Translation of Resolution Proofs into Short First-Order Proofs without Choice Axioms

Hans de Nivelle

Max Planck Institut für Informatik, Saarbrücken, Germany

14.01.2004

Abstract

I present a way of transforming a resolution proof on the first-order level (The CNF-transformation can contain Skolemization, subformula replacements. The clausal proof can contain resolution, paramodulation, factoring) into a purely first-order proof of the same formula. The resulting proof is short, i.e. polynomial in the size of the original resolution proof.

Motivation 1

Theorem provers are complicated pieces of software. Complicated software can be incorrect. Theorem provers are used for verification of critical software. Therefore, one needs a way of being able to trust theorem provers without hesitation. There are 3 approaches:

1. Verify the prover as it is. Difficult, because the prover is big and changing.
2. Make sure that the prover has a small 'trusted code base'. Only the trusted code base needs to be verified.
3. Let the prover produce explicit proofs which can be checked independently.

Approach 2 and 3 are closely related. For 2, one has to generate the proof 'on the fly'. For 3, one has to store it.

Motivation 2

I admit, the previous slide only partially motivates this work.

As far as Motivation 1 is concerned, it does not matter whether one uses choice axioms in the proof.

However, using choice axioms for proving first-order formulas is ugly. It is much nicer to use only first-order principles, when proving first-order formulas.

Resolution is disliked by many people because of the Skolemization.

Main Idea:

In general, functions can be replaced by relations in formulas: The equation $1 + 1 = 2$ can be replaced by $\forall\alpha\ S(1,1,\alpha) \rightarrow (\alpha = 2)$. The meaning of $S(x,y,z)$ is 'z is a sum of x and y.'

The same replacement can be made with Skolem functions in clauses. Let

$$\forall x\ p(x, f(x)) \vee q(f(x), x)$$

be a clause with f a Skolem function. It can be replaced by

$$\forall x\ \forall\alpha\ F(x,\alpha) \rightarrow p(x,\alpha) \vee q(\alpha, x).$$

Similarly, $\forall x\ p(x, f(f(x))\ )$ can be replaced by

$$\forall x\ \forall\alpha\beta\ F(x,\alpha) \rightarrow F(\alpha,\beta) \rightarrow p(x,\beta).$$

Surprising Fact:

It is possible to replace functions by relations in resolution proofs. The result is a valid first-order proof, on the condition that

1. the introduced relations are serial,
2. all paramodulation steps are non-separating.

Seriality means: $\forall x_1 \cdots x_n \; \exists y \; F(x_1, \ldots, x_n, y)$ is provable.

Paramodulation is the following rule: From $t_1 \approx t_2 \vee R_1$ and $A[t_1] \vee R_1$ derive $A[t_2] \vee R_1 \vee R_2$.

Non-separating means that one sometimes has to replace additional occurrences of t_1 as well. (An exact definition comes later)

Definition We write $\mathcal{F}_{\mathrm{Prob}}$ for the function symbols occurring in the original clauses.

We write $\mathcal{F}_{\mathrm{Repl}}$ for the function symbols that we want to replace. (The Skolem functions) We have $\mathcal{F}_{\mathrm{Repl}} \subseteq \mathcal{F}_{\mathrm{Prob}}$.

We write $\mathcal{F}_{\mathrm{Def}}$ for the variables that will act as definitions. It should be the case that $\mathcal{F}_{\mathrm{Def}} \cap \mathcal{F}_{\mathrm{Prob}} = \emptyset$.

Furthermore, we assume a function $[\ ]$, which

- assigns to each n-ary function symbol in $\mathcal{F}_{\mathrm{Repl}}$ an $(n+1)$-ary relation.
- assigns to each term $f(t_1, \ldots, t_n)$ with $f \in \mathcal{F}_{\mathrm{Repl}}$, a unique variable in $\mathcal{F}_{\mathrm{Def}}$.

Definition The function $[\ ]$ is recursively extended to all terms over $\mathcal{F}_{\mathrm{Prob}}$ as follows: For a term $f(t_1, \ldots, t_n)$ with $f \notin \mathcal{F}_{\mathrm{Repl}}$, $[f(t_1, \ldots, t_n)] = f(\ [t_1], \ldots, [t_n]\ )$.

Definition For a literal/quantifier free formula A, $[A]$ is obtained by replacing all terms t in A by their corresponding $[t]$.

For a term/literal/quantifier free formula, $\mathrm{Var}(A)$ is the set of variables introduced by $[A]$.

For a term/literal/quantifier free formula, $\mathrm{Def}(A)$ is the set of definitions characterizing the variables introduced by $[A]$. Formally

$$\mathrm{Def}(A) = \{\ \ [f](\ \ [t_1], \ldots, [t_n], [f(t_1, \ldots, t_n)]\ \ ) \mid$$

$$f(t_1, \ldots, t_n) \text{ occurs in } A \text{ and } f \in \mathcal{F}_{\mathrm{Repl}}\ \}.$$

Example

Assume that $f \in \mathcal{F}_{\mathrm{Repl}}$, $[f] = F$, $[\ f(x)\ ] = \alpha$, $[\ f(f(x))\ ] = \beta$.

Then:

$[\ p(x, f(x))\ ] = p(x, \alpha)$,

$[\ q(f(x), f(f(x)))\ ] = q(\alpha, \beta)$,

$\mathrm{Var}(\ p(x, f(x))\ ) = \{\alpha\}$,

$\mathrm{Var}(\ q(f(x), f(f(x)))\ ) = \{\alpha, \beta\}$.

$\mathrm{Def}(\ p(f(x))\ ) = \{F(x, \alpha)\}$.
$\mathrm{Def}(\ q(f(x), f(f(x)))\ ) = \{F(x, \alpha),\ F(\alpha, \beta)\}$.

Example (continued)

If one also assumes that $[\ f(s(x))\ ] = \gamma$, then

$[\ p(f(s(x)), s(f(x)))\ ] = p(\gamma, s(\alpha))$,

$\mathrm{Var}(\ p(f(s(x)), s(f(x)))\ ) = \{\gamma, \alpha\}$,

$\mathrm{Def}(\ p(f(s(x)), s(f(x)))\ ) = \{\ F(s(x), \gamma),\ \ F(x, \alpha)\ \}$.

Definition

A clause $\forall \overline{x}\ A_1 \vee \cdots \vee A_p$ is replaced by

$\forall \overline{x}\ \forall\ \mathrm{Var}(A_1, \ldots, A_p) \quad \mathrm{Def}(A_1, \ldots, A_p) \rightarrow [A_1] \vee \cdots \vee [A_p]$.

Theorem

Let the clause $\forall\overline{y}\ D$ be obtained the from clauses $\forall\overline{x}_1\ C_1, \ldots, \forall\overline{x}_n\ C_n$ by repeated resolution, non-separating paramodulation, factoring, equality factoring, and equality reflexivity.

Then the translation $\forall\overline{y}\ \ \forall\ \mathrm{Var}(D)\ \mathrm{Def}(D) \rightarrow [D]$ has a first-order proof from the translations
$\forall\overline{x}_1\ \forall\ \mathrm{Var}(C_1)\ \mathrm{Def}(C_1) \rightarrow [C_1], \ldots, \forall\overline{x}_n\ \forall\ \mathrm{Var}(C_n)\ \mathrm{Def}(C_n) \rightarrow [C_n]$.

The length of this proof is polynomial in the original proof.

proof: The proof steps can be translated one-by-one.

We first single out instantiation, so that the remaining rules can be treated without instantiation.

After that we single out equality reflexivity.

At this point, all remaining rules have become trivial, with the exception of paramodulation.

Proof (Instantiation)

Assume that $\forall\overline{x}\ C$ has instance $\forall\overline{y}\ D$. Then

$C' = \forall\overline{x}\ \ \forall\ \mathrm{Var}(C)\ \ \mathrm{Def}(C) \to [C]$ subsumes

$D' = \forall\overline{y}\ \ \forall\ \mathrm{Var}(D)\ \ \mathrm{Def}(D) \to [D]$.

Example

The clause $D = p(f(0), f(0))$ is an instance of $C = p(f(x), f(0))$.

Put $[f(0)] = \alpha,\ [f(x)] = \beta$. The formula

$$D' = \forall\alpha\ F(0,\alpha) \to p(\alpha,\alpha)$$

can be proven from

$$C' = \forall x\ \forall\alpha\beta\ \ F(0,\alpha) \to F(x,\beta) \to p(\alpha,\beta).$$

Example 2

The clause $D = p(s(f(0)), s(f(0)))$ is an instance of $C = \forall xy\ p(s(x), s(f(y)))$. Put $[f(0)] = \alpha,\ [f(y)] = \beta$.

The formula

$$D' = \forall\alpha\ F(0, \alpha) \to p(s(\alpha), s(\alpha))$$

can be proven from

$$C' = \forall xy\ \forall\beta\ F(y, \beta) \to p(s(x), s(\beta)).$$

(instantiate $x := \alpha,\ y := 0,\ \beta := \alpha$)

The general pattern is: In C' instantiate each variable into the corresponding term in D'

Proof (Equality Reflexivity)

Assume that $\forall \overline{x}\ C$ is obtained from $\forall \overline{x}\ \ t \not\approx t \vee C$ by equality reflexivity.

Then $\forall \overline{x}\ \forall\ \mathrm{Var}(t \not\approx t, C)\ \ \mathrm{Def}(t \not\approx t, C) \rightarrow [t \not\approx t \vee C]$ implies $\forall \overline{x}\ \forall\ \mathrm{Var}(C)\ \ \mathrm{Def}(C) \rightarrow [C]$.

If $\mathrm{Def}(C) \subset \mathrm{Def}(t \not\approx t, C)$, then there are terms in t that do not occur in C and which have a function $f \in \mathcal{F}_{\mathrm{Repl}}$ on top. Let u be an outermost such term. Let $[u] = \alpha$. Then $\mathrm{Def}(t \not\approx t, C) \backslash \mathrm{Def}(C)$ contains an atom $F(\beta_1, \ldots, \beta_n, \alpha)$ defining α. This atom can be removed through the seriality axiom for F.

Repeating this procedure, all atoms $\mathrm{Def}(t \not\approx t, C) \backslash \mathrm{Def}(C)$ can be removed.

Example

The clause $D = \forall x\ p(x, x)$ follows from from
$C = \forall x\ f(x) \not\approx f(x) \vee p(x, x)$ by equality reflexivity.

Put $[f(x)] = \alpha$. Then

$$D' = \forall x\ p(x, x),$$

and

$$C' = \forall x\ \forall \alpha\ F(x, \alpha) \to \alpha \not\approx \alpha \vee p(x, x).$$

In order to prove $C' \vdash D'$, one needs an instance of α, for which $F(x, \alpha)$ is provable. This can be obtained from the seriality axiom $\forall x \exists y\ F(x, y)$.

The equality reflexivity rule plays an important rôle, because it handles the disappearing terms in all rules.

The other rules can be modified in such a way that they have no disappearing terms. For example, the resolution rule can be modified into:

From $\forall \overline{x}_1 \quad A \vee R_1$ and $\forall \overline{x}_2 \quad \neg A \vee R_2$ derive $\forall \overline{y} \; t_1 \not\approx t_1 \vee t_2 \not\approx t_2 \vee \cdots \vee t_n \not\approx t_n \vee A_1 \vee A_2$, where $t_1, \ldots, t_n$ are the maximal terms that occur in A, but not in $R_1 \vee R_2$.

After that, n equality reflexivity steps are needed to obtain $\forall \overline{y} \; A_1 \vee A_2$.

Non-Separating Paramodulation

Definition

Assume that $C_1 = \forall\overline{x}\ t_1 \approx t_2 \vee R_1$ paramodulates into $C_2 = \forall\overline{x}\ R_2$. Then $D = \forall\overline{x}\ R_2'$ is obtained from C_1, C_2 by non-separating paramodulation if for every term $f(u_1, \ldots, u_n)$ that occurs in R_2 with $f \in \mathcal{F}_{\mathrm{Repl}}$, whenever inside $f(u_1, \ldots, u_n)$ some occurrences of t_1 have been replaced by t_2, then exactly the same replacements have to be made in every occurrence of $f(u_1, \ldots, u_n)$ in R_2.

Examples:

Assume that $\mathcal{F}_{\mathrm{Repl}} = \{f\}$. Let $C_1 = 0 \approx 1$, let $C_2 = p(f(0,0), f(0,0), 0)$. The following clauses can be obtained by non-separating paramodulation:

$$p(f(0,0), f(0,0), 0),$$
$$p(f(0,1), f(0,1), 0),$$
$$p(f(1,0), f(1,0), 0),$$
$$p(f(1,1), f(1,1), 0),$$

$$p(f(0,0), f(0,0), 1),$$
$$p(f(0,1), f(0,1), 1),$$
$$p(f(1,0), f(1,0), 1),$$
$$p(f(1,1), f(1,1), 1).$$

Assume that

$$[f(0,0)] = \alpha,\ [f(0,1)] = \beta,\ [f(1,0)] = \gamma,\ [f(1,1)] = \delta.$$

Let $C_2' = \forall\alpha\ F(0,0,\alpha) \to p(\alpha,\alpha,0)$.

Each of the following clauses is provable (assuming $0 \approx 1$)

$$\forall\alpha\ F(0,0,\alpha) \to p(\alpha,\alpha,0),$$

$$\forall\alpha\ F(0,1,\beta) \to p(\beta,\beta,0),$$

$$\forall\alpha\ F(1,0,\gamma) \to p(\gamma,\gamma,0),$$

$$\forall\alpha\ F(1,1,\delta) \to p(\delta,\delta,0),$$

$$\forall\alpha\ F(0,0,\alpha) \to p(\alpha,\alpha,1),$$

$$\forall\alpha\ F(0,1,\beta) \to p(\beta,\beta,1),$$

$$\forall\alpha\ F(1,0,\gamma) \to p(\gamma,\gamma,1),$$

$$\forall\alpha\ F(1,1,\delta) \to p(\delta,\delta,1).$$

More Examples

Let $C_1 = 0 \approx 1$, $C_2 = p(f(s(0), s(0)))$. Assume that

$$[f(s(0), s(0))] = \alpha, \ [f(s(0), s(1))] = \beta,$$

$$[f(s(1), s(0))] = \gamma, [f(s(1), s(1))] = \delta.$$

We have $C_1' = C_1$ and

$$C_2' = \forall\alpha \ F(s(0), s(0), \alpha) \to p(\alpha).$$

The following clauses are provable

$$\forall\beta \ F(s(0), s(1), \beta) \to p(\beta),$$

$$\forall\gamma \ F(s(1), s(0), \gamma) \to p(\gamma),$$

$$\forall\delta \ F(s(1), s(1), \delta) \to p(\delta).$$

The General Scheme

Assume that $D = \forall\overline{x}\ R_2'$ is obtained from $C_1 = \forall\overline{x}\ t_1 \approx t_2 \vee R_1$ and $C_2 = \forall\overline{x}\ R_2$ by non-separating paramodulation.

> Assume the variables $\overline{x}$.
>
> Assume the variables in $\mathrm{Def}(R_2', R_2, R_1, t_1, t_2)$.
>
> Instantiate $[C_1']$ to obtain $[t_1] \approx [t_2] \vee [R_1]$
>
> Instantiate $[C_2']$, but replace each variable in $\mathcal{F}_{\mathrm{Def}}$, which corresponds to a subterm of R_2, by the translation of its corresponding subterm of R_2'.
>
> At this point $[R_2']$ is provable.

Failure in the Case of Non-Separating Paramodulation

The proof method on the previous slide does not work in the case of non-simultaneous paramodulation.

This is not surprising:

Suppose that $a \approx b$ is applied on $f(a) \approx f(a)$ to obtain $f(a) \approx f(b)$.

The translation of $f(a) \approx f(a)$ equals $\forall \alpha \; F(a, \alpha) \rightarrow \alpha \approx \alpha$, which is a tautology.

The translation of $f(a) \approx f(b)$ equals
$\forall \alpha \beta \; F(a, \alpha) \rightarrow F(b, \beta) \rightarrow \alpha \approx \beta$, which is a choice axiom for F on $a (\approx b)$

Skolemization

We have shown that it is possible to replace functions by serial relations in resolution proofs. It remains to obtain the relations. They must have two properties:

1. The translated clauses must follow from the original formula.
2. Seriality of the relation must follow from the original formula.

A Skolem function f originates from Skolemizing some formula $\forall x_1 \cdots x_n \; \exists y \; A(x_1, \ldots, x_n, y)$.

Simply take A as the serial relation for f :

1. $\forall x_1 \cdots x_n \forall y \; A(x_1, \ldots, x_n, y) \rightarrow A(x_1, \ldots, x_n, y)$ is a tautology.
2. $\forall x_1 \cdots x_n \exists y \; A(x_1, \ldots, x_n, y)$ is the original formula.

Using this approach, there are two problems:

Small Problem: Skolemization in a context

Consider the formula
$\forall x_1 \cdots x_n \; A(x_1, \ldots, x_n) \rightarrow \exists y \; B(x_1, \ldots, x_n, y)$, which Skolemizes into $\forall x_1 \cdots x_n \; A(x_1, \ldots, x_n) \rightarrow B(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$.

The existential quantifier occurs in a conditional context.

Simply encode the context into the relation:
$F(x_1, \ldots, x_n, y) := A(x_1, \ldots, x_n) \rightarrow B(x_1, \ldots, x_n, y)$.

Then the translation
$\forall x_1 \cdots x_n \; A(x_1, \ldots, x_n) \rightarrow \forall y \; F(x_1, \ldots, x_n, y) \rightarrow B(x_1, \ldots, x_n, y)$
is a tautology, and $\forall x_1 \cdots x_n \; \exists y \; F(x_1, \ldots, x_n, y)$ is equivalent to the original formula.

Big Problem: Parallel Skolemization:

Consider the formula $\forall x \; \exists y_1 y_2 \; p(x, y_1, y_2)$,

Its Skolemization equals $\forall x \; p(x, f_1(x), f_2(x) \;)$.

This would translate into
$\forall x \; \forall y_1 y_2 \; F_1(x, y_1) \rightarrow F_2(x, y_2) \rightarrow p(x, y_1, y_2)$.

In the original formula, y_2 is chosen after y_1. In the Skolemization, f_1 and f_2 have to choose in parallel.

As a consequence, in the translation, F_1 and F_2 do not know about each other's choice. There seems to be no way to define F_1 and F_2 independently, such that they are serial and the translation of the Skolemization becomes provable.

Solution: Inner Skolemization

Skolemization is usually performed outside-inside, because this results in smaller Skolem terms. If one Skolemizes inside-outside, the 'lack-of-knowledge' problem disappears, because each Skolem-term receives all variables on which it depends.

$\forall x \; \exists y_1 y_2 \; p(x, y_1, y_2)$ Skolemizes into $\forall x \; \exists y_1 \; p(x, y_1, f_2(x, y_1)\ )$, which in turn Skolemizes into $\forall x \; p(x, f_1(x), f_2(x, f_1(x))\ )$.

Inside-outside Skolemization results in bigger Skolem terms. In the example $f_2(x, f_1(x))$ is bigger than $f_2(x)$.

However, although the Skolem terms are bigger, they do not depend on more variables.

Theorem: Let F be some first-order formula. Let F_1 be its inner Skolemization. Let F_2 be its outer Skolemization. Let y be an existential variable in F. Let t_1 be its Skolem term in F_1. Let t_2 be its Skolem term in F_2. Then t_1 and t_2 contain exactly the same variables.

essential property:

For both F_1, F_2, a variable x belonging to subformula $\forall x\ A$ of F, is in the Skolem term for y iff
there is a sequence of existentially quantified subformulas
$\exists y_1\ B_1,\ \exists y_2\ B_2, \ldots, \exists y_n\ B_n$ of F, such that **(1)** $\exists y_1\ B_1$ is a subformula of A and x occurs in B_1, **(2)** each $\exists y_{i+1}\ B_{i+1}$ is a subformula of B_i and y_i occurs in B_{i+1}, **(3)** $y_n = y$.

A resolution proof remains valid if one replaces some function symbols by more complicated terms depending on the same subterms.

Definition:
A function replacement Θ is a set of rules of the form

$$g(x_1, \ldots, x_n) \Rightarrow t[x_1, \ldots, x_n].$$

It must be the case that if some function symbol $f \in \mathcal{F}_{\mathrm{Repl}}$ occurring in some $t_1[x_1, \ldots, x_n]$ has an i-th argument which is not a variable x_k, then there is no occurrence of f in the right-hand-side of some rule $g_2(y_1, \ldots, y_m) \Rightarrow t_2[y_1, \ldots, y_m]$, s.t. f has a variable y_j as argument on the i-th subterm.

Theorem: A resolution proof (with non-separating paramodulation) remains valid if one replaces each term t that occurs in it, by $\Theta(t)$.

Therefore, a proof remains valid if one replaces all outermost Skolemizations by innermost Skolemizations.

Innermost Skolemization is bad for proof search, but it does not increase the proof length, once it is found.

The Complete Procedure:

1. Transform formula into Clausal Normal Form, using Skolemization as usual.
2. Let theorem prover run, make sure that is uses $\mathcal{F}_{\mathrm{Skol}}$-simultaneous paramodulation only.
3. When it has produced a proof, replace the innermost Skolem terms by outermost Skolem terms.
4. Replace the outermost Skolem functions by relations.

The result is a proof that is completely first-order.

Conclusions, Future Work

- I presented a way of removing Skolem functions from resolution proofs, without increasing their size significantly.
- The method can be adapted to handle the splitting rule as well.
- Try to extend the method to general LK-proofs?
- It would be nice to implement the procedure.