Equational Theories in CERES*

Alexander Leitsch and Clemens Richter

Institute of Computer Languages (E185),
Vienna University of Technology, Favoritenstraße 9,
1040 Vienna, Austria
{leitsch|richter}@logic.at

Abstract. Cut-elimination is the most important proof transformation in logic. Equality is a central paradigm in mathematics and plays a key role in automated deduction. Therefore its importance awakes the necessity of integrating equality into existing cut-elimination methods.
In this paper we extend the resolution-based method of cut-elimination CERES to CERES-e by adding equality (and paramodulation to LK), where all the advantages of CERES are preserved; in particular CERES-e is superior to Gentzen type methods, is flexible with respect to resolution, paramodulation and its refinements and admits a semantical use of cut. We go even further and combine CERES-e with equational theories yielding a system which adds simplicity of proof notation to the advantages gained from equality.

1 Introduction

Due to the central importance of equality in mathematical proofs an investigation of cut-elimination in proofs with equality is very important to the application of cut-elimination. Resolution, paramodulation and its refinements have been playing a central role in automated deduction for decades. Experiments with resolution refinements within cut-elimination by resolution (see [4]) promise even more rewarding results after extension of CERES by equality.

We extend the Gentzen calculus **LK** to **LK-e** by a paramodulation-type rule allowing for a comfortable use of equality. For **LK-e** we introduce CERES-e, a cut-elimination method based on resolution and paramodulation. Furthermore we extend **LK-e** by the rule of semantic cut and by built-in equational theories. Two non-trivial examples will outline the strength of the concepts and demonstrate the workflow of the methods.

2 Definitions and Notation

In the following $x, y, z, x_0, y_0, z_0, \ldots$ denote bound individual variables whereas $u, v, w, u_0, v_0, w_0, \ldots$ denote free individual variables.

* supported by the Austrian Science Fund (project no. P17995-N12)

Definition 1 (term, semi-term). *Terms and semi-terms are defined inductively in the following way:*

1. *Individual constants are (semi-)terms.*
2. *Free variables (and bound variables) are (semi-)terms.*
3. *If f is a function symbol of arity n and $t_1, \ldots, t_n$ are (semi-)terms then $f(t_1, \ldots, t_n)$ is a (semi-)term.*

Thus semi-terms are terms with bound variables, the analogous applies to formulas.

Definition 2 (formula, semi-formula). *Formulas and semi-formulas are inductively defined as follows:*

1. *If P is an n-ary predicate symbol and $t_1, \ldots, t_n$ are (semi-)terms, then $P(t_1, \ldots, t_n)$ is a (semi-)formula. It is called an atomic formula or an atom.*
2. *If A and B are (semi-)formulas, then $\neg A$, $A \wedge B$, $A \vee B$ and $A \supset B$ are (semi-)formulas.*
3. *If A is a (semi-)formula not containing the bound variable x, then $(\forall x)A(u/x)$ and $(\exists x)A(u/x)$ are (semi-)formulas.*

Obviously semi-formulas are formulas with free occurrences of bound variables. Note that our definition of terms and formulas is due to Takeuti [9].

We will later be in the need to determine precisely the position of a term or formula within another term or formula; the following definition helps in this matter.

Definition 3 (position). *Let t be a (semi-)term. We define positions within t inductively as follows:*

1. *If t is an individual constant or a variable then 0 is the position in t representing the entire term t, i.e. $t.0 = t$.*
2. *Let t be of the form $t = f(t_1, \ldots, t_n)$ then again 0 is the position in t representing the entire term t, i.e. $t.0 = t$. Let further $\xi_i : (k_l, \ldots, k_1, 0)$ be a position in a t_i, for $1 \leq i \leq n$, and $t_i.\xi_i = s$; then we define the nested position ξ in t such that $t.\xi = s$ as $\xi : (i, k_l, \ldots, k_1, 0)$.*

Let $t.\xi = s$ then $t[r]_\xi$ denotes the (semi-)term t after replacement of s on position ξ by r, in particular $t[r]_\xi = r$. Moreover if $\varXi$ is a set of positions in t then $t[r]_\varXi$ is defined by replacing all sub-(semi-)terms $t.\xi$, for $\xi \in \varXi$, in t by r. Positions in formulas are defined analogously (simply consider all formulas as terms).

Substitutions are defined as usual as functions from the set of variables to the set of terms. We write $A(u)$ to indicate (potential) free occurrences of the variable u in A. Let t be an arbitrary term, then $A(u/t)$ stands for the replacement of all free occurrences of u in A by t, i.e. $A[t]_\varXi$ where $\varXi = \{\xi \mid A.\xi = u\}$.

In the following $\varGamma, \varDelta, \varPi, \varLambda, \varGamma_0, \varDelta_0, \varPi_0, \varLambda_0, \ldots$ denote finite (possibly empty) sequences of formulas.

Definition 4 (sequent). *A finite sequence of formulas, separated by the auxiliary syntactic symbol* $\vdash$*, is called a sequent, symbolically* $S : \Gamma \vdash \Delta$*. The empty sequent is denoted by* $\vdash$*. A sequent* S *is called atomic if* Γ *and* Δ *are sequences of atomic formulas.*

Definition 5 (axiom set). *A (possibly infinite) set* $\mathcal{A}$ *of atomic sequents is called an axiom set if it is closed under substitution, i.e. for all* $S \in \mathcal{A}$ *and for all substitutions* σ *we have* $S\sigma \in \mathcal{A}$*.*

Definition 6 (LK). *An inference rule is of the form*

$$\frac{S_1}{S}\ \rho_u \qquad \textit{or} \qquad \frac{S_1 \quad S_2}{S}\ \rho_b$$

where the sequents S_1, S_2 *are called the premises and the sequent* S *is called the conclusion of the inference. Unlike Gentzen's version of* **LK** *(see [7]) we use the additive version of* **LK** *as Girard (see [8]). In the following definition the auxiliary formulas are put in bold face and the principal formulas are underlined, but usually these markings are avoided because the auxiliary and principal formulas are mostly uniquely identifiable by their outermost positioning (respectively the permutations are given explicitly where needed).*

1. *The structural rules of*
   (a) *Weakening:*

$$\frac{\Gamma \vdash \Delta}{\underline{A_1}, \ldots, \underline{A_n}, \Gamma \vdash \Delta}\ \mathrm{w:l} \qquad \frac{\Gamma \vdash \Delta}{\Gamma \vdash \Delta, \underline{A_1}, \ldots, \underline{A_n}}\ \mathrm{w:r}$$

   (b) *Contraction:*

$$\frac{\boldsymbol{A}, \ldots, \boldsymbol{A}, \Gamma \vdash \Delta}{\underline{A}, \Gamma \vdash \Delta}\ \mathrm{c:l} \qquad \frac{\Gamma \vdash \Delta, \boldsymbol{A}, \ldots, \boldsymbol{A}}{\Gamma \vdash \Delta, \underline{A}}\ \mathrm{c:r}$$

   (c) *Permutation:*

$$\frac{A_1, \ldots, A_n, \Gamma \vdash \Delta}{A_{\pi(1)}, \ldots, A_{\pi(n)}, \Gamma \vdash \Delta}\ \mathrm{p}(\pi):\mathrm{l} \qquad \frac{\Gamma \vdash \Delta, A_1, \ldots, A_n}{\Gamma \vdash \Delta, A_{\pi(1)}, \ldots, A_{\pi(n)}}\ \mathrm{p}(\pi):\mathrm{r}$$

   where π *is a permutation of* $\{1, \ldots, n\}$*. The auxiliary formulas respectively the principal formulas are those* A_i *of the premises respectively those* $A_{\pi(i)}$ *of the conclusions where* $i \neq \pi(i)$*,* $i \in \{1, \ldots, n\}$*, holds.*

   (d) *Cut:*

$$\frac{\Gamma \vdash \Delta, \boldsymbol{A} \quad \boldsymbol{A}, \Pi \vdash \Lambda}{\Gamma, \Pi \vdash \Delta, \Lambda}\ \mathrm{cut}$$

2. *The logical rules for*
   (a) $\neg$*-introduction:*

$$\frac{\Gamma \vdash \Delta, \boldsymbol{A}}{\underline{\neg A}, \Gamma \vdash \Delta}\ \neg:\mathrm{l} \qquad \frac{\boldsymbol{A}, \Gamma \vdash \Delta}{\Gamma \vdash \Delta, \underline{\neg A}}\ \neg:\mathrm{r}$$

(b) $\land$*-introduction:*

$$\frac{\boldsymbol{A}, \varGamma \vdash \varDelta}{\underline{A \land B}, \varGamma \vdash \varDelta}\ \land : \mathrm{l}_1 \qquad \frac{\boldsymbol{B}, \varGamma \vdash \varDelta}{\underline{A \land B}, \varGamma \vdash \varDelta}\ \land : \mathrm{l}_2 \qquad \frac{\varGamma \vdash \varDelta, \boldsymbol{A} \quad \varPi \vdash \varLambda, \boldsymbol{B}}{\varGamma, \varPi \vdash \varDelta, \varLambda, \underline{A \land B}}\ \land : \mathrm{r}$$

(c) $\lor$*-introduction:*

$$\frac{\boldsymbol{A}, \varGamma \vdash \varDelta \quad \boldsymbol{B}, \varPi \vdash \varLambda}{\underline{A \lor B}, \varGamma, \varPi \vdash \varDelta, \varLambda}\ \lor : \mathrm{l} \qquad \frac{\varGamma \vdash \varDelta, \boldsymbol{A}}{\varGamma \vdash \varDelta, \underline{A \lor B}}\ \lor : \mathrm{r}_1 \qquad \frac{\varGamma \vdash \varDelta, \boldsymbol{B}}{\varGamma \vdash \varDelta, \underline{A \lor B}}\ \lor : \mathrm{r}_2$$

(d) $\supset$*-introduction:*

$$\frac{\varGamma \vdash \varDelta, \boldsymbol{A} \quad \boldsymbol{B}, \varPi \vdash \varLambda}{\underline{A \supset B}, \varGamma, \varPi \vdash \varDelta, \varLambda}\ \supset : \mathrm{l} \qquad \frac{\boldsymbol{A}, \varGamma \vdash \varDelta, \boldsymbol{B}}{\varGamma \vdash \varDelta, \underline{A \supset B}}\ \supset : \mathrm{r}$$

(e) $\forall$*-introduction:*

$$\frac{\boldsymbol{A(x/t)}, \varGamma \vdash \varDelta}{\underline{(\forall x)A(x)}, \varGamma \vdash \varDelta}\ \forall : \mathrm{l} \qquad \frac{\varGamma \vdash \varDelta, \boldsymbol{A(x/u)}}{\varGamma \vdash \varDelta, \underline{(\forall x)A(x)}}\ \forall : \mathrm{r}$$

where t is an arbitrary term and u does not occur in the conclusion.

(f) $\exists$*-introduction:*

$$\frac{\boldsymbol{A(x/u)}, \varGamma \vdash \varDelta}{\underline{(\exists x)A(x)}, \varGamma \vdash \varDelta}\ \exists : \mathrm{l} \qquad \frac{\varGamma \vdash \varDelta, \boldsymbol{A(x/t)}}{\varGamma \vdash \varDelta, \underline{(\exists x)A(x)}}\ \exists : \mathrm{r}$$

where u does not occur in the conclusion and t is an arbitrary term.

Definition 7 (proof, LK-proof). *A proof φ of a sequent S from an axiom set $\mathcal{A}$ is a directed labelled tree, where the nodes represent occurrences of sequents and the edges are labelled according to the inference rule applications in the calculus $\mathcal{K}$. The root is labelled by the occurrence of the end-sequent S and the leaves are labelled by occurrences of axioms, i.e. elements of $\mathcal{A}$.*

An **LK***-proof is a proof where $\mathcal{A}$ consists of atomic sequents and the inference rules applied are those of* **LK***.*

Definition 8 (ancestor). *Let*

$$\frac{S_1 : \varPi_1, \varGamma_1 \vdash \varDelta_1, \varLambda_1 \quad S_2 : \varPi_2, \varGamma_2 \vdash \varDelta_2, \varLambda_2}{S : \varPi, \varGamma_1, \varGamma_2 \vdash \varDelta_1, \varDelta_2, \varLambda}\ \rho$$

be an inference rule in an **LK***-proof φ, where $\varPi_i$ and $\varLambda_i$ respectively $\varPi$ and $\varLambda$ denote the (possibly empty) sequences of auxiliary formulas of the (one or two) premises respectively principal formulas of the conclusion; let further μ_k be the occurrence of the k-th principal formula in S and ν_{ij} be the occurrence of the j-th auxiliary formula in S_i, $i \in 1, 2$ and $j, k \in \mathbb{N}$. Then all ν_{ij} are ancestors of all μ_k.*

The ancestor relation in φ is defined as the reflexive and transitive closure of the above relation.

If $\varOmega$ is a set of formula occurrences in φ then by $S(\nu, \varOmega)$ respectively $\bar{S}(\nu, \varOmega)$ we denote the subsequent of S at the node ν of the **LK***-proof φ consisting of all formulas which are respectively are not ancestors of a formula occurrence in $\varOmega$.*

Definition 9 (LK-p). **LK-p** *is the calculus obtained from* **LK** *by adding the semantic cut rule* (p − cut)*, also called pseudo-cut, to the existing rules of* **LK**.

$$\frac{\Gamma \vdash \Delta, \boldsymbol{A} \quad \boldsymbol{B}, \Pi \vdash \Lambda}{\Gamma, \Pi \vdash \Delta, \Lambda}\ \mathrm{p-cut}$$

if $A \supset B$ is valid.

By comparison of the semantic cut rule with an implication introduction rule

$$\frac{\Gamma \vdash \Delta, A \quad B, \Pi \vdash \Lambda}{A \supset B, \Gamma, \Pi \vdash \Delta, \Lambda}\ \supset: \mathrm{l}$$

it is easy to see that $A \supset B$ can only be cut out if it is valid or in other words, that $\Gamma, \Pi \vdash \Delta, \Lambda$ is equivalent to $A \supset B, \Gamma, \Pi \vdash \Delta, \Lambda$ if $A \supset B$ is valid.

The introduction of the semantic cut rule is more ore less a replacement of the existing cut rule, which is just the specific case of the semantic cut rule where A is syntactically equivalent to B, i.e. $A \equiv B$.

Definition 10 (LK-p-proof). *An* **LK-p***-proof is a proof where $\mathcal{A}$ consists of atomic sequents and the inference rules applied are those of* **LK-p**.

Definition 11 (clause). *A clause is an atomic sequent, i.e. a sequent of the form $\Gamma \vdash \Delta$, where Γ and Δ are sequences of atomic formulas.*

The paramodulation rule for the resolution calculus, which we will be using in this paper, is defined as follows.

Definition 12 (paramodulation). *Let $C : \Gamma \vdash \Delta, s = t$ and $D : \Gamma \vdash \Delta, A[s']_{\Xi}$, $D' : A[s']_{\Xi}, \Pi \vdash \Lambda, A[s']$ be variable disjoint clauses and σ be the most general unifier of $\{s, s'\}$ then the paramodulation rule for the resolution calculus is defined as follows:*

$$\frac{C \quad D}{(\Gamma, \Pi \vdash \Delta, \Lambda)\sigma}\ p:\sigma \qquad \textit{resp.} \qquad \frac{C \quad D'}{(\Gamma, \Pi \vdash \Delta, \Lambda)\sigma}\ p:\sigma$$

where s and t are arbitrary terms.

The flipped paramodulation rule (p′) is defined analogous by replacing C with the clause $C' : \Gamma \vdash \Delta, t = s$.

Note that the particular definition of the paramodulation rule does not matter, again all refinements of paramodulation such as ordered paramodulation or superposition might be used (see [5] or [6] for more details on paramodulation and its refinements).

Definition 13 (PR-deduction). *A deduction of a set of clauses using resolution and paramodulation is called a* **PR***-deduction.*

3 Extension of CERES to Equality

The cut-elimination method by resolution (CERES) will be defined roughly in this paper, in the sense that only some basic parts and the necessary parts for the extension by equality will appear in the following subsection. You can find in-depth further explanation of the method in [3], [2] and on the CERES web page[1].

In the second part of this section we will introduce the cut-elimination method by resolution using equality CERES-e. An extensive example will demonstrate the method CERES-e and mark out its advantages in comparison to CERES.

3.1 CERES

To facilitate understanding, we give an overview how cut-elimination is done in CERES. First of all we have to skolemize the proof, i.e. to replace all eigenvariables by suitable Skolem terms which occur in ancestors of the end-sequent. Then all occurrences of ancestors of cut-formulas of the proof are used to build a refutable clause set, called the characteristic clause set of the proof. In the next step the clause set is refuted by resolution using arbitrary refinements. The grounded refutation is nothing less than a proof of the empty sequent containing only atomic cuts where the initial sequents are elements of the characteristic clause set. Serving as a skeleton, CERES augments the grounded resolution proof with cut-free parts of the original proof related only to the end-sequent, that is why the initial skolemization is needed to avoid violations of eigenvariable conditions within these proof parts. Finally we obtain a proof of the original end-sequent, and after re-skolemization of the original statement, with atomic cuts only.

Details regarding skolemization and re-skolemization can be found here [1] and are omitted in this paper.

Definition 14 (characteristic clause set). *Let φ be an* **LK***-proof of the sequent S and let Ω be the set of all occurrences of cut formulas in φ. The characteristic clause set $\Theta(\varphi)$ is defined inductively as follows.*

If the proof-node ν is the occurrence of an initial sequent S in φ, then the characteristic clause set of S at the node ν corresponds to the subsequent of S consisting of all formulas which are ancestors of an occurrence in Ω, i.e.

$$\Theta(\varphi)/\nu = \{S(\nu, \Omega)\}.$$

Let us assume that the clause sets $\Theta(\varphi)/\nu$ are already constructed for all nodes ν in φ with $\mathrm{depth}(\nu) \leq n$. *Now let ν be a node with* $\mathrm{depth}(\nu) = n+1$. *We distinguish the following cases:*

[1] The documentation of the method respectively the system and an online version of the system CERES are available at `http://www.logic.at/ceres/`.

1. ν is the immediate consequent of μ, i.e. a unary rule applied to μ gives ν. Then we simply define

$$\Theta(\varphi)/\nu = \Theta(\varphi)/\mu.$$

2. ν is the immediate consequent of μ_1 and μ_2, i.e. a binary rule ρ applied to μ_1 and μ_2 gives ν. Then we distinguish between

(a) all of the auxiliary formulas of ρ are ancestors of Ω, i.e. the auxiliary formulas occur in $S(\mu_1, \Omega)$ respectively $S(\mu_2, \Omega)$. Then

$$\Theta(\varphi)/\nu = \Theta(\varphi)/\mu_1 \cup \Theta(\varphi)/\mu_2,$$

(b) none of the auxiliary formulas of ρ is an ancestor of Ω, i.e. the auxiliary formulas do not occur in $S(\mu_1, \Omega)$ and $S(\mu_2, \Omega)$. Then

$$\Theta(\varphi)/\nu = \Theta(\varphi)/\mu_1 \times \Theta(\varphi)/\mu_2$$

where

$$\{\Gamma_1 \vdash \Delta_1, \ldots, \Gamma_n \vdash \Delta_n\} \times \{\Pi_1 \vdash \Lambda_1, \ldots, \Pi_m \vdash \Lambda_m\} = \{\Gamma_i, \Pi_j \vdash \Delta_j, \Lambda_j \mid 1 \leq i \leq n, 1 \leq j \leq m\}.$$

Finally, the characteristic clause set $\Theta(\varphi)$ of φ is defined as $\Theta(\varphi)/\nu$ where ν is the root node of φ.

Note that in a binary **LK**-inference either all auxiliary formulas are ancestors of Ω or none of them.

Remark 1. If φ is a cut-free **LK**-proof then there are no occurrences of cut formulas in φ hence $\Theta(\varphi) = \{\vdash\}$.

Theorem 1. *Let φ be an* **LK***-proof. Then $\Theta(\varphi)$ is unsatisfiable, i.e. there exist a resolution refutation of $\Theta(\varphi)$.*

Proof. In [2] or [3].

Proposition 1. *CERES also eliminates semantic cuts, i.e. is a cut-elimination method for* **LK-p***.*

Proof. In [2].

Note that standard methods of cut-elimination (e.g. Gentzen's method) are not capable of eliminating semantic cuts.

3.2 Extension to Equality

There exist various substantial different approaches how to integrate equality into **LK**. Some of them are solely based on axiomatization, i.e. adding equality axioms to the existing axioms, without any extension of the rules (e.g. see Takeuti [9] for more details). This kind of equality integration would of course be possible in CERES without any changes (note that we allow arbitrary atomic sequents as axioms), but has at least two major drawbacks. On the one hand it is mathematically a very unnatural way of using equality within a proof, concerning formalization and interpretation of proofs. On the other hand the computational expense is much higher, e.g. by inheriting the used axioms throughout the entire proof to the end-sequent. In addition it would not be possible to use the existing paradigm of paramodulation within CERES-e which is especially designed to handle equality reasoning within resolution, which is a special design goal for us.

This disadvantages are overcome by introducing the theory of equality to **LK** by means of rules. Some might argue that this is a tradeoff against the loss of the sub-formula property and depending on the specific rules the introduction of implicit cuts. Loosing the sub-formula property is not avoidable if you intend to use equality in a mathematically natural and intuitive way. The argument of implicit cuts is immediately dismantled in CERES as we only intend to eliminate non-atomic cuts (since we are not using axioms of the form $A \vdash A$ - which of course also applies to the approach by axiomatization).

Again there are many different variants how to extend **LK** by equality with help of rules (e.g. see [10]). We will now define the best suitable version for our needs (similar to [6]).

Definition 15 (LK-e). **LK-e** *is the calculus obtained from* **LK** *by adding the following equality introduction (or paramodulation) rules to the existing rules of* **LK***:*

$$\frac{\Gamma \vdash \Delta, \boldsymbol{s = t} \quad \boldsymbol{A[s]_\Xi}, \Pi \vdash \Lambda}{\underline{A[t]_\Xi}, \Gamma, \Pi \vdash \Delta, \Lambda} =: \mathrm{l} \qquad \frac{\Gamma \vdash \Delta, \boldsymbol{s = t} \quad \Pi \vdash \Lambda, \boldsymbol{A[s]_\Xi}}{\Gamma, \Pi \vdash \Delta, \Lambda, \underline{A[t]_\Xi}} =: \mathrm{r}$$

where s and t are arbitrary terms.

For practical reasons we will additionally use the following rules in **LK-e**:

$$\frac{\Gamma \vdash \Delta, \boldsymbol{t = s} \quad \boldsymbol{A[s]_\Xi}, \Pi \vdash \Lambda}{\underline{A[t]_\Xi}, \Gamma, \Pi \vdash \Delta, \Lambda} =': \mathrm{l} \qquad \frac{\Gamma \vdash \Delta, \boldsymbol{t = s} \quad \Pi \vdash \Lambda, \boldsymbol{A[s]_\Xi}}{\Gamma, \Pi \vdash \Delta, \Lambda, \underline{A[t]_\Xi}} =': \mathrm{r}$$

Note that this rules could also be derived from the ones above using an additional paramodulation inference.

Definition 16 (LK-e-proof). *An* **LK-e***-proof is a proof where $\mathcal{A}$ consists of atomic sequents including the axiom set of reflexitivity, i.e.*

$$\{ \vdash t = t \mid t \textit{ a term} \},$$

and the inference rules applied are those of **LK-e***.*

Definition 17 (LK-ep). **LK-ep** *is the calculus obtained from* **LK-e** *by again adding the semantic cut (see definition 9) rule to the existing rules of* **LK-e***.*

For the extension of CERES to CERES-e no redefinition of the characteristic clause set and its computation is necessary; the paramodulation rule is treated as "ordinary" binary rule. The projections are built in exactly the same manner, hence also the necessity of skolemization a priori and re-skolemization a posteriori remains.

Therefore the only thing that remains to be shown is the following theorem.

Theorem 2. *Let φ be an* **LK-e***-proof. Then $\Theta(\varphi)$ is unsatisfiable, i.e. there exist a refutation with resolution and paramodulation of $\Theta(\varphi)$.*

Proof. As in [2] we show that, from the set $\Theta(\varphi)/\nu$ for any node ν in φ we can derive $S(\nu,\Omega)$ (the subsequent of S at ν consisting just of the ancestors of a cut). As there is no ancestor of a cut in the end sequent, we obtain an **LK-e** derivation of $\vdash$ from $\Theta(\varphi)$. The $=:$ l and $=:$ r rules behave like any other binary rule in **LK**, and the construction goes through like for **LK**. As **LK-e** is sound and we have derived $\vdash$, $\Theta(\varphi)$ must be unsatisfiable.

Definition 18 (LK-ep-proof). *An* **LK-ep***-proof is a proof where $\mathcal{A}$ consists of atomic sequents including the reflexitivity axiom and the inference rules applied are those of* **LK-ep***.*

Proposition 2. *CERES is a cut-elimination method for* **LK-ep***.*

Proof. By theorem 2 $\Theta(\varphi)$ is unsatisfiable. As **PR**-deduction is complete there exists a **PR**-refutation γ of $\Theta(\varphi)$. Let γ' be any ground projection of γ. Then γ' is an **LK-e** derivation of $\vdash$ from the axiom set defined by $\Theta(\varphi)$. γ' contains only atomic cuts. By inserting the proof projections on every leaf of γ' we obtain a proof of the original sequent with only atomic cuts.

Now we will demonstrate the strength of this method on a well-known example from group theory. The proof φ below verifies that a 2-nilpotent group is commutative using the cancellation principle as a lemma. Therefore we need to extend the set of axioms by all instances of the necessary group theoretic axioms:

$$\vdash (u \circ v) \circ w = u \circ (v \circ w), \tag{A}$$

$$\vdash e \circ u = u \qquad\qquad \vdash u \circ e = u, \qquad\qquad (E_l),\ (E_r)$$

$$\vdash u^{-1} \circ u = e \qquad\qquad \vdash u \circ u^{-1} = e, \qquad\qquad (I_l),\ (I_r)$$

$$\vdash u \circ u = e, \tag{N_2}$$

where u^{-1} denotes the inverse element of u.

Since the original proof of $\vdash (\forall x)(\forall y)x \circ y = y \circ x$ contains strong quantifiers it has to be skolemized in advance and the resulting cut-free proof re-skolemized afterwards.

We only give the skolemized proof (of the sequent $\vdash a \circ b = b \circ a$ for two individual constant symbols a and b). The proof of $\vdash (\forall x)(\forall y) x \circ y = y \circ x$ can be directly obtained by generalizing a to u and b to v and by afterwards applying $\forall : \mathrm{r}$ twice on $\vdash u \circ v = v \circ u$.

Within this section the following formula abbreviations are used:

$$
\begin{aligned}
P&\colon\ (a \circ b) \circ (a \circ b) = (b \circ a) \circ (a \circ b),\\
C&\colon\ a \circ b = b \circ a,\\
S&\colon\ u \circ w = v \circ w.
\end{aligned}
$$

Then, let the main proof φ be defined as follows.

$\varphi =$

$$
\dfrac{\vdash e = e \qquad \dfrac{\vdash b \circ b = e \qquad \dfrac{\vdash e \circ b = b \qquad \dfrac{\vdash a \circ a = e \qquad \dfrac{\vdash (a \circ a) \circ b = a \circ (a \circ b) \quad \varphi'}{e = b \circ ((a \circ a) \circ b) \vdash C}\ {=' : \mathrm{l}}}{e = b \circ (e \circ b) \vdash C}\ {= : \mathrm{l}}}{e = b \circ b \vdash C}\ {= : \mathrm{l}}}{e = e \vdash C}\ {= : \mathrm{l}}}{\vdash a \circ b = b \circ a}\ \mathrm{cut}
$$

$\varphi' =$

$$
\dfrac{\vdash (b \circ a) \circ (a \circ b) = b \circ (a \circ (a \circ b)) \qquad \dfrac{\vdash (a \circ b) \circ (a \circ b) = e \quad \varphi''}{e = (b \circ a) \circ (a \circ b) \vdash C}\ {= : \mathrm{l}}}{e = b \circ (a \circ (a \circ b)) \vdash C}\ {= : \mathrm{l}}
$$

$\varphi'' =$

$$
\dfrac{\varphi_c \quad \varphi_p}{(a \circ b) \circ (a \circ b) = (b \circ a) \circ (a \circ b) \vdash C}\ \mathrm{p-cut}
$$

This is the sub-proof of the cancellation lemma used in φ:

$\varphi_c =$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\vdash v \circ e = v \qquad \dfrac{\vdash u \circ e = u \qquad \dfrac{\vdash w \circ w^{-1} = e \qquad \dfrac{\vdash (v \circ w) \circ w^{-1} = v \circ (w \circ w^{-1}) \quad \varphi_c'}{S \vdash u \circ (w \circ w^{-1}) = v \circ (w \circ w^{-1})}\ {= : \mathrm{r}}}{S \vdash u \circ e = v \circ e}\ {= : \mathrm{r}}}{S \vdash u = v \circ e}\ {= : \mathrm{r}}}{u \circ w = v \circ w \vdash u = v}\ {= : \mathrm{r}}}{\vdash u \circ w = v \circ w \supset u = v}\ {\supset : \mathrm{r}}}{\vdash (\forall z)(u \circ z = v \circ z \supset u = v)}\ {\forall : \mathrm{r}}}{\vdash (\forall y)(\forall z)(u \circ z = y \circ z \supset u = y)}\ {\forall : \mathrm{r}}}{\vdash (\forall x)(\forall y)(\forall z)(x \circ z = y \circ z \supset x = y)}\ {\forall : \mathrm{r}}
$$

$\varphi'_c =$

$$\dfrac{\vdash (u\circ w)\circ w^{-1} = u\circ(w\circ w^{-1}) \qquad \dfrac{S\vdash u\circ w = v\circ w \quad \vdash (u\circ w)\circ w^{-1} = (u\circ w)\circ w^{-1}}{S\vdash (u\circ w)\circ w^{-1} = (v\circ w)\circ w^{-1}}\,=:\mathrm{r}}{S\vdash u\circ(w\circ w^{-1}) = (v\circ w)\circ w^{-1}}\,=:\mathrm{r}$$

$\varphi_p =$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{P\vdash (a\circ b)\circ(a\circ b) = (b\circ a)\circ(a\circ b) \quad a\circ b = b\circ a\vdash C}{(a\circ b)\circ(a\circ b) = (b\circ a)\circ(a\circ b)\supset a\circ b = b\circ a, P\vdash C}\,\supset:\mathrm{l}}{(\forall z_1)((a\circ b)\circ(a\circ z_1) = (b\circ a)\circ(a\circ z_1)\supset a\circ b = b\circ a), P\vdash C}\,\forall:\mathrm{l}}{(\forall z_0)(\forall z_1)((a\circ b)\circ(z_0\circ z_1) = (b\circ a)\circ(z_0\circ z_1)\supset a\circ b = b\circ a), P\vdash C}\,\forall:\mathrm{l}}{(\forall y)(\forall z_0)(\forall z_1)((a\circ b)\circ(z_0\circ z_1) = y\circ(z_0\circ z_1)\supset a\circ b = y), P\vdash C}\,\forall:\mathrm{l}}{(\forall x)(\forall y)(\forall z_0)(\forall z_1)(x\circ(z_0\circ z_1) = y\circ(z_0\circ z_1)\supset x = y), P\vdash C}\,\forall:\mathrm{l}$$

Hence for the characteristic clause set $\Theta(\varphi)$ of φ we obtain

$$\begin{aligned}\{\ & C_1 : a\circ b = b\circ a\vdash,\ C_2 : (a\circ b)\circ(a\circ b) = (b\circ a)\circ(a\circ b)\vdash (a\circ b)\circ(a\circ b) = (b\circ a)\circ(a\circ b),\\ & C_3 : \vdash (u\circ w)\circ w^{-1} = (u\circ w)\circ w^{-1},\ C_4 : u\circ w = v\circ w\vdash u\circ w = v\circ w,\\ & C_5 : \vdash (u\circ w)\circ w^{-1} = u\circ(w\circ w^{-1}),\ C_6 : \vdash (v\circ w)\circ w^{-1} = v\circ(w\circ w^{-1}),\\ & C_7 : \vdash w\circ w^{-1} = e,\ C_8 : \vdash u\circ e = u,\ C_9 : \vdash v\circ e = v,\ C_{10} : \vdash (a\circ b)\circ(a\circ b) = e,\\ & C_{11} : \vdash (b\circ a)\circ(a\circ b) = b\circ(a\circ(a\circ b)),\ C_{12} : \vdash (a\circ a)\circ b = a\circ(a\circ b),\\ & C_{13} : \vdash a\circ a = e,\ C_{14} : \vdash e\circ b = b,\ C_{15} : \vdash b\circ b = e,\ C_{16} : \vdash e = e\ \}.\end{aligned}$$

Since resolution admits subsumption and deletion of tautologies we obtain a reduced characteristic clause set by omitting the clauses C_2, C_4, C_6, C_9 and C_{16}:

$$\Theta(\varphi) = \{C_1, C_3, C_5, C_7, C_8, C_{10}, C_{11}, C_{12}, C_{13}, C_{14}, C_{15}\}.$$

A **PR**-refutation of $\Theta(\varphi)$ is given by the following derivations.

Derivation of C_{17}:

$$\dfrac{\dfrac{\overset{(C_7)}{\vdash w\circ w^{-1} = e} \quad \overset{(C_8)}{\vdash u\circ e = u}}{\vdash u\circ(w\circ w^{-1}) = u}\,\mathrm{p}' \qquad \overset{(C_5\sigma 1)}{\vdash (u'\circ w')\circ w'^{-1} = u'\circ(w'\circ w'^{-1})}}{\vdash (u\circ w)\circ w^{-1} = u}\,\mathrm{p}:\sigma_2 \qquad (C_{17})$$

where $\sigma_1 = \{u\mapsto u', w\mapsto w'\}$ and $\sigma_2 = \{u'\mapsto u, w'\mapsto w\}$.

Derivation of C_{18}:

$$\dfrac{\dfrac{\overset{(C_{13})}{\vdash a\circ a = e} \quad \overset{(C_{14})}{\vdash e\circ b = b}}{\vdash (a\circ a)\circ b = b}\,\mathrm{p}' \qquad \overset{(C_{12})}{\vdash (a\circ a)\circ b = a\circ(a\circ b)}}{\vdash b = a\circ(a\circ b)}\,\mathrm{p} \qquad (C_{18})$$

Derivation of C_{19}:

$$\dfrac{\dfrac{\overset{(C_{18})}{\vdash b = a \circ (a \circ b)} \quad \overset{(C_{15})}{\vdash b \circ b = e}}{\vdash b \circ (a \circ (a \circ b)) = e}\;\mathrm{p} \qquad \overset{(C_{11})}{\vdash (b \circ a) \circ (a \circ b) = b \circ (a \circ (a \circ b))}}{\vdash (b \circ a) \circ (a \circ b) = e}\;\mathrm{p} \qquad (C_{19})$$

Derivation of C_{20}:

$$\dfrac{\overset{(C_{17})}{\vdash (u \circ w) \circ w^{-1} = u} \quad \overset{(C_{17}\sigma_3)}{\vdash (u' \circ w') \circ w'^{-1} = u'}}{\vdash u \circ (w^{-1})^{-1} = u \circ w}\;\mathrm{p} : \sigma_4 \qquad (C_{20})$$

where $\sigma_3 = \{u \mapsto u', w \mapsto w'\}$ and $\sigma_4 = \{u' \mapsto u \circ w, w' \mapsto w^{-1}\}$.

Derivation of C_{21}:

$$\dfrac{\overset{(C_{10})}{\vdash (a \circ b) \circ (a \circ b) = e} \quad \overset{(C_{17})}{\vdash (u \circ w) \circ w^{-1} = u}}{\vdash e \circ (a \circ b)^{-1} = a \circ b}\;\mathrm{p} : \sigma_5 \qquad (C_{21})$$

where $\sigma_5 = \{u \mapsto a \circ b, w \mapsto a \circ b\}$.

Derivation of C_{22}:

$$\dfrac{\overset{(C_{20})}{\vdash u \circ (w^{-1})^{-1} = u \circ w} \quad \dfrac{\overset{(C_7\sigma_6)}{\vdash w' \circ w'^{-1} = e} \quad \overset{(C_{17})}{\vdash (u \circ w) \circ w^{-1} = u}}{\vdash e \circ (w'^{-1})^{-1} = w'}\;\mathrm{p} : \sigma_7}{\vdash e \circ w = w}\;\mathrm{p} : \sigma_8 \qquad (C_{22})$$

where $\sigma_6 = \{w \mapsto w'\}$, $\sigma_7 = \{u \mapsto w', w \mapsto w'^{-1}\}$ and $\sigma_8 = \{u \mapsto e, w' \mapsto w\}$.

Derivation of C_{23}:

$$\dfrac{\overset{(C_{22})}{\vdash e \circ w = w} \quad \overset{(C_{21})}{\vdash e \circ (a \circ b)^{-1} = a \circ b}}{\vdash (a \circ b)^{-1} = a \circ b}\;\mathrm{p} : \sigma_9 \qquad (C_{23})$$

where $\sigma_9 = \{w \mapsto (a \circ b)^{-1}\}$.

Derivation of C_{24}:

$$\dfrac{\overset{(C_{22})}{\vdash e \circ w = w} \quad \dfrac{\overset{(C_{23})}{\vdash (a \circ b)^{-1} = a \circ b} \quad \dfrac{\overset{(C_{19})}{\vdash (b \circ a) \circ (a \circ b) = e} \quad \overset{(C_{17})}{\vdash (u \circ w) \circ w^{-1} = u}}{\vdash e \circ (a \circ b)^{-1} = b \circ a}\;\mathrm{p} : \sigma_{10}}{\vdash e \circ (a \circ b) = b \circ a}\;\mathrm{p}}{\vdash a \circ b = b \circ a}\;\mathrm{p} : \sigma_{11} \qquad (C_{24})$$

where $\sigma_{10} = \{u \mapsto b \circ a, w \mapsto a \circ b\}$ and $\sigma_{11} = \{w \mapsto a \circ b\}$.

And finally we have a refutation:

$$\dfrac{\overset{(C_{24})}{\vdash a \circ b = b \circ a} \quad \overset{(C_1)}{a \circ b = b \circ a \vdash}}{\vdash} \; \mathrm{r}$$

Let ψ be the **PR**-refutation defined above (in form of a tree). By computing a ground projection ψ' of ψ we obtain a derivation of $\vdash$ in **LK-e** from instances of $\Theta(\varphi)$. There is only one non-trivial proof projection ψ_1' required, namely this to the clause C_1. The proof of $\vdash a \circ b = b \circ a$ with only atomic cuts φ^* is therefore:

$$\dfrac{a \circ b = b \circ a \vdash a \circ b = b \circ a \quad \overset{(\psi_1')}{a \circ b = b \circ a \vdash}}{\vdash a \circ b = b \circ a} \; \mathrm{cut}$$

By re-skolemizing φ^* we obtain a proof $\hat{\varphi}$ of $\vdash (\forall x)(\forall y) x \circ y = y \circ x$ with only atomic cuts (and clearly without use of the cancellation principle).

Note that the above **PR**-refutation was found by use of the automatic theorem prover Otter. Remarkable is also the fact that the left-neutrality appears as lemma within the resolution refutation based on the clause set $\Theta(\varphi)$.

4 Equational Theories

Definition 19 (equational axiom set). *A (possibly infinite) set $\mathcal{E}$ of term equations, i.e.*

$$\mathcal{E} = \{s_1 = t_1, s_2 = t_2, s_3 = t_3 \dots\},$$

is called an equational axiom set if it is closed under substitution, i.e. for all $E \in \mathcal{E}$ and for all substitutions σ we have $E\sigma \in \mathcal{E}$.

Definition 20 (equational theory). *Let $\mathcal{E}$ be an equational axiom set. An equational theory is defined as a congruence relation on $\mathcal{E}$ in the following way:*

$$s =_{\mathcal{E}} t \quad \Leftrightarrow \quad \mathcal{E} \models s = t.$$

Based on equational theories the presented calculi can be redefined to use the equational theory at every inference. We illustrate this principle by defining extended $\supset$ rules; the extension for the other rules is analogous.

$$\dfrac{\Gamma \vdash \Delta, \boldsymbol{A^*} \quad \boldsymbol{B^*}, \Pi \vdash \Lambda}{\underline{A \supset B}, \Gamma, \Pi \vdash \Delta, \Lambda} \supset : \mathrm{l} \qquad \dfrac{\boldsymbol{A^*}, \Gamma \vdash \Delta, \boldsymbol{B^*}}{\Gamma \vdash \Delta, \underline{A \supset B}} \supset : \mathrm{r}$$

if $A =_{\mathcal{E}} A^*$ and $B =_{\mathcal{E}} B^*$.

The following example demonstrates the usage of the concept of equational theories within **LK-ep**. In this example we will use **LK-ep**-inferences modulo groups, i.e. the underlying equational theory is the theory of groups G (the binary connective of G is $\circ$, the neutral element is e).

We define the following axioms:

$$\begin{aligned} A_1\colon &\ (\forall x)(x \neq s(x) \wedge x \neq s(s(x))),\\ A_2\colon &\ (\forall x)(\forall y)(x = s(y) \vee y = s(x) \vee x = y),\\ A_3\colon &\ (\forall x)(\forall y)(s(x) = s(y) \supset x = y) \text{ and the conclusion}\\ C\colon &\ (\forall x)(\forall y) x \circ y = y \circ x \end{aligned}$$

and use the following abbreviation $A_x\colon A_1, A_2, A_3$ for the entire axiom set.

We consider the following proof with G-pseudo-cuts:

$$\dfrac{\overset{\phi}{A_x \vdash (\forall y)(y = e \vee y = s(e) \vee y = s(s(e)))} \quad \overset{\psi}{(\exists z, z_1, z_2)(\forall y)(y = z \circ z_1 \vee y = z \circ z_2 \vee y = z) \vdash C}}{A_x \vdash C}\ \mathrm{p-cut}$$

Informal proof ϕ: From A_2 we get:

$$(\forall y)(e = s(y) \vee y = s(e) \vee e = y). \tag{*}$$

and setting y to $s(s(e))$ we get

$$e = s(s(s(e))) \vee s(s(e)) = s(e) \vee e = s(s(e)).$$

From A_1 we infer $e = s(s(s(e)))$. By $e = s(s(s(e)))$ and A_3 we get

$$(\forall y)(e = s(y) \leftrightarrow y = s(s(e))).$$

Therefore, from $(*)$ we obtain the left cut-formula

$$(\forall y)(y = e \vee y = s(e) \vee y = s(s(e))). \tag{I}$$

Note that no group theoretic inferences are required in ϕ.

Informal proof ψ: From the left cut-formula

$$(\exists z, z_1, z_2)(\forall y)(y = z \circ z_1 \vee y = z \circ z_2 \vee y = z) \tag{II}$$

it follows that the structure consists of three elements only. As the underlying structure is G we have a group with 3 elements. But there is only one such group and this is commutative. Therefore (II) and G imply C.

Moreover we have a pseudo-cut w.r.t. G. In fact (I) $\supset$ (II) is valid under G: just choose $z = e$, $z_1 = s(e)$, and $z_2 = s(s(e))$. This cut can be eliminated with

CERES-e under use of G.

Note that, within the example, the subproof ψ, i.e. the proof of the group containing only 3 elements, demonstrates also the expressional power of the combined method since it would not be possible to be proven without having equality also as a rule.

The example above should be treated as an appetizer for what is possible by means of the CERES-e method using equational theories.

5 Conclusion

The extensions of CERES described in this paper lay the foundations for cut-elimination of mathematically relevant proofs in **LK** with equality. The formalization of proofs containing cuts as well as the representation of proofs with only atomic cuts benefit from these extensions. Another important aspect of the results of this paper is a computational one. CERES has been shown to yield non-elementary speedups w.r.t. other cut-elimination methods like Gentzen or Schütte-Tait and is therefore the best choice for such extensions, especially from the implementational point of view. Additionally the use of numerous frequently used paradigms like paramodulation and its refinements admits the application of a variety of existing tools like Otter, SPASS or Vampire. Last but not least CERES-e, especially using the concept of equational theories, allows for handling (even more complex) proofs in a mathematically natural way.

References

1. M. Baaz, A. Leitsch: On skolemization and proof complexity, *Fundamenta Informaticae*, 20(4), pp. 353–379, 1994.
2. M. Baaz, A. Leitsch: Cut-Elimination and Redundancy-Elimination by Resolution, *Journal of Symbolic Computation*, 29, pp. 149-176, 2000.
3. M. Baaz, A. Leitsch: Towards a Clausal Analysis of Cut-Elimination, *Journal of Symbolic Computation* to appear.
4. M. Baaz, S. Hetzl, A. Leitsch, C. Richter and H. Spohr: Cut-Elimination: Experiments with CERES, *Proc. LPAR 2004*, pp. 481-495, Springer, 2004.
5. C.-L. Chang and R. C.-T. Lee: Symbolic Logic and Mechanical Theorem Proving, Academic Press, Boston, 1973.
6. A. Degtyarev and A. Voronkov: Equality Reasoning in Sequent-Based Calculi *Handbook of Automated Reasoning* vol. I, ed. by A. Robinson and A. Voronkov, chapter 10, pp. 611-706, Elsevier Science, 2001.
7. G. Gentzen: Untersuchungen über das logische Schließen, *Mathematische Zeitschrift* 39, pp. 405–431, 1934–1935.
8. J.Y. Girard: Proof Theory and Logical Complexity, in *Studies in Proof Theory*, Bibliopolis, Napoli, 1987.
9. G. Takeuti: Proof Theory, North-Holland, Amsterdam, 2nd edition, 1987.
10. A. S. Troelstra and H. Schwichtenberg: Basic Proof Theory, Cambridge University Press, 1996.