Strong Skolemization

Andreas Nonnengart*
Max-Planck-Institut für Informatik
and Saarbrücken University
Im Stadtwald, 66123 Saarbrücken, Germany

Abstract

Skolemization is a means to eliminate existential quantifiers within predicate logic sentences by replacing existentially quantified variables with Skolem function applications. The arguments of such Skolem functions are variables that are quantified outside the sub-formula under consideration. In this article a skolemization technique is introduced which abstracts from some of the Skolem function arguments. It turns out that the result obtained this way is usually more general than what can be achieved from standard (classical) Skolemization techniques. From a theorem prover's point of view such generalizations often lead to a reduction of both search space and proof length.

1 Introduction

In automated theorem proving much effort is spent on the development of more and more sophisticated reasoning calculi whereas the optimization of certain normal form transformations as they are, for example, necessary for resolution-based theorem proving, are fairly neglected.

One part of the usual clause normal form transformation for resolution is *Skolemization* (See the historical paper (Skolem 1920)). Its major effect lies in substituting so-called Skolem function applications for existentially quantified variables. Usually, there are two kinds of Skolemization techniques mentioned in the standard literature on automated theorem proving[1] which we call *Inner* and *Outer* Skolemization in this paper. The two differ mainly in the choice of the arguments for the Skolem functions (see below).

Strong Skolemization, as it is introduced here,[2]b can be viewed as a generalization of Inner Skolemization in the sense that some of the Skolem function arguments are replaced with new (universally quantified) variables. This often

*Andreas.Nonnengart@mpi-sb.mpg.de, http://www.mpi-sb.mpg.de/~nonnenga/
[1]see, e.g., (Loveland 1978, Chang and Lee 1973)
[2]A slightly less general version appeared in (Nonnengart 1996)

allows for considerable simplifications after generating normal forms for automated theorem proving and therefore leads to a reduction of both search space and proof length.

This article is organized as follows: First, the standard notion of Skolemization is recapitulated. Then the new Strong Skolemization technique is introduced which, finally, is to be compared with another, somewhat related, method developed recently, the so-called Optimized Skolemization.

2 Preliminaries

For simplicity, we assume formulas to be in negation normal form[3] and that any two different quantifiers inside a formula bind variables with different names. This is to ensure that the occurrence of a sub-formula $\exists x\, \Phi$ or $\forall x\, \Phi$ in a given formula Δ is unique.

By $\Delta\,[\Phi]$ we indicate that Φ is a sub-formula of Δ and $\Delta\,[\Phi \,/\, \Psi]$ is a representation of Δ with sub-formula Φ replaced by Ψ.[4]

We refer to the set of free (i.e. not bound by any quantifier) variables in Φ by $\mathcal{V}(\Phi)$. For any variable x occurring in Δ, the (unique) sub-formula starting with the quantifier which binds x is referred to by Δ^x. Thus, if $\exists x\, \Phi$ is a sub-formula of Δ then $\Delta^x = \exists x\, \Phi$.

For readability we abbreviate sequences of the form $u_1, \ldots, u_k$ of length k by $\overline{u}_k$.

A (first-order) interpretation $\mathcal{M} = (\mathcal{D}, \Im, \phi)$ consists of a non-empty domain of discourse $\mathcal{D}$, a mapping $\Im$ which associates functions and relations to function symbols and predicate symbols respectively, and a variable valuation ϕ which maps variables to elements of $\mathcal{D}$. Given a valuation ϕ we mean by $\phi[x/d]$ the variable valuation that is identical to ϕ except for the variable x which is to be interpreted as the domain value d. Similarly, by $\mathcal{M}[x/d]$ we understand the triple $(\mathcal{D}, \Im, \phi[x/d])$ provided $\mathcal{M} = (\mathcal{D}, \Im, \phi)$. Moreover, we use $\phi[\overline{x}_k/\overline{\alpha}_k]$ as short for $\phi[x_1/\alpha_1][x_2/\alpha_2]\cdots[x_k/\alpha_k]$.

Finally, whenever we write $\Phi\{x/t\}$ we mean Φ with every occurrence of the (free) variable x replaced by the term t.

Definition 2.1 (inner and outer skolemization)
We obtain the Skolemization *of a sentence Δ by replacing every occurrence of sub-formulas $\exists x\, \Phi$ in Δ with a corresponding $\Phi\{x/f(\overline{y}_n)\}$, where f is an n-place function symbol which is new to the problem under consideration. We speak of* Outer Skolemization *in case the variables $\overline{y}_n$ are all the universally quantified variables such that $\exists x\, \Phi$ is a sub-formula of Δ^{y_i} for each $1 \leq i \leq n$.*

[3] A formula is said to be in negation normal form if it doesn't contain implications, equivalences, and exclusive disjunctions, and all negations are moved inwards as far as possible (i.e., they only occur directly in front of predicate symbols).

[4] This definition won't be ambiguous since we are only going to replace sub-formulas of the form $\exists x\, \Phi$ in the sequel which, under the above assumption, are unique for the formulas we consider.

If $\{y_1, \ldots, y_n\} = \mathcal{V}(\exists x\, \Phi)$, *i.e. the variables* $\overline{y}_n$ *are* all *the free variables in the sub-formula* $\exists x\, \Phi$*, we speak of* Inner Skolemization.

Skolemization is not an equivalence transformation. It preserves satisfiability and unsatisfiability, though. This is stated in the following Lemma (proofs are omitted, for they can be found in most common textbooks and survey articles on first-order predicate logic theorem proving as, e.g., in (Chang and Lee 1973, Loveland 1978, Andrews 1981, Skolem 1920)).

LEMMA 2.2
Let $\Delta[\exists x\, \Phi]$ *be a sentence (in negation normal form) and let* $f, \overline{y}_n$ *as in Definition 2.1. Then*

$$\Delta\,[\exists x\, \Phi] \text{ is satisfiable}$$
$$\text{iff}$$
$$\Delta\,[\exists x\, \Phi \;/\; \Phi\{x/f(\overline{y}_n)\}] \text{ is satisfiable}$$

Outer Skolemization generates new Skolem functions which get all the universally quantified variables as arguments the formula under consideration depends on. There might be unnecessarily many such variable dependencies, however, if the formula to be skolemized is not in anti-prenex form.[5] As an example consider the formula $\forall x, y\; [\exists z\, P(x,z) \vee Q(x,y)]$ for which standard Outer Skolemization leads to $\forall x, y\; [P(x, f(x,y)) \vee Q(x,y)]$. If preceded with an anti-prenex transformation, however, we would end up with the simpler $\forall x\, [P(x, f(x)) \vee \forall y\, Q(x,y)]$ instead.

For Inner Skolemization, Skolem functions are introduced which take only "necessary" variables as arguments. However, the Skolemization result depends on the order of the Skolemization of the sub-formulas. As an example consider the formula $\forall x\, [\exists y\, (P(x,y) \wedge \exists z\, Q(x,y,z))]$. If we first skolemize on $\exists z\, Q(x,y,z)$ then we finally end up with $\forall x\; [P(x, f(x)) \wedge Q(x, f(x), g(x, f(x)))]$ whereas starting the Skolemization with the outer existentially quantified sub-formula yields $\forall x\, [P(x, f(x)) \wedge Q(x, f(x), g(x))]$. Evidently, the latter is to be preferred. Indeed, Inner Skolemization should always be performed from left to right if we want to avoid a nesting of Skolem functions.

Note that this still does not guarantee that the least possible variable dependencies are found. For instance, consider the formula $\forall x, y\, \exists z\, [P(z) \wedge (Q(x,z) \vee R(y,z))]$ which, after Skolemization, gets transformed into $\forall x, y\, [P(f(x,y)) \wedge (Q(x, f(x,y)) \vee R(y, f(x,y)))]$ and that regardless of which of the above Skolemization possibilities is performed. If we first distributed the $\wedge$ over the $\vee$, however, we would be able to perform an anti-prenex step and therefore skolemize the formula $\forall x, y\, [\exists z\, (P(z) \wedge Q(x,z)) \vee \exists z\, (P(z) \wedge R(y,z))]$. In this case we would end up with $\forall x, y\, [(P(f(x)) \wedge Q(x, f(x))) \vee (P(g(y)) \wedge R(y, g(y)))]$. Thus we have

[5] A formula is said to be in anti-prenex form if all quantifiers are moved inwards as far as possible. There is only one exception, namely, existential quantifiers are *not* distributed over disjunctions. It is even the case that formulas of the form $\exists x\, \Phi[x] \vee \exists y\, \Psi[y]$ are transformed into $\exists x\, (\Phi[x] \vee \Psi[x])$ in order to reduce the number of existential quantifiers. The value of anti-prenexing can be of quite some significance (see, e.g., (Egly 1994, Loveland 1978)) and that regardless of whether we perform Inner or Outer Skolemization.

fewer dependencies for the price of having to deal with more Skolem functions. Whether or not this is still of some advantage varies from case to case; we shall not discuss this issue in this paper. The interested reader is referred to (Baaz and Leitsch 1994).

3 The Strong Skolemization Technique

Strong Skolemization applies to sub-formulas of the form $\exists\,\overline{x}_k\,(\Phi_1 \wedge \ldots \wedge \Phi_n)$. Unlike in the definition of standard (Inner or Outer) Skolemization this rather general form is actually needed and cannot be simulated by several successive applications on cases of the form $\exists x\ (\Phi_1 \wedge \Phi_2)$ as will be seen in the examples to come. But before we formally define what Strong Skolemization is about let us first illustrate its behavior with the help of a little example.

Let $\mathcal{M} = (\mathcal{D}, \Im, \phi)$ be an interpretation that satisfies $\exists x\,(P(x) \wedge Q(x,y))$. The value of x may vary with the value of y, but, anyhow, we know that the set associated with the predicate symbol P, $\Im(P)$, is non-empty. Now let $\widehat{h}\colon \mathcal{D} \to \mathcal{D}$ be a total function whose codomain is just $\Im(P)$. Then

$$\begin{array}{ll} \mathcal{M}[x/\widehat{h}(\alpha)] \models P(x) & \text{for every } \alpha \in \mathcal{D} \text{ and} \\ \mathcal{M}[x/\widehat{h}(\alpha)] \models Q(x,y) & \text{for some } \alpha \in \mathcal{D} \end{array}$$

Therefore there exists an interpretation $\mathcal{M}'$ such that

$$\begin{array}{ll} \mathcal{M}'[x/\alpha] \models P(h(x)) & \text{for every } \alpha \in \mathcal{D} \text{ and} \\ \mathcal{M}'[x/\alpha] \models Q(h(x),y) & \text{for some } \alpha \in \mathcal{D} \end{array}$$

where $\mathcal{M}'$ which differs from $\mathcal{M}$ only in the interpretation of the new function symbol h which is to be interpreted as $\mathcal{M}'(h) = \widehat{h}$. Hence

$$\mathcal{M}' \models \forall x\, P(h(x)) \wedge \exists x\, Q(h(x),y)$$

and we know by standard (Inner) Skolemization that there exists an interpretation $\mathcal{M}''$ which extends $\mathcal{M}'$ only by a suitable interpretation for the new function symbol g such that

$$\mathcal{M}'' \models \forall x\, P(h(x)) \wedge Q(h(g(y)),y)$$

This implies (by instantiating the variable x) that

$$\mathcal{M}'' \models \forall w\, P(h(g(w))) \wedge Q(h(g(y)),y)$$

and thus finally (with $\mathcal{M}'''(f) \stackrel{\text{Def}}{=} \mathcal{M}'''(h) \circ \mathcal{M}'''(g)$)

$$\mathcal{M}''' \models \forall w\, P(f(w)) \wedge Q(f(y),y)$$

The sub-formula $\forall w\, P(f(w)) \wedge Q(f(y),y)$ denotes the Strong Skolemization result on the original sub-formula. Notice that standard Inner Skolemization applied to the very same input would result in $P(f(y)) \wedge Q(f(y),y)$, i.e., in a formula that is subsumed by the Strong Skolemization result; Strong Skolemization is thus more general.

Definition 3.1 (free variable splitting)
Given a formula $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$ *let*

$$\begin{array}{rcl} \{\overline{z_1}_{l_1}\} & = & \mathcal{V}(\Phi_1) \setminus \{\overline{x}_k\} \\ \{\overline{z_i}_{l_i}\} & = & (\mathcal{V}(\Phi_i) \setminus \bigcup_{1 \leq j < i} \mathcal{V}(\Phi_j)) \setminus \{\overline{x}_k\} \text{ for } 1 < i \leq n \end{array}$$

We call $\langle\{\overline{z_1}_{l_1}\}, \ldots, \{\overline{z_n}_{l_n}\}\rangle$ *the* free variable splitting *of* $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$.

Intuitively, each $\{\overline{z_i}_{l_i}\}$ contains the free variables of Φ_i (without the existentially quantified $\overline{x}_k$) that did not yet occur in an "earlier" Φ_j. Such a free variable splitting has some interesting features. So, e.g., its elements are pairwise disjoint and their overall union covers exactly the free variables of the formula $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$. An important property which will be used in the sequel is the following: Whenever $\langle\{\overline{z_1}_{l_1}\}, \ldots, \{\overline{z_n}_{l_n}\}\rangle$ is a free variable splitting of $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$ then $\langle\{\overline{z_1}_{l_1}, \overline{z_2}_{l_2}\}, \{\overline{z_3}_{l_3}\}, \ldots, \{\overline{z_n}_{l_n}\}\rangle$ is a free variable splitting of $\exists\,\overline{v}_k\ (\Phi_2 \wedge \ldots \wedge \Phi_n)\{x_i / f_i(\overline{z_1}_{l_1}, \overline{v}_k)\}$ provided the variables $\overline{v}_k$ are new to $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$.

With this free variable splitting standard (Inner) Skolemization can now be described as: Replace any occurrence of a sub-formula $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$ with

$$\begin{array}{c} \Phi_1\{x_i / f_i(\overline{z_1}_{l_1}, \ldots, \overline{z_n}_{l_n})\}\ \wedge \\ \Phi_2\{x_i / f_i(\overline{z_1}_{l_1}, \ldots, \overline{z_n}_{l_n})\}\ \wedge \\ \vdots \\ \Phi_n\{x_i / f_i(\overline{z_1}_{l_1}, \ldots, \overline{z_n}_{l_n})\} \end{array}$$

where $\langle\{\overline{z_1}_{l_1}\}, \ldots, \{\overline{z_n}_{l_n}\}\rangle$ is the free variable splitting of $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$. The effect of Strong Skolemization now lies in the replacement of some of these variable occurrences with fresh universally quantified variables. The following definition specifies this.

Definition 3.2 (strong skolemization)
Let Δ *be a first-order sentence in negation normal form. Strong Skolemization replaces existentially quantified sub-formulas of the form* $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$ *with*

$$\begin{array}{rl} \forall\,\overline{w_2}_{l_2}, \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n} & \Phi_1\{x_i / f_i(\overline{z_1}_{l_1}, \overline{w_2}_{l_2}, \ldots, \overline{w_n}_{l_n})\} \wedge \\ \forall\,\overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n} & \Phi_2\{x_i / f_i(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n})\} \wedge \\ & \vdots \\ \forall\,\overline{w_n}_{l_n} & \Phi_{n-1}\{x_i / f_i(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \ldots, \overline{z_{n-1}}_{l_{n-1}}, \overline{w_n}_{l_n})\} \wedge \\ & \Phi_n\{x_i / f_i(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \ldots, \overline{z_n}_{l_n})\} \end{array}$$

where $\langle\{\overline{z_1}_{l_1}\}, \ldots, \{\overline{z_n}_{l_n}\}\rangle$ *is the free variable splitting of* $\exists\,\overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$ *and the* $f_i, 1 \leq i \leq k$ *are new (Skolem) function symbols.*

Note that, in the worst case, each $l_2, \ldots, l_n$ equals 0, so that Strong Skolemization degenerates to standard Inner Skolemization.

Coming back to the example where we skolemized the sub-formula $\exists x\ (P(x) \wedge Q(x, y))$ we see that $\langle\{\}, \{y\}\rangle$ is the corresponding free variable splitting, that

$n = 2$, $l_1 = 0$, $l_2 = 1$, $\overline{x_2}_{l_2} = y$, and $k = 1$, and thus the Strong Skolemization results in (with $\overline{w_2}_1 = w$) $\forall w\ P(f(w)) \wedge Q(f(y), y)$.

Let us now come to the proof that Strong Skolemization behaves as desired, i.e., that it preserves both satisfiability and unsatisfiability. To this end we make use of some preliminary definitions as well as an important key lemma.

Definition 3.3 (selector functions, projections)
We call any total function $h : \mathcal{D}^k \to \mathcal{D}^k$ a selector function *for a set $\mathcal{S} \subseteq \mathcal{D}^k$ if codomain$(h) = \mathcal{S}$ provided $\mathcal{S} \neq \emptyset$. In case $\mathcal{S} = \emptyset$, h may be arbitrary.*
As usual, the projection functions *π_i allow us to refer to the i^{th} element of a given sequence, i.e., $\pi_i(e_1, \ldots, e_i, \ldots, e_n) = e_i$.*

Such selector functions evidently do exist. For example just take

$$h(\overline{x}_k) = \begin{cases} \overline{x}_k & \text{if } \overline{x}_k \in \mathcal{S} \\ \overline{a}_k & \text{for some arbitrary } \overline{a}_k \in \mathcal{S} \text{ if } \overline{x}_k \notin \mathcal{S} \text{ and } \mathcal{S} \neq \emptyset \\ \overline{b}_k & \text{for some arbitrary } \overline{b}_k \in \mathcal{D}^k \text{ if } \mathcal{S} = \emptyset \end{cases}$$

Lemma 3.4
If $\mathcal{M} \models \exists \overline{x}_k\ (\Phi \wedge \Psi)$ then $\mathcal{M}' \models \forall \overline{w}_k\ \Phi\{x_i/h_i(\overline{y}_n, \overline{w}_k)\} \wedge \exists \overline{v}_k\ \Psi\{x_i/h_i(\overline{y}_n, \overline{v}_k)\}$, where $\{\overline{y}_n\} = \mathcal{V}(\Phi) \setminus \{\overline{x}_k\}$ and $\mathcal{M}'$ extends $\mathcal{M}$ only by a suitable interpretation of the new function symbols $h_1, \ldots, h_k$.

Proof: Let $\mathcal{M} = (\mathcal{D}, \Im, \phi)$ and define

$$\mathbf{S}^\Phi(\overline{\beta}_n) = \{\overline{\alpha}_k \mid (\mathcal{D}, \Im, [\overline{y}_n/\overline{\beta}_n, \overline{x}_k/\overline{\alpha}_k]) \models \Phi\}$$

Since $\mathcal{M} \models \exists \overline{x}_k\ (\Phi \wedge \Psi)$, we know that $\mathbf{S}^\Phi(\phi(\overline{y}_n)) \neq \emptyset$, where $\phi(\overline{y}_n)$ is short for the sequence $\phi(y_1), \ldots, \phi(y_n)$. Thus we know that

$$\begin{array}{ll} \mathcal{M}[\overline{x}_k/\overline{\alpha}_k] \models \Phi & \text{for all } \overline{\alpha}_k \in \mathbf{S}^\Phi(\phi(\overline{y}_n)) \text{ and} \\ \mathcal{M}[\overline{x}_k/\overline{\alpha}_k] \models \Psi & \text{for some } \overline{\alpha}_k \in \mathbf{S}^\Phi(\phi(\overline{y}_n)) \end{array}$$

Now, for arbitrary $\overline{\gamma}_n \in \mathcal{D}^n$, let $h^\Phi(\overline{\gamma}_n) \colon \mathcal{D}^k \to \mathcal{D}^k$ be a *selector function* for $\mathbf{S}^\Phi(\overline{\gamma}_n) \subseteq \mathcal{D}^k$. Then

$$\begin{array}{ll} \mathcal{M}[\overline{x}_k/h^\Phi(\phi(\overline{y}_n))(\overline{\alpha}_k)] \models \Phi & \text{for all } \overline{\alpha}_k \in \mathcal{D}^k \text{ and} \\ \mathcal{M}[\overline{x}_k/h^\Phi(\phi(\overline{y}_n))(\overline{\alpha}_k)] \models \Psi & \text{for some } \overline{\alpha}_k \in \mathcal{D}^k \end{array}$$

$h^\Phi(\phi(\overline{y}_n))$ returns sequences of domain elements. Thus, in order to represent a substitute for each of the x_i, we have to refer to the respective projections on these sequences, i.e.

$$\begin{array}{ll} \mathcal{M}[\ldots, x_i/\pi_i(h^\Phi(\phi(\overline{y}_n))(\overline{\alpha}_k)), \ldots] \models \Phi & \text{for all } \overline{\alpha}_k \in \mathcal{D}^k \text{ and} \\ \mathcal{M}[\ldots, x_i/\pi_i(h^\Phi(\phi(\overline{y}_n))(\overline{\alpha}_k)), \ldots] \models \Psi & \text{for some } \overline{\alpha}_k \in \mathcal{D}^k \end{array}$$

This immediately leads to a suitable interpretation for the new function symbols $h_1, \ldots, h_k$ which extend $\mathcal{M}$ to $\mathcal{M}'$ by

$$\mathcal{M}'(h_i)(\overline{\beta}_n, \overline{\alpha}_k) = \pi_i(h^\Phi(\overline{\beta}_n)(\overline{\alpha}_k))$$

With this we get

$$\begin{array}{ll} \mathcal{M}'[\overline{w}_k/\overline{\alpha}_k] \models \Phi\{x_i/h_i(\overline{y}_n, \overline{w}_k)\} & \text{for all } \overline{\alpha}_k \in \mathcal{D}^k \text{ and} \\ \mathcal{M}'[\overline{v}_k/\overline{\alpha}_k] \models \Psi\{x_i/h_i(\overline{y}_n, \overline{v}_k)\} & \text{for some } \overline{\alpha}_k \in \mathcal{D}^k \end{array}$$

which finally can be simplified to

$$\mathcal{M}' \models \forall \overline{w}_k \ \Phi\{x_i/h_i(\overline{y}_n, \overline{w}_k)\} \wedge \exists \overline{v}_k \ \Psi\{x_i/h_i(\overline{y}_n, \overline{v}_k)\}$$

and this completes the proof. □

THEOREM 3.5
Strong Skolemization preserves both satisfiability and unsatisfiability.
Proof: Unsatisfiability preservation follows immediately from the fact that Strong Skolemization generalizes standard Skolemization.
For the other direction assume that

$$\mathcal{M} \models \Delta[\exists \overline{x}_k \ (\Phi_1 \wedge \ldots \wedge \Phi_n)]$$

where Δ is a formula in negation normal form. We prove by induction on the structure of Δ that there exists an interpretation $\mathcal{M}'$ that differs from $\mathcal{M}$ only in a suitable interpretation for the new Skolem function symbols $f_1, \ldots, f_k$ and that satisfies the Strong Skolemization result on the indicated sub-formula.
The base case (where $\Delta = \exists \overline{x}_k \ (\Phi_1 \wedge \ldots \wedge \Phi_n)$) is in fact the crucial part of the proof and requires another induction; this time on the number n.
For the case $n = 1$ Strong Skolemization does not differ from standard Inner Skolemization and we are done. For the induction step assume that $\mathcal{M} \models \exists \overline{x}_k \ (\Phi_1 \wedge \ldots \wedge \Phi_n)$. Then, by Lemma 3.4, we know that there exists an interpretation $\mathcal{M}'$ such that

$$\mathcal{M}' \models \forall \overline{w}_k \ \Phi_1\{x_i/h_i(\overline{z_1}_{l_1}, \overline{w}_k)\} \wedge \exists \overline{v}_k \ (\Phi_2 \wedge \ldots \wedge \Phi_n)\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}$$

where $\mathcal{M}'$ is like $\mathcal{M}$ except for the additional interpretation for the new function symbols h_i. Now, in order to apply the induction hypothesis on $\exists \overline{v}_k \ (\Phi_2 \wedge \ldots \wedge \Phi_n)\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}$ recall that the corresponding free variable splitting is given by $\langle\{\overline{z_1}_{l_1}, \overline{z_2}_{l_2}\}, \{\overline{z_3}_{l_3}\}, \ldots, \{\overline{z_n}_{l_n}\}\rangle$. Thus, an application of Strong Skolemization leads to the satisfiability of

$$\begin{array}{rl} \forall \overline{w}_k & \Phi_1\{x_i/h_i(\overline{z_1}_{l_1}, \overline{w}_k)\} \wedge \\ \forall \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n} & \Phi_2\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n})\} \wedge \\ \forall \overline{w_4}_{l_4}, \ldots, \overline{w_n}_{l_n} & \Phi_3\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \ldots, \overline{z_3}_{l_3}, \overline{w_4}_{l_4}, \ldots, \overline{w_n}_{l_n})\} \wedge \\ & \vdots \\ \forall \overline{w_n}_{l_n} & \Phi_{n-1}\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \ldots, \overline{z_{n-1}}_{l_{n-1}}, \overline{w_n}_{l_n})\} \wedge \\ & \Phi_n\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \ldots, \overline{z_n}_{l_n})\} \end{array}$$

Instantiating the variables w_i in the first conjunct with $g_i(\overline{z_1}_{l_1}, \overline{u_2}_{l_2}, \ldots, \overline{u_n}_{l_n})$ – where the elements in each $\overline{u_j}_{l_j}$ are fresh universally quantified variables – then reveals the satisfiability of

$$\begin{array}{rl}
\forall \overline{u_2}_{l_2}, \ldots, \overline{u_n}_{l_n} & \Phi_1\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \overline{u_2}_{l_2}, \ldots, \overline{u_n}_{l_n})\} \wedge \\
\forall \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n} & \Phi_2\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n})\} \wedge \\
& \vdots \\
\forall \overline{w_n}_{l_n} & \Phi_{n-1}\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \ldots, \overline{z_{n-1}}_{l_{n-1}}, \overline{w_n}_{l_n})\} \wedge \\
& \Phi_n\{x_i/h_i(\overline{z_1}_{l_1}, \overline{v}_k)\}\{v_j/g_j(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \ldots, \overline{z_n}_{l_n})\}
\end{array}$$

Now let $f_i(\overline{u_1}_{l_1}, \ldots, \overline{u_n}_{l_n}) = h_i(\overline{u_1}_{l_1}, g_1(\overline{u_1}_{l_1}, \ldots, \overline{u_n}_{l_n}), \ldots, g_k(\overline{u_1}_{l_1}, \ldots, \overline{u_n}_{l_n}))$ for arbitrary $\overline{u_1}_{l_1}, \ldots, \overline{u_n}_{l_n}$. This then leads to the satisfiability of

$$\begin{array}{rl}
\forall \overline{w_2}_{l_2}, \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n} & \Phi_1\{x_i/f_i(\overline{z_1}_{l_1}, \overline{w_2}_{l_2}, \ldots, \overline{w_n}_{l_n})\} \wedge \\
\forall \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n} & \Phi_2\{x_i/f_i(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \overline{w_3}_{l_3}, \ldots, \overline{w_n}_{l_n})\} \wedge \\
& \vdots \\
\forall \overline{w_n}_{l_n} & \Phi_{n-1}\{x_i/f_i(\overline{z_1}_{l_1}, \ldots, \overline{z_{n-1}}_{l_{n-1}}, \overline{w_n}_{l_n})\} \wedge \\
& \Phi_n\{x_i/f_i(\overline{z_1}_{l_1}, \overline{z_2}_{l_2}, \ldots, \overline{z_n}_{l_n})\}
\end{array}$$

and we are done with the base case of the outer induction proof.
For the induction step let us assume that $\Delta = \Delta_1 \wedge \Delta_2[\exists\, \overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)]$. Then $\mathcal{M} \models \Delta_1$ and $\mathcal{M} \models \Delta_2[\exists\, \overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)]$. By the induction hypothesis this means that $\mathcal{M} \models \Delta_1$ and there exists an $\mathcal{M}'$ which extends $\mathcal{M}$ only by a suitable interpretation for the function symbols f_i and which satisfies the Strong Skolemization result. Since in Δ_1 there is no mention at all of these function symbols f_i we also know that $\mathcal{M}' \models \Delta_1$ and thus $\mathcal{M}'$ is a model for the altered Δ as well.
The cases where Δ is a disjunction, an existential or a universal quantification work analogously so that the respective proofs can be omitted. Note that Δ cannot be of the form $\neg\Delta_1[\exists\, \overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)]$ for otherwise it would not be in negation normal form. □

4 Observations and Examples

Strong Skolemization can always be applied whenever standard Skolemization can. However, it does not necessarily produce a different result. For example, in the trivial case $n = 1$ Strong Skolemization evidently degenerates to Inner Skolemization. Also, if $\mathcal{V}(\Phi_1) = \mathcal{V}(\Phi_2) = \cdots = \mathcal{V}(\Phi_n)$ it is evident that the corresponding variable splitting is just of the form $\langle\{\overline{x}\}, \emptyset, \ldots, \emptyset\rangle$ and therefore again there is no difference to the corresponding Inner Skolemization result. As a matter of fact, it turns out that the order in which the $\Phi_1, \ldots, \Phi_n$ occur plays a crucial rôle for the final Strong Skolemization outcome. In order to obtain an optimal result it ought to be ensured that for no $i < j$: $\mathcal{V}(\Phi_i) \supset \mathcal{V}(\Phi_j)$. If necessary, the Φ_i should be rearranged accordingly. In the best case (viewed from the Strong Skolemization perspective) it is possible to rearrange the Φ_i

in a way such that $\mathcal{V}(\Phi_1) \subset \mathcal{V}(\Phi_2) \subset \cdots \subset \mathcal{V}(\Phi_n)$ for then each of the l_j (see Definition 3.2) is different from 0 and so a maximal number of variable abstractions takes place. The little example we examined in the last section was just of this kind. This leads to the first important observation:

Observation 4.1
Given $\exists \overline{x}_k \, (\Phi_1 \wedge \ldots \wedge \Phi_n)$, if at all possible, it should be tried to rearrange the inner conjunctions as $\exists \overline{x}_k \, (\Phi_{i_1} \wedge \ldots \wedge \Phi_{i_n})$ such that $\mathcal{V}(\Phi_{i_1}) \subseteq \cdots \subseteq \mathcal{V}(\Phi_{i_n})$.

Example 4.2
Consider the formula

$$\forall x \, [C(x) \vee \forall y \, \exists z \, (B(x,y,z) \wedge A(y,z))] \tag{1}$$

Since $\mathcal{V}(B(x,y,z)) \supset \mathcal{V}(A(y,z))$ we should exchange the two conjuncts and therefore consider the formula $\forall x \, [C(x) \vee \forall y \, \exists z \, (A(y,z) \wedge B(x,y,z))]$ instead. The corresponding free variable splitting then is $\langle \{y\}, \{x\} \rangle$ and thus the Strong Skolemization results in $\forall x \, [C(x) \vee \forall y \; (\forall w \; (A(y, f(y,w)) \wedge B(x,y,f(y,x))))]$ and thus eventually to the clause set

$$\begin{array}{l} C(x) \vee A(y, f(y,w)) \\ C(x) \vee B(x,y,f(y,x)) \end{array}$$

Note that an immediate Strong Skolemization step on (1) produces the same result as Inner Skolemization because then the free variable splitting is $\langle \{x,y\}, \emptyset \rangle$.

Such a suitable ordering of the conjuncts cannot always be achieved as shown in the following.

Example 4.3
Consider the formula

$$\forall x \, [C(x) \vee \forall y \, \exists z \, (A(y,z) \wedge B(x,z))] \tag{2}$$

Unfortunately, we cannot follow the above recommendation because there is no subsort relation between the free variables of the two conjuncts. Let us therefore consider both possibilities. Then the free variable splitting we are faced with is either $\langle \{y\}, \{x\} \rangle$, or $\langle \{x\}, \{y\} \rangle$. Now, in order to find out whether one is still to be preferred over the other let us have a look at the two (Strong) Skolemization results. In the first case we get (in clause form)

$$\begin{array}{l} C(x) \vee A(y, f(w,y)) \\ C(x) \vee B(x, f(x,y)) \end{array}$$

whereas in the second case we obtain

$$\begin{array}{l} C(x) \vee B(x, f(x,w)) \\ C(x) \vee A(y, f(x,y)) \end{array}$$

We can immediately see that for the second case we merely achieved a variant of the standard Skolemization outcome. For the first case, however, a real generalization takes place.
A little remark: the first case clauses would have been obtained from standard Skolemization applied to the formula $\forall x, y\ \exists z\ ((\forall u\ C(u) \vee A(y,z)) \wedge (C(x) \vee B(x,z)))$. The reader may check by her/himself that this is indeed logically equivalent to formula (2).

The above example shows that the "correct" order of the conjuncts can make a difference even if the conjuncts cannot be ordered by way of comparing their respective free variables. Hence, we can suppose that there are a further criteria which might help us to find a more complete ordering of the conjuncts.
Notice that in the above example the A-literal has no mention of the variable x and the B-literal has no mention of the variable y and that Δ^y is a sub-formula of Δ^x. This is no coincidence. The conjunct $A(y,z)$ should be considered first because the quantifier of the variable y is "closer" to $A(y,z)$ than the quantifier of the variable x to $B(x,z)$. There is thus a better chance that the other literals that possibly occur in the final clausification result in the same clause have no mention of y at all and so the possibility increases that some clauses consist of variable-disjoint sub-clauses.

Observation 4.4
Let Δ be a closed formula in negation normal form (with all universal quantifications moved inwards as far as possible)[6], and let $\exists \overline{x}_k\ (\Phi_1 \wedge \ldots \wedge \Phi_n)$ be a sub-formula of Δ (for simplicity let us assume that this sub-formula lies not within the scope of another existential quantifier). In case $\mathcal{V}(\Phi_i)$ and $\mathcal{V}(\Phi_j)$ are incomparable for some i, j then consider Φ_i first provided there exists a $v \in \mathcal{V}(\Phi_j)$ such that for all $w \in \mathcal{V}(\Phi_i)$: Δ^w is a sub-formula of Δ^v.

Evidently, the observations from above do not provide us with a total ordering on sub-formulas. Such a total ordering is in fact not possible as the following example illustrates.

Example 4.5
Consider the formula

$$\forall x, y\, [C(x,y) \vee \exists z\, (A(y,z) \wedge B(x,z))] \tag{3}$$

Unfortunately it is not possible to follow the recommendation of Observation 4.4 because there is no "nearer" quantification. We therefore obtain either

$$\begin{array}{l} C(x,y) \vee A(y, f(w,y)) \\ C(x,y) \vee B(x, f(x,y)) \end{array}$$

or, alternatively,

$$\begin{array}{l} C(x,y) \vee B(x, f(x,w)) \\ C(x,y) \vee A(y, f(x,y)) \end{array}$$

[6] This restriction is not necessary; however, it may help as Example 4.3 shows.

There is no way to decide which might be more valuable. This depends on the other clauses of the whole problem under consideration and not on this formula alone. For instance, if the rest of the problem implies that $\forall y\, \neg C(t, y)$ for some term t then the first possibility should be preferred. If, however, the rest of the problem implies $\forall x\, \neg C(x, t)$ for some term t then the second possibility seems more valuable. But note: we cannot choose a "wrong" one; both are more general than the corresponding standard Skolemization result.

The above example is not artificial. As an interesting instance consider the so-called "density property" of some binary relation R, i.e.,

$$\forall x, y\, [R(x, y) \supset \exists z\, (R(x, z) \wedge R(z, y))] \tag{4}$$

Here, as in the previous example, Strong Skolemization ends up with either

$$\begin{array}{c} \neg R(x, y) \vee R(x, f(x, w)) \\ \neg R(x, y) \vee R(f(x, y), y) \end{array} \quad \text{or} \quad \begin{array}{c} \neg R(x, y) \vee R(x, g(x, y)) \\ \neg R(x, y) \vee R(g(w, y), y) \end{array}$$

The former is particularly interesting if R can be proven to be forward serial, whereas the latter is more important in case backward seriality of R holds. Note that the results achieved from Strong Skolemization are identical to the application of standard Skolemization on

$$\forall x, y\, \exists z\, [\forall w\, (R(x, w) \supset R(x, z)) \wedge (R(x, y) \supset R(z, y))] \tag{5}$$

and

$$\forall x, y\, \exists z\, [(R(x, y) \supset R(x, z)) \wedge \forall w\, (R(w, y) \supset R(z, y))] \tag{6}$$

respectively. I leave it to the reader to verify that (4), (5), and (6) are pairwise equivalent.

The above examples show that the Strong Skolemization outcome can depend on the choice of the conjunct. Quite often, however, this choice turns out to be irrelevant because the possible cases are somewhat symmetric. One such example can be found in the following.

Example 4.6
Consider the formula

$$\forall x, y, z\, [(R(x, y) \wedge R(x, z)) \supset \exists u\, (R(y, u) \wedge R(z, u))] \tag{7}$$

which expresses something like a "confluence" property of the relation R. Standard Skolemization (plus clausification) results in

$$\begin{array}{c} \neg R(x, y) \vee \neg R(x, z) \vee R(y, f(y, z)) \\ \neg R(x, y) \vee \neg R(x, z) \vee R(z, f(y, z)) \end{array}$$

whereas Strong Skolemization yields

$$\begin{array}{c} \neg R(x, y) \vee \neg R(x, z) \vee R(y, f(y, w)) \\ \neg R(x, y) \vee \neg R(x, z) \vee R(z, f(y, z)) \end{array}$$

Here, condensing is possible on the literals $\neg R(x, y)$ and $\neg R(x, z)$ with substitution $\sigma = \{z/y\}$ and so we end up with

$$\begin{array}{l} \neg R(x,y) \vee R(y, f(y,w)) \\ \neg R(x,y) \vee \neg R(x,z) \vee R(z, f(y,z)) \end{array}$$

Similarly, if we exchange the two conjuncts, Strong Skolemization leads to (again after condensing)

$$\begin{array}{l} \neg R(x,y) \vee \neg R(x,z) \vee R(y, g(y,z)) \\ \neg R(x,z) \vee R(z, g(w,z)) \end{array}$$

In contrast to the earlier examples the two outcomes are mere variants of each other. Nevertheless, both subsume the standard Skolemization result.

The above examples are all of the form $\exists x\, (\Phi \wedge \Psi)$. However, the definition of Strong Skolemization also covers the more general cases in which there is more than just one existentially quantified variable and more than just two conjuncts. Such more general cases usually cannot be simulated by a sequence of Strong Skolemization steps of the above kind. It is even the case that it is sometimes preferable to avoid anti-prenexing as the following example shows.

Example 4.7
Let

$$\Delta = \forall x\, [A(x) \vee \exists u, v\, (B(u) \wedge C(u,v) \wedge D(x,u,v))] \tag{8}$$

Standard Skolemization then results in

$$\forall x\, [A(x) \vee (B(f(x)) \wedge C(f(x), g(x)) \wedge D(x, f(x), g(x)))]$$

whereas Strong Skolemization leads to

$$\forall x\, [A(x) \vee (\forall w\, B(f(w)) \wedge \forall w\, C(f(w), g(w)) \wedge D(x, f(x), g(x)))]$$

If we first performed an anti-prenexing on the input formula, however, we would have to deal with

$$\forall x\, [A(x) \vee \exists u\, (B(u) \wedge \exists v\, (C(u,v) \wedge D(x,u,v)))]$$

and thus obtained by Strong Skolemization

$$\forall x\, [A(x) \vee (\forall w\, B(f(w)) \wedge C(f(x), g(x)) \wedge D(x, f(x), g(x)))]$$

which, although being more general than the standard Skolemization result, is less general than what we would obtain when we did without an initial anti-prenexing.

Observation 4.8
Anti-prenex transformations cannot absolutely be recommend. Rather, existential quantifiers should be moved outwards over conjunctions in order to obtain more general results.

In the worst case Strong Skolemization behaves like standard Inner Skolemization. There are, however, sub-languages of first-order predicate logic for which Strong Skolemization is necessarily superior to standard Skolemization techniques. For instance, in the translation of modal or terminological logic formulas (like $\mathcal{ALC}$; see the paper by Schmidt-Schauß and Smolka (1991)) we have to deal with certain syntactic restrictions in the translation result which fit perfectly well to Strong Skolemization. I do not want to go into detail what these languages is concerned. What is important for our purpose, however, is that after performing a so-called relational translation into first-order predicate logic we are faced with formulas that contain sub-formulas of the kind $\exists x\,[R(y,x) \wedge \Phi[x]]$ where R is a binary predicate symbol and $\Phi[x]$ is a representative of a more or less complicated formula with x as its only free variable. Here Strong Skolemization nicely comes into play. Why this is so is justified by the following observation.

Observation 4.9
Let Δ be a sentence with sub-formula $\exists x\,(\Phi \wedge \Psi)$ such that x is the only free variable in Φ and let $\{y_1,\ldots,y_n\} = \mathcal{V}(\Psi) \setminus \{x\}$. Then Strong Skolemization transforms $\Delta[\exists x\,(\Phi \wedge \Psi)]$ into

$$\Delta[\exists x\,(\Phi \wedge \Psi)\,/\,\forall\,\overline{w}_n\,\Phi\{x/f(\overline{w}_n)\} \wedge \Psi\{x/f(\overline{y}_n)\}]$$

where f is a new $n-$place function symbol.
Note that the sub-formula $\forall\,\overline{w}_n\,\Phi\{x/f(\overline{w}_n)\}$ obtained this way has no free variables at all. This is of particular importance after clause normal form generation takes place, for it allows us to split clauses into variable disjoint sub-clauses and thus splits the reasoning process into several independent branches.

In running a huge $\mathcal{ALC}$ database we have found that in about half of the examples there was a significantly better performance after Strong Skolemization than after standard Inner Skolemization. In approximately one fourth of all examples Strong Skolemization made it at all possible for the theorem prover to stop its reasoning process, be it with a proof or with a saturated set of clauses.[7]

Such gains have been proved true by testing other collections of first-order theorems as, e.g., in (Pelletier 1986). For instance, the first-order presentation of the (in)famous "Pigeon-Hole Problems" allows only for one single application of Strong Skolemization, namely at the sub-formula

$$\forall x\ \mathrm{Object}(x) \rightarrow \exists y\,(\mathrm{Hole}(y) \wedge \mathrm{In}(x,y))$$

Standard Skolemization leads to $\forall x\ \mathrm{Object}(x) \rightarrow (\mathrm{Hole}(f(x)) \wedge \mathrm{In}(x,f(x)))$ whereas Strong Skolemization yields the more general

$$\forall x\ \mathrm{Object}(x) \rightarrow (\forall w\,\mathrm{Hole}(f(w)) \wedge \mathrm{In}(x,f(x)))$$

[7]Evidently, there are specially designed inference mechanisms and translation techniques whose performance is even better. It is certainly not our aim to compete with these methods; our aim was just to illustrate the superiority of Strong Skolemization over standard Skolemization.

Depending on the theorem prover at hand this generalization gives rise to a time gain of 10-40%.

Summarizing: Standard Inner Skolemization and Strong Skolemization can best be compared by having a closer look at the respective final clausification results. It turns out that the structure of the clauses is preserved; also the number of clauses and the number of literals within clauses remains as in the standard case. The difference lies in the abstraction from some of the variables in Skolem functions resulting in new universally quantified variables. This abstraction can help to simplify and/or to split clauses without jeopardizing soundness and completeness of the underlying calculus. And, moreover, its costs are essentially the costs of standard Skolemization; there is no high price to be paid for the gain.

5 Related Work

There is a somewhat related technique to Strong Skolemization, the so-called Optimized Skolemization, that has been developed by Christoph Weidenbach and Hans Jürgen Ohlbach in (1995). Slightly generalizing, their approach can briefly be described as follows:

DEFINITION 5.1 ((GENERALIZED) OPTIMIZED SKOLEMIZATION)
Let $\Delta[\exists\,\overline{x}_k\,(\Phi \wedge \Psi)]$ *be a sentence in negation normal form and let us assume that* $\Delta \models \forall\,\overline{y}_n\,\exists\,\overline{x}_k\,\Phi$ *where* $\{y_1,\ldots,y_n\} = \mathcal{V}(\exists\,\overline{x}_k\,(\Phi \wedge \Psi))$. *Then*

$$\Delta \text{ is satisfiable}$$
$$\text{iff}$$
$$\forall\,\overline{y}_n\,\Phi\{x_i/f_i(\overline{y}_n)\}\ \wedge \Delta[\exists\,\overline{x}_k\,(\Phi \wedge \Psi)/\Psi\{x_i/f_i(\overline{y}_n)\}] \text{ is satisfiable}$$

where $f_1,\ldots,f_k$ *are new (Skolem) function symbols.*

Note that Optimized Skolemization involves a theorem prover on its own to prove the intermediate goal $\Delta \models \forall\,\overline{y}_n\,\exists\,\overline{x}_k\,\Phi$. Since this can be arbitrarily complicated, an implementation of Optimized Skolemization has to somehow restrict the resources for this intermediate theorem prover in order to eventually come to a decision (Weidenbach, Gaede and Rock 1996).

As an example consider again the density property for R, i.e., $\forall u,v\,[R(u,v) \supset \exists w\,(R(u,w) \wedge R(w,v))]$. In case we know (or can easily prove) that the binary relation R is serial, i.e., $\forall x \exists y\, R(x,y)$, the Optimized Skolemization of the density property results in the clauses

$$R(x,f(x,y))$$
$$\neg R(u,v) \vee R(f(u,v),v)$$

The very same two clauses could be obtained by standard Skolemization from

$$\forall x,y\,\exists z\,[R(x,z) \wedge (R(x,y) \supset R(z,y))] \tag{9}$$

I leave it to the reader to verify that the formula (9) is logically equivalent to the conjunction of seriality and density for R.

Now let us compare this with the method proposed in this paper. Recall that Strong Skolemization finally results in the clauses

$$\begin{array}{l}\neg R(u,v) \vee R(u,f(u,w)) \\ \neg R(u,v) \vee R(f(u,v),v)\end{array}$$

Moreover, suppose that there is a further clause, say $R(x,g(x))$, which states the seriality of R. Then a resolution step with the first clause from above yields the unit clause $R(u,f(u,w))$ which even subsumes its two-literal parent clause. Thus, we finally end up with the same result as the Optimized Skolemization, i.e., Optimized Skolemization has in a sense been simulated by Strong Skolemization and Resolution.

The above shows some similarity between Strong Skolemization and Optimized Skolemization. However, the similarities are not always as strong as the above example might suggest.

Consider again formula (7) and assume that we are able to prove that $\forall y \exists u\, R(y,u)$, i.e., R is serial. Optimized Skolemization then results in the clauses

$$\begin{array}{l}R(y,f(y,z)) \\ \neg R(x,y) \vee \neg R(x,z) \vee R(z,f(y,z))\end{array}$$

whereas Strong Skolemization on $R(y,u)$ leads to

$$\begin{array}{l}\neg R(x,y) \vee R(y,f(y,w)) \\ \neg R(x,y) \vee \neg R(x,z) \vee R(z,f(y,z))\end{array}$$

Note that – in contrast to the "density example" – the seriality of R does not help here to entirely simulate Optimized Skolemization, although, somewhat surprisingly, backward seriality would; and, after all, seriality follows from backward seriality in the example under consideration (although not the other way round). A major reason for this is, that Optimized Skolemization takes the whole problem into account and not only the sub-problem in which the sub-formula to be skolemized occurs. In other words, Optimized Skolemization works globally on the problem under consideration whereas Strong Skolemization applies only locally to the various sub-problems. This might lead to the conclusion that, if Optimized Skolemization applies at all, then it produces a more general result than Strong Skolemization and thus in particular a more general result than standard Skolemization. This, however, is not quite true, for, consider again formula (8)

$$\forall x\, [A(x) \vee \exists y,z\, (B(y) \wedge C(y,z) \wedge D(x,y,z))]$$

There are three possible entry points for Optimized Skolemizations to start, one for each conjunct. I.e., it is tried to prove from the rest of the problem that either of the following holds: $\exists y\, B(y)$, $\exists y,z\, C(y,z)$, or, $\forall x \exists y,z\, D(x,y,z)$. If none can be proved then Optimized Skolemization returns the very same as standard Inner Skolemization. Otherwise it comes up with one the following clause sets

$$\begin{array}{lll} B(f(x)) & A(x) \vee B(f(x)) & A(x) \vee B(f(x)) \\ A(x) \vee C(f(x), g(x)) & C(f(x), g(x)) & A(x) \vee C(f(x), g(x)) \\ A(x) \vee D(x, f(x), g(x)) & A(x) \vee D(x, f(x), g(x)) & D(x, f(x), g(x)) \end{array}$$

Now let us compare these with the corresponding Strong Skolemization result, i.e., with the clauses

$$\begin{array}{l} A(x) \vee B(f(w)) \\ A(x) \vee C(f(w), g(w)) \\ A(x) \vee D(x, f(x), g(x)) \end{array}$$

The first and the second of the above three clauses can be split into two variable disjoint sub-clauses respectively. This leaves us with two branches of the reasoning process, namely one which contains merely the unit clause $A(x)$ and one which contains the clauses

$$\begin{array}{l} B(f(x)) \\ C(f(x), g(x)) \\ A(x) \vee D(x, f(x), g(x)) \end{array}$$

This clause set subsumes the first two possibilities from Optimized Skolemization. Only for the third possibility we cannot seriously make a prediction: It depends on the rest of the problem under consideration. At least the supposed superiority of Optimized Skolemization compared to Strong Skolemization begins to falter.

Now, one might object that we could apply Optimized Skolemization by checking a more complex conjunct first, namely $\exists y, z\,(A(y) \wedge B(y, z))$. After all, if this turns out to be provable from the rest of the problem then Optimized Skolemization would end up in exactly the same branch as Strong Skolemization, and that without having to consider the other branch as well, thus again showing its superiority. Fair enough, but we have to bear in mind that Optimized Skolemization involves theorems to be proven (automatically). And the more intermediate theorems there are, the more effort is to be spent for the clausification procedure alone. Thus, in order to reduce this effort, real implementations of the Optimized Skolemization method (see (Weidenbach et al. 1996)) avoid considering all possible combinations of conjuncts; rather they consider one conjunct after the other and as soon as one attempt turns out to be successful they take this one as *the* candidate. In the example above, for instance, this means that Optimized Skolemization performs at most three intermediate proofs, whereas in case we checked every combination there were many more such attempts.

Summarizing: In comparing standard Skolemization with Optimized Skolemization on the one hand and Strong Skolemization on the other hand we can see that Optimized Skolemization essentially cuts off literals from clauses of the standard clausification result whereas Strong Skolemization replaces some of the variables by new ones. It is not always clear which of the two is to be preferred, although it seems that in practice Optimized Skolemization comes up with slightly better results. The price to be paid for this, however, is that Optimized Skolemization requires a theorem prover on its own in order to prove its intermediate lemmas. If these proof attempts fail, be it because the goal cannot be proven or the necessary resource limitations on the prover prevent us to find a proof, then Optimized Skolemization performs a standard Inner Skolemization step and therefore shows no advantage at all in such cases. Latest at this stage, if not earlier simply because the implementer of the clause normal form generator is not willing to include a full theorem prover into his/her algorithm, Strong Skolemization can come into play. In fact, it can even be imagined that it would make sense to perform both Optimized and Strong Skolemization and to compare the respective results. Whatever seems more valuable should be chosen, and this could vary from case to case, i.e., from sub-formula to sub-formula.

It might even be possible to combine the two techniques. How this could be done is still an open question and requires some further work.

References

Andrews, P. B.: 1981, Theorem Proving via General Mating, *Journal of the Association for Computing Machinery* **28**(2), 193–214.

Baaz, M. and Leitsch, A.: 1994, On Skolemization and Proof Complexity, *Fundamenta Informaticae* **20**, 353–379.

Chang, C.-L. and Lee, R. C.: 1973, *Symbolic Logic and Mechanical Theorem Proving*, Computer Science and Applied Mathematics Series, Academic Press, New York.

Egly, U.: 1994, On the Value of Antiprenexing, *Proceedings of the LPAR'94*, Springer, LNAI 822.

Loveland, D.: 1978, *Automated Theorem Proving: A Logical Basis*, Fundamental Studies in Computer Science, North-Holland.

Nonnengart, A.: 1996, Strong Skolemization, *Technical Report MPI-I-96-2-010*, Max-Planck-Institut für Informatik, Saarbrücken, Germany. http://www.mpi-sb.mpg.de/~nonnenga/publications/.

Ohlbach, H. J. and Weidenbach, C.: 1995, A Note on Assumptions about Skolem Functions, *Journal of Automated Reasoning* **15**(2), 267–275.

Pelletier, F. J.: 1986, Seventy-Five Problems for Testing Automatc Theorem Provers, *Journal of Automated Reasoning* **2**(2), 191–216.

Schmidt-Schauß, M. and Smolka, G.: 1991, Attributive Concept Description with Complements, *Artificial Intelligence* **48**, 1–26.

Skolem, T.: 1920, Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze, nebst einem Theoreme über dichte Mengen, *Skrifter utgit av Videnskappsellkapet i Kristiania* **4**, 4–36. see also: (Skolem 1976).

Skolem, T.: 1976, Logico-combinatorial Investigations in the Satisfiability and Provability of Mathematical Propositions, *in* J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931*, Harvard University Press.

Weidenbach, C., Gaede, B. and Rock, G.: 1996, SPASS & FLOTTER Version 0.42, *Proceedings of the 13th International Conference on Automated Deduction, CADE 13*, Springer, LNAI 1104, pp. 141–145.