

Basis for automated proof: Skolemization

Stéphane Devismes Pascal Lafourcade Michel Lévy
Jean-François Monin (jean-francois.monin@imag.fr)

Université Joseph Fourier, Grenoble I

March 13, 2015

Overview

Introduction

Examples and properties

Skolemization

Conclusion

Introduction

Herbrand's theorem applies to the domain closure of a set of formulae **with no quantifier**.

For formulae with existential quantification, use **skolemization**.

This transformation was introduced by **Thoralf Albert Skolem** (1887 - 1963), Norwegian mathematician and logician.

General view

Skolemization

- ▶ transforms a set of closed formulae to the domain closure of a set of formulae with no quantifier.
- ▶ preserves the **existence** of a model.

Example 5.2.1

The formula $\exists xP(x)$ is **skolemized** as $P(a)$.

We note the following relations between the two formulae :

1. $\exists xP(x)$ is a consequence of $P(a)$
2. $P(a)$ is **not** a consequence of $\exists xP(x)$, but a model of $\exists x P(x)$ **« provides »** a model of $P(a)$.

Indeed, let I be a model of $\exists xP(x)$. Hence there exists $d \in P_I$.

Let J be the interpretation such that $P_J = P_I$ and $a_J = d$.

J is model of $P(a)$.

Example 5.2.2

The formula $\forall x \exists y Q(x, y)$ is **skolemized** as $\forall x Q(x, f(x))$.

Again :

1. $\forall x \exists y Q(x, y)$ is a consequence of $\forall x Q(x, f(x))$
2. $\forall x Q(x, f(x))$ is **not** a consequence of $\forall x \exists y Q(x, y)$; but a model of $\forall x \exists y Q(x, y)$ **< provides >** a model of $\forall x Q(x, f(x))$.

Let I be a model of $\forall x \exists y Q(x, y)$ and let D be the domain of I .

For every $d \in D$, the set $\{e \in D \mid (d, e) \in Q_I\}$ is not empty, hence there exists a function $g : D \rightarrow D$ such that for every $d \in D$, $g(d) \in \{e \in D \mid (d, e) \in Q_I\}$.

Let J be the interpretation J such that $Q_J = Q_I$ and $f_J = g$: **J is a model of $\forall x Q(x, f(x))$.**

Properties

Skolemization **eliminates existential quantifiers** and **transforms a closed formula A to a formula B** such that :

- ▶ A is a consequence of B , ($B \models A$)
- ▶ every model of A « provides » a model of B

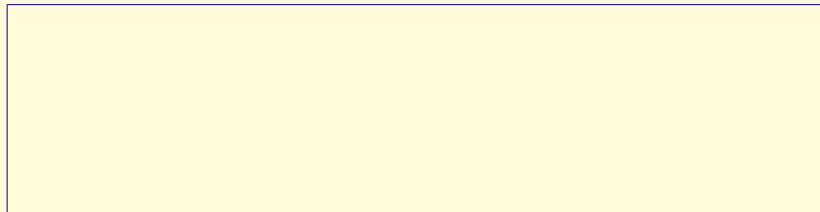
Hence, A **has a model if and only if B has a model** : **skolemization preserves the existence of a model**, in other words it **preserves satisfiability**.

Definitions

Definition 5.2.3

A closed formula is said to be **proper**, if it does not contain any variable which is bound by two distinct quantifiers.

Example 5.2.4



Definitions : generalized normal form

A first-order logic formula is in **normal** form if it does not contain equivalences, implications, and if negations only apply to atomic formulae.

How to skolemize a closed formula A ?

Definition 5.2.5 (skolemization)

Let A a closed formula and E the normal formula with no quantifier, obtained by the following transformation : E is the **Skolem form** of A .

1. B = normalization of A
2. C = make B proper
3. D = **Elimination of existential quantifiers from C .**
This transformation only preserves the existence of a model.
4. E = Transformation of **the closed, normal, proper formula with no existential quantifiers D** into **a normal formula without quantifiers.**

Normalization

1. Eliminate the equivalences
2. Eliminate the implications
3. Move the negations towards the atomic formulae

Rules

$$A \Leftrightarrow B \equiv (A \Rightarrow B) \wedge (B \Rightarrow A)$$

$$A \Rightarrow B \equiv \neg A \vee B$$

$$\neg \neg A \equiv A$$

$$\neg(A \wedge B) \equiv \neg A \vee \neg B$$

$$\neg(A \vee B) \equiv \neg A \wedge \neg B$$

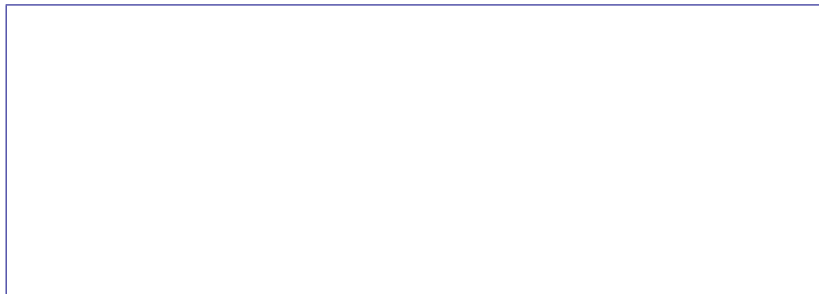
$$\neg \forall x A \equiv \exists x \neg A$$

$$\neg \exists x A \equiv \forall x \neg A$$

Hint : replace $\neg(A \Rightarrow B)$ by $A \wedge \neg B$

Example 5.2.7

The normal form of $\forall y(\forall xP(x,y) \Leftrightarrow Q(y))$ is :



Transformation to a proper formula

Change the name of correctly linked variables, e.g., by choosing new variables at every change of a name.

Example 5.2.8

- ▶ The formula $\forall xP(x) \vee \forall xQ(x)$ is changed to

- ▶ The formula $\forall x(P(x) \Rightarrow \exists xQ(x) \wedge \exists yR(x, y))$ is changed to

Elimination of existential quantifiers

Theorem 5.2.9

Let A be a closed normal and proper formula having one occurrence of the sub-formula $\exists yB$. Let x_1, \dots, x_n be the free variables of $\exists yB$, with $n \geq 0$. Let f be a symbol **not appearing in A** . Let A' be the formula obtained by replacing this occurrence of $\exists yB$ by $B < y := f(x_1, \dots, x_n) >$ (If $n = 0$, f is a constant).

The formula A' is a closed normal and proper formula satisfying :

1. A is a consequence of A'
2. If A has a model then A' has an identical model up to the truth value of f .

Theorem proof 5.2.9

Let us show that A is a consequence of A' .

Since the formula A is closed and proper, the free variables of $\exists yB$, which are bound outside $\exists yB$, are not bound by any quantifier in B (otherwise the proper property would not be respected), hence the term $f(x_1, \dots, x_n)$ is free for y in B .

According to corollary 4.3.38 : $B < y := f(x_1, \dots, x_n) >$ has as consequence $\exists yB$. Hence, we deduce that A is a consequence of A' .

Proof of theorem 5.2.9

Let us show that every model of A provides a model of A' .

Suppose that A has a model I where I is an interpretation with domain D . Let $c \in D$. For all $d_1, \dots, d_n \in D$, let E_{d_1, \dots, d_n} be the set of elements $d \in D$ such that the formula B equals 1 in the interpretation I and the state $x_1 = d_1, \dots, x_n = d_n, y = d$ of its free variables. Let $g : D^n \rightarrow D$ be a function such that if $E_{d_1, \dots, d_n} \neq \emptyset$ then $g(d_1, \dots, d_n) \in E_{d_1, \dots, d_n}$ else $g(d_1, \dots, d_n) = c$. Let J be the interpretation identical to I except that $f_J = g$. We have :

1. $[\exists y B]_{(I, e)} = [B \langle y := f(x_1, \dots, x_n) \rangle]_{(J, e)}$, according to the interpretation of f and of theorem 4.3.36, for every state e of the variables,
2. $[\exists y B]_{(I, e)} = [\exists y B]_{(J, e)}$, since the symbol f is new, the value of $\exists y B$ does not depend of the truth value of f .
3. $\exists y B \Leftrightarrow B \langle y := f(x_1, \dots, x_n) \rangle \models A \Leftrightarrow A'$, according to the property of replacement 1.3.10, which holds in first-order logic as well.

According to these three points, we obtain $[A]_{(J, e)} = [A']_{(J, e)}$ and since f is not in A and since the formulae A and A' do not contain free variables, we have $[A]_I = [A']_J$. Since I is model of A , J is model of A' .

Remark 5.2.10

In theorem 5.2.9, note that the formula A' obtained from formula A by elimination of a quantifier remains closed, normal and proper.

Hence, by « applying » the theorem repeatedly, **which implies choosing a new symbol for each eliminated quantifier**, one can transform a closed, normal and proper formula A into a closed, normal, proper and **without existential quantifier** formula B such that :

- ▶ A is a consequence of B
- ▶ If A has a model, then B has an identical model except for the truth value of the new symbols

Example 5.2.11

By eliminating existential quantifiers in the formula

$\exists x \forall y P(x, y) \wedge \exists z \forall u \neg P(z, u)$ we obtain $\forall y P(a, y) \wedge \forall u \neg P(b, u)$.



Therefore a new symbol must be used whenever an existential quantifier is eliminated.

Transformation in universal closure

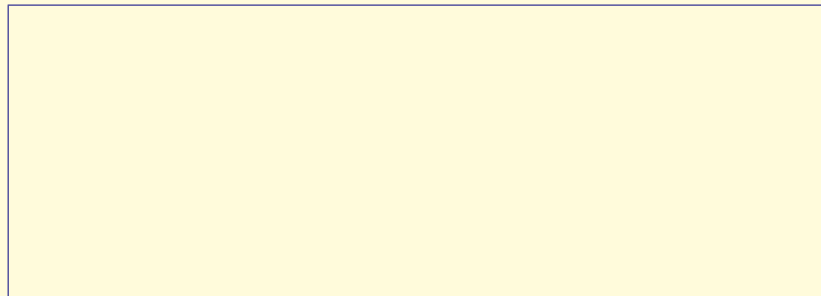
Theorem 5.2.13

Let A be a closed, normal, proper formula without existential quantifier.

Let B be the formula obtained by removing from A all the universal quantifiers (B is the Skolem form of A).

Formula A is equivalent to the domain closure of B .

Proof.



Property of skolemization

Property 5.2.14

Let A be a closed formula and B the Skolem form of A .

- ▶ The formula $\forall(B)$ has as consequence the formula A
- ▶ if A has a model then $\forall(B)$ has a model

Hence A has a model if and only if $\forall(B)$ has a model.

Proof.

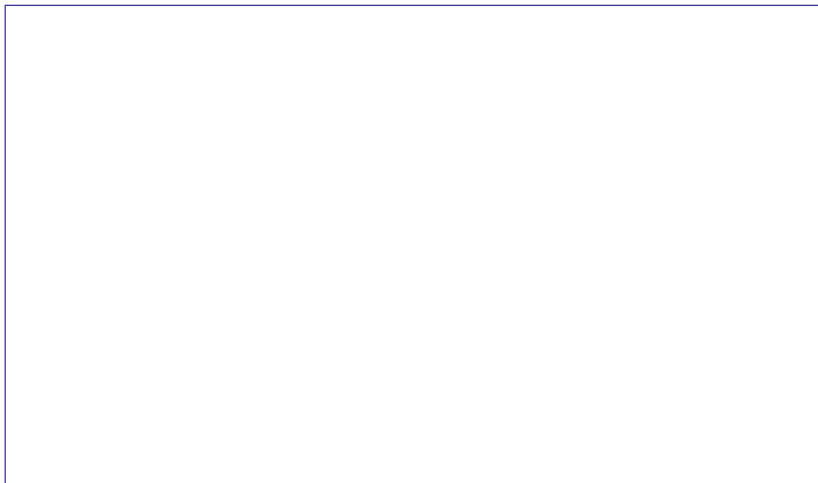
Let C be the closed proper formula in normal form, obtained at the end of the first two steps of skolemization of A . Let D be the result of the elimination of existential quantifiers applied to C . According to remark 5.2.10 we have :

- ▶ The formula D has as consequence the formula C
- ▶ if C has a model then D has a model.

Since the first two steps change the formulae into equivalent formulae, A and C are equivalent. According to theorem 5.2.13, D is equivalent to $\forall(B)$. Hence we can replace above D by $\forall(B)$ and C by A , QED. □

Example 5.2.15

Let $A = \forall x(P(x) \Rightarrow Q(x)) \Rightarrow (\forall xP(x) \Rightarrow \forall xQ(x))$. We skolemize $\neg A$.



Skolemizing a set of formulae

Corollary 5.2.16

Let Γ be a set of closed formulae. Skolemization of Γ consists in applying skolemization to all formulae of Γ , by selecting a new symbol for each existential quantifier eliminated in the third step of skolemization.

We obtain a set Δ of formulae without quantifiers such that :

- ▶ Every model of $\forall(\Delta)$ is model of Γ
- ▶ If Γ has a model then $\forall(\Delta)$ has a model which is the same as for Γ up to the truth value of new symbols.

Today

- ▶ Skolemization

Next course

- ▶ Clausal form
- ▶ Unification
- ▶ First-order resolution
- ▶ Consistency
- ▶ Completeness