Global Skolemization with grouped quantifiers

Domenico Cantone* Marianna Nicolosi Asmundo*
Eugenio Omodeo†

Abstract

When 'local' Skolemization treats proper axioms along with the premisses and the negated conclusion of a conjecture, each sentence originates a finite number of Skolem symbols, without any explicit connection among symbols originating from different sentences.

A global approach, proposed by Davis and Fechter, consists in the simultaneous introduction of infinitely many new functors, which eliminates all quantifiers of the language in a single shot. Two conflicting goals in this elimination process are: to keep as small as possible the collection of 'key' formulae that deserve their own Skolem functors, and to avoid time-consuming simplifications during Skolemization.

Some initial contribution is given here to this potentially open-ended research topic. A technique is proposed by which: contiguous alike quantifiers are treated as a single bunch, so as to lower the arity of Skolem functors; and their relative order —which is immaterial— does not affect the result.

1 Introduction

Automated deduction systems often embody a preprocessing phase, called *Skolemization*, which reshapes a given sentence χ in the convenient format

$$\forall\, x_1 \cdots \forall\, x_h\, \vartheta \text{ , with no quantifiers or descriptors inside } \vartheta.$$

There is no need to explicitly write down the quantifiers in front of such a *prenex Skolemized form* ϑ, as all variables are universally bound.

A usual preliminary, when one aims at obtaining a theorem $\beta_1 \wedge \cdots \wedge \beta_m \!\rightarrow\! \gamma$ from the axioms $\alpha_1, \ldots, \alpha_n$ of a first-order theory, is to Skolemize the α_is, the hypotheses β_j, and the negation $\neg\gamma$ of the thesis separately. (Here α_is, β_js and γ stand for sentences.) Then one conjoins together the φ_is, ψ_js, and χ resulting from the α_is, β_js, and $\neg\gamma$, thus obtaining a Skolemized form

$$\vartheta \;\equiv_{\text{Def}} \varphi_1 \wedge \cdots \wedge \varphi_n \wedge \psi_1 \wedge \cdots \wedge \psi_m \wedge \chi$$

of the negation of the conjectured implication $\alpha_1 \wedge \cdots \wedge \alpha_n \wedge \beta_1 \wedge \cdots \wedge \beta_m \!\rightarrow\! \gamma$. Skolemization preserves satisfiability; therefore, it makes sense to seek a refutation of ϑ in order to conclude that the conjectured implication is valid.

It must be stressed that Skolemization usually leads from the language $\mathcal{L}$ to which the sentences α_i, β_j, and γ belong to a richer language $\mathcal{L}^\star$, whose signature comprises new functors originating from the elimination of existential quantifiers.

Some difficulty with the above-described way of proceeding arises when one intends to work under infinitely many axioms. Significant cases of this are the induction axiom

*Università degli Studi di Catania, Dipartimento di Matematica, viale Andrea Doria 6, 95125 - Catania. cantone@dipmat.unict.it

†Università degli Studi di L'Aquila, Dipartimento di Matematica Pura ed Applicata, via Vetoio Loc. Coppito, 67010 - L'Aquila. omodeo@univaq.it

scheme in the Peano arithmetic, and the separation axiom scheme in the Zermelo set theory. Each of these schemes, which are respectively

$$\varphi_0^x \to \big(\forall x (\varphi \to \varphi_{\mathsf{s}(x)}^x) \to \forall x \, \varphi \big) \quad \text{and} \quad \forall y \, \exists x \, \forall z \big(z \in x \leftrightarrow (z \in y \wedge \varphi) \big) ,$$

has in fact a distinct instance corresponding to every formula φ. A similar, slightly more complex, case arising from elementary geometry is the continuity axiom scheme.

Luckily, as we will see, one can expand the signature of $\mathcal{L}$ in such a way as to obtain a simultaneous Skolemization of all of its formulae. This is to say that every formula χ of $\mathcal{L}$ translates into a formula ϑ over the richer signature so that no quantifiers occur in ϑ. Suitable axioms (to be introduced below with the name of *ε-formulae*), concerning the new functors, will ensure this translation to be faithful in all respects. This global treatment of quantifiers and descriptors, to be called *quantifier elimination*, removes a major hindrance from the design of derivation rules intended to serve as the syntactic counterpart of the semantic notion of logical consequence (cf. [5]).

A virtue of the global approach inspiring this paper is that it leads to structural, as opposed to prefix, Skolemization techniques. This is to say, quantifiers are treated in place instead of being moved to the front of the sentence preliminary to the elimination of the existential ones. In fact, as shown by [3], prefix techniques perform intrinsically worse than the structural ones.

For specific theories regarding, e.g., numbers or sets, *ad hoc* Skolemization techniques may perform better than the ones to be discussed below. Admittedly, this study is aimed at a syntactic method whereas in the original Skolem's conception quantifier elimination was mainly sensitive to the semantics.

2 First-order languages of predicate logic

In defining a language of *first-order predicate logic*, one begins with a signature Σ that can be broken into nine disjoint pieces, as follows:

$$\begin{aligned} \Sigma \;=\; & \{ \to_{/2}, \ldots, \mathbf{f}_{/0} \} \cup \{ = \} \cup \mathcal{P} \cup \mathcal{F} \cup \mathcal{V} \cup \{ \exists x_{/1} \,:\, x \text{ in } \mathcal{V} \} \cup \\ & \{ \forall x_{/1} \,:\, x \text{ in } \mathcal{V} \} \cup \{ \varepsilon x_{/1} \,:\, x \text{ in } \mathcal{V} \} \cup \{ \iota x_{/1} \,:\, x \text{ in } \mathcal{V} \} . \end{aligned}$$

Here $\to, \ldots, \mathbf{f}$ are the usual *connectives* and propositional constants and $=$ is the *equality sign*; $\mathcal{V}$ consists of an infinite sequence $\mathsf{v}_1, \mathsf{v}_2, \mathsf{v}_3, \ldots$ of symbols of degree 0, named *individual variables*. To each individual variable x there univocally corresponds an *existential quantifier* $\exists x$, as well as a *universal quantifier* $\forall x$, a *Hilbert descriptor* εx, and a *Peano descriptor* ιx (see below).

Connectives, the equality sign, variables, quantifiers, and descriptors constitute the fixed endowment of a first-order language. Unlike them, $\mathcal{P}$ and $\mathcal{F}$ differ from one language to another, depending on which domain, or domains, of interest one intends to describe through the logical formalism:

- $\mathcal{P}$ consists of the so-called *relators*, each of which has a specific degree > 0;
- $\mathcal{F}$ consists of the so-called *functors*; it may comprise symbols of degree 0, named *constants*, and symbols of degree > 0, named *function symbols*.

For example, when the domain of interest is a universe of sets, one will put in $\mathcal{P}$ the membership relator $\in_{/2}$ (and possibly others, such as $\subseteq_{/2}$, $Ord_{/1}$, etc.); one may decide also to put in $\mathcal{F}$ the constant $\emptyset$, the binary functors $\cap, \setminus, \cup$, the unary functors $\bigcup$, $\wp$, ${}^{-1}$ (for unionset, powerset, and map inverse operations), etc.

The first-order language $\mathcal{L}_\Sigma$ associated with a signature Σ as above is strictly contained in the collection $\tau(\Sigma)$ of all expressions (i.e., abstract syntax trees coherently labelled) over Σ; the reason why it does not span the whole of it is that the constructors are subject to a typing discipline. Only the *well-typed* expressions of $\tau(\Sigma)$, classified into two disjoint categories, *formulae* and *terms*, enter into the language. The reader is certainly familiar with these notions, save perhaps with the formation rule for descriptions (cf. [8]), which causes a mutual recursion between terms and formulae:

Definition 1 *A* TERM *is an expression t belonging to $\tau(\Sigma)$ such that either*

- $t \equiv g(\,t_1,\ldots,t_n\,)$*, where g stands for a functor of degree n and $t_1,\ldots,t_n$ are terms (in particular t might simply be a constant); or*
- $t \equiv \varepsilon x\,\psi$ *or* $t \equiv \iota x\,\psi$*, where x is a variable and ψ is a formula.*

A FORMULA *is an expression φ belonging to $\tau(\Sigma)$ such that either*

- $\varphi \equiv R(\,t_0,\ldots,t_n\,)$*, where R is a relator of degree $n+1$ and $t_0,\ldots,t_n$ are terms; or*
- $\varphi \equiv \exists x\,\psi$*,* $\varphi \equiv \forall x\,\psi$*,* $\varphi \equiv \neg\,\psi$*, or* $\varphi \equiv \psi \star \chi$*, where x, ψ, χ, and $\star$ stand for a variable, two formulae (not necessarily distinct from each other), and a binary propositional connective, respectively.* □

Throughout this paper, the Greek letters $\varphi, \psi, \chi, \alpha, \beta, \gamma, \vartheta$ are used as *metavariables* ranging over the formulae of a first-order language. Moreover, Q, R and g, f are used as metavariables ranging over relators and functors, respectively; t, s, d as metavariables ranging over first-order terms; and x, y, z as metavariables ranging over the individual variables. Metavariables enable one to schematize formulae; e.g., $\neg\varphi \wedge Q(\,x,y,z\,)$ encompasses all conjunctions whose two conjuncts are, respectively, a negated formula and an atomic one consisting of a relator of degree 3 with distinct individual variables as arguments.

Shorthands will be introduced by means of the metasymbol $\equiv_{\mathrm{Def}}$, through which the language can be enriched with new constructs, e.g., by defining

$$\exists\,!\,x\,\varphi \equiv_{\mathrm{Def}} \exists\,y\,\forall\,x\,(\,\varphi \leftrightarrow y = x\,)\,.$$

Another possible exploitation of them is in *eliminating* inessential constructions from the language; for instance, one might state that $\iota\,x\,\varphi \equiv_{\mathrm{Def}} \varepsilon\,y\,\forall\,x\,(\,\varphi \rightarrow y = x\,)\,.$

It is well-known —in the light of the semantics to be highlighted soon— that through slightly more complex rewriting mechanisms one can get rid of both forms ε, ι of descriptors retaining but one form of quantifiers —this explains, by the way, why so many presentations of first-order logic do without them. Our attitude will be, against the main stream, to eliminate quantifiers of both forms retaining ε-descriptors only.

An *interpretation* $\Im$ of $\mathcal{L}_\Sigma$ (cf., e.g., [6]) assigns a constrained value to every term or formula in which any occurrence of each variable x falls within the scope of one of the constructors $\forall\,x$, $\exists\,x$, $\varepsilon\,x$, $\iota\,x$. The process of determining such value, abstractly specifiable through the designation rules below (whose contents are seldom algorithmic), is named *evaluation*; it presupposes that a structure of the following kind is given:

(A) a nonnull *domain of discourse* $\forall^\Im$, whose entities are temporarily regarded as new symbols of degree 0 ('new' in the sense that $\Sigma \cap \forall^\Im = \emptyset$);

(B) an association of a relation $Q^\Im \subseteq \underbrace{\forall^\Im \times \cdots \times \forall^\Im}_{n\ \text{times}}$ with every degree n relator Q;

(C) an association of a function $g^{\Im}\ : \underbrace{\forall^{\Im}\times\cdots\times\forall^{\Im}}_{n\ \text{times}}\ \longrightarrow\ \forall^{\Im}$ (which is a value drawn from $\forall^{\Im}$ if $n=0$) with every degree n functor g.

It is convenient to refer the evaluation process to the first-order language over the enriched signature $\Sigma\cup\forall^{\Im}$, after extending the association (C) to these new constants in the most obvious fashion: for all $k\in\forall^{\Im}$, $\quad k^{\Im}\ \equiv_{\text{Def}} k$.

Definition 2 (Value-designation rules) *For* terms*:*

(1) $g(\,t_1,\ldots,t_n\,)^{\Im}$ *is the value* $g^{\Im}(\,t_1^{\Im},\ldots,t_n^{\Im}\,)$*;*

(2) $(\,\varepsilon\,x\,\varphi\,)^{\Im}$ *is an element* k *of the domain* $\forall^{\Im}$*, possibly such that* $(\,\varphi^{x}_{k}\,)^{\Im}=\mathbf{t}$*.*
(This is to say: if there is at least one k *for which* $(\,\varphi^{x}_{k}\,)^{\Im}$ *holds, one such* k *is arbitrarily selected as* $(\,\varepsilon\,x\,\varphi\,)^{\Im}$*; otherwise,* $(\,\varepsilon\,x\,\varphi\,)^{\Im}$ *is an arbitrary member of* $\forall^{\Im}$*).*

For sentences*:*

(3) $Q(\,t_1,\ldots,t_n\,)^{\Im}$ *is the value* $\begin{cases}\mathbf{t} & \textit{if}\ \ \textit{the ordered } n\textit{-tuple } t_1^{\Im},\ldots,t_n^{\Im} \textit{ belongs to } Q^{\Im}\,,\\ \mathbf{f} & \textit{otherwise}\,;\end{cases}$

(4) $(\,\neg\alpha\,)^{\Im}$ *is the value* $\begin{cases}\mathbf{t} & \textit{if}\ \ \alpha^{\Im}=\mathbf{f}\,,\\ \mathbf{f} & \textit{otherwise}\,;\end{cases}$ *analogously one defines* $(\,\alpha\star\beta\,)^{\Im}$ *for each binary propositional connective* $\star$*, on the basis of the appertaining truth-table;*

(5) $(\,\forall\,x\,\varphi\,)^{\Im}$ *is the value* $\begin{cases}\mathbf{t} & \textit{if}\ \ (\,\varphi^{x}_{k}\,)^{\Im}=\mathbf{t} \textit{ for all } k \textit{ in } \forall^{\Im}\,,\\ \mathbf{f} & \textit{otherwise}\,;\end{cases}$

$(\,\exists\,x\,\varphi\,)^{\Im}$ *is the value* $\begin{cases}\mathbf{t} & \textit{if}\ \ (\,\varphi^{x}_{k}\,)^{\Im}=\mathbf{t} \textit{ for some } k \textit{ in } \forall^{\Im}\,,\\ \mathbf{f} & \textit{otherwise}\,.\end{cases}$ □

To illustrate the usage of a first-order language, we exploit in the following example two binary relators: ϱ and λ.

Example 1 Given a relation ϱ, a *bisimulation* over ϱ is, by definition, a symmetric relation λ enjoying the following property: when $u\lambda v$ and $v\varrho w$, there is a z fulfilling $u\varrho z$ and $z\lambda w$. Formally stated, the properties λ is to meet are:

$$\begin{array}{ll}\forall\,x\,\forall\,y\,(\quad x\lambda y \quad\rightarrow\quad y\lambda x \quad)\,,\\ \forall\,x\,\forall\,y\,(\quad \exists\,z\,(\,z\varrho y\wedge x\lambda z\,)\quad\rightarrow\quad \exists\,z\,(\,z\lambda y\wedge x\varrho z\,)\quad)\,.\end{array}$$

3 Examples of quantifier elimination

For ease of presentation, let us momentarily add a new symbol v_0 to our former supply $\mathcal{V}$ of variables. In a sense, any formula φ devoid of quantifiers that involves only the variables $\mathsf{v}_0,\mathsf{v}_1,\ldots,\mathsf{v}_n$ (for some $n\geq 0$) expresses a relationship between an input tuple $e_1,\ldots,e_n$ and a result e_0. Indeed, given a tuple $e_1,\ldots,e_n$ of values drawn from the domain of discourse, either there is at least one e_0 making φ true under the assignment $\mathsf{v}_0\mapsto e_0,\mathsf{v}_1\mapsto e_1,\ldots,\mathsf{v}_n\mapsto e_n$ or no such value can be found. When deemed useful, one can create a functor h_φ to designate a (total) function $[e_1,\ldots,e_n]\overset{h_\varphi}{\mapsto} e_0$ so as to meet φ, if possible, for any given n-tuple $e_1,\ldots,e_n$. (The degree n of h_φ may of course be 0.)

Through a suitable (infinite) provision of functors h_φ generated in this fashion, one can get rid of quantifiers and descriptors altogether, as the following handful of tiny examples is intended to suggest. For now we will be essentially complying with the quantifier-elimination technique proposed in [5].

Example 2 Let $\varphi_0 \equiv_{\text{Def}} \mathsf{v}_0 \in \mathsf{v}_1$. The functor h_{φ_0} has degree 1, and it designates a function selecting a member v_0 of its input parameter v_1 whenever possible. In particular, by binding the actual parameter v_0 to the formal parameter v_1 of h_{φ_0} one obtains $h_{\varphi_0}(\mathsf{v}_0)$, which is to designate a member of v_0 unless v_0 is 'empty', i.e. devoid of members. Anyway $h_{\varphi_0}(\mathsf{v}_0)$ designates something in the domain of discourse, and consequently, the quantifier-free formula $\varphi_1 \equiv_{\text{Def}} h_{\varphi_0}(\mathsf{v}_0) \notin \mathsf{v}_0$ expresses what one usually writes as $\forall \mathsf{v}_1(\mathsf{v}_1 \notin \mathsf{v}_0)$.

The functor h_{φ_1} has degree 0, because v_0 is the only variable in φ_1; it designates an empty entity provided one exists. Hence the free-variable term $t \equiv_{\text{Def}} h_{\varphi_1}(\ )$ designates what one might also designate by the description $\varepsilon \mathsf{v}_0\, \varphi_1$. One can state the fact that an empty entity exists by saying that if t gets bound to the formal parameter of h_{φ_0}, the result will actually fulfill φ_1: otherwise stated, $h_{\varphi_0}(t) \notin t$. This is a captious formulation of what is usually written as $\exists \mathsf{v}_0 \forall \mathsf{v}_1(\mathsf{v}_1 \notin \mathsf{v}_0)$, as if one were stating in natural language the existence of an empty set by saying "Pick a set e possibly devoid of members, select a member m of e if this is possible: then m will not be a member of e". $\Box$

Our next example brings to light some drawbacks of the elimination technique just illustrated:

Example 3 By an analysis similar to the one carried out in Example 2, one realizes that the term $d_1 \equiv_{\text{Def}} h_{Q(\mathsf{v}_0, h_{Q(\mathsf{v}_1,\mathsf{v}_0)}(\mathsf{v}_0))}(\ )$ designates the first component e_1 of a pair belonging to the relation E designated by Q —if any. Likewise, $d_2 \equiv_{\text{Def}} h_{Q(\mathsf{v}_1,\mathsf{v}_0)}(d_1)$ designates the second component of a pair in E whose first component is e_1. The sentence $Q(d_1, d_2)$ hence expresses the existence of at least one pair in E: namely, $\exists \mathsf{v}_1 \exists \mathsf{v}_2\, Q(\mathsf{v}_1, \mathsf{v}_2)$.

Let us remark that in spite of the similarity of their rôles, the two quantifiers have been translated somewhat differently: $\exists \mathsf{v}_2$ led to a functor of degree 1, whereas $\exists \mathsf{v}_1$ led to a constant. However, two constants would have done to the case very well.

The anomaly becomes more apparent if we do the similar exercise of expressing $\exists \mathsf{v}_2 \exists \mathsf{v}_1\, Q(\mathsf{v}_1, \mathsf{v}_2)$ without quantifiers. Still following [5], we obtain a sentence with a somewhat different structure, namely $Q(d_4, d_3)$ where $d_3 \equiv_{\text{Def}} h_{Q(h_{Q(\mathsf{v}_0,\mathsf{v}_1)}(\mathsf{v}_0),\mathsf{v}_0)}(\ )$ is the constant corresponding to $\exists \mathsf{v}_2$ and $d_4 =_{\text{Def}} h_{Q(\mathsf{v}_0,\mathsf{v}_1)}(d_3)$ involves the function symbol $h_{Q(\mathsf{v}_0,\mathsf{v}_1)}$ corresponding to $\exists \mathsf{v}_1$.

This dissymmetry is unpleasant, because the two quantified sentences differ only by an interchange of quantifiers, which has no significance whatsoever. We will discuss in the next section how to remedy such awkwardnesses by a refined quantifier-elimination policy. $\Box$

4 Elimination of quantifiers grouped in bunches

Let us consider a signature Σ_0 consisting of the same ingredients used to define a first-order predicate language, save quantifiers and descriptors:

$$\Sigma_0 = \{ \rightarrow_{/2}, \ldots, \mathbf{f}_{/0} \} \cup \{ = \} \cup \mathcal{P} \cup \mathcal{F} \cup \mathcal{V}.$$

A convenient practice, in schematizing terms and formulae of Σ_0 at the metalevel, is to adopt a special symbol called the *anonymous variable*, $\_$, in place of any variable that occurs only once. For us, $g(\_,\_)$ stands for the term $g(\mathsf{v}_1,\mathsf{v}_2)$ when taken in isolation, but it stands for $g(\mathsf{v}_2,\mathsf{v}_3)$ inside the formula $Q\big(x,\_,h(g(\_,\_),x)\big)$, where, reasonably, v_1 must take the place of the first $\_$, and v_4 the place of both occurrences of x.

To abridge the formulae of Σ_0 we hence utilize another signature $\breve{\Sigma}_0$ that differs from Σ_0 for having a countable infinity of new symbols in place of $\mathcal{V}$: $\breve{\Sigma}_0 \equiv_{\mathrm{Def}} (\Sigma_0 \setminus \mathcal{V}) \cup \{\_\} \cup \{x,y,z,\dots\}$. Each one of $x,y,z,\dots$ and $\_$ acts as a *metavariable* for a variable in $\mathcal{V}$, as in the example just given. Inside $\tau(\breve{\Sigma}_0)$ one finds syntactic entities analogous to terms and formulae: it is precisely these entities that stand for terms and formulae of Σ_0.

Key formulae will play a crucial rôle in this section, inasmuch as they will be treated as the only formulae deserving their own Skolem functors:

Definition 3 *Let Σ_0 and $\breve{\Sigma}_0$ be signatures related to each other as said above.*

A formula χ of $\breve{\Sigma}_0$ is said to be a KEY FORMULA *if it meets the following two conditions:*

- *every term inside χ either* is *$\_$ (the anonymous variable) or* contains *an occurrence of some other metavariable;*
- *at least one metavariable other than $\_$ occurs in φ.*

A formula φ of Σ_0 is said to be a KEY FORMULA *of* INDEGREE *A if there is a key formula χ of $\breve{\Sigma}_0$, standing for φ, where $\_$ occurs A times: the number of distinct metavariables other than $\_$ in χ is regarded as the* OUTDEGREE *of φ.* □

Recalling our previous conventions about schematizing first-order formulae at the metalevel, the structure of the key formulae of Σ_0 can be described much more explicitly: to be a key formula, φ must be such that

- each one of the 'input parameters' $\mathsf{v}_1,\mathsf{v}_2,\dots,\mathsf{v}_A$ appears exactly once in φ and
- the occurrence of v_j is on the left of v_{j+1} for $j=1,\dots,A-1$;
- the 'output parameter' v_{A+1} occurs in φ;
- every term appearing in φ which is none of the variables $\mathsf{v}_1,\dots,\mathsf{v}_A$ contains at least one 'output parameter' v_j with $j>A$;
- if a variable v_j with $j>A+1$ appears in φ, its first occurrence must be preceded by at least one occurrence of v_{j-1} in φ.

Unlike the outdegree, the indegree of a key formula can be 0. In Σ_0 it may happen that φ be key formula of indegree A for more than one A; anyway, its oudegree will depend on A univocally.

Example 4 Of the two key formulae $\varphi \equiv_{\mathrm{Def}} R\big(g(\mathsf{v}_1),\mathsf{v}_2\big)$ and $\psi \equiv_{\mathrm{Def}} R(\mathsf{v}_1,\mathsf{v}_2)$, the former has outdegree 2, while the latter can have outdegree 2 or 1. Corresponding key formulae of $\breve{\Sigma}_0$ are

$$\chi \equiv_{\mathrm{Def}} R\big(g(x),y\big) \text{ and } \chi_0 \equiv_{\mathrm{Def}} R(x,y),\ \chi_1 \equiv_{\mathrm{Def}} R(\_,y).$$

Skolem functors $h_1^{\chi} \equiv h_{1,2}^{\varphi}$, $h_2^{\chi} \equiv h_{2,2}^{\varphi}$, $h_1^{\chi_0} \equiv h_{1,2}^{\psi}$, $h_2^{\chi_0} \equiv h_{2,2}^{\psi}$, $h_1^{\chi_1} \equiv h_{1,1}^{\psi}$ associated with these can be exploited to rewrite

$\exists\mathsf{v}_1\exists\mathsf{v}_2 R\big(g(\mathsf{v}_1),\mathsf{v}_2\big)$, $\exists\mathsf{v}_1\exists\mathsf{v}_2 R(\mathsf{v}_1,\mathsf{v}_2)$, $\forall\mathsf{v}_1\exists\mathsf{v}_2 R(\mathsf{v}_1,\mathsf{v}_2)$, $\forall\mathsf{v}_1\exists\mathsf{v}_2 R\big(g(\mathsf{v}_1),\mathsf{v}_2\big)$ <u>as</u>
$R\big(g(h_1^{\chi}()),h_2^{\chi}()\big)$, $R\big(h_1^{\chi_0}(),h_2^{\chi_0}()\big)$, $\forall\mathsf{v}_1 R\big(\mathsf{v}_1,h_1^{\chi_1}(\mathsf{v}_1)\big)$, $\forall\mathsf{v}_1 R\big(g(\mathsf{v}_1),h_1^{\chi_1}(g(\mathsf{v}_1))\big)$.

Notice that φ borrows a Skolem functor from ψ in the last case; similarly, the key formula $R(x,\_)$, which stands for $R(\mathsf{v}_2,\mathsf{v}_1)$, will serve the elimination of $\exists\mathsf{v}_2$ from $\forall\mathsf{v}_1\exists\mathsf{v}_2 R(\mathsf{v}_2,\mathsf{v}_1)$ and from $\forall\mathsf{v}_1\exists\mathsf{v}_2 R(\mathsf{v}_2,g(\mathsf{v}_1))$ alike. □

Lemma 1 *Given a signature $\breve{\Sigma}_0$ as above, there is a signature $\breve{\Sigma}_\infty$ which in addition to the symbols of $\breve{\Sigma}_0$ comprises enough new functors $\mathcal{S}_\infty$ for a one-to-one correspondence $[\chi, j] \mapsto h(\chi, j)$ to exist between the pairs χ, j with*

$$\chi \textit{ key formula of } \breve{\Sigma}_\infty, \quad 0 < j \leq \textit{ outdegree}(\chi)$$

and $\mathcal{S}_\infty$, the equality $$\textit{degree}(h(\chi, j)) = \textit{number of occurrences of } \_ \textit{ in } \chi$$ *holding for all χ and all j.*

Proof. Let $\mathcal{S}_0 \equiv_{\mathrm{Def}} \emptyset$. For $i = 0, 1, 2, \ldots$, let $\breve{\Sigma}_{i+1} \equiv_{\mathrm{Def}} \breve{\Sigma}_0 \cup \mathcal{S}_{i+1}$, where

- $\mathcal{S}_{i+1}$ is the superset of $\mathcal{S}_i$ comprising exactly one symbol $h(\chi, j)$ for each pair χ, j with χ key formula of $\breve{\Sigma}_i$, and
- the degree of each symbol $h(\chi, j)$ is as required by the thesis.

It is almost obvious that by putting $\mathcal{S}_\infty \equiv_{\mathrm{Def}} \bigcup_{i=0}^{\infty} \mathcal{S}_i$ and $\breve{\Sigma}_\infty \equiv_{\mathrm{Def}} \breve{\Sigma}_0 \cup \mathcal{S}_\infty$ one obtains the desired signature. It goes without saying that the key formulae of $\breve{\Sigma}_i$, for $i = 1, 2, \ldots, \infty$, are defined exactly like those of $\breve{\Sigma}_0$. ■

Before returning to the signature Σ_0, to whose service $\breve{\Sigma}_0$ was created, we note that every $\breve{\Sigma}_i$ serves in the same fashion the signature $\Sigma_i \equiv_{\mathrm{Def}} \Sigma_0 \cup \mathcal{S}_i$: the equality $\breve{\Sigma}_i = (\Sigma_i \setminus \mathcal{V}) \cup \{\_\} \cup \{x, y, z, \ldots\}$ holds invariably, and the key formulae of $\breve{\Sigma}_i$ correspond to those of Σ_i.

Definition 4 *The symbols $h(\chi, j)$ forming $\mathcal{S}_\infty$ are called* SKOLEM FUNCTORS *of Σ_0. Each of them will also be indicated as $h^{\varphi}_{j,C}$ (and even as h^{φ}_{j}), where φ is the formula indicated by χ and C is the related outdegree.*

The DAVIS-FECHTER CLOSURE *of Σ_0 is the signature $\Sigma_\infty \equiv_{\mathrm{Def}} \Sigma_0 \cup \mathcal{S}_\infty$.* □

Note that quantifiers are still missing from the latter signature; as a matter of fact, our aim is to introduce these symbols with the status of mere abbreviating conveniences. In sight of that, one proves the following proposition:

Lemma 2 *Let ψ and $\mathsf{v}_{j_1}, \ldots, \mathsf{v}_{j_C}$ be a formula of Σ_i and a list, with $C \neq 0$, of distinct variables which occur in ψ. Then one can uniquely determine*

- *an integer $A \geq 0$,*
- *a key formula φ of indegree A and outdegree C;*
- *a list $t_1, \ldots, t_A$ of terms not containing any of variables $\mathsf{v}_{j_1}, \ldots, \mathsf{v}_{j_C}$, and*
- *a permutation π of $1, \ldots, C$*

so that $$\psi \equiv \varphi[t_1, \ldots, t_A, \mathsf{v}_{j_{\pi_1}}, \ldots, \mathsf{v}_{j_{\pi_C}}],$$ *i.e., ψ results from φ by simultaneous replacement of each v_q by the q-th term in the list $t_1, \ldots, t_A, \mathsf{v}_{j_{\pi_1}}, \ldots, \mathsf{v}_{j_{\pi_C}}$ $(q = 1, \ldots, A + C)$.*

Proof. Let us indicate: • by $\mathsf{v}_{h_1}, \ldots, \mathsf{v}_{h_C}$, the variables $\mathsf{v}_{j_1}, \ldots, \mathsf{v}_{j_C}$ so arranged that the first occurrence of v_{h_k} precedes the first occurrence of $\mathsf{v}_{h_{k+1}}$ in ψ; • by π, the permutation sending each g to the subscript $\pi_g = k$ for which $j_k = h_g$.

An algorithm to determine φ, together with the desired list of t_hs, can be described as follows. Locate the leftmost term t_1 not containing any of the variables $\mathsf{v}_{j_1}, \ldots, \mathsf{v}_{j_C}$ inside ψ. If no such term can be found, put $A = 0$, and replace everywhere each v_{h_k} by the corresponding v_k in ψ to obtain φ. Otherwise, do the provisional assignment $A = 1$ and then, as long as there is a term containing none of the v_{j_h}s after the end of the occurrence of t_A just located in ψ, locate the leftmost such term t_{A+1} and increase A by 1. Now, set φ equal to the result of simultaneously replacing, in ψ: • each v_{h_k} by v_{A+k} everywhere, and • each one of the occurrences t_i that have been singled out, by the corresponding variable v_i. Little reflection shows that $\psi \equiv \varphi[t_1, \dots, t_A, \mathsf{v}_{h_1}, \dots, \mathsf{v}_{h_C}]$ as desired.

To see that this representation of ψ is unique, suppose that γ is a key formula such that $\psi \equiv \gamma[d_1, \dots, d_m, \mathsf{v}_{h_1}, \dots, \mathsf{v}_{h_C}]$ (where no d_i contains any v_{j_k} and m is the largest number such that v_{m+C} occurs in γ). A straightforward inductive argument shows that each d_k must coïncide with t_k, so that $m = A$; therefore γ results from ψ in the same manner as φ, and, accordingly, $\gamma \equiv \varphi$. ■

This point having been reached, we can define quantifiers as follows. Under the same hypotheses of the preceding lemma, once φ and $t_1, \dots, t_A$ have been determined in the way it indicates, we put

$$(\exists \mathsf{v}_{j_1}, \dots, \mathsf{v}_{j_C})(\psi) \quad \equiv_{\text{Def}} \quad \varphi[t_1, \dots, t_A, h_1^{\varphi}(t_1, \dots, t_A), \dots, h_C^{\varphi}(t_1, \dots, t_A)]$$
$$(\forall \mathsf{v}_{j_1}, \dots, \mathsf{v}_{j_C})(\psi) \quad \equiv_{\text{Def}} \quad \varphi[t_1, \dots, t_A, h_1^{\neg\varphi}(t_1, \dots, t_A), \dots, h_C^{\neg\varphi}(t_1, \dots, t_A)].$$

As for the Hilbert descriptor ε, which is a formal analogue of the indefinite article "an", it can be eliminated as follows. Let ψ and v_j be a formula and a variable of Σ_∞. If v_j appears in ψ, we can relate the list $\varphi, t_1, \dots, t_A$ to ψ and v_j as in the preceding lemma, and then put

$$\varepsilon\, \mathsf{v}_j\, \psi \ \equiv_{\text{Def}} h_1^{\varphi}(t_1, \dots, t_A).$$

If v_j instead does not occur in ψ, we put

$$\varepsilon\, \mathsf{v}_j\, \psi \ \equiv_{\text{Def}} h_1^{\mathsf{v}_1 = \mathsf{v}_1}(\ ).$$

A price was paid for eliminating descriptors and quantifiers in this fashion; as a matter of fact, we had to shift from the original signature Σ_0 to a somewhat richer signature Σ_∞ (which, anyway, we could do in a very constructive manner). As a compensation for this, we can now characterize a useful deductive machinery for Σ_∞ in extremely straightforward terms, as we are about to see.

Let us consider the following categories of formulae in Σ_∞:

- *tautologies*;
- *identity axioms*, of the form $t = t$;
- *congruency axioms*, of the two kinds:

$$s_0 = d_0 \wedge \cdots \wedge s_n = d_n \quad \rightarrow \quad g(s_0, \dots, s_n) = g(d_0, \dots, d_n)\,,$$
$$s_0 = d_0 \wedge \cdots \wedge s_n = d_n \quad \rightarrow \quad \big(Q(s_0, \dots, s_n) \rightarrow Q(d_0, \dots, d_n)\big)\,;$$

- *ε-formulae*, of the form

$$\varphi[d_1, \dots, d_A, t_1, \dots, t_C] \rightarrow \varphi[d_1, \dots, d_A, h_1^{\varphi}(d_1, \dots, d_A), \dots, h_C^{\varphi}(d_1, \dots, d_A)],$$

where t, s_i, d_i stand for arbitrary terms, g for a functor with positive degree (in $\mathcal{S}_\infty \cup \mathcal{F}$), Q for a predicate symbol, and φ for a key formula of in-/out-degree A/C.

Note that all of these are universally valid formulae, save the ε-formulae, which are nevertheless essential to reflect the intended meaning of the Skolem functors. They enter into the notion of *derivability*, which is introduced as follows:

Definition 5 *The collection of all tautologies, identity and congruency axioms, and ε-formulae, is denoted Λ, and its elements are called* LOGICAL AXIOMS.

Let Γ be a collection of formulae of Σ_∞, and let Θ be the smallest superset of $\Gamma \cup \Lambda$ which is closed with respect to MODUS PONENS*, in the following sense:*

when two formulae of respective forms ψ and $\psi \rightarrow \vartheta$ belong to Θ
(where ψ indicates the same formula twice), also ϑ belongs to Θ.

We will indicate that a formula ϑ belongs to Θ by the notation $\Gamma \vdash \vartheta$, which reads as: Γ YIELDS *(or* ENTAILS*) ϑ, or as: ϑ is* DERIVABLE *from* Γ. □

The equipollence of this notion of derivability with more classical ones, and hence its completeness with respect to the semantics, is proved in [5].

5 Conclusions

Currently, two implementations exist of the quantifier elimination techniques discussed above. One was carried out in LPA-Prolog, along the original lines of [5]; the other was carried out in SETL and offers the option to treat bunches of quantifiers along the lines of this paper.

These implementations will evolve in the sense of embodying —at least as options— simplification rules reflecting, e.g., the associative-commutative-idempotency properties of $\wedge$ and $\vee$.

We intend to exploit these techniques in a tableau-based proof-system (cf. [4]). We expect that our sparing economy in introducing Skolem functors will be rewarded by the reduction of redundancies in the tableau development (cf. [2, 7, 1]).

Acknowledgments

This work enjoyed support of the C.N.R. (National Research Council) of Italy, Research Project No. 95.00411.CT12 (S.E.T.A.), of MURST 40% (*Calcolo algebrico e simbolico*, *Modelli della computazione e dei linguaggi di programmazione*) and of MURST 60% projects.

References

[1] M. Baaz and Ch. G. Fermüller. Non-elementary speedups between different versions of tableaux. In *Proceedings of the 4th International Workshop TABLEAUX'95*, pages 217–230. Springer-Verlag LNAI, 918, 1995.

[2] B. Beckert, R. Hähnle, and P. H. Schmitt. The even more liberalized δ-rule in free variable semantic tableaux. In D. Mundici, G. Gottlob, and A. Leitsch, editors, *Computational Logic and Proof Theory, Proc. 3rd Kurt Gödel Colloquium, Brno*, pages 108–119. Springer-Verlag LNCS, 713, 1993.

[3] M. Baaz and A. Leitsch. On Skolemization and proof complexity. *Fundamenta Informaticae*, 20(4):353–379, 1994.

[4] D. Cantone. A fast saturation strategy for set-theoretic tableaux. In D. Galmiche, editor, *Proc. of Automated Reasoning with Analytic Tableaux and Related Methods International Conf., TABLEAUX'97*, pages 122–137. Springer-Verlag LNAI, 1227, 1997.

[5] M.D. Davis and R. Fechter. A free variable version of the first-order predicate calculus. *Journal of Logic and Computation*, 1(14):431–451, 1991.

[6] Herbert B. Enderton. *A Mathematical Introduction to Logic.* Academic Press, New York and London, 1972.

[7] R. Hähnle and P. H. Schmitt. The liberalized δ-rule in free variable semantic tableaux. *Journal of Automated Reasoning*, 13:211–221, 1994.

[8] A.C. Leisenring. *Mathematical Logic and Hilbert's ε-symbol.* MacDonald & Co., 1969.