Effective Skolemization*

Matthias Baaz[1] [0000-0002-7815-2501] and Anela Lolić[2] [0000-0002-4753-7302]

[1] Institute of Discrete Mathematics and Geometry, TU Wien, Vienna, Austria
baaz@logic.at
[2] Kurt Gödel Society, Institute of Logic and Computation, TU Wien, Vienna, Austria
anela@logic.at

Abstract. We define a new relatively simple Skolemization method called atomic Skolemization which allows for a non-elementarily bounded speed-up of cut-free **LK**-proofs and resolution proofs w.r.t. the standard Skolemization and Andrews Skolemization.

Keywords: Skolemization · Cut-free Proofs · Resolution Proofs

1 Introduction

Skolem functions are one of the most important features of classical and related first-order logics. They represent quantifiers within the term language, similar to epsilon calculus. A Skolemization is a functional from a closed formula with distinct bound variables to a closed formula with distinct bound variables, which replaces some occurrences of bound variables by Skolem terms (terms of bound variables and new functions) such that all bound variables in the Skolem term belong to not replaced quantifiers where the term is in the scope.

For satisfiability of formulas, the main precondition of the introduction of Skolem functions is the preservation of soundness. For validity of formulas the dual main precondition is that the original formula is valid when the Skolemized formula is valid. In this contribution we work with Skolemization in the sense of satisfiability.

The standard Skolemization in the satisfiability case is based on the replacement of positive existential and negative universal quantifiers by Skolem functions depending on all negative existential and positive universal quantifiers where the replaced quantifier is in the scope.

Example 1. Consider the formula

$$\forall x(\exists y P(y) \lor \forall u \exists v(R(x,u) \lor Q(x,v)).$$

Then its Skolemization is

$$\forall x(P(f(x)) \lor \forall u(R(x,u) \lor Q(x,g(x,u)).$$

The quantified variable y is replaced by $f(x)$, where f is a fresh Skolem function symbol, and the quantified variable v is replaced by $g(x,u)$, for the fresh Skolem function symbol g.

* Supported by FWF projects I 5848, I 4427, and P 36571.

The standard Skolemization is sound because the addition of Skolem axioms

$$\forall \overline{x}(\exists y A(y,\overline{x})) \supset A(f(\overline{x}),\overline{x}) \quad \text{and} \quad \forall \overline{x}(A(f(\overline{x}),\overline{x}) \supset \forall y A(y,\overline{x}))$$

to a satisfiable set of sentences is conservative (it is possible to argue also directly replacing quantifiers within the formulas). The conservativity of Skolem axioms corresponds to the fact that in the case of validity Skolemized formulas are not weaker than the original ones. The introduction of Skolem formulas by projection of positive universal and negative existential quantifiers is always possible.[3]

Andrews Skolemization [2, 3] is an optimized form of standard Skolemization, where positive existential and negative universal quantifiers are replaced by Skolem functions depending only on the negative existential or positive universal quantifiers which bind in the subformula that begins with the quantifier to be replaced.

Example 2. Consider the formula

$$\forall x(\exists y P(y) \vee \forall u \exists v(R(x,u) \vee Q(x,v)).$$

Then its Andrews Skolemization is

$$\forall x(P(c) \vee \forall u(R(x,u) \vee Q(x,g(x,u)).$$

Here, the quantified variable y is replaced by the Skolem constant c (as x does not occur in $P(y)$, and the quantified variable v is replaced by $g(x,u)$, as x and u occur in $R(x,u) \vee Q(x,v)$.

To refute a formula in theorem proving based on resolution refutation, the formula has first to be Skolemized, then transformed into its clause form, and finally refuted with the resolution method. It was shown that Andrews Skolemization allows for a non-elementarily[4] bounded speed-up of the resolution proofs with regard to standard Skolemization [8]. In this contribution we present a simple algorithm for a Skolemization method, which is more effective than Andrews Skolemization: There is a speed-up even over Andrews Skolemization.

[3] It is obvious that the validity of the argument for the conservativity of Skolem axioms is equivalent to the validity of the full axiom of choice. To demonstrate that valid Skolemized formulas can be retransferred to their original form needs at most the completeness of first-order logic, i.e. the validity of König's lemma, which is much weaker than the axiom of choice. This difference can be explained as follows: The argument for conservativity of Skolem axioms validitates automatically the Skolem functions as functions, i.e. their identity axioms $\overline{x} = \overline{y} \supset f(\overline{x}) = f(\overline{y})$. Such axioms are not automatically eliminated when resetting Skolemized formulas in the validity sense.

[4] A primitive recursive function $f(x)$ is elementary if it is bound by a fix stack of 2: $2^{2^{\cdot^{\cdot^{x}}}}$.

2 Standard Skolemization and Andrews Skolemization

In this section the standard Skolemization method and the Andrews Skolemization method are introduced and compared.

Definition 1 (standard Skolem form w.r.t. satisfiability). *Let A be a closed first-order formula. If A does not contain positive existential or negative universal quantifiers, we define its standard Skolemization as* $\mathrm{sk}(A) = A$.

Suppose now that A contains positive existential or negative universal quantifiers and (Qy) is the first positive existential or negative universal quantifier occurring in A. If (Qy) is not in the scope of negative existential or positive universal quantifiers, then its standard Skolemization is

$$\mathrm{sk}(A) = \mathrm{sk}(A\backslash(Qy)\{y \leftarrow c\}),$$

where $A\backslash(Qy)$ denotes the formula A after omission of (Qy) and c is a constant symbol not occurring in A. If (Qy) is in the scope of the negative existential or positive universal quantifiers $(Q_1x_1)\dots(Q_nx_n)$, then its standard Skolemization is

$$\mathrm{sk}(A) = sk(A\backslash(Qy)\{y \leftarrow f(x_1,\dots,x_n)\}),$$

where f is a function symbol (Skolem function) not occurring in A.

In Andrews' method the introduced Skolem functions do not depend on the positive existential or negative universal quantifiers $(Q_1x_1)\dots(Q_nx_n)$ dominating the positive universal or negative existential quantifier (Qx), but on the subset of $\{x_,\dots,x_n\}$ appearing (free) in the subformula dominated by (Qx). In general, this method leads to smaller Skolem terms.

Definition 2 (Andrews Skolem form w.r.t. satisfiability). *Let A be a closed first-order formula. If A does not contain positive existential or negative universal quantifiers, we define its Andrews Skolemization as* $\mathrm{sk}_A(A) = A$.

Suppose now that A contains positive existential or negative universal quantifiers, $(Qy)B$ is a subformula of A and (Qy) is the first positive existential or negative universal quantifier occurring in A (in a tree-like ordering). If $(Qy)B$ has no free variables which are quantified by a negative existential or positive universal quantifier, then its Andrews Skolemization is

$$\mathrm{sk}_A(A) = \mathrm{sk}_A(A\backslash(Qy)\{y \leftarrow c\}),$$

where $A\backslash(Qy)$ denotes the formula A after omission of (Qy) and c is a constant symbol not occurring in A. If $(Qy)B$ has n variables $x_1,\dots,x_n$ which are quantified by a negative existential or positive universal quantifier from outside, then its Andrews Skolemization is

$$\mathrm{sk}_A(A) = \mathrm{sk}_A(A\backslash(Qy)\{y \leftarrow f(x_1,\dots,x_n)\}),$$

where f is a function symbol not occurring in A.

Let $\Gamma \to \Delta$ be a sequent, and let $F = \bigwedge \Gamma \supset \bigvee \Delta$ and $\mathrm{sk}_A(F) = \bigwedge \Pi \supset \bigvee \Lambda$, then we define the Andrews Skolemization of the sequent $\Gamma \to \Delta$ as

$$\mathrm{sk}_A(\Gamma \to \Delta) = \Pi \to \Lambda.$$

The usual Skolemizations are outside-in. This uses the global knowledge which of the bound variables are bound by positive universal or negative existential quantifiers. If we define standard Skolemization locally (i.e. inside-out), the result is an iteration of the Skolem functions within the Skolem semi-terms[5].

Example 3. Consider the formula

$$\exists x \forall y \exists u \forall v A(x, y, u, v).$$

Following the standard Skolemization (outside-in) we obtain

$$\forall y \forall v A(c, y, f(y), v),$$

and following the standard Skolemization inside-out we obtain

$$\forall y \forall v A(c, y, g(c, y), v).$$

The Skolem functions in the Skolem semi-terms are ordered in occurrence. Let g and h be Skolem function symbols that occur in a Skolem semi-term as $h(\ldots g(\ldots)\ldots)$, then we say that $h < g$. The iteration of the Skolem terms poses no problem by the following proposition which allow their elimination. We call such Skolem terms normalized.

Proposition 1. *The formulas A and A' are equi-satisfiable, where A is obtained from A' by replacing different iterated Skolem semi-terms $h(\ldots g(\ldots)\ldots)$ by Skolem semi-terms $f_i(\ldots)$ with new function symbols.*

Proof. $\Rightarrow$: obvious.

$\Leftarrow$: A $<$-minimal Skolem semi-term $g(\ldots)$ corresponds directly to one $f_i(\ldots)$ w.r.t. satisfiability. In the iterated case the Skolem semi-term $h(\ldots g(\ldots)\ldots)$ w.r.t. satisfiability corresponds also directly to a $f_j(\ldots)$, as $g(\ldots)$ is already determined.

From now on we will denote with # the operator that normalizes Skolem semi-terms according to Proposition 1.

Theorem 1. *The Andrews Skolemization preserves soundness.*

Proof. Proposition 1 allows us to argue locally, i.e. to replace positive existential or negative universal quantifiers inside-out. Assume the innermost still existing such quantifier is existential (analogously for the case of an universal quantifier). Then

$A(\ldots \exists x B(x, \overline{y}) \ldots)$ is satisfiable, where the occurrence of $\exists x B(x, \overline{y})$ is positive

[5] Semi-terms are terms that might contain bound variables.

$$\Downarrow$$

$$A(\ldots F(\overline{y})\ldots) \wedge \forall \overline{y}(F(\overline{y}) \supset \exists x B(x,\overline{y})) \wedge \forall \overline{y}(\exists x B(x,\overline{y}) \supset F(\overline{y})) \text{ is satisfiable}$$

$$\Downarrow$$

$$A(\ldots F(\overline{y})\ldots) \wedge \forall \overline{y}(F(\overline{y}) \supset B(f(\overline{y}),\overline{y})) \wedge \forall \overline{y}(B(f(\overline{y}),\overline{y}) \supset F(\overline{y})) \text{ is satisfiable}$$

by standard Skolemization with f and instantiation

$$\Downarrow$$

$$A(\ldots B(f(\overline{y}),\overline{y})\ldots) \text{ is satisfiable}$$

$$\Downarrow$$

$$A(\ldots \exists x B(x,\overline{y})\ldots) \text{ is satisfiable.}$$

Theorem 2 ([8]). *There is a sequence of refutable formulas $A_1, A_2, \ldots$ such that the length of the shortest resolution refutations of their standard clause forms*[6] *with standard Skolemization cannot be elementarily bounded in the length of the shortest resolution refutations of their standard clause forms with Andrews Skolemization.*

Proof (Sketch). The validity variant of standard Skolemization, i.e. the replacement of positive universal and negative existential quantifiers by Skolem terms corresponds exponentially in the length of cut-free proofs to usual sequent calculus **LK**, whereas Andrews Skolemization corresponds exponentially in the length of cut-free proofs to sequent calculus $\mathbf{LK}^+$ [1]. $\mathbf{LK}^+$ is obtained from **LK** by weakening the eigenvariable condition. The resulting calculus is therefore globally but possibly not locally sound. This means that all derived statements are true but that not every sub-derivation is meaningful. $\mathbf{LK}^+$-proofs are based on the side variable relation $<_{\varphi,\mathbf{LK}}$. We say b is a side variable of a in φ (written $a <_{\varphi,\mathbf{LK}} b$) if φ contains a positive universal or negative existential quantifier inference of the form

$$\frac{\Gamma \rightarrow \Delta, A(a,b,\overline{c})}{\Gamma \rightarrow \Delta, \forall x A(x,b,\overline{c})}\ \forall_r$$

or of the form

$$\frac{A(a,b,\overline{c}), \Gamma \rightarrow \Delta}{\exists x A(x,b,\overline{c}), \Gamma \rightarrow \Delta}\ \exists_l$$

Proofs are determined by $\mathbf{LK}^+$-suitable quantifier inferences. We say a quantifier inference is suitable for a proof φ if either it is a positive existential or negative universal quantifier inference, or the following three conditions are satisfied:

- (substitutability) the eigenvariable does not appear in the conclusion of φ.
- (side variable condition) the relation $<_{\varphi,\mathbf{LK}}$ is acyclic.

[6] See Definition 7.

– (weak regularity) the eigenvariable of an inference is not the eigenvariable of another positive universal or negative existential quantifier inference in φ.

$\mathbf{LK}^+$ is obtained from **LK** by replacing the usual eigenvariable conditions by $\mathbf{LK}^+$-suitable ones. $\mathbf{LK}^+$ admits cut-elimination and there is a non-elementary speed-up of cut-free $\mathbf{LK}^+$-proofs w.r.t. cut-free **LK**-proofs.

The following proposition is obvious.

Proposition 2. *Standard Skolemization and Andrews Skolemization coincide on prenex formulas.*

3 Atomic Skolemization

For simplicity we define the new algorithm for satisfiability and closed formulas with distinct bound variables in negation normal form (NNF). Therefore, existential quantifiers are replaced by Skolem terms.

Similar to Andrews Skolemization, atomic Skolemization is based on the elimination of the innermost quantifiers, i.e. generating iterated Skolem semi-terms in principle. This situation can be stratified using Proposition 1.

Definition 3. *Let F be a closed NNF formula with distinct bound variables. Then $<_F$ is a total order of the bound variables occurring in F, such that whenever Qx occurs in the scope of $Q'y$, we have that $x <_F y$, where $Q, Q' \in \{\forall, \exists\}$ and x, y are bound variables in F.*

Note that we might omit the subscript F in $<_F$ whenever it is clear from the context. For simplicity reasons in the Skolemization procedure, we will introduce the notion of *corresponding quantifier* of a bound variable.

Definition 4. *Let F be a closed NNF formula with distinct bound variables. Let x be such a bound variable. Then its corresponding quantifier is denoted by $\Psi(x)$, i.e.*

$$\Psi(x) = \begin{cases} \exists & \text{if } x \text{ is bound by } \exists, \\ \forall & \text{if } x \text{ is bound by } \forall. \end{cases}$$

The atomic Skolemization of a closed NNF formula F with distinct bound variables is computed based on the set of atomic semi-formulas occurring in F and containing the bound variables, and on the substitutions of Skolem semi-terms for these bound variables. We first give a description of the procedure, and then a formal definition of the algorithm for atomic Skolemization.

In a first step we consider all the atoms of the formula F and construct a set of sets of bound variables by collecting all the bound variables occurring in each of the atoms, which are not empty (this set will later be denoted with L_n). The substitution is initialized with the identity substitution. As long as L_n is not empty, we pick the $<_F$-minimal bound variable x and the corresponding sets in L_n containing x. Note that these sets might contain also other variables, which we denote by $\overline{y}$. In case the corresponding quantifier of x is existential, i.e. $\Psi(x) = \exists$, we delete all sets $\{x, \overline{y}_i\}$ from L_n and add $\{\overline{y}\}$ to the remaining variables. Furthermore, we add $\{x \leftarrow f(\overline{y})\}$, where f is a new function symbol to the set of substitutions. Alternatively, in case $\Psi(x) = \forall$, the sets $\{x, \overline{y}_i\}$ are again deleted from L_n and we add a set $\{\overline{y}\}$ to the remaining variables, but the set of substitutions is not updated. ($\{\overline{y}\}$ is only added when it is maximal under inclusion and the initial L_0 is stratified in this respect.) Finally, the iterated Skolem terms are replaced by uniterated ones according to Proposition 1.

Definition 5. *Let F be a closed NNF formula with distinct bound variables $V(F)$. Then its atomic Skolemization* $\mathrm{AS}(F)$ *is computed by the following steps:*

1. $L_0 = \{\{\gamma_1, \ldots, \gamma_n\} \mid \{\gamma_1, \ldots, \gamma_n\} \in V(F)$ *(and* $\neq \emptyset$*) which occur jointly in an atom of* $F\}$.
2. $\sigma_0 = id$ (σ_n *will substitute Skolem semi-terms for bound variables*).
3. $L_n = L_n \backslash \{\gamma_1, \ldots, \gamma_n\}$ *if* $\{\gamma_1, \ldots, \gamma_n\}$ *is not maximal in* L_n *w.r.t. inclusion.*
4. *while* $L_n \neq \emptyset$
   6. *Let* x *be the* $<_F$*-minimal variable in* L_n *and* $\Delta_{n+1} = \{\{\gamma_1, \ldots, \gamma_n\} \mid \{\gamma_1, \ldots, \gamma_n\}$ *in* L_n *containing* $x\}$. *Let* $x, \overline{y}$ *all the variables in* Δ_{n+1}.
   7. *If* $\Psi(x) = \exists$*:*
      $L_{n+1} = L_n \backslash \Delta_n \cup \{\overline{y}\}$ *if* $\{\overline{y}\}$ *is maximal in* $L_n \backslash \Delta_n$*,* $L_n \backslash \Delta_n$ *otherwise,*
      $\sigma_{n+1} = \sigma_n \cup \{x \leftarrow f(\overline{y})\}$, *where* f *a new function symbol.*
   8. *If* $\Psi(x) = \forall$*:*
      $L_{n+1} = L_n \backslash \Delta_n \cup \{\overline{y}\}$ *if* $\{\overline{y}\}$ *is maximal in* $L_n \backslash \Delta_n$*,* $L_n \backslash \Delta_n$ *otherwise.*
9. $L_n = \emptyset \Rightarrow \sigma = \sigma_n$.
10. *Let* F' *be* F *after deletion of* $\exists$. *Then* $\mathrm{AS}(F) = \#F'\sigma$.

Note that this algorithm is at most quadratic in the number of symbols of the original formula. However, its verification will need exponentially many steps.

Example 4. Let F be the formula

$$\forall x(\exists y P(y) \vee \forall u \exists v(R(x, u) \vee Q(x, v))).$$

We calculate its atomic Skolemization $\mathrm{AS}(F)$. To start, we initialize the set $L_0 = \{\{y\}, \{x, u\}, \{x, v\}\}$, with the ordering $v <_F u <_F y <_F x$.

As $\Psi(v) = \exists$ we obtain

$$L_1 = \{L_0 \backslash \{x, v\}\} \cup \{x\}, \quad \sigma_1 = \sigma_0 \cup \{v \leftarrow h(x)\}.$$

A $<_F$-minimal variable is now u. Then, as $\Psi(u) = \forall$, we obtain

$$L_2 = \{L_1 \backslash \{x, u\}\}, \quad \sigma_2 = \sigma_1$$

as $\{x\}$ is already in L_1. Now y is $<_F$-minimal. As $\Psi(y) = \exists$ we obtain in a next step

$$L_3 = L_2 \backslash \{y\}, \quad \sigma_3 = \sigma_2 \cup \{y \leftarrow c\}.$$

In a last step, as $\Psi(x) = \forall$, we obtain

$$L_4 = L_3 \backslash \{x\} = L_3 \backslash L_3 = \emptyset, \quad \sigma_4 = \sigma_3$$

F' is F after deletion of all occurrences of $\exists$, and $F'\sigma_4$ is

$$\forall x(P(c) \vee \forall u(R(x,u) \vee Q(x,h(x))))$$

which is also $\#F'\sigma = \mathrm{AS}(F)$ as no iterated Skolem terms occur.

Proposition 3. *Skolem functions can be combined over disjunctions. Let* $\overline{x}_i \in \overline{x}$

$$\forall \overline{x} \bigvee_i A_i(f_i(\overline{x}_i)) \supset \forall \overline{x} \bigvee_i A_i(f(\overline{x}))$$

is satisfiable, where f *is a new function symbol.*

Theorem 3 (Soundness of atomic Skolemization).

Proof. Consider step 3. in the AS-algorithm given in Definition 5. We have $L_n \neq 0$ and x the $<_F$-minimal variable.

$$\Delta_{n+1} = \{\{\gamma_1, \ldots, \gamma_n\} \mid \{\gamma_1, \ldots, \gamma_n\} \text{ in } L_n \text{ containing } x\},$$

$x, \overline{y}$ all the bound variables in Δ_n. Let $\exists x A(x, \overline{y})$ be the corresponding subformula.

$$\models \forall \overline{y} \forall \overline{z} (\exists x A(x, \overline{y}) \leftrightarrow \exists x \overbrace{\bigvee_i (\bigwedge_j B_{i,j}(x, \overline{y}_{i,j})) \wedge C_i(\overline{y}, \overline{z}))}^{(\times)},$$

where $\overline{y}_i = \cup_j(\overline{y}_{i,j})$, $(\times)$ is a suitable CNF where the $B_{i,j}$ atomic contain x and the C_i atomic do not.

$$\models \forall \overline{y} \forall \overline{z} (\exists x (\times) \leftrightarrow \overbrace{\bigvee_i (\exists x \bigwedge_j B_{i,j}(x, \overline{y}_{i,j})) \wedge C_i(\overline{y}, \overline{z}))}^{(\times\times)}, \quad \overline{y}_{i,j} \subseteq \overline{y}$$

$$\models \forall \overline{y} \forall \overline{z} ((\times\times) \to \overbrace{\bigvee_i \bigwedge_j B_{i,j}(f_i(\overline{y}), \overline{y}_{i,j})) \wedge C_i(\overline{y}, \overline{z}))}^{(\times\times\times)}$$

by Andrews Skolemization

$$\models \forall \overline{x} \forall \overline{z} ((\times \times \times) \to \overbrace{\bigvee_i (\bigwedge_j B_{i,j}(f(\overline{y}), \overline{y})) \wedge C_i(\overline{y}, \overline{z}))}^{(\times\times\times\times)}$$

by Proposition 3

$$\models \forall \overline{x} \forall \overline{z} ((\times \times \times\times) \to \exists x \overbrace{\bigvee_i \bigwedge_j B_{i,j}(x, \overline{y}) \wedge C_i(\overline{y}, \overline{z}))}^{(\times)}$$

Now let $\forall x A(x,\overline{y})$ be the corresponding subformula.

$$\models \forall\overline{y}\forall\overline{z}(\forall x A(x,\overline{y}) \leftrightarrow \forall x(\overbrace{\bigwedge_i(\bigvee_j B_{i,j}(x,\overline{y}_{i,j}) \wedge C_i(\overline{y},\overline{z}))}^{(\circ)})),$$

where $\overline{y}_i = \cup_j(\overline{y}_{i,j})$, $(\circ)$ is a suitable CNF where the $B_{i,j}$ contain x and the $C_{i,j}$ do not.

$$\models \forall\overline{y}\forall\overline{z}(\forall x(\circ) \leftrightarrow \bigwedge_i(\forall x \bigvee_j B_{i,j}(x,\overline{y}_{i,j})) \wedge C_i(\overline{y},\overline{z}))).$$

Now introduce new predicates F_i and add suitable

$$\forall\overline{y}(F(\overline{y}_{i,j}) \leftrightarrow \forall x \bigvee_j B_{i,j}(x,\overline{y}_{i,j}))$$

and continue to work with the formula after replacement. Semi-subformulas containing x disappear from the main formula. The consideration to work with $\overline{y}$ instead of the subsets $\overline{y}_i$ might lead to larger dependencies, but not incorrect ones as all relevant variables are contained in $\overline{y}$.

As an application we obtain:

Corollary 1. *The monadic fragment of classical first-order logic is decidable.*

Proof. For a monadic function-free formula A, AS(A) contains only constants as Skolem functions, and therefore it is decidable whether a Herbrand expansion for AS(A) exists.

Proposition 4. *The arity of the Skolem function symbols w.r.t. atomic Skolemization is less or equal to the arity of the Skolem function symbols w.r.t. Andrews Skolemization which is less or equal to the arity of Skolem function symbols in standard Skolemization. The number of introduced Skolem function symbols is not increased.*

4 Speed-up Result for Cut-Free Proofs

In this section we demonstrate that there is a non-elementary speed-up for cut-free proofs of atomic Skolemization w.r.t. standard Skolemization and Andrews Skolemization. Let $\tau = \{Q\overline{x}A(\overline{x}) \vee Q^D\overline{x}A(\overline{x})$ closed $\mid Q$ quantifier string, Q^D dual quantifier sequence, A atomic$\}$. Our argument is based on the following theorem.

Theorem 4. *There is a sequence of sequents*

$$A_1 \rightarrow, A_2 \rightarrow, \ldots, A_i \rightarrow,$$

where $A_1, \ldots, A_i$ are in NNF containing universal quantifiers only such that

1. *there is a bound for a sequence of cut-free* **LK**-*proofs for*

$$\Delta_1, A_1 \rightarrow, \Delta_2, A_2 \rightarrow, \ldots$$

elementary in the complexity of $A_1 \rightarrow, A_2 \rightarrow, \ldots$ *for suitable* $\Delta_i \subseteq \tau$.
2. *there is no elementary bound for any sequence of cut-free proofs for*

$$A_1 \rightarrow, A_2 \rightarrow, \ldots$$

in the complexity of $A_1 \rightarrow, A_2 \rightarrow, \ldots, A_i \rightarrow$.

Proof. Consider Statman's sequence of provable quantifier-free statements following from universal formulas where the cut-free proofs grow non-elementarily versus the proofs with cuts, which are elementarily bounded [10, 7]. Cuts can be closed by inferring $A \supset A$ on the left side instead of the cut, closing $A \supset A$ with universal quantifiers and cutting it. Replace all cuts by prenex cuts in an elementary way [6]. Code the matrices of the cuts by using coding formulas

$$\forall x(F(\overline{x}) \leftrightarrow M(\overline{x}))$$

added to the antecedents and replace the cuts:

$$\dfrac{\begin{matrix} \Pi_i \rightarrow \Gamma_i, M(\overline{s}_i) & & M(\overline{s}_i), \Lambda_j \rightarrow \Delta_j \\ \vdots\ \vdots\ \vdots & & \vdots\ \vdots\ \vdots \\ \Pi \rightarrow \Gamma, Q\overline{x}M(\overline{x}) & & Q\overline{x}M(\overline{x}), \Lambda \rightarrow \Delta \end{matrix}}{\Pi, \Lambda \rightarrow \Gamma, \Delta}$$

$$\Downarrow$$

$$\dfrac{\Pi_i \rightarrow \Gamma_i, M(\overline{s}_i) \qquad F(\overline{s}_i) \rightarrow F(\overline{s}_i)}{M(\overline{s}_i) \supset F(\overline{s}_i), \Pi_i \rightarrow \Gamma_i, F(\overline{s}_i)} \qquad \dfrac{F(\overline{s}_i) \rightarrow F(\overline{s}_i) \qquad M(\overline{s}_i), \Lambda_j \rightarrow \Delta_j}{F(\overline{s}_i) \supset M(\overline{s}_i), \Lambda_j \rightarrow \Delta_j, F(\overline{s}_i)}$$

Apply $\wedge : l$ and $\forall : l$ to infer the equivalence $\forall x(F_i(\overline{x}) \leftrightarrow M_i(\overline{x}))$.

$$\Downarrow$$

$$\dfrac{\Pi \rightarrow \Gamma, Q\overline{x}F(\overline{x}) \qquad Q\overline{x}F(\overline{x}), \Lambda \rightarrow \Delta}{\forall x(F(\overline{x}) \leftrightarrow M(\overline{x})), \Pi, \Lambda \rightarrow \Gamma, \Delta}$$

These codings do not shorten the cut-free proofs much, as they can be immediately eliminated by replacing F by M and eliminating $\forall \overline{x}(M(\overline{x}) \leftrightarrow M(\overline{x}))$ by universal cuts whose elimination is at most double exponential. By an easy transformation we obtain cut-free proofs by adding $Q\overline{x}F(\overline{x}) \vee Q^D\overline{x}\neg F(\overline{x})$.

Note that for standard, Andrews, and atomic Skolemization it holds that the Skolemization of A w.r.t. satisfiability corresponds to the Skolemization of $A \rightarrow$ w.r.t. validity.

Definition 6. *$H(A)$, where $A \in \tau$ ($A = Q\overline{x}A(x) \vee Q^D\overline{x}A(\overline{x})$) is the prenexification of A such that $\forall$ always stands in front of the dual $\exists$, and $H(\Delta)$, where $\Delta \subseteq \tau$, is $\{H(A) \mid A \in \Delta\}$.*

Example 5. $H(\exists x\forall yB(x,y) \vee \forall u\exists v\neg B(u,v)) = \forall u\exists x\forall y\exists v(B(x,y) \vee \neg B(u,v))$.

Theorem 5. *There is a sequence of formulas $B_1, B_2 \ldots$ such that*

1. *there is a bound for a sequence of cut-free proofs for*

$$\mathrm{AS}(B_1) \rightarrow, \mathrm{AS}(B_2) \rightarrow, \ldots$$

   elementary in the complexity of $B_1, B_2 \ldots$.
2. *there is no elementary bound for any sequence of cut-free proofs for*

$$sk(B_1) \rightarrow, sk(B_2) \rightarrow, \ldots$$

   in the complexity of $B_1, B_2 \ldots$.
3. *there is no elementary bound for any sequence of cut-free proofs for*

$$sk_A(B_1) \rightarrow, sk_A(B_2) \rightarrow, \ldots$$

   in the complexity of $B_1, B_2 \ldots$.

Proof. By Proposition 2 standard Skolemization and Andrews Skolemization coincide for prenex formulas. Therefore, we argue only for standard Skolemization. Let $B_i = \bigwedge_{H(\Delta_i)\wedge A_i}$ from Theorem 4 (note that B_i is in NNF). Assume that there is an elementary bound for the cut-free proofs of

$$sk(B_1) \rightarrow,\ sk(B_2) \rightarrow,\ \ldots.$$

Therefore, there is an elementary bound for cut-free proofs of

$$sk(C_1^1), \ldots sk(C_n^1), sk(A_1') \rightarrow,\ sk(C_1^2), \ldots sk(C_n^2), sk(A_2') \rightarrow,\ \ldots,$$

where Δ_i is $C_1^i, \ldots C_n^i$ and A_i' is obtained from A_i by shifting the universal quantifiers outside. By [5] there is an elementary bound for the corresponding Herbrand sequent. Note that the Skolem terms always depend on the dual position, w.l.o.g.

$$D(\ldots t_j \ldots) \vee \neg D(\ldots f_i(\ldots t_j \ldots) \ldots).$$

Now replace all occurrences of $f_i(\ldots t_j \ldots)$ inside-out by t_j. As the Herbrand expansion is propositionally valid, and the term is replaced on all positions by the same term, the result remains valid. Finally all Skolem terms disappear, and the original Skolemized formulas in $H(\Delta)$ are transformed into formulas of the form $E_i \vee \neg E_i$, which do not influence the validity of the remaining sequent. Hence, the size of the remaining sequents is elementarily bounded and therefore the cut-free proofs are elementarily bounded. Contradiction to Theorem 4.

Now consider

$$\mathrm{AS}(B_1), \mathrm{AS}(B_2), \ldots.$$

Note that the bound variables in $Q\overline{x}A(\overline{x})$ and $Q^D\overline{x}A(\overline{x})$ in $Q\overline{x}A(\overline{x}) \vee Q^D\overline{x}A(\overline{x}) \in \Delta_i$ are distinct, which remains invariant w.r.t. any prenexation. Therefore, the atomic Skolemization of

$$H(Q\overline{x}A(\overline{x}) \vee Q^D\overline{x}A(\overline{x}))$$

is the standard Skolemization of $Q\overline{x}A(\overline{x}) \vee Q^D\overline{x}A(\overline{x})$. Deskolemization of cut-free proofs is exponential [4], therefore the cut-free proofs of

$$\mathrm{AS}(B_1) \to, \mathrm{AS}(B_2) \to, \dots$$

are elementarily bounded.

5 Cut-Free LK-Proofs With Positive Existential / Negative Universal Quantifiers and Resolution

As we are interested in this paper mainly in the impact of different forms of Skolemization we allow any elementary form of clause form constructions (for the purpose of this paper it is not necessary to specify the exact form of resolution proofs, as they simulate each other within elementary bounds in the complexity of the proofs). This leads to a non-elementary speed-up of resolution proofs presupposing atomic Skolemization w.r.t. resolution proofs presupposing standard Skolemization or Andrews Skolemization.

Definition 7. *Let A be a formula which contains only positive existential or negative universal quantifiers when written on the left side of the sequent sign and therefore only positive universal or negative existential quantifiers when written on the right side of the sequent sign. An admissible clause form construction consists of sequents $A \to C$ and $C \to A$ elementary in the complexity of A, where*

1. *C (the clause form) is a conjunction of universally quantified disjunctions of literals (negated or unnegated atomic formulas),*
2. *$A \to C$ and $C \to A$ are cut-free elementary derivable in the complexity A.*

Note that both, structural clause forms and standard clause forms fall under this definition, together with clause forms which allow for atom evaluation etc. [9].

Theorem 6.

1. *Let φ be a cut-free* **LK***-proof of the sequent*

$$A_1, \dots, A_n \to B_1, \dots, B_m$$

with positive existential or negative universal quantifiers only. Then there is a resolution refutation of an admissible clause form of

$$A_1 \wedge \ldots \wedge A_n \wedge \neg B_1 \wedge \ldots \wedge \neg B_m$$

elementary in the complexity of φ.

2. *Let φ' be a resolution refutation of an admissible clause form of*

$$A_1 \wedge \ldots \wedge A_n \wedge \neg B_1 \wedge \ldots \wedge \neg B_m.$$

Then there is a cut-free **LK***-proof of*

$$A_1, \ldots, A_n \rightarrow B_1, \ldots, B_m$$

with positive existential or negative universal quantifiers only elementary in the complexity of φ'.

Proof. See [8, 9].

The next theorem follows directly from the theorem above.

Theorem 7. *There is a sequence of formulas $B_1, B_2 \ldots$ such that*

1. *there is a bound for a sequence of resolution refutation of standard clause forms of*
$$\mathrm{AS}(B_1) \rightarrow,\ \mathrm{AS}(B_2) \rightarrow,\ \ldots$$
elementary in the complexity of $B_1, B_2 \ldots$.
2. *there is no elementary bound for any sequence of resolution refutations of standard clause forms of*
$$sk(B_1) \rightarrow,\ sk(B_2) \rightarrow,\ \ldots$$
in the complexity of B_1, B_2, ...
3. *there is no elementary bound for any sequence of resolution refutations of standard clause forms of*
$$sk_A(B_1) \rightarrow,\ sk_A(B_2) \rightarrow,\ \ldots$$
in the complexity of $B_1, B_2, \ldots$.

6 Conclusion

The worst case sequences constructed in this paper are highly artificial. It might be asked if they have an impact in the real world. It is however a known fact that worst case examples with extreme complexities correspond to practical examples which are not that bad, but bad enough.

References

1. Aguilera, J.P., Baaz, M.: Unsound inferences make proofs shorter. J. Symb. Log. **84**(1), 102–122 (2019)
2. Andrews, P.B.: Resolution in type theory. J. Symb. Log. **36**(3), 414–432 (1971)
3. Andrews, P.B.: Theorem proving via general matings. J. ACM **28**(2), 193–214 (1981)

4. Baaz, M., Hetzl, S., Weller, D.: On the complexity of proof deskolemization. J. Symb. Log. **77**(2), 669–686 (2012)
5. Baaz, M., Leitsch, A.: On skolemization and proof complexity. Fundam. Informaticae **20**(4), 353–379 (1994)
6. Baaz, M., Leitsch, A.: Cut normal forms and proof complexity. Ann. Pure Appl. Log. **97**(1-3), 127–177 (1999)
7. Baaz, M., Leitsch, A.: Methods of Cut-elimination, vol. 34. Springer Science & Business Media (2011)
8. Baaz, M., Lolic, A.: Andrews skolemization may shorten resolution proofs non-elementarily. In: LFCS. Lecture Notes in Computer Science, vol. 13137, pp. 9–24. Springer (2022)
9. Eder, E.: Relative complexities of first order calculi. Springer-Verlag (2013)
10. Statman, R.: Lower bounds on Herbrand's theorem. Proceedings of the American Mathematical Society pp. 104–107 (1979)