Effective Skolemization

Matthias Baaz and Anela Lolić
TU Wien

WoLLIC 2023

July 11, 2023

Skolemization in theorem proving: resolution refutation

- Skolemize the formula,
- transform the Skolemized formula into clause form,
- refute the clause form with the resolution method.

Most prominent method of Skolemization: standard (structural) Skolemization

We present different Skolemizations that are more effective than standard Skolemization.

Standard Skolemization

Example

Consider the formula

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

Then its standard Skolemization is

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

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

Andrews Skolemization

A ... closed first-order formula.

- A does not contain positive existential or negative universal quantifiers: $\mathrm{sk}_A(A) = A$.
- 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:
  - $(Qy)B$ has no free variables which are quantified by a negative existential or positive universal quantifier:

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

    c is a constant symbol not occurring in A.
  - $(Qy)B$ has n variables $x_1, \ldots, x_n$ which are quantified by a negative existential or positive universal quantifier from outside:

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

    f is a function symbol not occurring in A.

Example

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 $R(x,u) \vee Q(x,v)$.

Recall its standard Skolemization

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

Theorem

Andrews Skolemization preserves soundness.

Proof.

Assume the innermost still existing quantifier is existential (analogously for the case of an universal quantifier). Then

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

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

$\Rightarrow A(\dots F(\overline{y}) \dots) \wedge \forall \overline{y}(F(\overline{y}) \supset B(f(\overline{y}),\overline{y})) \wedge \forall \overline{y}(B(f(\overline{y}),\overline{y}) \supset F(\overline{y}))$ is satisfiable by standard Skolemization with f and instantiation

$\Rightarrow A(\dots B(f(\overline{y}),\overline{y}) \dots)$ is satisfiable

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

Theorem (First speed-up result, [B&L22])

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 with standard Skolemization cannot be elementarily bounded in the length of the shortest resolution refutations of their standard clause forms with Andrews Skolemization.

Theorem (First speed-up result, [B&L22])

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 with standard Skolemization cannot be elementarily bounded in the length of the shortest resolution refutations of their standard clause forms with Andrews Skolemization.

Proposition

Standard Skolemization and Andrews Skolemization coincide on prenex formulas.

Atomic Skolemization

F ... closed NNF formula with distinct bound variables $V(F)$. 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. $\sigma_0 = \mathrm{id}$ (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 x is existentially quanified: $L_{n+1} = L_n \backslash \Delta_n \cup \{\overline{y}\}$ 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 x is universally quantified: $L_{n+1} = L_n \backslash \Delta_n \cup \{\overline{y}\}$ 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 $\mathrm{AS}(F) = F'\sigma$.

Example

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

$L_0 = \{\{y\},\{x,u\},\{x,v\}\}$, with the ordering $v <_F u <_F y <_F x$.

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

$u\colon \forall \Rightarrow \quad L_2 = \{L_1 \backslash \{x,u\}\}, \quad \sigma_2 = \sigma_1$ ($\{x\}$ is already in L_1).

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

$x\colon \forall \Rightarrow \quad L_4 = L_3 \backslash \{x\} = L_3 \backslash L_3 = \emptyset, \quad \sigma_4 = \sigma_3$.

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 $\mathrm{AS}(F)$.

Theorem

Atomic Skolemization preserves soundness.

Proof.

Step 4. in the AS-algorithm: $L_n \neq 0$ 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

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

as Skolem functions can be combined over disjunctions

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

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.

As an application we obtain:

Collorary

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

Proof.

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

Proposition

The arity of the Skolem function symbols w.r.t. atomic Skolemization $\leq$ the arity of the Skolem function symbols w.r.t. Andrews Skolemization $\leq$ the arity of Skolem function symbols in standard Skolemization. The number of introduced Skolem function symbols is not increased.

Speed-up Result for Cut-Free Proofs

$\tau = \{Q\overline{x}A(\overline{x}) \vee Q^D\overline{x}A(\overline{x})$ closed $\mid Q$ quantifier string, Q^D dual quantifier sequence, A atomic$\}$.

Theorem

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

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

Theorem

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.

Standard Skolemization and Andrews Skolemization coincide for prenex formulas $\Rightarrow$ we argue only for standard Skolemization.

$B_i = \bigwedge_{H(\Delta_i) \wedge A_i}$ from the last theorem. 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 [B&Leitsch1994] there is an elementary bound for the corresponding Herbrand sequent.

Skolem terms always depend on the dual position, w.l.o.g.

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

Replace all occurrences of $f_i(\ldots t_j \ldots)$ inside-out by t_j. The Herbrand expansion is propositionally valid, and the term is replaced on all positions by the same term $\Rightarrow$ the result remains valid.

All Skolem terms disappear, and the original Skolemized formulas in $H(\Delta)$ are transformed into formulas of the form $E_i \vee \neg E_i$.

The size of the remaining sequents is elementarily bounded and therefore the cut-free proofs are elementarily bounded.

$\Rightarrow$ Contradiction to the last theorem.

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

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 does not change when in prenex form.

$\Rightarrow$ 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 [B&Hetzl&Weller2012]

$\Rightarrow$ the cut-free proofs of

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

are elementarily bounded.

Cut-free Proofs and Resolution

A ... formula containing only positive existential or negative universal quantifiers when written on the left side of the sequent sign (analogously when written on the right side)

An admissible clause form construction consists of sequents $A \rightarrow C$ and $C \rightarrow A$ elementary in the complexity of A, where

1. C (the clause form) is a conjunction of universally quantified disjunctions of literals,
2. $A \rightarrow C$ and $C \rightarrow A$ are cut-free elementary derivable in the complexity A.

Theorem

1. Let φ be a cut-free LK-proof of the sequent $A_1, \ldots, A_n \rightarrow B_1, \ldots, B_m$ 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 [B&L2022] or [Eder2013].

Theorem

There is a sequence of formulas $B_1, B_2 \ldots$ such that

1. there is a bound for a sequence of resolution refutations 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 $B_1, B_2, \ldots$.