text{carry}_{n-1}) \end{aligned}$$

and, for $i \geq 0$:

$$\begin{aligned} \text{carry}_0 &= \top \\ \text{carry}_{i+1} &= ((\neg \text{bit_of}_i \ x) \wedge \perp) \vee (((\neg \text{bit_of}_i \ x) \text{ xor } \perp) \wedge \text{carry}_i) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x is a “**bbT**” application. So given that:

$$x = \text{bbterm } x_0 \dots x_i \dots x_j \dots x_n$$

then, the rule can be alternatively phrased as:

$i. \triangleright$ bitblast_neg
 $\text{bvneg } xy \approx A_2$
 with $A_2 := \text{bbterm } ((\neg x_0 \text{ xor } \perp) \text{ xor } \text{carry}_0) \dots ((\neg x_{n-1} \text{ xor } \perp) \text{ xor } \text{carry}_{n-1})$, and carry_k being defined, for $k \geq 0$, as:

$$\begin{aligned} \text{carry}_0 &= \perp \\ \text{carry}_{i+1} &= (\neg x_i \wedge \perp) \vee ((\neg x_i \text{ xor } \perp) \wedge \text{carry}_i) \end{aligned}$$

Rule 106: bitblast_mult

$i. \triangleright$ bitblast_mult
 $\text{bvmult } x \ y \approx A_1$
 in which both x and y must have the same type **BitVec** n . We perform bit-blasting following the *shift add multiplier* way of computing. That is, the term “ A_1 ” is:

$$\text{bbterm } \text{res}_0^{n-1} \dots \text{res}_{n-1}^{n-1}$$

in which, for $i, j \geq 0$,

$$\begin{aligned} \text{res}_i^0 &= \text{sh}_i^0 \\ \text{res}_0^j &= \text{sh}_0^j \\ \text{res}_i^{j+1} &= \text{res}_i^j \text{ xor } \text{sh}_i^{j+1} \text{ xor } \text{carry}_i^{j+1} \\ \text{carry}_i^0 &= \perp \\ \text{carry}_{i+1}^{j+1} &= \begin{cases} \text{res}_i^j \wedge \text{sh}_i^{j+1} \vee ((\text{res}_i^j \text{ xor } \text{sh}_i^{j+1}) \wedge \text{carry}_i^{j+1}) & \text{if } j < i \\ \perp & \text{otherwise} \end{cases} \\ \text{sh}_i^j &= \begin{cases} x_{i-j} \wedge y_j & \text{if } j \leq i \\ \perp & \text{otherwise} \end{cases} \end{aligned}$$

Example mult for $n = 4$:

$$\left[\begin{array}{lcl}
 \text{res}_0^3 & = & (x_0 \wedge y_0); \\
 \text{res}_1^3 & = & \underbrace{(x_1 \wedge y_0) \text{ xor } (x_0 \wedge y_1)}_{\text{res}_1^0 \text{ xor sh}_1^1} \text{ xor } \underbrace{\text{carry}_1^3}_{\perp}; \\
 \text{res}_2^3 & = & \underbrace{(x_2 \wedge y_0) \text{ xor } (x_1 \wedge y_1)}_{\text{res}_2^0 \text{ xor sh}_2^1} \text{ xor } \underbrace{\text{carry}_2^1}_{\text{res}_1^0 \wedge \text{sh}_1^1 \dots \text{sh}_2^2} \text{ xor } \underbrace{(x_0 \wedge y_2) \text{ xor } \text{carry}_2^3}_{\perp}; \\
 \text{res}_3^3 & = & \underbrace{(x_3 \wedge y_0) \text{ xor } (x_2 \wedge y_1)}_{\text{res}_3^0 \text{ xor sh}_3^1} \text{ xor } \underbrace{\text{carry}_3^1}_{\dots} \text{ xor } \underbrace{(x_1 \wedge y_2) \text{ xor } \text{carry}_3^2}_{\text{res}_2^1 \wedge \text{sh}_2^2 \dots \text{sh}_3^3} \text{ xor } \underbrace{(x_0 \wedge y_3) \text{ xor } \text{carry}_3^3}_{\perp};
 \end{array} \right]$$

with

$$\begin{aligned}
 \text{carry}_3^1 &= \text{res}_2^0 \wedge \text{sh}_2^1 \vee ((\text{res}_2^0 \text{ xor sh}_2^1) \wedge \text{carry}_2^1) \\
 \text{carry}_2^1 &= \text{res}_1^0 \wedge \text{sh}_1^1 \vee ((\text{res}_1^0 \text{ xor sh}_1^1) \wedge \underbrace{\text{carry}_1^1}_{\perp}) \\
 \text{carry}_3^2 &= \text{res}_2^1 \wedge \text{sh}_2^2 \vee ((\text{res}_2^1 \text{ xor sh}_2^2) \wedge \underbrace{\text{carry}_2^2}_{\perp})
 \end{aligned}$$

Rule 107: bitblast_and

$i. \triangleright$ $\text{bvand } x \ y \approx A_1$ bitblast_and
in which both x and y must have the same type **BitVec** n . The term “ A_1 ” is

$$\begin{aligned}
 &\text{bbterm } (\text{bit_of}_0 \ x \wedge \text{bit_of}_0 \ y) \\
 &\quad (\text{bit_of}_1 \ x \wedge \text{bit_of}_1 \ y) \\
 &\quad \dots \\
 &\quad (\text{bit_of}_{n-1} \ x \wedge \text{bit_of}_{n-1} \ y)
 \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “bbterm” applications. So given that:

$$\begin{aligned}
 x &= \text{bbterm } x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n \\
 y &= \text{bbterm } y_0 \ \dots \ y_i \ \dots \ y_j \ \dots \ y_n
 \end{aligned}$$

then, the rule can be alternatively phrased as:

$i. \triangleright$ $\text{bvand } x \ y \approx A_2$ bitblast_and
with $A_2 := \text{bbterm } (x_0 \wedge y_0) \ \dots \ (x_{n-1} \wedge y_{n-1})$.

Rule 108: bitblast_or

i. $\triangleright$ bitblast_or
 $\mathbf{bvor} \ x \ y \approx A_1$
 in which both x and y must have the same type $\mathbf{BitVec} \ n$. The term “ A_1 ” is

$$\begin{aligned} & \mathbf{bbterm} \ (\mathbf{bit_of}_0 \ x \vee \mathbf{bit_of}_0 \ y) \\ & \quad (\mathbf{bit_of}_1 \ x \vee \mathbf{bit_of}_1 \ y) \\ & \quad \dots \\ & \quad (\mathbf{bit_of}_{n-1} \ x \vee \mathbf{bit_of}_{n-1} \ y) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that:

$$\begin{aligned} x &= \mathbf{bbterm} \ x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n \\ y &= \mathbf{bbterm} \ y_0 \ \dots \ y_i \ \dots \ y_j \ \dots \ y_n \end{aligned}$$

then, the rule can be alternatively phrased as:

i. $\triangleright$ bitblast_or
 $\mathbf{bvor} \ x \ y \approx A_2$
 with $A_2 := \mathbf{bbterm} \ (x_0 \vee y_0) \ \dots \ (x_{n-1} \vee y_{n-1})$.

Rule 109: bitblast_xor

i. $\triangleright$ bitblast_xor
 $\mathbf{bvxor} \ x \ y \approx A_1$
 in which both x and y must have the same type $\mathbf{BitVec} \ n$. The term “ A_1 ” is

$$\begin{aligned} & \mathbf{bbterm} \ (\mathbf{bit_of}_0 \ x \mathbf{xor} \mathbf{bit_of}_0 \ y) \\ & \quad (\mathbf{bit_of}_1 \ x \mathbf{xor} \mathbf{bit_of}_1 \ y) \\ & \quad \dots \\ & \quad (\mathbf{bit_of}_{n-1} \ x \mathbf{xor} \mathbf{bit_of}_{n-1} \ y) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that:

$$\begin{aligned} x &= \mathbf{bbterm} \ x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n \\ y &= \mathbf{bbterm} \ y_0 \ \dots \ y_i \ \dots \ y_j \ \dots \ y_n \end{aligned}$$

then, the rule can be alternatively phrased as:

i. $\triangleright$ bitblast_xor
 $\mathbf{bvxor} \ x \ y \approx A_2$
 with $A_2 := \mathbf{bbterm} \ (x_0 \mathbf{xor} y_0) \ \dots \ (x_{n-1} \mathbf{xor} y_{n-1})$.

Rule 110: bitblast_xnor

i. $\triangleright$ bitblast_xnor
 $\mathbf{bvxnor} \ x \ y \approx A_1$

in which both x and y must have the same type **BitVec** n . The term “ A_1 ” is

$$\begin{aligned} & \mathbf{bbterm} \ (\mathbf{bit_of}_0 \ x \approx \mathbf{bit_of}_0 \ y) \\ & \quad (\mathbf{bit_of}_1 \ x \approx \mathbf{bit_of}_1 \ y) \\ & \quad \dots \\ & \quad (\mathbf{bit_of}_{n-1} \ x \approx \mathbf{bit_of}_{n-1} \ y) \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x and y are “**bbterm**” applications. So given that:

$$\begin{aligned} x &= \mathbf{bbterm} \ x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n \\ y &= \mathbf{bbterm} \ y_0 \ \dots \ y_i \ \dots \ y_j \ \dots \ y_n \end{aligned}$$

then, the rule can be alternatively phrased as:

$$\begin{aligned} i. \triangleright & \quad \mathbf{bvxnor} \ x \ y \approx A_2 & \text{bitblast_xnor} \\ \text{with } A_2 &:= \mathbf{bbterm} \ (x_0 \approx y_0) \ \dots \ (x_{n-1} \approx y_{n-1}). \end{aligned}$$

Rule 111: bitblast_not

$$i. \triangleright \quad \mathbf{bvnot} \ x \approx A_1 \quad \text{bitblast_not}$$

in which both x and y must have the same type **BitVec** n . The term “ A_1 ” is

$$\begin{aligned} & \mathbf{bbterm} \ \neg \mathbf{bit_of}_0 \ x \\ & \quad \neg \mathbf{bit_of}_1 \ x \\ & \quad \dots \\ & \quad \neg \mathbf{bit_of}_{n-1} \ x \end{aligned}$$

Alternatively, the rule may also be phrased as a “short-circuiting” of the above when x is a “**bbterm**” application. So given that:

$$x = \mathbf{bbterm} \ x_0 \ \dots \ x_i \ \dots \ x_j \ \dots \ x_n$$

then, the rule can be alternatively phrased as:

$$\begin{aligned} i. \triangleright & \quad \mathbf{bvnot} \ x \approx A_2 & \text{bitblast_not} \\ \text{with } A_2 &:= \mathbf{bbterm} \ \neg x_0 \ \dots \ \neg x_{n-1}. \end{aligned}$$

Rule 112: symm

$$\begin{aligned} i. \triangleright & \quad \varphi \approx \psi & (\dots) \\ j. \triangleright & \quad \psi \approx \varphi & (\text{symm } i) \end{aligned}$$

If $\varphi \neq \psi$ then the conclusion *must not* be $\varphi \approx \psi$.

Note that since Alethe allows the implicit reordering of equalities, this rule is technically superfluous. However, the rule is useful to indicate an explicit usage of the symmetry of equality to aid proof reconstruction.

Example 112.1: The side condition ensures that the following example is not a valid application of the rule. Without this condition, this derivation could be obtained by applying symmetry of equality implicitly to the conclusion.

$$\begin{array}{ll} 10. \triangleright & P(a) \approx Q(b) \quad (\dots) \\ 11. \triangleright & P(a) \approx Q(b) \quad (\text{symm } 10) \end{array}$$

Rule 113: not_symm

$$\begin{array}{ll} i. \triangleright & \neg(\varphi \approx \psi) \quad (\dots) \\ j. \triangleright & \neg(\psi \approx \varphi) \quad (\text{not_symm } i) \end{array}$$

If $\varphi \neq \psi$ then the conclusion *must not* be $\neg(\varphi \approx \psi)$.

See `symm` for an explanation of this rule.

Rule 114: eq_symmetric

$$i. \triangleright \quad (t_1 \approx t_2) \approx (t_2 \approx t_1) \quad (\text{eq_symmetric})$$

Note that since Alethe allows the implicit reordering of equalities, this rule is technically superfluous. However, the rule is useful to state an explicit usage of symmetry of equality reasoning to aid proof reconstruction.

Rule 115: miniscope_distribute

$$\begin{array}{ll} i. \triangleright & \forall x_1, \dots, x_n. (\varphi_1 \wedge \dots \wedge \varphi_m) \approx \quad (\text{miniscope_distribute}) \\ & (\forall x_1, \dots, x_n. \varphi_1) \wedge \dots \wedge (\forall x_1, \dots, x_n. \varphi_m) \\ & \text{or} \\ i. \triangleright & \exists x_1, \dots, x_n. (\varphi_1 \vee \dots \vee \varphi_m) \approx \quad (\text{miniscope_distribute}) \\ & (\exists x_1, \dots, x_n. \varphi_1) \vee \dots \vee (\exists x_1, \dots, x_n. \varphi_m) \end{array}$$

Rule 116: miniscope_split

$$\begin{array}{ll} i. \triangleright & \forall x_1, \dots, x_n. (\varphi_1 \vee \dots \vee \varphi_m) \approx \quad (\text{miniscope_split}) \\ & (\forall x_{1,1}, \dots, x_{1,n_1}. \varphi_1) \vee \dots \vee (\forall x_{m,1}, \dots, x_{m,n_m}. \varphi_m) \\ & \text{or} \\ i. \triangleright & \exists x_1, \dots, x_n. (\varphi_1 \wedge \dots \wedge \varphi_m) \approx \quad (\text{miniscope_split}) \\ & (\exists x_{1,1}, \dots, x_{1,n_1}. \varphi_1) \wedge \dots \wedge (\exists x_{m,1}, \dots, x_{m,n_m}. \varphi_m) \end{array}$$

where $\{x_{1,1}, \dots, x_{m,n_m}\} \subseteq \{x_1, \dots, x_n\}$ and the right side of the equality has no free variables in $\{x_1, \dots, x_n\}$.

Rule 117: miniscope_ite

$$i. \triangleright \quad \forall x_1, \dots, x_n. (\text{ite } \varphi_1 \varphi_2 \varphi_3) \approx \quad (\text{miniscope_ite})$$

$$\text{ite } \varphi_1 (\forall x_1, \dots, x_n. \varphi_2) (\forall x_1, \dots, x_n. \varphi_3)$$

where φ_1 has no free variable from $\{x_1, \dots, x_n\}$.

Rule 118: aci_simp

$$i. \triangleright \quad t_1 \approx t_2 \quad (\text{aci_simp})$$

where t_1 and t_2 are equivalent modulo associativity, commutativity, identity elements and idempotency. Checking this rule involves a normalization step of t_1 and t_2 , that exploits

the previous properties of the operations involved in both terms. The rule applies to SMT-LIB ACI operators (that is, those that satisfy associativity, commutativity, idempotency and that have an identity element), AC operators (associativity, commutativity and identity elements), as well to operators that satisfy only associativity and that have an identity element. The actual way in which the previous properties are exploited might be solver-dependent, which means that there is no such thing as an “ACI normal form”, standard among any solver producing Alethe certificates.

The operators that can trigger the previous process include:

- Logical operations: `or`, `and`.
- Bit vectors: `bvor`, `bvand`, `bvadd`, `bvmul`, `concat`.
- Arithmetic: `-` (binary subtraction), `+`, `*`.

The process itself can be described with a set of rewriting rules, which capture a relation $\Rightarrow$ between terms representing the application of operators and their simplified form. In what follows, op will represent an operator with the appropriate property exploited in the rewriting rule being applied; id will represent the op ’s identity element (if appropriate); we will indicate with $x \nRightarrow$ that term x cannot be further reduced through $\Rightarrow$ (which, in this case, reduces to check that x is not an application of some ACI operator).

- $op\ x\ id \Rightarrow x$, for $x \nRightarrow$.
- $op\ id\ x \Rightarrow x$, for $x \nRightarrow$.
- $op\ x\ x \Rightarrow x$, for $x \nRightarrow$.
- $op\ x_1\ x_2 \Rightarrow op\ x_{1'}\ x_2$, where $x_1 \Rightarrow x_{1'}$.
- $op\ x\ x_2 \Rightarrow op\ x\ x_{2'}$, where $x_2 \Rightarrow x_{2'}$ and $x \nRightarrow$.
- $op\ x_1\ x_2 \Rightarrow op\ x_{1'}\ x_{2'}$, where
 - $x_1 \neq x_2$
 - $x_{1'}, x_{2'} \in \{x_1, x_2\}$
 - Multi-sets $\{x_{1'}, x_{2'}\}$ and $\{x_1, x_2\}$ are equal.
 - $x_1 \nRightarrow$ and $x_2 \nRightarrow$.

In the last rule, note that we do not impose some particular order among sub-terms. The intention of the rule is to allow for a particular implementation of a checker for the rule, to take advantage of the commutativity of op , if desired.

Also, note that no evaluation of the application of operators (*constant folding*) is performed by the rule.

Example 118.1. The following simple `aci_simp` step, taken from a `QF_UF` problem, serves to illustrate the application of a single reduction rule from $\Rightarrow$:

```
(step t1 (cl (= (and d d) d)) :rule aci_simp)
```

The application of the rule exploits the idempotency property of the `and` operator (third reduction rule from $\Rightarrow$), to conclude that $(= (\text{and } d \ d) \ d)$ holds.

Example 118.2. The following example application of the `aci_simp` rule illustrates a situation where the reduction of expressions, using $\Rightarrow$, also involves *aci-reduction* of sub-expressions:

```
(step t1 (cl (= (or false (or true true)) (or true false))) :rule aci_simp)
```

As in the previous example, the application of the rule exploits an idempotency property, in this case of the `or` operator, but over the sub-expression $(\text{or } \text{true } \text{true})$. That is, the reduction also employs the fifth reduction rule defining $\Rightarrow$, that allows for the reduction of sub-expressions. Finally, in order to be able to conclude that both terms are *aci-compatible*, the reasoning employed exploits the last reduction rule of $\Rightarrow$, that allows for an arbitrary order of operands of applications of commutative operators. This allows us to conclude that $(\text{or } \text{false } \text{true})$ (the result of the previous reduction step) is *aci-compatible* with $(\text{or } \text{true } \text{false})$. Note that, as explained in the specification of $\Rightarrow$, there is no concrete order of operands specified, leaving some degree of freedom for the implementation of checkers of the rule. This, as long as the resulting $\Rightarrow$ relation ends up relating expressions that are equivalent modulo commutativity.

Rule 119: rare_rewrite

$i_1.$	$\triangleright$	$t_1 \approx t'_1$	$(\dots)$
$i_2.$	$\triangleright$	$t_2 \approx t'_2$	$(\dots)$
		$\vdots$	
$i_n.$	$\triangleright$	$t_n \approx t'_n$	$(\dots)$
$j.$	$\triangleright$	$t \approx s$	$(\text{rare_rewrite } i_1, \dots, i_n) [\text{rule_id}, a_1, \dots, a_m]$

This rule refers to the application of the RARE 2.0 rewrite identified by the string literal `rule_id` and which justifies concluding $t \approx s$. As described in Section 4.2, a RARE rewrite rule is defined using a domain-specific language, called RARE. The `rare_rewrite` step is correct when its premises $i_1, \dots, i_n$ match the conditions established in the specification of the RARE rewrite, if any, and its conclusion matches the conclusion established in the specification instantiated with the given arguments $a_1 \dots, a_m$, if any.

An important detail is that RARE admits a list semantics for associative operators, so some a_k may correspond to a list of terms, which is represented as $(@list \ a_{k,1}, \dots, a_{k,l})$.

Rule 120: multi_rare_rewrite

$i.$	$\triangleright$	$t \approx s$	$(\text{multi_rare_rewrite } rule_1 \dots rule_n)$
------	------------------	---------------	--

This rule refers to the repeated application of the RARE rewrite rules $rule_1 \dots rule_n$ over the left-hand side of the conclusion, term t , until it can be successfully rewritten as the right-hand side term s . That is, among rules $rule_1 \dots rule_n$ there should be one rule whose conclusion's left-hand side matches against term t . By applying it, we should be able to rewrite term t as some other term t' , such that it matches against the conclusion's

left-hand side of some rule in $rule_1 \dots rule_n$. This process should be repeated until some intermediate term can be rewritten as term s (that is, by applying a rewrite rule whose conclusion's right-hand side matches s).

The instantiation of some of the referred rewrite rules could require arbitrary arguments, with rule-dependent semantics. These arguments can be passed after the name of the each invoked rewrite rule, using the list constructor **@list**, as in the case of rule **rare_rewrite** (rule 119). In this case, the rule invocation would be as follows:

i. $\triangleright \quad t \approx s \quad (\text{multi_rare_rewrite } rule_1, (@\text{list } x_1, \dots, x_n), \dots, rule_n)$

Where, after each rewrite rule that requires arguments, a list **@list** with the corresponding arguments is indicated. In the previous invocation of the **multi_rare_rewrite** rule, we are assuming that $rule_1$ requires arguments, which are referred through the list $(@\text{list } x_1, \dots, x_n)$. It is possible to interleave with invocations of rules that do not require arguments. In that case, only the rule name is mentioned. That would be the case of rule $rule_n$ in our previous example.

5.3 Index of Rules

ac_simp, 44
aci_simp, 58
and, 38
and_neg, 40
and_pos, 40
and_simplify, 42
assume, 30

bfun_elim, 50
bind, 36
bind_let, 49
bitblast_add, 53
bitblast_and, 55
bitblast_concat, 51
bitblast_eq, 52
bitblast_extract, 50
bitblast_mult, 54
bitblast_neg, 53
bitblast_not, 57
bitblast_or, 56
bitblast_sext, 51
bitblast_slt, 52
bitblast_ult, 52
bitblast_xnor, 56
bitblast_xor, 56
bool_simplify, 44

comp_simplify, 48
cong, 37
connective_def, 41
contraction, 32

distinct_elim, 49
div_simplify, 47

eq_congruent, 37
eq_congruent_pred, 37
eq_reflexive, 37
eq_simplify, 46
eq_symmetric, 58
eq_transitive, 37
equiv1, 40

equiv2, 40
equiv_neg1, 41
equiv_neg2, 41
equiv_pos1, 41
equiv_pos2, 41
equiv_simplify, 43

false, 31
forall_inst, 36

hole, 31

implies, 39
implies_neg1, 41
implies_neg2, 41
implies_pos, 41
implies_simplify, 43
ite1, 41
ite2, 41
ite_intro, 50
ite_neg1, 41
ite_neg2, 41
ite_pos1, 41
ite_pos2, 41
ite_simplify, 45

la_disequality, 35
la_generic, 32
la_mult_neg, 36
la_mult_pos, 35
la_rw_eq, 49
la_tautology, 35
la_totality, 35
let, 48
lia_generic, 35

miniscope_distribute, 58
miniscope_ite, 58
miniscope_split, 58
minus_simplify, 47
multi_rare_rewrite, 60

nary_elim, 50

- not_and, 39
- not_equiv1, 40
- not_equiv2, 40
- not_implies1, 40
- not_implies2, 40
- not_ite1, 41
- not_ite2, 41
- not_not, 31
- not_or, 38
- not_simplify, 43
- not_symm, 58
- not_xor1, 39
- not_xor2, 39

- onepoint, 45
- or, 38
- or_neg, 40
- or_pos, 40
- or_simplify, 42

- prod_simplify, 47

- qnt_cnf, 38
- qnt_join, 46
- qnt_rm_unused, 46
- qnt_simplify, 45

- rare_rewrite, 60
- refl, 37
- reordering, 39
- resolution, 32

- shuffle, 39
- sko_ex, 36
- sko_forall, 36
- subproof, 32
- sum_simplify, 48
- symm, 57

- tautology, 32
- th_resolution, 31
- trans, 37
- true, 31

- unary_minus_simplify, 47

- weakening, 39

- xor1, 39
- xor2, 39
- xor_neg1, 40
- xor_neg2, 40
- xor_pos1, 40
- xor_pos2, 40

Changelog

Unreleased

Proof rules:

- General rewriting of rule `la_generic`, mainly focused on presentation. The algorithm stays the same.
- Added rules `rare_rewrite` and `multi_rare_rewrite`, which serves as an interface with rewriting rules written in the RARE language.
- Added rule `aci_simp`, that describe ACI reasoning performed to simplify terms using SMT-LIB ACI operators.
- Rule `connective_def` extended to complement the case where an existential quantification is rewritten as a negated universal quantification.
- Addition of a section describing bitvector proofs.
- Bitblasting rules: `bitblast_extract`, `bitblast_add`, `bitblast_ult`.
- Addition of rules `la_mult_pos` and `la_mult_neg` to describe multiplication with a positive or negative factor.
- Addition of rules `symm` and `not_symm` to express symmetry of equality explicitly.
- Addition of the rule `eq_symmetric` to express symmetry of equality explicitly but as an equivalence. Note that in principle this could be done with the rule `symm` above, but would require a long and tedious use of `subproof` for each direction of the equivalence.
- Addition of the rule `weakening` to express weakening of a clause.
- Addition of the `reordering` rule to represent reordering of the literals in a clause.
- Addition of the `bind_let` rule. This rule can be used to preprocess `let` expressions similar to the `bind` rule used with ordinary quantifiers.
- Addition of miniscoping rules: `miniscope_split`, `miniscope_distribute`, `miniscope_ite`
- Addition of the `shuffle` to permute the arguments of a commutative operator.
- Clarify that `distinct_elim` cannot be applied to only one argument due to SMT-LIB restrictions.

Breaking changes:

- Allow arbitrary extra annotations in `assume` commands.

- Add the sort to all variables in contexts. Before, the context of a bind could be $(x\ S)\ (:=\ y\ x)$. Now it must be $(x\ S)\ (:=\ (y\ S)\ x)$.
- The arguments for `forall_inst` have been changed to no longer take the shape of bindings using $(:=\ x\ c)$. Instead, the list of instantiation terms must follow the variable order and cover all the respective bound variables.
- The rules `and_pos`, `or_neg`, `and`, `not_or` now have one argument that indicates which subterm they use.
- Add new syntax for decimal and negative numbers and use it for `la_generic`.
- Restrict sorting of numbers such that the sort of a constant is only determined by its syntactic category.
- Add new syntax for decimal and negative numbers and clarify that the set of allowed symbols is restricted to not conflict with this new syntax.
- Always use decimals for the constants in `la_generic`.
- Restrict sorting of numbers such that the sort of a constant is only determined by its syntactic category.
- Make `cong` more flexible. It is now possible to skip a prefix of the arguments if they are already the same on both sides. This is intended to future-proof the format for support of partial application.

Clarifications and corrected errors:

- Clarify that the `:args` annotation in `anchor` can be omitted if the list is empty.
- Clarify that `bind` can work on existential and universal quantifiers.
- Fix mistake in proof grammar. It now uses the `context_annotation` non-terminal in the rule for the `anchor` command.
- Fix the example of `onepoint`.
- Add the missing context to the conclusion of `bind`, `sco_ex`, `sco_forall`, `onepoint`.
- Highlight the equality used with context by using the $\approx$ symbol.
- Specify that printing a proof in SMT-LIB mode requires the printing of extra parentheses around the proof.
- Clarify the handling of variadic operators.

0.3 — 2023-02-10

This release overhauls the entire document, but introduces only few changes to the proof format itself.

The standard now specifies that `assume` commands can only be issued at the start of the proof or right after an `anchor` command.

Beyond many smaller clarifications and typographic improvements, the following changes are implemented in this release.

- Add an abstract proof checking procedure to clarify the semantics of the proof format.
- Add a description of the semantics of contexts based on λ -terms.
- List all transformations that are implicit in Alethe proofs.
- Change the notation used for terms from first-order style (e.g., $f(x, g(y))$) to higher-order style (e.g., $(f\ x\ (g\ y))$). This is only a change in notation – the used logic remains many-sorted first-order logic.
- Eliminate the distinction between if-and-only-if and equality. Instead, use equality (the symbol $\approx$) with appropriate sorts.
- Add an index that lists all proof rules.

Proof rules:

- The rule `implies_simplify` is no longer allowed to perform the simplification $(\varphi_1 \rightarrow \varphi_2) \rightarrow \varphi_2 \Rightarrow \varphi_1 \vee \varphi_2$. This is now covered by `bool_simplify`.

0.2 — 2022-12-19

This is an intermediate release. It collects all changes to the original specification document before the major changes that were implemented as part of Hans-Jörg Schurr's PhD thesis. These changes will be reflected in release 0.3.

This release implements major changes to the structure of the document to clarify the difference between the *language* and the *rules*. The language has a formal definition and a proof of soundness. The syntax describes how proofs are encoded in the text file.

The syntax was extended to allow extra annotations. Tools consuming Alethe proofs must be able to ignore such extra annotations.

List of rules:

- Improve description of `sko_ex`.
- Add `hole` rule to allow holes. A proof that contains steps that use this rule is not valid.

Corrections:

- Grammar: the `choice` binder can only bind one variable.

Clarifications:

- Clarify functionality of choice in introduction.
- Add illustrating example to introduction.
- Normalize printing of (variable, term) arguments in the abstract notation.
- Fix linear arithmetic example in introduction.
- Change syntax of abstract proof steps to be clearer.

0.1 – 2021-07-10

This is the first public release of this document. It coincides with the seventh PxTP Workshop.

Index

abstraction
 lambda, 7, 18
Alethe, 4
anchor, 11
assumption, 5

binder, 17
bitblasting, 23
bitvector, 23

choice, 10
context, 5, 6, 12, 17
 calculated, 15

lambda calculus, 17
lemma, 5

metaterms, 17

proof, 5
 valid, 17
 well-formed, 17
proof checker, 3

rule
 concluding, 11

S-expression, 3
step, 5
 outermost, 17
subproof, 5
 first-innermost, 15
 valid, 16
substitution, 4

term
 lambda, 17
 meta, 17

well-formed, 14

References

- [1] Haniel Barbosa, Clark Barrett, Martin Brain, Gereon Kremer, Hanna Lachnitt, Makai Mann, Abdalrhman Mohamed, Mudathir Mohamed, Aina Niemetz, Andres Nötzli, Alex Ozdemir, Mathias Preiner, Andrew Reynolds, Ying Sheng, Cesare Tinelli, and Yoni Zohar. *cvc5: A versatile and industrial-strength SMT solver*. In Dana Fisman and Grigore Rosu, editors, *TACAS 28*, pages 415–442, Cham, 2022. Springer International Publishing.
- [2] Haniel Barbosa, Jasmin C. Blanchette, Mathias Fleury, and Pascal Fontaine. Scalable fine-grained proofs for formula processing. *Journal of Automated Reasoning*, 64(3):485–510, January 2019.
- [3] Clark Barrett, Pascal Fontaine, and Cesare Tinelli. The SMT-LIB Standard: Version 2.6. Technical report, Department of Computer Science, The University of Iowa, 2017. Available at www.SMT-LIB.org.
- [4] Frédéric Besson, Pascal Fontaine, and Laurent Théry. A flexible proof format for SMT: A proposal. In Pascal Fontaine and Aaron Stump, editors, *PxTP 1*, pages 15–26, August 2011.

- [5] David Déharbe, Pascal Fontaine, and Bruno Woltzenlogel Paleo. Quantifier inference rules for SMT proofs. In Pascal Fontaine and Aaron Stump, editors, *PxTP 1*, pages 33–39, August 2011.
- [6] Burak Ekici, Alain Mebsout, Cesare Tinelli, Chantal Keller, Guy Katz, Andrew Reynolds, and Clark Barrett. SMTCoq: A plug-in for integrating SMT solvers into Coq. In Rupak Majumdar and Viktor Kunčák, editors, *CAV 29*, volume 10426 of *Lecture Notes in Computer Science*, pages 126–133. Springer, 2017.
- [7] Mathias Fleury and Hans-Jörg Schurr. Reconstructing veriT proofs in Isabelle/HOL. In Giselle Reis and Haniel Barbosa, editors, *PxTP 6*, volume 301 of *EPTCS*, pages 36–50, 2019.
- [8] Hans-Jörg Schurr, Mathias Fleury, and Martin Desharnais. Reliable reconstruction of fine-grained proofs in a proof assistant. In André Platzer and Geoff Sutcliffe, editors, *CADE 28*, Lecture Notes in Computer Science, pages 450–467, Cham, 2021. Springer International Publishing.
- [9] Wikipedia contributors. Alethe (bird) – Wikipedia, the free encyclopedia, 2022. Online; accessed 2022-09-02.