Deskolemization, Equality and Logical Complexity

Daniel Weller
joint work with Matthias Baaz and Stefan Hetzl

TU Vienna

February 2, 2012
Workshop ,,Logical Models of Reasoning and Computation"

The power of functions

- Setting of this talk: classical first-order logic.[1]
- ,,Quantifiers can be eliminated by introduction of fresh functions".
- Known as Skolemization, Herbrandization.

[1] For simplicity, officially consider only formulas in NNF.

The power of functions

Example

Consider an assumption

$$\exists z \forall x \exists y. y > x \wedge y \neq z$$

Its Skolemization is

$$\forall x. s(x) > x \wedge s(x) \neq 0$$

where s is an uninterpreted function symbol, 0 an uninterpreted constant.

The power of functions

Proposition

For every formula φ,

- *its Skolemization $\mathrm{sk}(\varphi)$ does not contain $\forall$ quantifiers, and*
- *φ is valid iff $\mathrm{sk}(\varphi)$ is.*

The power of functions

Proposition

For every formula φ,

- *its Skolemization $\mathrm{sk}(\varphi)$ does not contain $\forall$ quantifiers, and*
- *φ is valid iff $\mathrm{sk}(\varphi)$ is.*

- Useful when working with (cut-free) proof systems: only have to consider one type of quantifier.

The power of functions

Proposition

For every formula φ,

- *its Skolemization* $\mathrm{sk}(\varphi)$ *does not contain* $\forall$ *quantifiers, and*
- φ *is valid iff* $\mathrm{sk}(\varphi)$ *is.*

- Useful when working with (cut-free) proof systems: only have to consider one type of quantifier.
- In general, $\varphi \to \mathrm{sk}(\varphi)$ but not vice-versa.

The power of functions

Proposition

For every formula φ,

- *its Skolemization* $\mathrm{sk}(\varphi)$ *does not contain* $\forall$ *quantifiers, and*
- φ *is valid iff* $\mathrm{sk}(\varphi)$ *is.*

- Useful when working with (cut-free) proof systems: only have to consider one type of quantifier.
- In general, $\varphi \rightarrow \mathrm{sk}(\varphi)$ but not vice-versa.
- We call this *proof-theoretic Skolemization*.

The power of functions

Proposition

The theory $T \cup \{\forall \mathbf{x}.\exists y \varphi(\mathbf{x}, y) \rightarrow \varphi(\mathbf{x}, f(\mathbf{x}))\}$ *is a conservative extension of* T *(where the language of* T *does not contain* f*).*

- We call this *model-theoretic Skolemization*.

The power of functions

- In a sense, Skolem functions have no power:
  - $\mathrm{sk}(\varphi)$ is valid iff φ is valid, and
  - adding Skolem axioms yields a conservative extension.
- In another sense, they may have power: How expensive is it to go from a proof with Skolem functions to a proof without?

The power of functions

- With a focus on logical complexity (number of nodes in the proof-tree), we discuss various results concerning the *deskolemization problem*:

How can Skolem functions be removed from proofs?
How does this affect the length of proofs?

The rest of this talk

1. A first approach
2. Cut-free tree-like proofs
3. Adding equality

A first approach

- The first algorithm to remove (both model- and proof-theoretic) Skolem functions from proofs (known to me) was given in D. Hilbert and P. Bernays, *Grundlagen der Mathematik II*, Springer, 1939.

Hilbert's ε-calculus

Predicate calculus + ε-symbol + ε-formulas

Example

ε-term: $\varepsilon_x \, \forall y \; x \neq s(y)$.

Hilbert's ε-calculus

Predicate calculus + ε-symbol + ε-formulas

Example

ε-term: $\varepsilon_x \, \forall y \; x \neq s(y)$.

ε-formulas:

$$\exists x \varphi(x) \rightarrow \varphi(\varepsilon_x \varphi(x))$$

The problem

We will look for an algorithm solving the following

Problem

Given a proof of φ using model-theoretic Skolem axioms, find a proof of φ that does not use these axioms.

Proof: Some proof system with cut (Hilbert-style, sequent calculus, ...)

A solution

Problem

Given a proof of φ using Skolem axioms, find a proof of φ that does not use Skolem axioms.

1. Consider only $\varphi = \exists \mathbf{x} \forall \mathbf{y} \psi(\mathbf{x}, \mathbf{y})$ with ψ quantifier-free.
2. Replace Skolem axioms and terms by ε-formulas and ε-terms.
3. Apply proof-theoretic Skolemization: $\exists \mathbf{x} \psi(\mathbf{x}, f_1(\mathbf{x}), \ldots, f_n(\mathbf{x}))$.
4. Apply the *extended first ε-Theorem*. Obtain proof of *Herbrand disjunction* $\bigvee_{1 \leqslant i \leqslant \ell} \psi(\mathbf{t_i}, f_1(\mathbf{t_i}), \ldots, f_n(\mathbf{t_i}))$.
5. This proof *does not use* ε-formulas.
6. Replace Skolem terms by variables and introduce quantifiers.

Summary of the approach

- Eliminating model-theoretic Skolem functions from proofs reduces to eliminating ε-terms from proofs.
- In the process, we introduce and eliminate proof-theoretic Skolem functions.
- All of this can be done, but the approach uses the extended first ε-theorem, which has *non-elementary worst-case complexity*.

Summary of the approach

- Complexity of algorithm due to the fact that an "essentially cut-free" proof is produced.
- But removal of the proof-theoretic Skolem functions from this proof is polynomial.
- In practice, Skolem functions are used differently.
- The rest of this talk[2] will deal with these uses.

[2] Except for the next slide.

Further results on model-theoretic Skolemization

- An algorithm due to Maehara (1955), based on cut-elimination.
- An algorithm due to Shoenfield (2001), based on Herbrand's theorem.
- A better algorithm for a subproblem due to Avigad (2003).

Methodology

- We are interested in the difference between natural pairs of systems: One with Skolem functions, one without.
- Two kinds of results:
  - *Transformations*: Showing how to go from one system to the other, along with complexity.
  - *Speed-ups*: Lower bounds on the complexity of transformation algorithms.

The rest of this talk

1. A first approach
2. Cut-free tree-like proofs
3. Adding equality

Cut-free proofs

- We are interested in the effect of Skolem functions on *cut-free proofs*.
- Cut-free proofs are interesting:
  - Usually generated by automated theorem provers.
  - Efficient extraction of data: Interpolants, Herbrand sequents.

Cut-free proofs

- We are interested in the effect of Skolem functions on *cut-free proofs*.
- Cut-free proofs are interesting:
  - Usually generated by automated theorem provers.
  - Efficient extraction of data: Interpolants, Herbrand sequents.
- Here, we look at cut-free *tree-like* proofs.

Cut-free proofs

We consider the "proof-theoretic deskolemization problem":

Problem

Input: φ, *proof of* $\mathrm{sk}(\varphi)$.
Output: *Proof of* φ.

A trivial example

$$\varphi = \forall x Px \to \forall x Px, \qquad \mathrm{sk}(\varphi) = \forall x Px \to Pc$$

$$\dfrac{Pc \vdash Pc}{\vdash \forall x Px \to Pc} \to_r, \forall_l$$

A trivial example

$$\varphi = \forall x P x \rightarrow \forall x P x, \qquad \mathrm{sk}(\varphi) = \forall x P x \rightarrow P c$$

$$\frac{P\alpha \vdash P\alpha}{\vdash \forall x P x \rightarrow \forall x P x} \rightarrow_r, \forall_r, \forall_l$$

The prenex case

- We have seen: The cut-free prenex case can be solved with polynomial expense.
- But in the cut-free case, requiring prenex formulas can have bad consequences.

The first speed-up

Theorem (Baaz, Leitsch 1994)

There exists a family of formulas $(\varphi_i)_{i \in \mathbb{N}}$ *(of elementary size) such that*

1. $\mathrm{sk}(\varphi_i)$ *have cut-free proofs of elementary length, but*
2. *there exist prefix forms* ψ_i *of* φ_i *such that all cut-free proofs of* $\mathrm{sk}(\psi_i)$ *have non-elementary length.*

Another example

$$\varphi = (\exists x Px \vee Q) \rightarrow \exists x.(Px \vee Q) \wedge T,$$
$$\mathrm{sk}(\varphi) = (Pc \vee Q) \rightarrow \exists x.(Px \vee Q) \wedge T$$

$$\frac{Pc \vee Q \vdash Pc \vee Q \quad \vdash T}{\vdash (Pc \vee Q) \rightarrow \exists x.(Px \vee Q) \wedge T} \rightarrow_r, \exists_r, \wedge_r$$

Another example

$$\varphi = (\exists x Px \vee Q) \to \exists x.(Px \vee Q) \wedge T,$$
$$\mathrm{sk}(\varphi) = (Pc \vee Q) \to \exists x.(Px \vee Q) \wedge T$$

$$\frac{\exists x Px \vee Q \vdash Pt \vee Q \qquad \vdash T}{\vdash (\exists x Px \vee Q) \to \exists x.(Px \vee Q) \wedge T} \to_r, \exists_r, \wedge_r$$

Another example

$$\varphi = (\exists x Px \lor Q) \to \exists x.(Px \lor Q) \land T,$$
$$\mathrm{sk}(\varphi) = (Pc \lor Q) \to \exists x.(Px \lor Q) \land T$$

$$\frac{\exists x Px \vdash \exists x.(Px \lor Q) \land {\color{red}T} \quad Q \vdash \exists x.(Px \lor Q) \land {\color{red}T}}{\vdash (\exists x Px \lor Q) \to \exists x.(Px \lor Q) \land T} \to_r, \lor_l,$$

The second speed-up

Proposition

There exists a family of formulas $(\varphi_k)_{k\in\mathbb{N}}$ *(of polynomial size) such that*

1. *there exist polynomial-length proofs of* $\mathrm{sk}(\varphi_i)$ *but*
2. *all proofs of* φ_i *have exponential length.*

A matching upper bound

- But this is the worst that can happen.

Proposition

Let π *be a proof of* $\mathrm{sk}(\varphi)$*. Then there exists a proof* λ *of* φ *such that* $|\lambda| \leqslant 2^{p(|\pi|)}$ *for some polynomial* p*.*

A matching upper bound

- But this is the worst that can happen.

Proposition

Let π be a proof of $\mathrm{sk}(\varphi)$. Then there exists a proof λ of φ such that $|\lambda| \leqslant 2^{p(|\pi|)}$ for some polynomial p.

Proof.

The idea is to extract a Herbrand disjunction D from π which is polynomial in $|\pi|$. Since D is propositional, it has an exponential proof. This proof can be converted to a proof of φ with polynomial expense. □

A matching upper bound

- But this is the worst that can happen.

Proposition

Let π be a proof of $\mathrm{sk}(\varphi)$. Then there exists a proof λ of φ such that $|\lambda| \leqslant 2^{p(|\pi|)}$ for some polynomial p.

Proof.

The idea is to extract a Herbrand disjunction D from π which is polynomial in $|\pi|$. Since D is propositional, it has an exponential proof. This proof can be converted to a proof of φ with polynomial expense. $\square$

Since φ is infix, one has to use a more involved data structure (*expansion tree proofs* due to Miller 1983) instead of Herbrand disjunctions.

Extending upper bounds

- This result can be lifted to some proofs with cut:

Proposition

Let π be a proof of $\mathrm{sk}(\varphi)$ such that for all Skolem terms $f(t_1, \ldots, t_n)$ occurring in cut-formulas, no t_i contains a bound variable. Then there exists a proof λ of φ such that $|\lambda| \leqslant 2^{p(|\pi|)}$.

Extending upper bounds

- This result can be lifted to some proofs with cut:

Proposition

Let π be a proof of $\mathrm{sk}(\varphi)$ such that for all Skolem terms $f(t_1, \ldots, t_n)$ occurring in cut-formulas, no t_i contains a bound variable. Then there exists a proof λ of φ such that $|\lambda| \leqslant 2^{p(|\pi|)}$.

Proof.

$$\dfrac{\Gamma \vdash \Delta, C(f(t)) \quad C(f(t)), \Gamma \vdash \Delta}{\Gamma \vdash \Delta}\ cut$$

is replaced by

$$\dfrac{\dfrac{\Gamma \vdash \Delta, C(f(t)) \quad C(f(t)), \Gamma \vdash \Delta}{C(f(t)) \to C(f(t)), \Gamma \vdash \Delta}\ \to_l}{\forall x. C(x) \to C(x), \Gamma \vdash \Delta}\ \forall_l$$

□

Extending upper bounds

- This result can be lifted to some proofs with cut:

Proposition

Let π *be a proof of* $\mathrm{sk}(\varphi)$ *such that for all Skolem terms* $f(t_1, \ldots, t_n)$ *occurring in cut-formulas, no* t_i *contains a bound variable. Then there exists a proof* λ *of* φ *such that* $|\lambda| \leqslant 2^{p(|\pi|)}$.

- Any cut-free deskolemization algorithm can be lifted to this class of proofs.
- One is reminded of the restriction imposed by (Miller 1983) to obtain *soundness* of Skolemization in higher-order logic.

The rest of this talk

1. A first approach
2. Cut-free tree-like proofs
3. Adding equality

Adding $=$

- We distinguish the signatures Σ (original) and $\mathcal{SK}$ (skolem functions).
- We work in proof systems with cut. Analogous results hold in the cut-free case.

Adding $=$

- Consider **LK**-proofs with $=$:
- Allow axioms $\forall x.x = x$ and
- the *equality schema*

$$\forall \mathbf{x}, \mathbf{y}. \mathbf{x} = \mathbf{y} \wedge A(\mathbf{x}) \rightarrow A(\mathbf{y})$$

where $A(\mathbf{x})$ is a formula over Σ or $\Sigma \cup \mathcal{SK}$.

Adding $=$

Problem

Input: *Proof of* $\mathrm{sk}(\varphi)$ *using the equality schema for* $\Sigma \cup \mathcal{SK}$.
Output: *Proof of* φ *using the equality schema for* Σ.

The third speed-up

Proposition

There exists a family $(\varphi_n)_{n\in\mathbb{N}}$ *of formulas, such that*

1. *the length of proofs of* φ_n *necessarily grows in* n*, but*
2. $\mathrm{sk}(\varphi_n)$ *have proofs using the equality schema for* $\Sigma \cup \mathcal{SK}$ *of constant length.*

Arbitrary speed-up in the presence of $=$

- For the proof, we use results on the generalization of proofs from (Baaz, Wojtylak 2008).
- This requires the notion of *proof skeleton*.

Proof Skeletons

Definition

The *skeleton* of a proof is obtained from a proof by dropping all sequents.

Hence a proof skeleton is a tree labelled with rules of proof. A formula φ is *derivable with a proof skeleton* Π if Π can be extended to a proof of φ.

Proof Generalization

Theorem (Baaz, Wojtylak 2008)

Let T be a finite extension of $\mathbf{LK}_{es}$ containing $\forall x.s(x) = x$. If $A(s^n(0))$ is derivable in T with a proof skeleton Π and n is large relative to A and Π, then $A(s^n(a))$ with a fresh free variable a is derivable in T.

Proof Generalization

We use the following consequence of this theorem:

Lemma

Let T be a finite extension of $\mathbf{LK}_{es}$ containing $\forall x.s(x) = x$ and assume that there exists a constant c such that $s^n(0) = 0$ has a proof π with length $\leqslant c$ for all n. Then $a = 0$ is derivable in T.

Proof Generalization

We use the following consequence of this theorem:

Lemma

Let T be a finite extension of $\mathbf{LK}_{es}$ containing $\forall x.s(x) = x$ and assume that there exists a constant c such that $s^n(0) = 0$ has a proof π with length $\leqslant c$ for all n. Then $a = 0$ is derivable in T.

Proof.

By the previous theorem, there exists an n such that $s^n(a) = 0$ is provable. Using $\forall x.s(x) = x$ we prove $a = 0$. $\square$

Speed-up

Now we can prove

Proposition

There exists a family $(\varphi_n)_{n\in\mathbb{N}}$ of formulas, such that

1. *the length of proofs of φ_n necessarily grows in n, but*
2. $\mathrm{sk}(\varphi_n)$ *have proofs using the equality schema for $\Sigma \cup \mathcal{SK}$ of constant length.*

Speed-up

Consider

$$\begin{aligned}\varphi_n \equiv & \forall x (s(x) = x) \wedge \\ & \forall xy \exists z.(x = 0 \wedge y = 0 \rightarrow z = 0) \wedge (z = x \rightarrow y = 0) \\ & \rightarrow s^n(0) = 0\end{aligned}$$

Speed-up

Consider

$$\begin{aligned}\varphi_n \equiv & \forall x(s(x) = x) \wedge \\ & \forall xy \exists z.(x = 0 \wedge y = 0 \rightarrow z = 0) \wedge (z = x \rightarrow y = 0) \\ & \rightarrow s^n(0) = 0\end{aligned}$$

Since $a = 0$ is not provable, by the previous lemma we know that $\forall x(s(x) = x) \rightarrow s^n(0) = 0$ do not have proofs of constant length.

Speed-up

Consider

$$\begin{aligned}\varphi_n \equiv & \forall x(s(x) = x) \wedge \\ & \forall xy \exists z.(x = 0 \wedge y = 0 \rightarrow z = 0) \wedge (z = x \rightarrow y = 0) \\ & \rightarrow s^n(0) = 0\end{aligned}$$

Since $a = 0$ is not provable, by the previous lemma we know that $\forall x(s(x) = x) \rightarrow s^n(0) = 0$ do not have proofs of constant length. Assume that φ_n has a constant length proof.

Speed-up

Consider

$$\begin{aligned}\varphi_n \equiv & \forall x(s(x) = x) \wedge \\ & \forall xy \exists z.(x = 0 \wedge y = 0 \to z = 0) \wedge (z = x \to y = 0) \\ & \to s^n(0) = 0\end{aligned}$$

Since $a = 0$ is not provable, by the previous lemma we know that $\forall x(s(x) = x) \to s^n(0) = 0$ do not have proofs of constant length. Assume that φ_n has a constant length proof. Note that $\forall xy \exists z.(x = 0 \wedge y = 0 \to z = 0) \wedge (z = x \to y = 0))$ is valid, hence $s(0) = 0 \to s^n(0) = 0$ has a constant length proof, contradiction.

Speed-up

It remains to prove

Proposition

There exists a family $(\varphi_n)_{n\in\mathbb{N}}$ of formulas, such that

1. *the length of proofs of φ_n necessarily grows in n, but*
2. $\mathrm{sk}(\varphi_n)$ *have proofs using the equality schema for $\Sigma \cup \mathcal{SK}$ of constant length.*

Speed-up

We have

$$
\begin{aligned}
\mathrm{sk}(\varphi_n) \equiv & \forall x(s(x) = x) \wedge \\
& \forall xy(x = 0 \wedge y = 0 \to f(x, y) = 0) \wedge (f(x, y) = x \to y = 0) \\
& \to s^n(0) = 0
\end{aligned}
$$

Note that the only binary symbol is $f \in \mathcal{SK}$.

Proving $\mathrm{sk}(\varphi_n)$

Define (Yukami's trick)

$$t_0(x, y) \equiv x, \qquad t_{k+1}(x, y) \equiv f(t_k(x, y), s^k(y)).$$

The proof of $\mathrm{sk}(\varphi_n)$ uses the equality schema

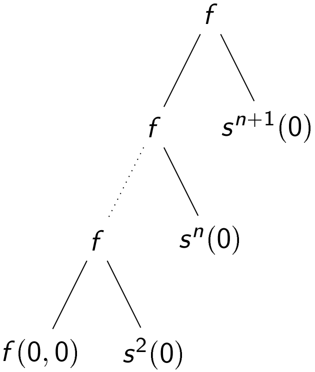
$$f(0, 0) = 0 \wedge s(0) = 0 \rightarrow t_n(f(0, 0), s(0)) = t_n(0, 0)$$

and the reflexivity axiom

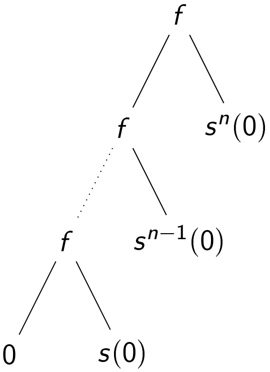
$$t_n(f(0, 0), s(0)) = f(t_n(0, 0), s^n(0)).$$

Proving $\mathrm{sk}(\varphi_n)$

It helps to visualize this. If $f(0,0) = 0$ and $s(0) = 0$ then the equality schema allows us to go from



to



The idea is that the equality schema allows replacement on all leaves in the term tree in a single step.

Sharpening the result

- Is this necessary? To find out, we can try to sharpen the result.
- Distinguish a natural subclass of the equality schema:
- An equality schema

$$\forall \mathbf{x}, \mathbf{y} . \mathbf{x} = \mathbf{y} \rightarrow f(\mathbf{x}) = f(\mathbf{y}),$$

where f is a function symbol is called an *equality axiom*.

Sharpening the result

Proposition

There exists a family $(\varphi_n)_{n\in\mathbb{N}}$ *of formulas, such that*

1. *the length of proofs of* φ_n *necessarily grows in* n, *but*
2. $\mathrm{sk}(\varphi_n)$ *have proofs using the equality schema for* $\Sigma \cup \mathcal{SK}$ *of constant length.*

Sharpening the result

Proposition

There exists a family $(\varphi_n)_{n\in\mathbb{N}}$ of formulas, such that

1. *the length of proofs of* $\mathrm{sk}(\varphi_n)$ *using the equality schema for* Σ*, but the equality* axioms *for* $\mathcal{SK}$*, necessarily grows in* n*, and*
2. $\mathrm{sk}(\varphi_n)$ *have proofs using the equality schema for* $\Sigma \cup \mathcal{SK}$ *of constant length.*

Sharpening the result

- The old proof does not work here: the proof generalization result does not apply in the presence of f.
- Another way: Find an algorithm that removes Skolem axioms from π such that the length increase is bounded by the length of π.

Sharpening the result

- This is work in progress.
- Idea for an algorithm:
- Translate a proof of

$$\forall x, y.x = y \to f(x) = f(y) \vdash \exists x.\varphi(x, f(x))$$

to a proof of

$$\bigwedge_{s,t} s = t \to f(s) = f(t) \vdash \bigvee_{r} \varphi(r, f(r)).$$

Sharpening the result

- This is work in progress.
- Idea for an algorithm:
- Translate a proof of

$$\forall x, y. x = y \rightarrow f(x) = f(y) \vdash \exists x. \varphi(x, f(x))$$

  to a proof of

$$\bigwedge_{s,t} s = t \rightarrow f(s) = f(t) \vdash \bigvee_{r} \varphi(r, f(r)).$$

- Then replace $f(s)$ by $f(t)$ or vice-versa.

Sharpening the result

- This is work in progress.
- Idea for an algorithm:
- Translate a proof of

$$\forall x, y. x = y \to f(x) = f(y) \vdash \exists x. \varphi(x, f(x))$$

  to a proof of

$$\bigwedge_{s,t} s = t \to f(s) = f(t) \vdash \bigvee_{r} \varphi(r, f(r)).$$

- Then replace $f(s)$ by $f(t)$ or vice-versa.
- The problem is that when both s and t are witnesses of $\exists x$ the result is not a Herbrand disjunction anymore.

A similar result for **LJ**

- We haven shown that there exists an arbitrary speed-up when using equality schemata for Skolem functions.
- There is a somewhat analogous result in the setting of intuitionistic logic due to (Mints 1998).

A similar result for **LJ**

- Skolemization for **LJ** is in general not complete:
- There exists unprovable φ such that $\mathrm{sk}(\varphi)$ is provable.

A similar result for **LJ**

- Skolemization for **LJ** is in general not complete:
- There exists unprovable φ such that $\mathrm{sk}(\varphi)$ is provable.

Proposition (Mints)

*Let S be a prenex sequent. Then $\mathrm{sk}(S)$ is **LJ**-provable iff S is.*

A similar result for **LJ**

- Again, the situation changes when $=$ for Skolem functions is added:

Proposition (Mints)

There exists a prenex sequent S such that $\mathrm{sk}(S)$ has an $\mathbf{LJ}_=$ proof, but S is not provable in $\mathbf{LJ}_=$.

A similar result for **LJ**

- Again, the situation changes when $=$ for Skolem functions is added:

Proposition (Mints)

There exists a prenex sequent S such that $\mathrm{sk}(S)$ has an $\mathbf{LJ}_=$ proof, but S is not provable in $\mathbf{LJ}_=$.

Proof.

Take

$$\forall z \exists x P(z,x) \vdash \forall z_1 \exists x_1 \forall z_2 \exists x_2 . P(z_1,x_1) \wedge P(z_2,x_2) \wedge (z_1 = z_2 \rightarrow x_1 = x_2)$$

then $\mathrm{sk}(S)$ is

$$\forall z P(z, f(x)) \vdash \exists x_1 \exists x_2 . P(c, x_1) \wedge P(g(x_2), x_2) \wedge (c = g(x_1) \rightarrow x_1 = x_2).$$

$\square$

Conclusion

- There are many interesting open problems regarding Skolem functions and proofs.
- The general algorithms are of non-elementary symbolic complexity.
- The use of the equality schema for Skolem functions has an interesting effect on logical complexity.
- Some open problems:
  - Complexity of the cut-free case for DAG-like proofs.
  - Sharpening the speed-up on equality schemata.
  - Consider systems with more assumptions on Skolem functions.