Epsilon, Delta, and Speed-Ups
WORK in PROGRESS

Georg Moser

WWU Münster, Germany*

Abstract. This paper studies the applicability of Hilbert's ε-calculus in automated reasoning, in particular its incorporation in free-variable tableaux is investigated. Based on this study a careful comparison between δ^{ε}-tableaux [1] and $\delta^{\star^{\star}}$-tableaux [8] becomes possible. Finally, in the spirit of [8], a new liberalization of the $\delta^{\star}$-rule [3] is defined. This liberalization allows a non-elementary speed-up over $\delta^{\star^{\star}}$-tableaux.

1 Introduction

In proof theory and automated reasoning Skolemization techniques play a vital rôle. I.e. in the context of refutational theorem proving the elimination of *weak* quantifiers in a formula that is to be refuted, is of utmost importance. Here we are mainly concerned with tableaux based proof procedures. In recent years major improvements upon the basic calculus of Beth and Hintikka have been described, which focus on different versions of δ-expansion, i.e. on different Skolemization techniques, cf. [10, 6, 3, 1, 8].

A very interesting stance has been taken by W. Ahrendt, and M. Giese in [1]. They use the computational power of Hilbert's ε-calculus (see Section 3 for a formal definition) and employ ε-terms in the rôle of Skolem functions. The authors (at least partly) motivate their interest in the ε-formalism, by the fact that the established δ^{ε}-tableaux system allows exponential speed-up over previously known versions. This work motivates our interest in studying the possibility to employ Hilbert's ε-calculus successfully in *automated* reasoning.

Hilbert's ε-formalism [11], is a formalism that (employed on first order logic) provides an effective formalization of first-order logic. This formalism arose from the foundational works of David Hilbert and his colleagues. However our interest in the ε-calculus has nothing to do with foundational debates, but is based on the observation that via the ε-calculus the logical complexity of formulas can be (almost) completely coded into the term-structure.

Before we can asses the applicability of the ε-calculus to automated reasoning, we have to provide suitable information on the ε-calculus. In Section 3 we will establish some basic and some less basic facts upon this formalism. To interpret these facts correctly, we have to be more precise on the sort of *application* of Hilbert's ε-calculus we are interested in.

* Institut für Mathematische Logik und Grundlagenforschung, Einsteinstrasse 62, Münster
moserg@math.uni-muenster.de

We take the point of view that we want to employ the ε-calculus as a sort of *black box* in first-order automated theorem proving: Firstly we want to know whether it is possible to render a formalization of Hilbert's ε-calculus that is suitable and effective for proof search. The best solution would be to come up with a (suitable) tableaux or resolution based proof procedure. Second we want to make sure that an arbitrary first-order problem e.g. in the TPTP-library can be translated effectively in the language of the formalization obtained. Clearly we cannot assume that a given proof attempt will always succeed. Hence it is necessary to understand the output of the broken proof attempt. Therefore we demand (a weak form of) the reverse translation as well. (The qualifier 'a weak form of' will be discussed later, when more details have been presented.)

Note that we are not particularly interested in obtaining a formalization of Hilbert's ε-calculus for its own sake. More precisely we do not want to extend the basic language by the ε-symbol to express *new* things, but only to express *old* and reliable things as quantifiers in a (hopefully) more efficient way. Based on the machinery established in Section 3 we take a careful look at the result in [1]. This study will take place in Section 4.

The aim of the remaining part of the paper is mainly to relate the results obtained in [1] to those obtained in [8]. Such a comparison is hindered as the results of the former paper are expressed in a tableaux calculus that is based on the ε-formalism, while the latter results are completely based on 'usual' first-order tableaux. To overcome this problem, we will define in Section 6 a tableaux calculus, based on a liberalization of the δ-rule, called δ'-rule, that allows the same speed-up results as δ^{ε}-tableaux. The δ'-tableaux rules are simply rendered from the δ^{ε}-rules by consequently interpreting ε-terms as Skolem functions.

The relation of δ'-tableaux (and hence δ^{ε}-tableaux) with the further improved tableaux variant $\delta^{\star^\star}$-rule introduced in [8] is presented (together with further results) in Section 8.

Finally, in the spirit of [8], a further liberalization of the $\delta^{\star}$-rule [3] is defined. This liberalization allows a non-elementary speed-up over $\delta^{\star^\star}$-tableaux, cf. Section 9

2 Preliminaries

We assume familiarity with the basic concepts of tableau and resolution based theorem proving. However, we fix some notations. Recall that Smullyan introduced a classification scheme for logical operators (and thus for tableaux rules). Within this scheme there are four different types, α (conjunctive propositional), β (disjunctive propositional), γ (universal), and finally δ (existential). Given a δ-formula δ, the notation $\delta_0(x)$ will be used to denote the (semi-)formula $\phi(x)$, if $\delta = \exists x \phi(x)$ or $\neg\phi(x)$, if $\delta = \forall x \neg\phi(x)$. (We refer to [9] for a nice introduction into tableaux and [13] for definitions of clauses, substitutions, ground instances, factors etc.)

Below we present a number of different free tableaux calculi which all differ only in the definition of the δ-expansion rule. To be concise, let us call the proof procedures that are based on the usual tableaux rules and the different versions δ' of the δ-rule as δ'-tableaux. (We refer to [9] for a nice introduction into tableaux and to[13] for definitions of clauses, substitutions, ground instances, factors etc.) The *size*, $\|\tau\|$, of a resolution proof τ (or tableaux τ), is the number of clauses or formulas occurring in it.

It is notationally convenient to distinguish between *bound* and *free* variables. Bound variables will be denoted by lower-case letters from the end of the alphabet, while free variables will be denoted by upper-case letters from the beginning of the alphabet. *Terms* are constructed as usual from constants, variables, and function symbols; *semi-terms* are like terms but may also contain bound variables. *Formulas* are defined as usual with the proviso that only bound variables are allowed to be quantified and only free variables may occur free. *Semi-formulas* are defined as formulas but may also contain free occurrences of bound variables. W.l.o.g. we assume throughout this paper that each quantifier occurrence in a formula is unequivocally associated with a unique variable name. As usual, the set of *free* variables of a term or formula E is denoted as $\mathrm{Var}(E)$.

The *polarity* of sub-formulae be defined as usual. An occurrence of $\exists$ in a formula ϕ is called *weak* (*strong*) if it is the leading symbol of a positive (negative) sub-formula of ϕ. Dually for $\forall$.

We want to compare different versions of tableaux calculi on the minimal size of proofs. A natural way to compare the proof complexity of different proof systems is to relate them to the calculus-independent measure of *Herbrand complexity*. Note that in *refutational* proof systems, as e.g. tableaux, *Skolemization* eliminates *weak* quantifier occurrences.

Definition 1. *Let ϕ be an unsatisfiable first-order formula that only contains strong quantifier occurrences.Then* $\mathrm{HC}(\phi)$ *denotes the minimal cardinality over all unsatisfiable sets of ground instances of ϕ. (Herbrand's Theorem assures the existence and finiteness of such a set.)*

For every clause C we assign a closed formula $\mathrm{Plf}(C)$ by

$$\mathrm{Plf}(C) = \forall x_1 \cdots \forall x_n C\{A_1 \mapsto x_1, \ldots, A_n \mapsto x_n\} \text{ for } \mathrm{Var}(C) = \{A_1, \ldots, A_n\}$$

For sets of clauses $\mathcal{C}$, we define $\mathrm{HC}(\mathcal{C})$ as the minimal cardinality over all unsatisfiable sets of ground instances of clauses in $\mathcal{C}$.

Finally we define the *non-elementary function* 2_x as $2_0 = 2$ and $2_{x+1} = 2^{2_x}$.

3 Hilbert's ε-calculus

We extend the definition of terms to include ε-terms.

Definition 2. *If $\phi(A)$ is a formula, not containing the bound variable x, then $\varepsilon_x\phi(x)$ is a term; this term is called* ε-term.

If on the other hand x does occur at positions $p_1, \ldots, p_k$ in $\phi(a)$, we obtain a variant *ϕ' by replacing x at p_i for all $1 \leq i \leq k$ by some other bound variable y not already occurring in ϕ. The variant ϕ' is then used to form the ε-term $\varepsilon_x\phi'(x)$.*

Although the ε-symbol ε makes use of the *quantified* variable x, it should not be considered as a quantifier: Quantifiers are logical operators on formulas, while the ε-symbol acts on terms. Note that ε-terms $\varepsilon_x\phi(x), \varepsilon_z\phi(z)$ which are equal upto renaming of bound variables are considered as equal.

We consider instances of the following axiom schema

$$\exists x\phi(x) \supset \phi(\varepsilon_x\phi(x)) \tag{1}$$

The instances of this schema are sometimes called *critical axioms*. In [11], p. 12, an informal explanation of the ε-symbol is given. Roughly, the value of $\varepsilon_x\phi(x)$ is supposed to be a witness of $\exists x\phi(x)$, if there exists such a witness at all.[1]

Let **PL** denote some arbitrary Hilbert-type proof system for first-order logic with language $\mathcal{L}$. We extend **PL** with the axiom schema (1); the obtained proof system is called $\mathbf{T}'$. The presence of ε-terms in the language clearly extends the expressibility of the given language $\mathcal{L}$; conclusively the language of $\mathbf{T}'$ is denoted as $\mathcal{L}'$.

We write $A \leftrightarrow B$ to abbreviate the conjunction $A \supset B \wedge B \supset A$. For an arbitrary proof system **T**, we write $\mathbf{T} \vdash \phi$ to denote the derivability of ϕ in **T**. The *length*, $|\Pi|$, of a proof in **T** counts the number of steps used.

Lemma 1.
- $\mathbf{T}' \vdash \exists x\phi(x) \leftrightarrow \phi(\varepsilon_x\phi(x))$, *and*
- $\mathbf{T}' \vdash \forall x\phi(x) \leftrightarrow \phi(\varepsilon_x\neg\phi(x))$

Proof. We show only the latter equivalence, as the proof of the former is similar. The direction $\rightarrow$ follows from the substitution axiom $\forall x\phi(x) \supset \phi(t)$. While the other follows by an application of the axiom

$$\exists x\neg\phi(x) \supset \neg\phi(\varepsilon_x\neg\phi(x))$$

together with contraposition. □

Hence in $\mathbf{T}'$ quantifiers become definable; this suggest a (Hilbert-type) proof system, obtained from $\mathbf{T}'$ by deleting all quantifier rules. In addition we have to change the shape of the axiom schema to the logical equivalent

$$\phi(t) \supset \phi(\varepsilon_x\phi(x)) \tag{2}$$

where t denotes an arbitrary term. The new proof system is called **T**; its language is denoted as $\mathcal{L}^\star$. Following the convention in [11], we sometimes call the system **T** *(Hilbert's) ε-calculus.*

Definition 3. *For any formula ϕ in $\mathcal{L}'$, we define a formula $\phi^\star$ in $\mathcal{L}^\star$:*
- $\phi^\star = \phi$, *if ϕ is an atomic formula.*
- $(\phi \vee \psi)^\star = \phi^\star \vee \psi^\star$. *(Similar for $(\phi \wedge \psi)^\star$ and $(\phi \supset \psi)^\star$.*
- $(\exists x\, \phi(x))^\star = \phi^\star(\varepsilon_x\phi^\star(x))$.
- $(\forall x\, \phi(x))^\star = \phi^\star(\varepsilon_x\neg\phi^\star(x))$.

[1] It is quite interesting to note that the explanation given by Hilbert and his colleagues is different form the one given later in the literature, compare [2, 12].

Remark 1. In particular for any formula ϕ in $\mathcal{L}$, we can transform the formula $\phi^\star$ in $\mathcal{L}'$ back to ϕ. Simply employ the given transformation from right to left.

Based on this definition it is easy to prove an embedding of $\mathbf{T}'$ into $\mathbf{T}$.

Proposition 1. – *If* $\mathbf{T}' \vdash \phi$*, then* $\mathbf{T} \vdash \phi^\star$-

– *If* $\mathbf{T}' \vdash \phi$ *for closed* ϕ*, then there exists a derivation* Π *of* $\phi^\star$ *in* $\mathbf{T}$ *such that all formulas in* Π *are closed.*

Note that the proposition remains valid if '$\mathbf{T}'$' is replaced by '**PL**'. Hence, as immediate corollary we see that for any formula ϕ valid in first-order logic, there exists a number l, a sequence of formulas $\psi_1(t_1), \ldots, \psi_l(t_l)$, and a sequence of terms $t_1, \ldots, t_l$ s.t.

$$\bigwedge_{i=1}^{l} (\psi_i(t_i) \supset \psi_i(\varepsilon_x \psi_i(x))) \supset \phi^\star$$

is a tautology.

3.1 Semantics

We have not yet given any semantical interpretation of the Hilbert-type proof system $\mathbf{T}$. It may appear that the above equivalences, together with the informal explanation stated, naturally induce a reasonable semantical interpretation. However, to our best knowledge no commonly accepted semantics for Hilbert's ε-calculus exist. See e.g. [12, 1, 7, 15] for some attempts in this direction. In the following, we will present an example, which (at least partly) explains the difficulty of this task.

Example 1. Consider

$$\phi(\varepsilon_x \neg\phi(x,x), \varepsilon_x \neg\phi(x,x)) \tag{3}$$

$$\phi(\varepsilon_x \neg\phi(x,x), \varepsilon_y \phi(\varepsilon_x \neg\phi(x,x), y)) \tag{4}$$

where ϕ is atomic. Note that the implication (3) $\supset$ (4) is an instance of the critical axiom

$$\phi(\varepsilon_x \neg\phi(x,x), A) \supset \phi(\varepsilon_x \neg\phi(x,x), \varepsilon_y \phi(\varepsilon_x \neg\phi(x,x), y))$$

Hence there exists a derivation in $\mathbf{T}$ of (4) from (3). Observe that (3) equals $(\forall x\, \phi(x,x))^\star$. $\square$

We assume a 'suitable' semantic interpretation of the ε-symbol, s.t. $\mathbf{T}$ is actually sound (wrt. this semantics). Let $\mathcal{M}$ be a structure over the signature $\mathcal{L}^\star$, which is in accordance with the assumed semantics; As usual we define an evaluation function $v_{(\mathcal{M})}$ on M. Furthermore we assume that $v_{\mathcal{M}}((\forall x(\phi(x,x))^\star) = \mathbf{true}$.

Hence

$$\phi(\varepsilon_x \neg\phi(x,x), \varepsilon_y \phi(\varepsilon_x \neg\phi(x,x), y))$$

has to be true in $\mathcal{M}$. This induces the question: *Is it possible to define a formula* ψ *in* $\mathcal{L}$ *such that either* $\psi^\star = (4)$ *or at least* $\mathbf{T} \vdash \psi^\star \leftrightarrow (4)$.

We consider the question wrt. a seemingly good candidate: Let ψ be equal to $\forall x \exists y\ \phi(x,y)$. We apply the transformation $()^\star$ from inside to outside. The formula $\exists y\ \phi(A,y)$ becomes $\phi(A,\varepsilon_y\phi(A,y))$; while $\forall x\ \phi(x,\varepsilon_y\phi(x,y))$ becomes

$$\phi(\varepsilon_x\neg\phi(x,\varepsilon_y\phi(x,y)),\varepsilon_y\phi(\varepsilon_x\neg\phi(x,\varepsilon_z\phi(x,z)),y)) \tag{5}$$

Unfortunately this formula is different from (4): Every occurrence of the semi-term $\varepsilon_y\phi(x,y)$ in (5) is replaced in (4) by the (bound) variable x. Furthermore, the following lemma shows that we cannot even prove the equivalence of (5) and (3) in **T**.

Lemma 2. – $\mathbf{T} \vdash (5) \supset (4)$, *but*
– $\mathbf{T} \not\vdash (4) \supset (5)$

Proof. The first part of the lemma is easy. The implication follows by a single application of the schema (2) together with contraposition.

The second part is less trivial. To show the unprovability we make use of the assumed semantics. We only assert the (rather natural) requirement, that the given semantics interprets an ε-term $\varepsilon_x\phi(x)$ as a witness of $\exists x\phi(x)$, if there exists such a witness at all. Now, we argue indirectly. Assume $(4) \supset (5)$ is derivable in **T**; let a structure $\mathcal{N}$ (according to the assumed semantics) is given. Assume further the interpretation $\phi^{\mathcal{N}}$ of ϕ is chosen, s.t. $\phi^{\mathcal{N}}(c_{\mathcal{N}},c_{\mathcal{N}})$ is true, for some element $c_{\mathcal{N}}$ of the domain, and false everywhere else. (The cardinality of the domain of N is >1.)

By assumption and elimination of constants

$$\phi(c,\varepsilon_y\phi(c,y)) \supset \phi(\varepsilon_x\neg\phi(x,\varepsilon_y\phi(x,y)),\varepsilon_y\phi(\varepsilon_x(\neg\phi(x,\varepsilon_z\phi(x,z))),y))$$

becomes derivable. Set $v_{\mathcal{N}}(c) = c_{\mathcal{N}}$, set $v_{\mathcal{N}}(\varepsilon_x\phi(c,y)) = c_{\mathcal{N}}$, and assigned an arbitrary element in the domain of $\mathcal{N}$, different from $c_{\mathcal{N}}$, otherwise. With this assignment we obtain

$$v_{\mathcal{N}}(\varepsilon_x\neg\phi(x,\varepsilon_y\phi(x,y))) = d_{\mathcal{N}} \tag{6}$$

$$v_{\mathcal{N}}(\varepsilon_y\phi(\varepsilon_x\neg\phi(x,\varepsilon_z\phi(x,z)),y)) = d_{\mathcal{N}} \tag{7}$$

for an arbitrary element $d_{\mathcal{N}}$ in the domain. Note that $\neg\phi(d,\varepsilon_y\phi(d,y))$ is true if the constant d is interpreted as $d_{\mathcal{N}}$. Hence we conclude $v_{\mathcal{N}}(\phi(c,\varepsilon_y\phi(c,y))) = \mathbf{true}$, and

$$v_{\mathcal{N}}(\phi(\varepsilon_x\neg\phi(x,\varepsilon_y\phi(x,y)),\varepsilon_y\phi(\varepsilon_x\neg\phi(x,\varepsilon_z\phi(x,z)),y))) = \mathbf{false}$$

□

We have shown that the setting $\psi = \forall x\exists y\ \phi(x,y)$ is not suitable to answer our question. In a similar, but simpler way, the other possibilities for the formula ψ are excluded. We conclude that the ε-calculus is more expressible than standard Hilbert-type proof systems for first-order logic. In summary we have the following proposition.

Proposition 2. *Let $\mathcal{M}$ be defined as above. Set θ equal to*

$$\phi(\varepsilon_x\neg\phi(x,x),\varepsilon_y\phi(\varepsilon_x\neg\phi(x,x),y))$$

Then no formula ψ in $\mathcal{L}$ exists such that either $\psi^\star = \theta$ or at least $\mathbf{T} \vdash \psi^\star \leftrightarrow \theta$.

The reason for this seems to be that the axiom schema (2) allows the replacement of *semi-terms* by bound variables. (In our example the $\varepsilon_y\phi(x,y)$ is replaced by the bound variable x.) In standard first-order logic, such behavior is prohibited.

This shows that the reversed implication of Proposition 1 fails. However, we have the following fact, expressing that no problem arises as long as we are only employing $\mathbf{T}'$ wrt. formulas free of ε-terms.

Proposition 3 ([11]). *Let the proof systems* $\mathbf{T}'$ *and* **PL** *be defined as before. If* $\mathbf{T}' \vdash \phi$ *with a deduction* Π *such that* ϕ *is in* $\mathcal{L}$*, then we can transform the proof* Π *into a* **PL**-*derivation of* ϕ.

3.2 The Epsilon Calculus with Equality

In this section we want to comment on the computational effects of axiomatizing equality in the presence of the ε-rule. It is clear that we will need the following principle, which is usually called ε-*equality axiom*.

$$s = t \supset \varepsilon_x\phi(x,s) = \varepsilon_x\phi(x,t)$$

We extend the system **T** by this axiom schema. The obtained proof system is denoted as $\mathbf{T}^\star$. (Note that the ε-equality axiom is restricted form of the ε-*extensionality axiom*, see e.g. [1, 12].) We employ Yukami's trick [16]; set $0^k \stackrel{\text{def}}{=} \underbrace{0 + (0 + \cdots (0 + 0))}_{k \text{ times } 0}$.

Proposition 4. *Using two instances of the following restricted scheme of identity*

$$s = 0 \supset g(s) = g(0) \tag{8}$$

we can derive $0^k = 0$ *from (i)* $0 + 0 = 0$*, (ii)* $\forall x,y,z\ x = y \wedge y = z \supset x = z$*, and (iii)* $\forall x,y\ x + y = y \supset x = 0$ *in constant length for any k.*

Proof. The following equalities can be derived if we employ suitable instances of (8) together with additional instances of the transitivity axiom (ii) and axiom (i).

$$\begin{aligned} 0^n + (0^{n-1} + \cdots + (0^2 + 0)) &= \\ 0^{n-1} + (0^{n-2} + \cdots + (0 + 0)) &= \\ 0^{n-1} + (0^{n-2} + \cdots + (0^2 + 0)) & \end{aligned}$$

Hence we have derived $0^n + A = A$, where $A = 0^{n-1} + (0^{n-2} + \cdots + (0^2 + 0))$; we employ axiom (iii) to obtain the desired result. □

Proposition 5. *There exists an existential formula* ϕ*, derivable in* $\mathbf{T}^\star$*, such that no function can exist which limits* $\mathrm{HC}(\neg\phi)$ *in the length of the proof.*

Proof. First we derive the restricted scheme of identity (8), by the use of ε-equality axiom. We start by deriving $\varepsilon_x(x = g(A)) = g(A)$ from $g(A) = g(A)$ and $g(A) = g(A) \supset$

$\varepsilon_x(x = g(A)) = g(A)$. Now using ε-equality, $s = t \supset \varepsilon_x(x = g(s)) = \varepsilon_x(x = g(t))$ together with reflexivity and transitivity, we obtain $s = t \supset g(s) = g(t)$ and hence (setting $t = 0$) $s = 0 \supset g(s) = g(0)$ becomes derivable.

Now we prove the proposition. We use the same notation as in Proposition 4. Due to this proposition and the fact that the restricted identity schema is derivable in $\mathbf{T}^\star$, we conclude the existence of *uniform* proofs, i.e. proofs with fixed length N of

$$(\forall x(x = x))^\star \wedge (i)^\star \wedge (ii)^\star \wedge (iii)^\star \supset 0^k = 0$$

for any k.

Now assume the existence of a bound on HC(¬ϕ) in the length of these proofs N. Hence, by completeness of **PL**, the formula

$$\forall x(x = x) \wedge (i) \wedge (ii) \wedge (iii) \supset 0 + (0 + \cdots A \cdots) = 0$$

for some free variable has to be **PL**-provable, too. This is absurd. □

Remark 2. Let ϕ be defined as in the proposition. Then the proposition actually shows that the proof system $\mathbf{T}^\star$ allows *unbounded* speed-up over the minimal size of proofs of ϕ either in tableaux or resolution proof procedures, compare Section 7.

Note that 'full first-order calculi, like Gentzen's sequent calculus with cuts, *only* allow non-elementary speed-up over tableaux or resolution proof procedures.

The unbounded speed-up is gained as the restricted scheme of identity (8) is derivable in $\mathbf{T}^\star$. Which on the other hand implies an explosion of the search space, if we attempt to use the proof system $\mathbf{T}^\star$ for proof search. Note that this effect has nothing to do with the fact that we have formalized equality through axioms. If we would use a more elegant formalizations for equality, as e.g. paramodulation, we still have to take care of the presence of the ε-symbol.

From this we see, how difficult, if not impossible to define a calculus suitable for proof-search that formalizes the ε-calculus plus ε-equality.

4 Automated Theorem Proving and the Epsilon Calculus

In this section we reassess the δ^ε-tableaux introduced in [1]. W. Ahrendt, and M. Giese introduce a free-variable tableaux, denoted δ^ε-tableaux, which includes an ε-*expansion rule*. This expansion rule formalizes the axiom schema (1) presented in Section 3. The tableaux rules are given in Table 1 and Table 2.

Table 1. δ^ε-tableaux rules for quantified formulas

$$\frac{\gamma}{\gamma_0(A)}\ (\gamma - \text{rule}) \qquad \frac{\exists x \phi(x)}{\phi(\varepsilon_x \phi(x))}\ (\delta^\varepsilon - \text{rule}) \qquad \frac{\neg\forall x \phi(x)}{\neg\phi(\varepsilon_x \neg\phi(x))}\ (\delta^\varepsilon - \text{rule})$$

where A is a free variable.

Table 2. ε-expansion rule
$\overline{\forall x \neg\phi(x) \quad | \quad \phi(\varepsilon_x \phi(x))}$

It is easy to see that the δ^ε-tableaux is equivalent to the Hilbert-type proof system $\mathbf{T}'$ introduced in Section 3. Note that for a closed formula ϕ the ε-expansion rule actually amounts to the *cut*-rule. Hence in [1] further restrictions are introduced to control the application of this rule. Before we go into further details we want to study the unrestricted tableaux calculus some more.

First it comes a surprise that the critical axioms are directly formalized as a separate rule. One is tempted to think that the same effect should have been possible by a suitable modification of the closure rule. Recall the definition of closed tableau: A tableau T is closed, if there is a substitution σ such that every branch in $T\sigma$ contains a complementary pair of formulas. We alter this to take care of the ε-expansion rule.

Definition 4. *A tableau T is* ε-closed*, if there is a substitution σ such that there exists formulas ϕ', ϕ'' and $\phi'\sigma = \phi\{A \mapsto \varepsilon_x \phi(x)\}$ with $\phi''\sigma = \neg\phi\{A \mapsto t\}$, where A is a free variable (not necessarily occurring in ϕ', ϕ'') and t any term.*

(It is obvious that the given closure rule can be simulated by one application of the ε-expansion rule.) Unfortunately this revised definition of closure is defective, as it encodes full second-order unification, which is well-known to be undecidable, compare [5].

Now we employ the machinery presented in the previous section. For clarity reasons we denote the δ^ε-tableaux minus γ- and δ^ε-rule by a separate name.

Definition 5. *Consider a free-variable tableaux consisting of α and β rules, together with the following variant translation of the ε-expansion rule*

$$\overline{\neg\phi(A) \quad | \quad \phi(\varepsilon_x \phi(x))}$$

where A is a free variable. (Note that this rule is equivalent to the axiom schema (2).) This tableaux system is called ε-tableaux*.*

An ε-tableaux T is closed, if there is a substitution σ such that every branch in $T\sigma$ contains a complementary pair of formulas.

The following proposition is an easy corollary to Proposition 1 and shows that δ^ε-tableaux (without restrictions on the ε-rule) is equivalent to ε-tableaux.

Proposition 6. *Let ϕ be an arbitrary formula in $\mathcal{L}'$, in particular ϕ may contain quantifiers together with ε-terms.*

Assume ϕ is δ^ε-tableaux provable, then $\phi^\star$ is ε-tableaux provable and vice versa. In particular let T be a closed δ^ε-tableaux for ϕ. Then we can define a linear transformation of T into a closed ε-tableaux T' for $\phi^\star$ (and vice versa).

Proof. It suffices to show one direction, the other one is easy. We conclude from Proposition 1 that any closed tableaux for a formula ϕ in δ^ε-tableaux, can be transformed to a closed ε-tableaux T' of $\phi^\star$. The second part of the proposition follows directly from the proof of Proposition 1. □

Now we complete the presentation of the δ^{ε}-tableaux: The application of this rule is only allowed if for the ε-term $\varepsilon_x\phi(x)$

1. the branch contains an atomic formula $(\neg)\phi(t_1,\ldots,t_n)$ such that
   (a) $\varepsilon_x\phi(x)$ is a subterm of one of the t_i and
   (b) $\varepsilon_x\phi(x)$ is a term, not a semi-term,
2. $\varepsilon_x\phi(x)$ was not previously introduced by the δ^{ε}-rule, and
3. the ε-rule was not previously applied for $\varepsilon_x\phi(x)$ on this branch.

Even with this restriction δ^{ε}-tableaux turns out to be sound and complete for *substitutive structures*. (For our purpose, we need not be too precise on this semantics. We only want to point out that the *substitutive semantics* defined is a restriction of the *extensional semantics* described in [12].) It is not difficult to see that the introduced restrictions do not considerably reduce the power of the Hilbert's ε-calculus.

Proposition 7. *Let ϕ be an arbitrary formula in $\mathcal{L}'$, in particular it may contain quantifiers together with ε-terms.*

Assume ϕ is δ^{ε}-tableaux provable (with restrictions), then there exists ψ, s.t. ϕ is logically equivalent to ψ, and $\psi^{\star}$ is ε-tableaux provable and vice versa.

Proof. Let Π be a δ^{ε}-tableaux proof of ϕ. Assume w.l.o.g only one ε-expansion rule is employed in Π which doesn't fit the restrictions given above. Denote the critical ε-term as $\varepsilon_x\varphi(x)$. Assume that the restriction (1a) is violated. Then define

$$\psi = \phi \wedge (p(\varepsilon_x\varphi(x)) \vee \neg p(\varepsilon_x\varphi(x)))$$

Clearly ψ is logical equivalent to ϕ. Hence this takes care of restriction (1a).

Now consider the other restrictions: The case (1b) is trivially fulfilled for any ε-tableaux by definition. The case (2) is void, as ε-tableaux does not admit δ-rules. The last case, is a little bit more difficult, but it is easy to see that this restriction can only occur if the ε-rule is applied wrt. the *same* formula ϕ. Hence the branch contains a cycle that can either be eliminated or otherwise Π cannot be a proof. □

Remark 3. Note that our argumentation to prove the propositions is purely proof-theoretically, in particular no use of any semantics of Hilbert's ε-calculus or any form of completeness is made.

Remark 4. It was already noted in [1] that this schema is at least as powerful as a *cut*-rule. However, it is an easy corollary of the proposition, which seems to have been overlooked by W. Ahrendt and M. Giese, that even with the introduced restrictions, there exists a $\delta^{\star}$-tableaux proof of Statman's example [14]. Hence $\delta^{\star}$-tableaux (with restrictions) admits non-elementary speed-up over any 'usual' tableaux.

Now we consider an ε-tableaux proof of

$$(\forall x\, \phi(x,x))^{\star} \supset \phi(\varepsilon_x\neg\phi(x,x), \varepsilon_y\phi(\varepsilon_x\neg\phi(x,x),y))$$

Example 1 (continued).

$$(1)\ (\forall x\, \phi(x,x))^\star$$

$$(2)\ \neg\phi(\varepsilon_x\neg\phi(x,x),\varepsilon_y\phi(\varepsilon_x\neg\phi(x,x),y))$$

$$(3,\varepsilon)\neg\phi(\varepsilon_x\neg\phi(x,x),\varepsilon_x\neg\phi(x,x)) \qquad (6,\varepsilon)\phi(\varepsilon_x\neg\phi(x,x),\varepsilon_y\phi(\varepsilon_x\neg\phi(x,x),y))$$

$$\times \qquad\qquad \times$$

Remark 5. Almost the same tableaux is obtained if we consider the restricted version of δ^ε-tableaux with the formula

$$\forall x\, \phi(x,x) \supset \phi(\varepsilon_x\neg\phi(x,x),\varepsilon_y\phi(\varepsilon_x\neg\phi(x,x),y))$$

Hence, Proposition 2 remains valid, if we replace the reference to the proof system **T** either by the ε-tableaux system or the restricted version of δ^ε-tableaux: We are faced with sentences, which cannot be (re-)translated into formulas in $\mathcal{L}$ employing the $()^\star$ transformation.

We collect the pieces. The above proposition destroys our hopes that we can use Hilbert's ε-calculus as a *black box*, as motivated in the Introduction. Therefore, the translation of formulas in $\mathcal{L}^\star$ back to formulas in $\mathcal{L}$ fails, at least with respect to the (quite natural) transformation $()^\star$. Furthermore this translation even fails if we restrict our attention to 'interesting' formulas in $\mathcal{L}^\star$, namely those that can occur as (satisfiable) formulas in (failed) proof attempts.

Leaving this problem aside, we conclude that any computational calculus which proves the same formulas as the Hilbert-type system **T** will need to incorporate the full power of the schema of critical axioms. This follows by Proposition 7 together with the statement previous to Definition 5. Furthermore, if we want to reason about equality, any computational calculus based on Hilbert's ε-formalism will feature the same *unbound* speed-up over 'usual' variants of tableaux; clearly this speed-up comes with the cost of an explosion in search space.

These facts should raise some doubts whether we should employ the ε-formalism in *automated* reasoning, at all. Clearly this does not imply that the ε-calculus may not be fruitfully employed in *interactive* theorem proving, compare [1].

5 Variants of Tableaux

In recent years major improvements upon the basic calculus of Beth and Hintikka have been described, which focus on different versions of δ-expansion, i.e. on different Skolemization techniques.

In [10] a liberalization of the original δ-rule is presented: Only the variables occurring in δ are taken as arguments of the introduced Skolem function. This version of δ-expansion is called δ^+-rule. A further liberalization of the δ-rule is formulated in [6]. Now we may assign a Skolem function $f_{[\delta]}$ to each equivalence class $[\delta]$, containing all formulas that are identical to δ up to renaming of variables. (See [6] for a formal definition.)

Recall that δ^{+^+}-tableaux admit a *non-elementary* speed-up over δ-tableaux, cf. [3]. In [3] a further liberalization of δ-expansion, called $\delta^\star$-rule, is defined. This rule is based on the notion of *relevant* variables. (See [3] for a formal definition. A very similar, but sligthly improved variant is defined in Section 9.) The $\delta^\star$-rule is presented in Table 3.

Note that these liberalizations form a sort of hierarchy: Less and less variables are taken into account in the definition of Skolem functions. It is quite reasonable to argue that the real nature of these speed-ups comes about by the elimination of irrelevant variables in the definition of Skolem functions, compare [3].

Table 3. $\delta^\star$-rule

$$\frac{\delta}{\delta_0(f_{[\delta]}(A_1,\dots,A_n))}\ (\delta^\star\text{-rule})$$

where $A_1,\dots,A_n$ are the *relevant variables* of δ and $f_{[\delta]}$ is defined as in [3].

In [3] it is demonstrated that $\delta^\star$-tableaux allows a further non-elementary speed-up over δ^{+^+}-tableaux. Arguably these speed-ups are the result of ignoring irrelevant variables as arguments to Skolem functions.

The rules of δ^ε-tableaux [1] have already be presented in Section 4. The cited paper contains an example of an exponential speed-up over δ^{+^+}-tableaux. The authors claim that the δ^ε-tableaux may be strengthened, in a way such that it is strictly stronger than $\delta^\star$-tableaux. Furthermore the question is raised whether the method of [3] can be employed successfully for δ^ε-tableaux. We will comment on the raised questions below.

Finally in [8] yet another liberalization of the δ-rules is presented, called $\delta^{\star^\star}$-rule which combines the ideas present in [1] and in [3]. Contrary to [3] a *recursive* definition of relevance is defined together with the notion of a *key formula*. (See [8] for a formal definition.) Here the first concept allows the reduction argument places that need to be taken into account, the second idea allows the reduction of different Skolem function that need to be introduced. The cited paper presents (i) an example of an exponential speed-up over $\delta^\star$-tableaux which comes about as the introduction of different Skolem functions is prevented and (ii) a non-elementary speed-up over $\delta^\star$-tableaux which is mainly due to the recursive definition of relevance.

We have not yet given a definition of HC for arbitrary formulas; instead $\mathrm{HC}(\phi)$ is only defined if ϕ does not contain weak quantifier occurrences. For δ^{+^+}, $\delta^\star$-tableaux, and $\delta^{\star^\star}$-tableaux there is a natural way to evade this problem: Since there is a one-one correspondence between introduced Skolem functions and occurrences of weak quantifiers in the refuted formula every such tableau τ determines a Skolemized form of the formula: Traverse τ bottom-up and replace each occurrence of a δ-formula by the corresponding δ_0 formula. Let us call the resulting formula on the root of τ the *τ-Skolemization of F*, $\mathrm{Sk}_\tau(F)$. Based on that, the following proposition is easy to see.

Proposition 8. *For any closed δ^{+^+}-,$\delta^\star$, or $\delta^{\star^\star}$-tableaux τ for F we have* $\|\tau\| > \mathrm{HC}(\mathrm{Sk}_\tau(F))$.

Note that this proposition doesn't carry over easily to δ^{ε}-tableaux as there we introduce ε-terms instead of Skolem terms. In the case of δ- and δ^{+}-tableaux no unique Skolemization is defined, hence (the proof of) Proposition 8 doesn't carry over. However, it is clear that employing δ- and δ^{+}-tableaux (instead of δ^{++}- or $\delta^{\star}$-tableaux) can only increase the size of the tableaux.

6 From Epsilon Terms to Skolem Functions

In this section we show how we can preserve the exponential speed-up of δ^{ε}-tableaux over δ^{++}-tableau, without the need to axiomatize the full ε-calculus.

Definition 6. *Assume a formula ϕ can be written as $\psi(t_1,\dots,t_n)$ such that all (maximal) occurring terms are indicated. Then $\psi(A_1,\dots,A_n)$ is called an* abstraction *of ϕ.*

Note that the notion of an abstraction is similar to the concept of a *key formula* as defined in [8]. We define for each δ-formula in $\mathcal{L}$ an equivalence class $[\![\delta]\!]$ such that $[\![\delta]\!]$ consists of all formulas whose abstractions are equal upto renaming of free variables. All formulas in an equivalence class $[\![\delta]\!]$ are assigned the same new Skolem function $f_{[\![\phi]\!]}$. We define for the language $\mathcal{L}$

$$\mathrm{Sk}(\mathcal{L}) = \mathcal{L} \cup \{ f_{[\![\phi]\!]} : \phi \in \mathcal{L}, \phi \text{ is an abstraction of a } \delta\text{-formula}\}$$

Based on that we can inductively define an extension $\mathcal{L}'$ of $\mathcal{L}$ including all Skolem functions obtained by the elimination of weak quantifiers in a given formula F, compare [6]: $\mathcal{L}^0 \stackrel{\text{def}}{=} \mathcal{L}$ and $\mathcal{L}^{n+1} \stackrel{\text{def}}{=} \mathrm{Sk}(\mathcal{L}^n)$. Finally we set $\mathcal{L}' = \bigcup_i \mathcal{L}^i$.

Example 2. The formulas $\exists x\phi(x,A,B)$ and $\exists z\phi(z,z,z)$ do not belong to the same equivalence class. On the other hand the formulas $\exists x\phi(x,A,b)$, $\exists x\phi(x,a,B)$, a,b constants do belong to the same class and therefore are assigned the same function symbol $f_{[\![\exists x\phi(x,A,B)]\!]}$. □

Table 4. δ'-rule

$$\frac{\delta}{\delta_0(f_{[\![\phi]\!]}(A_1,\dots,A_n)\sigma} \ (\delta'\text{-rule})$$

Let $\phi = \psi(A_1,\dots,A_n)$ denote an abstraction of δ_0, where $A_1,\dots,A_n$ are the fresh free variables. Then there exists σ, s.t. the domain of σ is a subset of $\{A_1,\dots,A_n\}$, s.t. $\delta = \psi(A_1,\dots,A_n)\sigma$. The Skolem function $f_{[\![\phi]\!]}$ is defined as above.

Before we show soundness and completeness, we restate an example from from [1]; in the presentation of [1] f actually denotes an ε-term.

Example 3.

$$(1)\ \forall x p(x,a,b)$$

$$(2)\ \forall u \exists x \neg p(x,u,b) \vee \forall v \exists z \neg p(z,a,v)$$

$$(3,\gamma(1))\ \ p(A,a,b)$$

$(4,\beta(2))\ \forall u \exists x \neg p(x,u,b)$	$(7,\beta(2))\ \forall v \exists z \neg p(z,a,v)$
$(5,\gamma(4))\ \exists x \neg p(x,B,b)$	$(8,\gamma(7))\ \exists x \neg p(x,a,C)$
$(6,\delta'(5))\ \neg p(f(B,b),B,b)$	$(9,\delta'(8))\ \neg p(f(a,C),a,C)$
$\times$	$\times$

The tableaux becomes closed with the substitution $\{A \mapsto f(a,b), B \mapsto a, C \mapsto b\}$. The interesting aspect of this example is that in both branches the *same* function f is used. Clearly 'ordinary' tableaux would need two different Skolem functions, therefore increasing the size of the minimal proof. We will take up this example in Theorem 4.

Now we show soundness and completeness. It suffices to consider soundness as the completeness of standard tableaux can be easily adapted to δ'-tableaux, compare [9]. Let $\mathcal{M}$ denote a first-order interpretation, and μ a variable assignment. (The evaluation function $v_{(\mathcal{M},\mu)}$ is defined as usual.) A branch B of a δ'-tableau is called *satisfied* by $(\mathcal{M},\mu)$ if $v_{(\mathcal{M},\mu)}(\phi) = \textbf{true}$ for all formulas ϕ in B. Alternatively, we write $(\mathcal{M},\mu) \models B$. A tableau T is satisfiable if there exists an interpretation $\mathcal{M}$ and for all variable assignment μ there exists a branch B in T such that $(\mathcal{M},\mu) \models B$ holds.

Lemma 3. *If T is a tableaux whose root is labeled by a satisfiable closed formula, then T is satisfiable.*

Proof. First we define a partial interpretation $\mathcal{M}$ over the signature $\mathcal{L}'$. If simplifies our definition if we introduce the notion of *rank* of a Skolem-function f: The rank of f is the least number n such that $f \in \mathcal{L}^n$, where $\mathcal{L}'$ equals $\bigcup_i \mathcal{L}^i$ as above.

Let f denote a Skolem function of rank $n+1$. We fix a certain variable assignment μ; now suppose $(\mathcal{M},\mu) \models \delta(\overline{A},\overline{t})$ ($\overline{A}$ denotes the tuple of free variables in δ, while $\overline{t}$ denotes the tuple of closed terms therein.) By definition there exists an element $c_{\mathcal{M}}$ of the domain of $\mathcal{M}$ such that[2] $(\mathcal{M},\mu\{B \mapsto c_{\mathcal{M}}\}) \models \delta_0(B,\overline{A},\overline{t})$. Therefore we set

$$v_{(\mathcal{M},\mu)}(f(v_{(\mathcal{M},\mu)}(\overline{A}), v_{(\mathcal{M},\mu)}(\overline{t}))) = c_{\mathcal{M}}$$

Otherwise the evaluation of f remains undefined. This completes the definition of the partial interpretation $\mathcal{M}$.

Based on $\mathcal{M}$, the lemma is proven by induction on the expansion rules. The proof is essentially the same as the proof given in [6] and hence omitted. $\square$

Lemma 4. *Let T be a satisfiable tableau and σ a substitution that associates with every free variable in T a term in the language of T, then $T\sigma$ is satisfiable.*

[2] Since f is of rank $n+1$ the symbols in δ are from the language $\mathcal{L}^n$.

Proof. The claim follows as in [6].

In summary we have obtained the following theorem.

Theorem 1. *If T is a closed tableau, whose root is labeled by the closed formula $\neg\phi$, then ϕ is universally valid.*

7 Resolution, Functional Extension, and Tableaux

Recall that [14] demonstrate that proofs in 'full' logical calculi (like Gentzen- or Hilbert-style calculi) can be non-elementarily shorter than the Herbrand complexity of the proven formula. However, if we consider resolution, the size of the resolution refutation of an unsatisfiable clause set $\mathcal{C}$ can be at most exponentially smaller than $\mathrm{HC}(\mathcal{C})$, compare [4].

We review some concepts and results from [4]. Consider the valid first order schema $(\forall x)(\phi(x) \vee \psi(x)) \supset (\mathsf{Q}x)\phi(x) \vee (\mathsf{Q}^d x)\psi(x)$, where Q and Q^d are dual quantifiers.

Definition 7. *Let $\mathcal{C}$ be a set of clauses and $C_1 \vee C_2 \in \mathcal{C}$. Let $A_1,\ldots A_n, B$ be the (free) variables that occur in both, C_1 and C_2. Then*

$$\mathcal{C} \cup \{C_1 \vee C_2\{B \to f(A_1,\ldots A_n)\}\},$$

where f is a new function symbol, is called a 1-F-extension *of $\mathcal{C}$.*

If the context uniquely determines the clause set in question we will call $C_1 \vee C_2\{B \to f(A_1,\ldots A_n)\}$ a 1-F-extension of the clause $C_1 \vee C_2$. Clearly the 1-F-extension of a clause-set $\mathcal{C}$ cannot be logically equivalent to $\mathcal{C}$. However, the extension rule preserves satisfiability and if we augment a (refutationally) complete resolution calculus by this extension rule we obtain a complete and correct calculus. (See [4] for further details.) We denote the resulting calculus as *1-F-resolution*.

Proposition 9. *There exist a set of clauses $\mathcal{C}$ and a sequence of literals C_n such that*

- *the Herbrand complexities as well as the size of the shortest resolution refutations of $\mathcal{C} \cup \{C_n\}$ are $\geq 2_n \cdot c$ for some constant c (and the non-elementary function 2_n defined in Section 2).*
- *but there are 1-F-resolution refutations of $\mathcal{C} \cup \{C_n\}$ of size $\leq 2^{dn}$ for some constant d.*

To simulate Proposition 9 in the context of tableau proofs we define for any 1-F-extension step $C_1 \vee C_2 \Rightarrow C_1 \vee C_2\{B \mapsto f(A_1,\ldots A_n)\}$ of a 1-F-resolution proof the corresponding *justifying formula*: $(\forall x_1)\cdots(\forall x_n)(\forall y)[(C_1 \vee C_2) \supset (C_1 \vee (\exists z)C_2\{y \mapsto z\})]$. where $B, A_1,\ldots A_n$ are all free variables occurring in C_1 and C_2.

Proposition 10. *Any 1-F-resolution refutation ρ of a clause set $\mathcal{C}$ can be translated into a closed δ^{++}-, $\delta^{\star}$, $\delta^{\star^\star}$, or δ'-tableau τ_ρ for $\mathrm{Plf}(\mathcal{C}) \wedge J$, where J is the conjunction of all justifying formulae corresponding to 1-F-extension steps in ρ. Moreover, $\|\tau_\rho\| < (d+1)^{c\|\rho\|}$, for some constant c, where d is the maximal number of literals occurring in a clause of ρ.*

8 Some Speed-up results

To show a non-elementary speed-up of $\delta^{\star}$-tableaux over δ^{++}-tableaux [3] define the following *variant* of a justifying formula:

$$(\forall x_1)\cdots(\forall x_n)(\forall y)[(C_1 \vee C_2) \supset (C_1 \vee (\exists z)[C_2\{y \to z\} \vee (p(y) \wedge \neg p(y))])] \quad (9)$$

Observe that these variants are logically equivalent to the original formulas.

Lemma 5. *Let ϕ be an quantifier-free formula; let $\phi' = \forall \bar{y}\phi(\bar{y}) \wedge \psi_1 \wedge \cdots \wedge \psi_n$ where $\psi_i = \forall \bar{x}, y[(C_1^i \vee C_2^i) \supset (C_1^i \vee [C_2^i\{z \to f^i(\bar{x},y)\} \vee \varphi^i])]$, such that the φ^i are contradictory formulas and the f^i do not occur in ϕ nor in any ψ_j, if $j < i$. Then $\mathrm{HC}(\phi') \geq \mathrm{HC}(\phi)$.*

Proof. It suffices to prove the following claim; the lemma will then follow by inductively applying the claim.

Claim. Let $\phi' = \forall \bar{y}\phi(\bar{y}) \wedge \forall \bar{x}, y[(C_1 \vee C_2) \supset (C_1 \vee [C_2\{z \to f(\bar{x},y)\} \vee \varphi]$ for some contradictory formula φ. Then $\mathrm{HC}(\phi') \geq \mathrm{HC}(\phi)$.

We consider the smallest unsatisfiable set of instances $\mathcal{F}' \stackrel{\text{def}}{=} \{(\forall \bar{z}\phi(\bar{z}) \wedge \psi)\rho_i : 1 \leq i \leq n\}$ of ϕ', where ψ is defined as in the statement of the claim. We will show that already the set $\{(\forall \bar{z}\phi(\bar{y}))\rho_i : 1 \leq i \leq n\}$ is unsatisfiable. First we order the set of instances $f(\bar{t}, t')$ of $f(\bar{z}, y)$ occurring in $\mathcal{F}'$ according to the sub-term relation. Now we replace each maximal instance $f(\bar{t}, t')$ by t'. It is easy to see two facts: Firstly, as only maximal instances (in the defined partial ordering) are replaced, other instances of $f(\bar{z}, y)$ are unaffected. Secondly the formulas $\psi\rho_i$ either keep their shape or become tautologies. We repeat this procedure till all the $\psi\rho_i$ become tautologies. Now, we observe that we may remove these tautologies. Hence we obtain that the set $\{(\forall \bar{z}\phi(\bar{z}))\rho_i : 1 \leq i \leq n\}$ is unsatisfiable. This completely proves the claim. □

Theorem 2. *There exists a sequence of unsatisfiable formulae ϕ_n s.t.*

- *the smallest closed δ'-tableaux for ϕ_n are of size $\geq 2_n \cdot c$ for some constant c,*
- *but there are closed $\delta^{\star}$-tableaux for ϕ_n of size $\leq e(n)$ for some elementary function e.*

Proof. We take $\phi_n = \mathrm{Plf}(\mathcal{C} \cup \{C_n\}) \wedge (\psi_1 \wedge \cdots \wedge \psi_n)$ where $\mathcal{C} \cup \{C_n\}$ are defined as in Proposition 9 and $\psi_1 \wedge \cdots \wedge \psi_n$ denotes the conjunction of employed variants of the justifying formulas. Due to Proposition 10 we can translate this 'short' 1-F-resolution resolution proof into a $\delta^{\star}$-tableaux proof. Therefore we obtain the second assertion.

Now we consider the first assertion. Recall that $\mathrm{HC}(\mathcal{C} \cup \{C_n\})$ has a non-elementary lower bound, due to Proposition 10. However, in general $\mathrm{HC}(\mathcal{C}) \neq \mathrm{HC}(\mathrm{Plf}(\mathcal{C}))$, but we have $\mathrm{HC}(\mathrm{Plf}(\mathcal{C})) < \mathrm{HC}(\mathcal{C}) \leq |\mathcal{C}| \cdot \mathrm{HC}(\mathrm{Plf}(\mathcal{C}))$. This together with the above lemma shows that $\mathrm{HC}(\mathcal{C} \cup \{C_n\})$ is a lower-bound for $\mathrm{HC}(\mathrm{Sk}_\tau(\phi_n))$. Finally we employ Proposition 8 to observe that $\|\tau\| > \mathrm{HC}(\mathrm{Sk}_\tau(\phi_n))$. □

Theorem 3. *There exists a sequence of unsatisfiable formulas ϕ'_n s.t.*

- *the smallest closed δ-tableaux for ϕ'_n are of size $\geq 2_n \cdot c$ for some constant c,*
- *but there are closed δ'-tableaux for ϕ'_n of size $\leq e(n)$ for some elementary function e.*

Proof. Let the definition of a sequence of unsatisfiable formulas F_n be altered such that the original justifying formulas are employed instead of the D-variants. Then the liberalized δ'-rule allows us to ignore the free variable B, while the δ-rule forces its incorporation as argument of the introduced Skolem function. This suffices to see that the just given proof can be re-applied. □

Theorem 4. *There exists a sequence of unsatisfiable formulae ψ_n s.t.*

- *the smallest closed δ^{++}-tableaux for ψ_n are of size $\geq 2^n \cdot c$ for some constant c,*
- *but there are closed δ'-tableaux for ψ_n of size $\leq n \cdot d$ for some constant d.*

Proof. Define

$$\psi_0 = \mathbf{true}$$
$$\psi_{n+1} = \exists x (\psi_n \wedge p_n(x,a,b) \supset [\exists u \forall x p_n(x,u,b) \wedge \exists v \forall z p_n(z,a,v)])$$

Now the proof proceeds similar as in [6]. □

The following proposition is an easy corollary to the obtained theorems and answers the related questions posed in [1].

Corollary 1.
- *$\delta^\star$-tableaux admits non-elementary speed-up over the δ^ε-tableaux,*
- *δ^ε-tableaux admit non-elementary speed-up over δ-tableaux, and*
- *$\delta^\star$-tableaux admit exponential speed-up over δ^{++}-tableaux.*

Proof. We only sketch the proof for the first case. The formulas ϕ_n employed in the theorem are free of ε-terms, hence any δ^ε-tableaux does not admit the use of the ε-rule. Hence the introduced ε-terms are theory-free and can simply be interpreted as Skolem functions. This allows an argumentation as in Lemma 5. Now the further proof follows the argumentation in the theorem. □

Remark 6. Note in particular that the non-elementary speed of $\delta^\star$-tableaux over δ^{++}-tableaux cannot carry over to δ^ε-tableaux. This would only be possible if we extend δ^ε-tableaux by further axiom schemata.

9 A new notion of relevance

In this section we define a further liberalization of the δ'-rule, denoted as δ''-rule that allows non-elementary speed-up over $\delta^\star$-tableaux, and $\delta^{\star^\star}$-tableaux. The main novelty is a refined notion of *relevance*.

We define, for any formula ϕ, the set $\mathrm{Rel}(\phi,A)$ of free variables that *occur relevantly* w.r.t. to a free variable A. Note that the given definition is a straightforward extension of the one given in [3].

Definition 8. *If A does not occur in ϕ then $\mathrm{Rel}(\phi,A) = \emptyset$. Otherwise:*

- *If ϕ is an atomic formula, then $\mathrm{Rel}(\phi,A)$ is the set of all variables in ϕ except A.*
- *If $\phi = \neg\psi$, then $\mathrm{Rel}(\phi,A) = \mathrm{Rel}(\psi,A)$.*

– *If* $\phi = \psi_1 \vee \psi_2$ *or* $\psi_2 \vee \psi_1$ *and A does not occur in* ψ_2 *(but in* ψ_1*) then* $\mathrm{Rel}(\phi, A) = \mathrm{Rel}(\psi_1, A)$. *If A does occur in* ψ_2 *but this occurrence is restricted to at most one literal in* ψ_2*, and* ψ_2 *is contradictory, then* $\mathrm{Rel}(\phi, A) = \mathrm{Rel}(\psi_1, A)$. *If A occurs in both,* ψ_1 *and* ψ_2*, then* $\mathrm{Rel}(\phi, A)$ *is the set of all free variables in* ϕ *except A. (Similarly for* $\phi = \psi_1 \wedge \psi_2$ *and* $\phi = \psi_1 \supset \psi_2$*.)*
– *If* $\phi = (\mathsf{Q}y)\psi$ *then* $\mathrm{Rel}(\phi, A) = \mathrm{Rel}(\psi\{y \to B\}, A) - \{B\}$.

If ϕ is a δ-formula, i.e. $\phi = (\exists x)\psi$ or $\phi = \neg(\forall x)\psi$ then we define the set of *relevant variables* to be $\mathrm{Rel}(\psi\{x \to A\}, A)$.

The rules for quantified formulas are given in Table 5. The soundness and completeness of δ''-tableaux follows, easily.

Table 5. δ''-rule

$$\frac{\delta}{\delta_0(f_{[\![\phi]\!]}(B_1,\ldots,B_m)\sigma} \ (\delta''\text{-rule})$$

Let $\phi = \psi(A_1,\ldots,A_n)$ denote the abstraction of δ_0,
where $A_1,\ldots,A_n$ are fresh free variables. Then there exists a σ, domain of σ is a subset of $\{A_1,\ldots,A_n\}$, s.t. $\delta = \psi(A_1,\ldots,A_n)\sigma$. Finally the set $\{B_1,\ldots,B_m\} \subseteq \{A_1,\ldots,A_n\}$ denotes the *relevant* variables of ϕ, according to the above definition.

Using the same example that was employed in the proof of Theorem 4, we can now conclude the following theorem.

Theorem 5. *There exists a sequence of unsatisfiable formulae* ψ_n *s.t.*
– *the smallest closed* $\delta^\star$*-tableaux for* ψ_n *are of size* $\geq 2^n \cdot c$ *for some constant c,*
– *but there are closed* δ''*-tableaux for* ψ_n *of size* $\leq n \cdot d$ *for some constant d.*

Using a similar example an exponential speed-up of the $\delta^{\star^\star}$-tableaux over $\delta^\star$-tableaux has been demonstrated in [8]. Furthermore a non-elementary speed-up of $\delta^{\star^\star}$-tableaux over $\delta^\star$-tableaux has been established. It is important to note that the second result is due to the recursive definition of relevance, employed in the $\delta^{\star^\star}$-tableaux, hence comes about by the removal of “irrelevant” argument places in the definition of Skolem functions, while the first employs the fact that through the $\delta^{\star^\star}$-rule less Skolem functions need to be introduced, then via the $\delta^\star$-rule.

Employing Definition 8 we obtain the following theorem. We define the following variant of a justifying formula:

$$(\forall x_1)\cdots(\forall x_n)(\forall y)[(C_1 \vee C_2) \supset (C_1 \vee (\exists z)[C_2\{y \to z\} \vee (p(y) \wedge \neg p(y) \wedge q(z,y))])] \quad (10)$$

Again this variant is logically equivalent to the original formula.

Theorem 6. *There exists a sequence of unsatisfiable formulae* ψ_n *s.t.*
– *the smallest closed* $\delta^{\star^\star}$*-tableaux for* ψ_n *are of size* $\geq 2_n \cdot c$ *for some constant c,*

– *but there are closed δ''-tableaux for ψ_n of size $\leq e(n)$ for some elementary function e.*

Proof. The proof proceeds as the proof of Theorem 2. We only have to employ the above variant (10) of the justifying formulas. □

10 Conclusion and Future Work

In this paper we have studied the possibility of using Hilbert's ε-calculus in first-order automated reasoning and the relation between Skolem functions and ε-terms.

Based on the results in Section 3 and Section 4, we conclude that the ε-calculus (without drastic alterations) is a too strong tool, that should not be considered as appropriate for *automated* reasoning.

On the other hand, if adjust our aims and seek for elegant liberalizations of the δ-rule, we have seen (in Section 6) that we need not really consider the 'full' ε-formalism, but can indeed simulate its effect by suitable defined Skolem functions. Thereby showing how we can achieve a similar effect as in [1] replacing ε-terms by Skolem functions.

Using this result we can compare δ^{ε}-tableaux with $\delta^{\star^\star}$-tableaux and conclude that the later allows non-elementary speed-up over the former. Finally we have defined a further liberalizations of the δ-expansion rule, called δ''-rule which is similar to the $\delta^{\star^\star}$-rule but admits non-elementary speed-up over the latter.

With respect to the last result a warning seems appropriate: Note that the proof of Theorem 6 depends exclusively on the special form of *relevance* employed. Admittingly this form is neither natural, nor efficient. However, our point here is something else. In effect we can define a whole zoo of different notions of relevance (even efficient ones), s.t. each of these triggers a non-elementary speed-up over the other. With this fact in mind, it seems save to say that $\delta^{\star^\star}$-tableaux and δ''-tableaux should be considered as equal.

It still remains to give a more detailed analysis of the exponential speed-up of the δ''-tableaux over the $\delta^{\star}$-tableaux, as reported in Theorem 5.

> *Is it possible to improve this exponential speed-up, following the argument of the theorem, to a non-elementary speed-up?*

Unfortunately we cannot answer this question at the moment.

References

1. W. Ahrendt and M. Giese. Hilbert's ε-terms in automated theorem proving. In *Proc. TABLEAUX'99*, LNCS 1617, pages 171–185. Springer Verlag, 1999.
2. G. Asser. Theorie der logischen Auswahlfunktionen. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 3:30–68, 1957.
3. M. Baaz and C. Fermüller. Non-elementary speedups between different versions of tableaux. In *Proc. of TABLEAUX'95*, pages 217–230, 1995.
4. M. Baaz and A. Leitsch. Complexity of resolution proofs and function introduction. *Ann. Pure Appl. Logic*, 57:181–215, 1992.

5. M. Baaz and A. Leitsch. Methods of function extensions. In *Collegium Logicum—Annals of the Kurt Gödel Society*, volume 1, pages 87–122. Springer Verlag, 1995.
6. B. Beckert, R. Hähnle, and P. Schmitt. The even more liberalized δ-rule in free variable semantic tableaux. In *Proc. of KGC'93*, LNCS 713, pages 108–119. Springer, 1993.
7. A. Blass and Y. Gurevich. The logic of choice. *JSL*, pages 1264–1310, 2000.
8. D. Cantone and M. Nicolosi Asmundo. A further and effective liberalizatin of the δ-rule. In *Automated Deduction in Classical and Non-Classical Logics*, LNCS 1761, pages 109–125. Springer, 2000.
9. M. Fitting. *First-Order Logic and Automated Theorem Proving*. Springer, 1990.
10. R. Hähnle and P. Schmitt. The liberalized delta-rule in free variable semantic tableaux. *Journal of Automated Reasoning*, 13:211–222, 1994.
11. D. Hilbert and P. Bernays. *Grundlagen der Mathematik II*. Springer Verlag, Second Edition, 1970.
12. A.C. Leisenring. *Mathematical Logic and Hilbert's ε-symbol*. MacDonald Technical and Scientific, London, 1969.
13. Alexander Leitsch. *The Resolution Calculus*. EATCS - Texts in Theoretical Computer Science. Springer, 1997.
14. R. Statman. Lower bounds on Herbrand's theorem. In *Proc. Math. Soc.*, volume 75, pages 104–107, 1979.
15. Claus-Peter Wirth. A new indefinite semantics for hilbert's epsilon. Accepted for publication in Proc. of TABLEAUX'02., 2002.
16. T. Yukami. Some results on speed-up. *Ann. Japan Assoc. Philos. Sci.*, 6:195–205, 1984.