Extensional Higher-Order Paramodulation and RUE-Resolution

Christoph Benzmüller

Fachbereich Informatik, Universität des Saarlandes
chris@ags.uni-sb.de

Abstract. This paper presents two approaches to primitive equality treatment in higher-order (HO) automated theorem proving: a calculus $\mathcal{EP}$ adapting traditional first-order (FO) paramodulation [RW69] , and a calculus $\mathcal{ERUE}$ adapting FO RUE-Resolution [Dig79] to classical type theory, i.e., HO logic based on Church's simply typed λ-calculus. $\mathcal{EP}$ and $\mathcal{ERUE}$ extend the extensional HO resolution approach $\mathcal{ER}$ [BK98a]. In order to reach Henkin completeness without the need for additional extensionality axioms both calculi employ new, positive extensionality rules analogously to the respective negative ones provided by $\mathcal{ER}$ that operate on unification constraints. As the extensionality rules have an intrinsic and unavoidable difference-reducing character the HO paramodulation approach loses its pure term-rewriting character. On the other hand examples demonstrate that the extensionality rules harmonise quite well with the difference-reducing HO RUE-resolution idea.

1 Introduction

Higher-Order (HO) Theorem Proving based on the resolution method has been first examined by Andrews [And71] and Huet [Hue72]. Whereas the former avoids unification the latter generally delays the computation of unifiers and instead adds unification constraints to the clauses in order to tackle the undecidability problem of HO unification. More recent papers concentrate on the adaption of sorts [Koh94] or theory unification [Wol93] to HO logic. Common to all these approaches is that they do not sufficiently solve the extensionality problem in HO automated theorem proving, i.e., all these approaches require the extensionality axioms to be added into the search space in order to reach Henkin completeness (which is the most general notion of semantics that allows complete calculi [Hen50]). This leads to a search space explosion that is awkward to manage in practice. A solution to the problem is provided by the extensional HO resolution calculus $\mathcal{ER}$ [BK98a]. This approach avoids the extensionality axioms and instead extends the syntactical (pre-)unification process by additional extensionality rules. These new rules allow for recursive calls during the (pre-) unification process to the overall refutation search whenever pure syntactical HO unification is too weak to show that two terms can be equalised modulo the extensionality principles. $\mathcal{ER}$ has been implemented in Leo [BK98b] and case studies have demonstrated its suitability, especially for reasoning about sets.

There are many possibilities to improve the extensional HO resolution approach and the probably most promising one concerns the treatment of equality. $\mathcal{ER}$ assumes that equality is defined by the Leibniz principle (two things

H. Ganzinger (Ed.): CADE-16, LNAI 1632, pp. 399–413, 1999.

are equal iff they have the same properties) or by any other valid definition principle, and thus provides no support for primitive equality. But a primitive equality treatment seems to be more appropriate as it avoids the many flexible literals introduced when using defined equality, which unfortunately increase the amount of blind search with $\mathcal{ER}$'s primitive substitution rule *Prim*. Therefore we adapt two well known first-order (FO) approaches to primitive equality: the paramodulation approach [RW69] (the basis of many successful refinements such as the superposition approach) and the RUE-resolution approach [Dig79] (a generalisation of E-resolution [Dar68]). The main goal thereby is to preserve Henkin completeness. We will show that therefore positive extensionality rules are needed (which operate on positive equation literals) as in contrast to FO logic single positive equations can be contradictory by themselves in HO logic.[1]

This paper summarises the Chapt. 6, 7, and 8 of [Ben99] and because of lack of space the preliminaries and the formal proofs can only be sketched here.

The preliminaries are concisely presented in Sect. 2 and calculus $\mathcal{ER}$ is reviewed in 3. Section 4 discusses interesting aspects on primitive and defined equality, before the extensional HO paramodulation calculus $\mathcal{EP}$ and the extensional HO RUE-resolution approach $\mathcal{ERUE}$ are discussed in 5 and 6. Both approaches are briefly compared by examples in 7 and the conclusion is presented in 8.

2 Higher-Order (HO) Logic

We consider a HO logic based on Church's simply typed λ-calculus [Chu40] and choose $BT := \{\iota, o\}$ as *base types*, where ι denotes the set of individuals and o the set of truth values. *Functional types* are inductively defined over BT. A *signature* Σ ($\Sigma^=$) contains for each type an infinite set of variables and constants and provides the *logical connectives* $\neg_{o\to o}$, $\vee_{o\to o\to o}$, and $\Pi_{(\alpha\to o)\to o}$ (additionally $=^\alpha \; := \; =_{\alpha\to\alpha\to o}$) for every type α. The set of all Σ-terms (closed Σ-terms) of type α is denoted by wff_α (cwff_α). Variables are printed as uppercase (e.g. X_α), constants as lower-case letters (e.g. c_α) and arbitrary terms appear as bold capital letters (e.g. $\mathbf{T}_\alpha$). If the type of a symbol is uniquely determined by the given context we omit it. We abbreviate function applications by $h_{\alpha_1\to\cdots\to\alpha_n\to\beta}\,\overline{\mathbf{U}^n_{\alpha_n}}$, which stands for $(\cdots(h_{\alpha_1\to\cdots\to\alpha_n\to\beta}\,\mathbf{U}^1_{\alpha_1})\cdots\mathbf{U}^n_{\alpha_n})$. For α-, β-, η-, $\beta\eta$-*conversion* and the definition of $\beta\eta$- and *head-normal form* (*hnf*) for a term $\mathbf{T}$ we refer to [Bar84] as well as for the definition of *free* variables, *closed* formulas, and *substitutions*. *Unification* and sets of *partial bindings* $\mathcal{AB}^h_\gamma$ are well explained in [SG89]. An example for a *pre-clause*, i.e., not in proper clause normal form, consisting of a positive literal, a negative literal, and a special negative equation literal (also called *unification constraint*) is $\mathcal{C} : [\neg(P_{\iota\to o}\;\mathbf{T}_\iota)]^T \vee [h_{\overline{\gamma}\to o}\;\overline{Y^n_{\gamma_n}}]^F \vee [Q_{\iota\to\iota}\;a_\iota = Y_{\iota\to\iota}\;b_\iota]^F$. The corresponding *proper clause*, i.e., properly normalised, is $\mathcal{C}' : [P\;\mathbf{T}]^F \vee [h\;\overline{Y^n}]^F \vee [Q\;a = Y\;b]^F$. The unification constraint in $\mathcal{C}$ and $\mathcal{C}'$ is called a *flex-flex* pair as both unification terms have *flexible* heads. A clause is called *empty*, denoted by $\Box$, if it consists

[1] Consider, e.g. the positive literal $[a_o = \neg a_o]^T$ or $[G\;X = f]^T$ (resulting from the following formulation of Cantor's theorem: $\neg\exists G_{\iota\to\iota\to o}.\,\forall P_{\iota\to o}.\,\exists X_\iota.\,G\;X = P$).

only of *flex-flex* unification constraints. A clause $\mathcal{C}_1$ generalises a clause $\mathcal{C}_2$, iff there is a substitution σ, such that the $\beta\eta$-normal form of $\sigma(\mathcal{C}_1)$ is an α-variant of the $\beta\eta$-normal form of $\mathcal{C}_2$.

A calculus R provides a set of rules $\{r_n |\ 0 < n \leq i\}$ defined on clauses. We write $\Phi \vdash^{r_n} \mathcal{C}$ ($\mathcal{C}' \vdash^{r_n} \mathcal{C}$) iff clause $\mathcal{C}$ is the result of an one step application of rule $r_n \in R$ to premise clauses $\mathcal{C}'_i \in \Phi$ (to $\mathcal{C}'$ respectively). Multiple step derivations in calculus R are abbreviated by $\Phi_1 \vdash_R \Phi_k$ (or $\mathcal{C}_1 \vdash_R \mathcal{C}_k$).

A *standard model* for $\mathcal{HOL}$ provides a fixed set $\mathcal{D}_\iota$ of individuals, and a set $\mathcal{D}_o := \{\top, \bot\}$ of truth values. The domains for functional types are defined inductively: $\mathcal{D}_{\alpha\to\beta}$ is the set of all functions $f\colon \mathcal{D}_\alpha \to \mathcal{D}_\beta$. *Henkin models* only require that $\mathcal{D}_{\alpha\to\beta}$ has enough members that any well-formed formula can be evaluated. Thus, the generalisation to Henkin models restricts the set of valid formulae sufficiently, such that complete calculi are possible. In Henkin and standard semantics Leibniz equality ($\doteq^\alpha := \lambda X_\alpha.\ \lambda Y_\alpha.\ \forall P_{\alpha\to o}.\ PX \Rightarrow PY$) denotes the intuitive equality relation and the functional extensionality principles $(\forall M_{\alpha\to\beta}.\ \forall N_{\alpha\to\beta}.\ (\forall X.\ (MX) = (NX)) \Leftrightarrow (M = N))$ as well as the Boolean extensionality principle $(\forall P_o.\ \forall Q_o.\ (P = Q) \Leftrightarrow (P \Leftrightarrow Q))$ are valid (see [Ben99,BK97]). *Satisfiability* and *validity* ($\mathcal{M} \models \mathbf{F}$ or $\mathcal{M} \models \Phi$) of a formula $\mathbf{F}$ or set of formulas Φ in a model $\mathcal{M}$ is defined as usual.

The completeness proofs employ the abstract consistency method of [BK97]& [Ben99] which extends Andrews' HO adaptation [And71] of Smullyan's approach [Smu63] to Henkin semantics. Here we only mention the two main aspects:

Definition 1 (*Acc* **for Henkin Models).** *Let Σ be a signature and Γ_Σ a class of sets of Σ-sentences. If the following conditions (all but ∇_e^*) hold for all $\mathbf{A}, \mathbf{B} \in cwff_o$, $\mathbf{F}, \mathbf{G} \in cwff_{\alpha\to\beta}$, and $\Phi \in \Gamma_\Sigma$, then we call Γ_Σ an* **abstract consistency class for Henkin models with primitive equality**, *abbreviated by $Acc^=$ (resp.* **abstract consistency class for Henkin models**, *abbreviated by Acc).*

Saturated $\Phi \cup \{\mathbf{A}\} \in \Gamma_\Sigma$ *or* $\Phi \cup \{\neg\mathbf{A}\} \in \Gamma_\Sigma$.

∇_c *If $\mathbf{A}$ is atomic, then $\mathbf{A} \notin \Phi$ or $\neg\mathbf{A} \notin \Phi$.*

$\nabla_\neg$ *If $\neg\neg\mathbf{A} \in \Phi$, then $\Phi \cup \{\mathbf{A}\} \in \Gamma_\Sigma$.*

∇_f *If $\mathbf{A} \in \Phi$ and $\mathbf{B}$ is the $\beta\eta$-normal form of $\mathbf{A}$, then $\Phi \cup \{\mathbf{B}\} \in \Gamma_\Sigma$.*

$\nabla_\vee$ *If $\mathbf{A} \vee \mathbf{B} \in \Phi$, then $\Phi \cup \{\mathbf{A}\} \in \Gamma_\Sigma$ or $\Phi \cup \{\mathbf{B}\} \in \Gamma_\Sigma$.*

$\nabla_\wedge$ *If $\neg(\mathbf{A} \vee \mathbf{B}) \in \Phi$, then $\Phi \cup \{\neg\mathbf{A}, \neg\mathbf{B}\} \in \Gamma_\Sigma$.*

$\nabla_\forall$ *If $\Pi^\alpha\mathbf{F} \in \Phi$, then $\Phi \cup \{\mathbf{F}\ \mathbf{W}\} \in \Gamma_\Sigma$ for each $\mathbf{W} \in cwff_\alpha$.*

$\nabla_\exists$ *If $\neg\Pi^\alpha\mathbf{F} \in \Phi$, then $\Phi \cup \{\neg(\mathbf{F}\ w)\} \in \Gamma_\Sigma$ for any new constant $w \in \Sigma_\alpha$.*

∇_b *If $\neg(\mathbf{A} \doteq^o \mathbf{B}) \in \Phi$, then $\Phi \cup \{\mathbf{A}, \neg\mathbf{B}\} \in \Gamma_\Sigma$ or $\Phi \cup \{\neg\mathbf{A}, \mathbf{B}\} \in \Gamma_\Sigma$.*

∇_q *If $\neg(\mathbf{F} \doteq^{\alpha\to\beta} \mathbf{G}) \in \Phi$, then $\Phi \cup \{\neg(\mathbf{F}\ w \doteq^\beta \mathbf{G}\ w)\} \in \Gamma_\Sigma$ for any new constant $w \in \Sigma_\alpha$.*

∇_e^r $\neg(\mathbf{A}_\alpha = \mathbf{A}) \notin \Phi$. ∇_e^s *If $\mathbf{F}[\mathbf{A}]_p \in \Phi$ and $\mathbf{A} = \mathbf{B} \in \Phi$, then $\Phi \cup \{\mathbf{F}[\mathbf{B}]_p\} \in \Gamma_\Sigma$.*[2]

Theorem 1 (Henkin Model Existence). *Let Φ be a set of closed Σ-formulas ($\Sigma^=$-formulas), Γ_Σ ($\Gamma_{\Sigma^=}$) be an Acc ($Acc^=$) and $\Phi \in \Gamma_\Sigma$. There exists a Henkin model $\mathcal{M}$, such that $\mathcal{M} \models \Phi$.*

$$\frac{\mathbf{C} \vee [\mathbf{A} \vee \mathbf{B}]^T}{\mathbf{C} \vee [\mathbf{A}]^T \vee [\mathbf{B}]^T} \vee^T \quad \frac{\mathbf{C} \vee [\mathbf{A} \vee \mathbf{B}]^F}{\mathbf{C} \vee [\mathbf{A}]^F} \vee_l^F \quad \frac{\mathbf{C} \vee [\mathbf{A} \vee \mathbf{B}]^F}{\mathbf{C} \vee [\mathbf{B}]^F} \vee_r^F$$

$$\frac{\mathbf{C} \vee [\neg \mathbf{A}]^T}{\mathbf{C} \vee [\mathbf{A}]^F} \neg^T \quad \frac{\mathbf{C} \vee [\neg \mathbf{A}]^F}{\mathbf{C} \vee [\mathbf{A}]^T} \neg^F \quad \frac{\mathbf{C} \vee [\Pi^\gamma \mathbf{A}]^T \quad X_\gamma \text{ new variable}}{\mathbf{C} \vee [\mathbf{A}\ X]^T} \Pi^T$$

$$\frac{\mathbf{C} \vee [\Pi^\gamma \mathbf{A}]^F \quad \mathrm{sk}_\gamma \text{ is a Skolem term for this clause}}{\mathbf{C} \vee [\mathbf{A}\ \mathrm{sk}_\gamma]^F} \Pi^F$$

Fig. 1. Clause Normalisation Calculus $\mathcal{CNF}$

$\boxed{\text{Clause Normalisation:}}$

$$\frac{\mathcal{D} \quad \mathcal{C} \in \mathcal{CNF}(\mathcal{D})}{\mathcal{C}} \; Cnf$$

$\boxed{\text{Resolution:}}$ (defined on proper clauses only)

$$\frac{[\mathbf{A}]^\alpha \vee \mathbf{C} \quad [\mathbf{B}]^\beta \vee \mathbf{D} \quad \alpha \neq \beta}{\mathbf{C} \vee \mathbf{D} \vee [\mathbf{A} = \mathbf{B}]^F} \; Res \qquad \frac{[\mathbf{A}]^\alpha \vee [\mathbf{B}]^\alpha \vee \mathbf{C} \quad \alpha \in \{T, F\}}{[\mathbf{A}]^\alpha \vee \mathbf{C} \vee [\mathbf{A} \ = \mathbf{B}]^F} \; Fac$$

$$\frac{[Q_\gamma\ \overline{\mathbf{U}^k}]^\alpha \vee \mathbf{C} \quad \mathbf{P} \in \mathcal{AB}_\gamma^{\{\neg,\vee\}\cup\{\Pi^\beta|\beta\in\mathcal{T}^k\}},\ \alpha \in \{T, F\}}{[Q_\gamma\ \overline{\mathbf{U}^k}]^\alpha \vee \mathbf{C} \vee [Q = \mathbf{P}]^F} \; Prim$$

$\boxed{\text{Extensional (Pre-)Unification:}}$

$$\frac{\mathbf{C} \vee [\mathbf{M}_{\gamma\to\beta} = \mathbf{N}_{\gamma\to\beta}]^F \quad s_\gamma \text{ Skolem term for this clause}}{\mathbf{C} \vee [\mathbf{M}\ s = \mathbf{N}\ s]^F} \; Func$$

$$\frac{\mathbf{C} \vee [\mathbf{A}_{\gamma\to\beta}\ \mathbf{C}_\gamma = \mathbf{B}_{\gamma\to\beta}\ \mathbf{D}_\gamma]^F}{\mathbf{C} \vee [\mathbf{A} = \mathbf{B}]^F \vee [\mathbf{C} = \mathbf{D}]^F} \; Dec$$

$$\frac{\mathbf{C} \vee [\mathbf{A} = \mathbf{A}]^F}{\mathbf{C}} \; Triv \qquad \frac{\mathbf{C} \vee [X = \mathbf{A}]^F \quad X \text{ does not occur in } \mathbf{A}}{\mathbf{C}_{\{\mathbf{A}/X\}}} \; Subst$$

$$\frac{\mathbf{C} \vee [F_\gamma\ \overline{\mathbf{U}^n} = h\ \overline{\mathbf{V}^m}]^F \quad \mathbf{G} \in \mathcal{AB}_\gamma^h}{\mathbf{C} \vee [F\ \overline{\mathbf{U}^n} = h\ \overline{\mathbf{V}^m} \vee [F = \mathbf{G}]^F]^F} \; FlexRigid$$

$$\frac{\mathbf{C} \vee [\mathbf{M}_o = \mathbf{N}_o]^F}{\mathbf{C} \vee [\mathbf{M}_o \Leftrightarrow \mathbf{N}_o]^F} \; Equiv \qquad \frac{\mathbf{C} \vee [\mathbf{M}_\gamma = \mathbf{N}_\gamma]^F}{\mathbf{C} \vee [\forall P_{\gamma\to o}\text{.}\ P\ \mathbf{M} \Rightarrow P\ \mathbf{N}]^F} \; Leib$$

$$\left(\frac{\mathbf{C} \vee [F_{\overline{\gamma^n}\to\gamma}\ \overline{\mathbf{U}^n} = H_{\overline{\delta^m}\to\gamma}\ \overline{\mathbf{V}^m}]^F \quad \mathbf{G} \in \mathcal{AB}_{\overline{\gamma^n}\to\gamma}^h \text{ for a constant } h_\tau}{\mathbf{C} \vee [F\ \overline{\mathbf{U}^n} = H\ \overline{\mathbf{V}^m}]^F \vee [F = \mathbf{G}]^F} \; FlexFlex\right)$$

$\mathcal{AB}_\gamma^h$ specifies the set of partial bindings of type γ for head h as defined in [SG89]

Fig. 2. Extensional HO Resolution Calculus $\mathcal{ER}$

3 $\mathcal{ER}$: Extensional HO Resolution

Figure 1 presents calculus $\mathcal{CNF} := \{\vee^T, \vee^F_l, \vee^F_r, \neg^T, \neg^F, \Pi^T, \Pi^F\}$ for clause normalisation. These rules are defined on (pre-)clauses and are known to preserve validity or satisfiability with respect to standard semantics.[3]

The syntactical unification rules (cf. Fig. 2) provided by $\mathcal{ER}$ which operate on unification constraints are $\mathcal{UNI} := \{Func, Dec, Triv, Subst,\ FlexRigid\}$. These rules realise a sound and complete approach to HO pre-unification. Note the double role of extensionality rule *Func*: on the one hand this rule works as a syntactical unification rule and subsumes the α- and η-rule as, e.g. presented in [BK98a]; on the other hand *Func* applies the functional extensionality principle if none of the two terms is a λ-abstraction. Apart from rule *Func*, $\mathcal{ER}$ provides the extensionality rules *Equiv* and *Leib* (cf. Fig. 2). The former applies the Boolean extensionality principle and the latter simply replaces a negative unification constraint (encoded as a negative equation) by a negative Leibniz equation. The extensionality rules operate on unification constraints only and do in contrast to the respective axioms not introduce flexible heads into the search space.

The main proof search is performed by the resolution rule *Res* and the factorisation rule *Fac*. It is furthermore well known for HO resolution, that the primitive substitution rule *Prim* is needed to ensure Henkin completeness.

For the calculi presented in this paper we assume that the result of each rule application is transformed into hnf[4], where the hnf of unification constraints is defined special and requires both unification terms to be reduced to hnf. A set of formulas Φ is refutable in calculus R, iff there is a derivation $\Delta : \Phi_{cl} \vdash_R \Box$, where $\Phi_{cl} := \{[\mathbf{F}']^T | \mathbf{F}' \text{ hnf of } \mathbf{F} \in \Phi\}$ is the clause-set obtained from Φ by simple pre-clausification. More details on $\mathcal{ER}$ are provided by [BK98a,Ben99].

Whereas completeness of $\mathcal{ER}$ has already been analysed in [BK98a] this paper (and [Ben99]) presents an alternative completeness proof for a slightly extended version of $\mathcal{ER}$ (this version, e.g. employs the instantiation guessing *FlexFlex*-rule). The new proof is motivated as follows: (i) it eases the proof of the lifting lemma and avoids the quite complicated notion of clause isomorphisms as used in [BK98a,Koh94], (ii) it can be reused to show the completeness for calculi $\mathcal{EP}$ and $\mathcal{ERUE}$ as well, (iii) it prepares the analysis of non-normal form resolution calculi, and (iv) it emphasises interesting aspects on rule *FlexFlex*, unification, and clause normalisation wrt. $\mathcal{ER}$, $\mathcal{EP}$, and $\mathcal{ERUE}$.

One such interesting aspect is that different to Huet [Hue72] eager unification is essential within our approach. This is illustrated by the argumentations for ∇_b and ∇_q in the completeness proofs (cf. [Ben99,BK98a]) as well as the examples presented in Sec. 7 or [Ben99]. However, we claim that rule *FlexFlex* can still be delayed until the end of a refutation, i.e., *FlexFlex* can be completely avoided.

The author has not been able to prove the latter claim yet. And thus the completeness proofs for $\mathcal{ER}$ (and $\mathcal{EP}$, $\mathcal{ERUE}$) still depends on the *FlexFlex*-rule.

[2] $\mathbf{A}$ does not contain free variables.
[3] For Skolemisation we employ Miller's sound HO correction [Mil83].
[4] One may also $\beta\eta$-normal form here.

We now sketch the main results on $\mathcal{ER}$ as discussed in detail in [Ben99].

Definition 2 (Extensional HO Resolution). *We define three calculi:*

$\mathcal{ER} := \{Cnf, Res, Fac, Prim\} \cup \mathcal{UNI} \cup \{Equiv, Leib\}$ *employs all rules (except* FlexFlex*) displayed in Fig. 2.*

$\mathcal{ER}_f := \mathcal{ER} \cup \{FlexFlex\}$ *uses full HO unification instead of pre-unification.*

$\mathcal{ER}_{fc} := (\mathcal{ER}_f \backslash \{Cnf\}) \cup \mathcal{CNF}$ *employs unfolded and stepwise clause normalisation instead of exhaustive normalisations with rule* Cnf.

These calculi treat equality as a defined notion only (e.g. by Leibniz equality) and primitive equations are not allowed in problem formulations. Although unification constraints are encoded as negative equation literals, no rule but the unification rules are allowed to operate on them.

Theorem 2 (Soundness). *The calculi $\mathcal{ER}$, $\mathcal{ER}_f$, and $\mathcal{ER}_{fc}$ are Henkin-sound (H-sound).*

Proof. Preservation of validity or satisfiability with respect to Henkin semantics is proven analogously to the standard FO argumentation. For Skolemisation (employed in rule Π^F and *Func*) we use Miller's sound HO version [Mil83]. Soundness of the extensionality rules *Equiv, Func*, and *Leib* is obvious as they simply apply the valid extensionality principles.

Lemma 1 (Lifting of $\mathcal{ER}_{fc}$). *Let Φ be clause set, $\mathcal{D}_1$ a clause, and σ a substitution. If $\sigma(\Phi) \vdash_{\mathcal{ER}_{fc}} \mathcal{D}_1$, then $\Phi \vdash_{\mathcal{ER}_{fc}} \mathcal{D}_2$ for a clause $\mathcal{D}_2$ generalising $\mathcal{D}_1$.*

Proof. One can easily show that each instantiated derivation can be reused on the uninstantiated level as well. In blocking situations caused by free variables at literal head position or at unification term head position, either rule *Prim* or rule *FlexFlex* can be employed in connection with rule *Subst* to introduce the missing term structure. The rather unintuitive clause isomorphisms of [BK98a] or [Koh94] are thereby avoided.

Theorem 3 (Completeness). *Calculus $\mathcal{ER}_{fc}$ is Henkin complete.*

Proof. Analogously to the proof in [BK98a] we show that the set of closed formulas that are not refutable in $\mathcal{ER}_{fc}$ (i.e., $\Gamma_\Sigma := \{\Phi \subseteq cwff_o | \Phi_{cl} \not\vdash_{\mathcal{ER}_{fc}} \Box\}$) is a saturated abstract consistency class for Henkin models (cf. Def. 1). This entails Henkin completeness for $\mathcal{ER}_{fc}$ by Thm. 1.

Lemma 2 (Theorem Equivalence). *The calculi $\mathcal{ER}_{fc}$ and $\mathcal{ER}_f$ are theorem equivalent, i.e., for each clause set Φ holds that $\Phi \vdash_{\mathcal{ER}_{fc}} \Box$ iff $\Phi \vdash_{\mathcal{ER}_f} \Box$.*

Proof. We can even prove a more general property: For each proper clause $\mathcal{C}$ holds $\Phi \vdash_{\mathcal{ER}_{fc}} \mathcal{C}$ implies $\Phi \vdash_{\mathcal{ER}_f} \mathcal{C}$. The proof idea is to show that the unfolded clause normalisations can be grouped together and then replaced by rule *Cnf*.

Question 1 (Theorem Equivalence). The author claims that the calculi $\mathcal{ER}$ and $\mathcal{ER}_{fc}$ (or $\mathcal{ER}_f$) are theorem equivalent. A formal proof has not been carried out yet. Some evidence is given by the case studies carried out with the LEO-prover [BK98b] and the direct completeness proof for $\mathcal{ER}$ in [BK98a].

4 Primitive Equality

Treating equality as a defined notion in HO logic (e.g. by the Leibniz principle) is convenient in theory, but often inefficient and unintuitive in practical applications as many free literal heads are introduced into the search space, which increases the degree of blind search with primitive substitution rule *Prim*.[5] This is the main motivation for the two approaches to primitive equality presented in the next sections. Before we discuss these approaches in detail we point to the following interesting aspects of defined equality in HO logic:

- There are infinitely many different valid definitions of equality in HO logic.[6] For instance: Leibniz equality ($\doteq^{\alpha} := \lambda X_{\alpha}.\ \lambda Y_{\alpha}.\ \forall P_{\alpha\to o}.\ PX \Rightarrow PY$), Reflexivity definition[7] ($\ddot{=}^{\alpha} := \lambda X_{\alpha}.\ \lambda Y_{\alpha}.\ \forall Q_{\alpha\to\alpha\to o}.\ (\forall Z_{\alpha}.\ (Q\ Z\ Z)) \Rightarrow (Q\ X\ Y)$), and infinitely many artificial modifications to all valid definitions (e.g., $\dddot{=}^{\alpha} := \lambda X_{\alpha}.\ \lambda Y_{\alpha}.\ \forall P_{\alpha\to o}.\ ((a_o \vee \neg\ a_o) \wedge P\ X) \Rightarrow ((b_o \vee \neg\ b_o) \wedge P\ Y)$). The latter definition is obviously equivalent to Leibniz definition as it just adds some tautologies to the embedded formulas.
- The artificially modified definitions demonstrate, that it is generally not decidable whether a formula is a valid definition of equality (as the set of tautologies is not decidable). Hence, it is not decidable whether an input problem to one of our proof procedures contains a valid definition of equality, and we cannot simply replace all valid definitions embedded in a problem formulation by primitive equations as one might wish to.

If we are interested in Henkin completeness, we therefore have to ensure that the paramodulation and RUE-resolution approaches presented in the next sections can handle all forms of defined equality (like the underlying calculus $\mathcal{ER}$) and can additionally handle primitive equality.[8]

5 $\mathcal{EP}$: Extensional HO Paramodulation

In this section we adapt the well known FO paramodulation approach [RW69] to our HO setting and examine Henkin completeness. A straightforward adaptation of the traditional FO paramodulation rule is given by rule *Para* in Fig. 3. Analogous to the $\mathcal{ER}$ rules *Res* and *Fac*, (pre-)unification is delayed by encoding the respective unification problem (its solvability justifies the rewriting step) as

[5] This is illustrated by the examples that employ defined equality in [BK98a] and the examples that employ primitive equality in Sect. 7.

[6] For this statement we assume Henkin or standard semantics as underlying semantical notion. In weaker semantics things get even more complicated as, e.g., Leibniz equality does not necessary denote the intended equality relation. For a detailed discussion see [Ben99,BK97].

[7] As presented in Andrews textbook [And86], p. 155.

[8] The author admits, that in practice one is mainly interested in finding proofs rather than in the theoretical notion of Henkin completeness. Anyhow, our motivation in this paper is to clarify the theoretical properties of our approaches.

$$\frac{[\mathbf{A}[\mathbf{T}_\gamma]]^\alpha \vee C \quad [\mathbf{L}_\gamma = \mathbf{R}_\gamma]^T \vee D}{[\mathbf{A}[\mathbf{R}]]^\alpha \vee C \vee D \vee [\mathbf{T} = \mathbf{L}]^F}\ Para \qquad \frac{[\mathbf{A}]^\alpha \vee C \quad [\mathbf{L}_\gamma = \mathbf{R}_\gamma]^T \vee D}{[P_{\gamma\to o}\ \mathbf{R}]^\alpha \vee C \vee D \vee [\mathbf{A} = P\ \mathbf{L}]^F}\ Para'$$

We implicitly assume the symmetric application of $[\mathbf{L}_\gamma = \mathbf{R}_\gamma]^T$.
$\mathbf{T}$ (in *Para*) does not contain free variables which are bound outside of $\mathbf{T}$.

Fig. 3. Adapted Paramodulation Rule and a HO specific reformulation

a unification constraint. Rule *Para'* is an elegant HO specific reformulation[9] of paramodulation that has a very simple motivation: It describes the resolution step with the clause $[P\ \mathbf{L}]^F \vee [P\ \mathbf{R}]^T \vee D$, i.e., the clause obtained when replacing the primitive equation $[\mathbf{L} = \mathbf{R}]^T$ by its Leibniz definition. Note that the paramodulant of *Para'* encodes all possible single rewrite steps, all simultaneous rewrite-steps with rule *Para*, and in some sense even the left premise clause itself. This is nicely illustrated by the following example: $\mathcal{C}_1 : [p\ (f\ (f\ a))]^T$ and $\mathcal{C}_2 : [f = h]^T$, where $p_{\iota\to o}, f_{\iota\to\iota}, h_{\iota\to\iota}$ are constants. Applying rule *Para'* to $\mathcal{C}_1$ and $\mathcal{C}_2$ from left to right leads to $\mathcal{C}_3 : [P_{(\iota\to\iota)\to\iota}\ h]^T \vee [p\ (f\ (f\ a)) = P_{(\iota\to\iota)\to\iota}\ f]^F$. Eager unification computes the following four solutions for P, which can be backpropagated to literal $[P\ h]^T$ with rule *Subst*:

$[\lambda Z_{\iota\to\iota}.\ p\ (f\ (f\ a))/P]$ the pure imitation solution encodes $\mathcal{C}_1$ itself.
$[\lambda Z_{\iota\to\iota}.\ p\ (Z\ (f\ a))/P]$ encodes the rewriting of the first f ($[p\ (h\ (f\ a))]^T$).
$[\lambda Z_{\iota\to\iota}.\ p\ (f\ (Z\ a))/P]$ encodes the rewriting of the second f ($[p\ (f\ (h\ a))]^T$).
$[\lambda Z_{\iota\to\iota}.\ p\ (Z\ (Z\ a))/P]$ encodes the simult. rewr. of both f ($[p\ (h\ (h\ a))]^T$).

Rule *Para'* introduces flexible literal heads into the search space such that rule *Prim* becomes applicable. Thus, a probably suitable heuristics in practice is to avoid all primitive substitution steps on flexible heads generated by rule *Para'*.

Note that reflexivity resolution[10] and paramodulation into unification constraints[11] are derivable in our approach and can thus be avoided.

[9] This rule was first suggested by Michael Kohlhase.

[10] In FO a reflexivity resolution rule is needed to refute negative equation literals $[\mathbf{T}_1 = \mathbf{T}_2]^F$ if $\mathbf{T}_1$ and $\mathbf{T}_2$ are unifiable. As such literals are automatically treated as unification constraints reflexivity resolution is not needed in our approach.

[11] Let $\mathcal{C}_1 : C \vee [\mathbf{A}[\mathbf{T}] = \mathbf{B}]^F$ and $\mathcal{C}_2 : [\mathbf{L} = \mathbf{R}]^T \vee D$. The rewriting step $Para(\mathcal{C}_1, \mathcal{C}_2) : \mathcal{C}_3 : C \vee D \vee [\mathbf{A}[\mathbf{R}] = \mathbf{B}]^F \vee [\mathbf{L} = \mathbf{T}]^F$ can be replaced by derivation $Leib(\mathcal{C}_1) : \mathcal{C}_4 : [p\ \mathbf{A}[\mathbf{T}]]^T \vee C,\ \mathcal{C}_5 : [p\ \mathbf{B}]^F \vee C;\ Para(\mathcal{C}_4, \mathcal{C}_2) : \mathcal{C}_6 : [p\ \mathbf{A}[\mathbf{R}]]^T \vee C \vee D \vee [\mathbf{L} = \mathbf{T}]^F;\ Res(\mathcal{C}_6, \mathcal{C}_5), Fac, Triv : \mathcal{C}_7 : C \vee D \vee [p\ \mathbf{A}[\mathbf{R}] = p\ \mathbf{B}]^F \vee [\mathbf{L} = \mathbf{T}]^F;\ Dec(\mathcal{C}_7) : \mathcal{C}_3$. Notational remark: $Res(\mathcal{C}_6, \mathcal{C}_5), Fac, Triv$ describes the application of rule *Res* to $\mathcal{C}_6$ and $\mathcal{C}_5$, followed by applications of *Fac* and *Triv* to the subsequent results.

In the following discussion we will use the traditional paramodulation rule *Para* only.[12] As *Para'* is obviously more general than *Para* we obtain analogous completeness results if we employ *Para'* instead.

Definition 3 (Simple HO Paramodulation). $\mathcal{EP}_{naive} := \mathcal{ER} \cup \{Para\}$ *extends the extensional HO resolution approach by rule* Para. *Primitive equations in input problems are no longer expanded by Leibniz definition.* Para *operates on proper clause only and omits paramodulation into unification constraints.*

Whereas soundness of rule *Para* can be shown analogously to the FO case, it turns out that our simple HO paramodulation approach is incomplete:

Theorem 4 (Incompleteness). *Calculus $\mathcal{EP}_{naive}$ is Henkin incomplete.*

Proof. Consider the following counterexamples: $\mathbf{E}_1^{Para}$: $\neg\exists X_o. (X = \neg X)$, i.e., the negation operator is fix-point free, which is obviously the case in Henkin semantics. Negation and clause normalisation leads to clause $\mathcal{C}_1 : [a = \neg a]^T$, where a_o is a new Skolem constant. The only rule that is applicable is self-paramodulation at positions $\langle 1\rangle$, $\langle 2\rangle$, and $\langle\rangle$, leading to the following clauses (including the symmetric rewrite steps):

$$
\begin{array}{lll}
Para(\mathcal{C}_1,\mathcal{C}_1)\ at\ \langle 1\rangle : \mathcal{C}_2 : & [a = \neg a]^T \vee [\neg a = a]^F, & \mathcal{C}_3 : [\neg a = \neg a]^T \vee [a = a]^F \\
Para(\mathcal{C}_1,\mathcal{C}_1)\ at\ \langle 2\rangle : \mathcal{C}_4 : & [a = \neg a]^T \vee [a = \neg a]^F, & \mathcal{C}_5 : [a = a]^T \vee [\neg a = \neg a]^F \\
Para(\mathcal{C}_1,\mathcal{C}_1)\ at\ \langle\rangle : \mathcal{C}_6 : & [a]^T \vee [\neg a = (a = \neg a)]^F, & \mathcal{C}_7 : [a]^F \vee [a = (a = \neg a)]^F
\end{array}
$$

A case distinction on the possible denotations $\{\top, \bot\}$ for a shows that all clauses are tautologous, such that no refutation is possible in $\mathcal{EP}_{naive}$. Additional examples are discussed in [Ben99], e.g. $\mathbf{E}_2^{Para}$: $[G\ X = p]^T$, which stems from a simple version of cantor's theorem $\neg\exists G_{\iota\to\iota\to o}. \forall P_{\iota\to o}. \exists X_\iota. G\ X = P$, or example $\mathbf{E}_3^{Para}$: $[M = \lambda X_o. \bot]^T$, which stems from $\exists M_{o\to o}. M \neq \emptyset$.

The problem is that in HO logic even single positive equation literals can be contradictory. And the incompleteness is caused as the extensionality principles are now also needed to refute such positive equation literals.[13] Hence, we add the positive counterparts *Func'* and *Equiv'* (cf. Fig. 4) to the already present negative extensionality rules *Func* and *Equiv*. The completeness proof and the examples show that a positive counterpart for rule *Leib* can be avoided.

Definition 4 (Extensional HO Paramodulation). *Analogously to the extensional HO resolution case we define the calculi $\mathcal{EP} := \mathcal{ER} \cup \{Para, Equiv', Func'\}$, $\mathcal{EP}_f := \mathcal{EP} \cup \{FlexFlex\}$, and $\mathcal{EP}_{fc} := (\mathcal{EP}_f \backslash \{Cnf\}) \cup \mathcal{CNF}$.*

Theorem 5 (Soundness). *The calculi $\mathcal{EP}$, $\mathcal{EP}_f$, and $\mathcal{EP}_{fc}$ are H-sound.*

[12] It has been pointed out by a unknown referee of this paper that rule *Para'* already captures full functional extensionality and should therefore be preferred over *Para*. Example $\mathbf{E}_1^{func}$ discussed in Sec. 10.6 of [Ben99] illustrates that this is generally not true.

[13] In contrast to $\mathcal{EP}$, the underlying calculus $\mathcal{ER}$ does not allow positive equation literals as the equality symbol is only used to encode unification constraints. Therefore the pure extensional HO resolution approach $\mathcal{ER}$ does not require a positive extensionality treatment.

$$\frac{\mathbf{C} \vee [\mathbf{M}_o = \mathbf{N}_o]^T}{\mathbf{C} \vee [\mathbf{M}_o \Leftrightarrow \mathbf{N}_o]^T} \; Equiv' \qquad \frac{\mathbf{C} \vee [\mathbf{M}_{\gamma\to\beta} = \mathbf{N}_{\gamma\to\beta}]^T \quad X \text{ new variable}}{\mathbf{C} \vee [\mathbf{M}\ X_\gamma = \mathbf{N}\ X_\gamma]^T} \; Func'$$

Fig. 4. Positive Extensionality Rules

Proof. Soundness of rule *Para* with respect to Henkin semantics can be proven analogously to the FO case and soundness of *Equiv'* and *Func'* is obvious, as they simply apply the extensionality principles, which are valid in Henkin semantics.

Lemma 3 (Lifting of $\mathcal{EP}_{fc}$). *Let Φ be a clause set, $\mathcal{D}_1$ a clause, and σ a substitution. If $\sigma(\Phi) \vdash_{\mathcal{ER}_{fc}} \mathcal{D}_1$, then $\Phi \vdash_{\mathcal{EP}_{fc}} \mathcal{D}_2$ for a clause $\mathcal{D}_2$ generalising $\mathcal{D}_1$.*

Proof. Analogous to Lemma 1. The additional rules do not cause any problems.

The main completeness theorem 6 for $\mathcal{EP}_{fc}$ below is proven analogously to Thm. 3, i.e., we employ the model existence theorem for Henkin models with primitive equality (cf. Thm. 1). As primitive equality is involved, we additionally have to ensure the abstract consistency properties ∇_e^r and ∇_e^s (cf. Def. 1), i.e., the reflexivity and substitutivity property of primitive equality. Whereas the reflexivity property is trivially met, we employ the following admissible[14] — and moreover even weakly derivable (i.e., modulo clause normalisation and lifting) — paramodulation rule to verify the substitutivity property.

Definition 5 (Generalised Paramodulation). *The generalised paramodulation rule* GPara *is defined as follows:*

$$\frac{[\mathbf{T}[\mathbf{A}_\beta]]^\alpha \vee C \quad [\mathbf{A}_\beta = \mathbf{B}_\beta]^T}{[\mathbf{T}[\mathbf{B}]]^\alpha \vee C} \; GPara$$

This rule extends Para *as it can be applied to non-proper clauses and it restricts* Para *as it can only be applied in special clause contexts, e.g. the second clause has to be a unit clause.* GPara *is especially designed to verify the substitutivity property of primitive equality ∇_e^s in the main completeness theorem 6.*

Weak derivability (which obviously implies admissibility) of *GPara* is shown with the help of the following weakly derivable generalised resolution rules.

Definition 6 (Generalised Resolution). *The generalised resolution rules $GRes_1$, $GRes_2$, and $GRes_3$ are defined as follows (for all rules we assume $\alpha, \beta \in \{T, F\}$ with $\alpha \neq \beta$, and for $GRes_2$ we assume that $\overline{Y^n} \notin$ **free**(**A**)):*

[14] Rule r is called admissible (derivable) in R, iff adding rule r to calculus R does not increase the set of refutable formulas (iff each application of rule r can be replaced by an alternative derivation in calculus R).

$$\frac{[\mathbf{A}_{\overline{\gamma}\to o}\ \overline{\mathbf{T}_\gamma^n}]^\alpha \vee C \quad [\mathbf{A}_{\overline{\gamma}\to o}\ \overline{X_\gamma^n}]^\beta \vee D}{(C \vee D)_{[\overline{T^n}/\overline{X^n}]}}\ GRes_1 \qquad \frac{[\mathbf{A}_\gamma\ \overline{Y^n}]^\alpha \vee C \quad [X_\gamma\ \overline{\mathbf{T}^n}]^\beta \vee D}{(C \vee D)_{[\mathbf{A}/X,\overline{\mathbf{T}^n}/\overline{Y^n}]}}\ GRes_2$$

$$\frac{[\mathbf{A}_\gamma\ \overline{\mathbf{T}^n}]^\alpha \vee C \quad [X_\gamma\ \overline{Y^n}]^\beta \vee D}{(C \vee D)_{[\mathbf{A}/X,\overline{\mathbf{T}^n}/\overline{Y^n}]}}\ GRes_3$$

These rules extend Res *as they can be applied to non-proper clauses, and they restrict* Res *as they are only defined for special clause contexts. The rules are designed just strong enough to prove weak derivability of* GPara*.*

Lemma 4 (Weak Derivability of $GRes_{1,2,3}$**).** *Let $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ be clauses and $r \in \{GRes_1, GRes_2, GRes_3\}$. If $\{\mathcal{C}_1, \mathcal{C}_2\} \vdash^r \mathcal{C}_3 \vdash_{\mathcal{CNF}} \mathcal{C}_4$ for a proper clause $\mathcal{C}_4$, then $\{\mathcal{C}_1, \mathcal{C}_2\} \vdash_{\mathcal{EP}_{fc}} \mathcal{C}_5$ for a clause $\mathcal{C}_5$ which generalises $\mathcal{C}_4$.*

Proof. The proof is by induction on the number of logical connectives in the resolution literals. It employs generalised (and weakly derivable) versions of the factorisation rule *Fac* and primitive substitution rule *Prim* (see [Ben99]), which are not presented here because lack of space. $GRes_2$ and $GRes_3$ are needed to prove weak derivability for $GRes_1$. As the rules *Para, Equiv', Func'* are not employed in the proof, this lemma analogously holds for calculus $\mathcal{ER}_{fc}$.

Lemma 5 (Weak Derivability of $GPara$**).** *Let $\mathcal{C}_1 : [\mathbf{T}[\mathbf{A}]_p]^\alpha \vee D_1, \mathcal{C}_2 : [\mathbf{A} = \mathbf{B}]^T, \mathcal{C}_3 : [\mathbf{T}[\mathbf{B}]_p]^\alpha \vee D_1$ be clauses. If $\Delta : \{\mathcal{C}_1, \mathcal{C}_2\} \vdash^{GPara} \mathcal{C}_3 \vdash_{\mathcal{CNF}} \mathcal{C}_4$ for a proper clause $\mathcal{C}_4$, then $\{\mathcal{C}_1, \mathcal{C}_2\} \vdash_{\mathcal{EP}_{fc}} \mathcal{C}_5$ for a clause $\mathcal{C}_5$ generalising $\mathcal{C}_4$.*

Proof. The proof is by induction on the length of Δ and employs the (weakly derivable) generalised resolution rule $GRes_1$ and the standard paramodulation rule *Para* in the quite complicated base case.

Theorem 6 (Completeness). *Calculus $\mathcal{EP}_{fc}$ is Henkin complete.*

Proof. Let Γ_Σ be the set of closed Σ-formulas that cannot be refuted with calculus $\mathcal{EP}_{fc}$ (i.e., $\Gamma_\Sigma := \{\Phi \subseteq cwff_o | \Phi_{cl} \not\vdash_{\mathcal{EP}_{fc}} \Box\}$). We show that Γ_Σ is a saturated abstract consistency class for Henkin models with primitive equality (cf. Def. 1). This entails completeness by the model existence theorem for Henkin models with primitive equality (cf. Thm. 1).

First we have to verify that Γ_Σ validates the abstract consistency properties ∇_c, $\nabla_\neg$, ∇_β, $\nabla_\vee$, $\nabla_\wedge$, $\nabla_\forall$, $\nabla_\exists$, ∇_b, ∇_q and that Γ_Σ is saturated. In all of these cases the proofs are identical to the corresponding argumentations in Thm. 3.

Thus, all we need to ensure is the validity of the additional abstract consistency properties ∇_e^r and ∇_e^s for primitive equality:
(∇_e^r) We have that $[\mathbf{A} =^\alpha \mathbf{A}]^F \vdash^{Triv} \Box$, and thus $\neg(\mathbf{A} =^\alpha \mathbf{A})$ cannot be in Φ.
(∇_e^s) Analogously to the cases in Sec. 3 we show the contrapositive of the assertion, and thus we assume that there is derivation $\Delta : \Phi_{cl} \cup \{[\mathbf{F}[\mathbf{B}]]^T\} \vdash_{\mathcal{EP}_{fc}} \Box$. Now consider the following $\mathcal{EP}_{fc}$-derivation: $\Delta' : \Phi_{cl} \cup \{[\mathbf{F}[\mathbf{A}]]^T, [\mathbf{A} = \mathbf{B}]^T\} \vdash^{GPara} \Phi_{cl} \cup \{[\mathbf{F}[\mathbf{A}]]^T, [\mathbf{A} = \mathbf{B}]^T, [\mathbf{F}[\mathbf{B}]]^T\} \vdash_{\mathcal{EP}_{fc}} \Box$. By Lemma 5 *GPara* is weakly derivable (hence admissible) for calculus $\mathcal{EP}_{fc}$, such that there is a $\mathcal{EP}_{fc}$-derivation $\Delta'' : \Phi_{cl} \cup \{[\mathbf{F}[\mathbf{A}]]^T, [\mathbf{A} = \mathbf{B}]^T\} \vdash_{\mathcal{EP}_{fc}} \Phi_{cl} \cup \{[\mathbf{F}[\mathbf{A}]]^T, [\mathbf{A} = \mathbf{B}]^T, [\mathbf{F}[\mathbf{B}]]^T\} \vdash_{\mathcal{EP}_{fc}} \Box$ which completes the proof.

Lemma 6 (Theorem Equivalence). *$\mathcal{EP}_{fc}$ and $\mathcal{EP}_f$ are theorem equivalent.*

Proof. Analogous to Lemma 2. The additional rules do not cause any problems.

Question 2 (Theorem Equivalence). The author claims that the calculi $\mathcal{EP}$ and $\mathcal{EP}_{fc}$ (or $\mathcal{EP}_f$) are theorem equivalent. The formal proof will most likely be analogous to the one for question 1.

6 $\mathcal{ERUE}$: Extensional HO RUE-Resolution

In this section we will adapt the **R**esolution by **U**nification and **E**quality approach [Dig79] to our higher-order setting. The key idea is to allow the resolution and factorisation rules also to operate on unification constraints (which is forbidden in $\mathcal{ER}$ and $\mathcal{EP}$). This implements the main ideas of FO RUE-resolution directly in our higher-order calculus. More precisely our approach allows to compute partial E-unifiers with respect to a specified theory E by resolution on unification constraints within the calculus itself (if we assume that E is specified in form of an available set of unitary or even conditional equations in clause form). This is due to the fact that the extensional higher-order resolution approach already realises a test calculus for general higher-order E-pre-unification (or higher-order E-unification in case we also add the rule *FlexFlex*). Furthermore, each partial E-(pre-)unifier can be applied to a clause with rule *Subst*, and, like in the traditional FO RUE-resolution approach, the non-solved unification constraints are encoded as (still open) unification constraints, i.e., negative equations, within the particular clauses.

Definition 7 (Extensional HO RUE-Resolution). *We now allow the factorisation rule* Fac *and resolution rule* Res *to operate also on unification constraints and define the calculi $\mathcal{ERUE} := \mathcal{ER} \cup \{Equiv', Func'\}$, $\mathcal{ERUE}_f := \mathcal{ERUE} \cup \{FlexFlex\}$, and $\mathcal{ERUE}_{fc} := (\mathcal{ERUE}_f \backslash \{Cnf\}) \cup \mathcal{CNF}$.*

Theorem 7 (Soundness). *The calculi $\mathcal{ERUE}_{fc}$, $\mathcal{ERUE}_f$, and $\mathcal{ERUE}$ are H-sound.*

Proof. Unification constraints are encoded as negative literals, such that soundness of the extended resolution and factorisation rules with respect to Henkin semantics is obvious.

Lemma 7 (Lifting of $\mathcal{ERUE}_{fc}$). *Let Φ be a clause set, $\mathcal{D}_1$ a clause, and σ a substitution. If $\sigma(\Phi) \vdash_{\mathcal{ER}_{fc}} \mathcal{D}_1$, then $\Phi \vdash_{\mathcal{ERUE}_{fc}} \mathcal{D}_2$ for a clause $\mathcal{D}_2$ generalising $\mathcal{D}_1$.*

Proof. Analogous to Lemmata 1 and 3.

Within the main completeness proof we proceed analogously to previous section and employ the generalised paramodulation rule *GPara* to verify the crucial substitutivity property ∇_e^s. Thus, we need to show that *GPara* is admissible in calculus $\mathcal{ERUE}_{fc}$. Note that in Lemma 5 we were even able to show a weak derivability property of rule *GPara* for calculus $\mathcal{EP}_{fc}$. Whereas *GPara* is not weakly derivability for calculus $\mathcal{ERUE}_{fc}$, we can still prove admissibility of this rule here. As in Lemma 5, we employ the generalised resolution rules which are weakly derivable in $\mathcal{ERUE}_{fc}$ as well.

Lemma 8 (Weak Derivability of $GRes_{1,2,3}$). *Let $\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3$ be clauses and $r \in \{GRes_1, GRes_2, GRes_3\}$. If $\{\mathcal{C}_1, \mathcal{C}_2\} \vdash^r \mathcal{C}_3 \vdash_{\mathcal{CNF}} \mathcal{C}_4$ for a proper clause $\mathcal{C}_4$, then $\{\mathcal{C}_1, \mathcal{C}_2\} \vdash_{\mathcal{ERUE}_{fc}} \mathcal{C}_5$ for a clause $\mathcal{C}_5$ which generalises $\mathcal{C}_4$.*

Proof. Analogous to Lemma 4.

Lemma 9 (Admissibility of $GPara$). *Let Φ be a clause set, such that $\Delta : \Phi \vdash^{GPara} \Phi' \vdash_{\mathcal{ERUE}_{fc}} \Box$, then there exists a refutation $\Phi \vdash_{\mathcal{ERUE}_{fc}} \Box$.*

Proof. The proof is (analogous to Lemma 5) by induction on the length of Δ and employs the weakly derivable generalised resolution rule $GRes_1$. The applications of rule *Para* in the proof of Lemma 5 are replaced by corresponding derivations employing resolution and factorisation on unification constraints. The latter causes the loss of the weak derivability property.

Theorem 8 (Completeness). *Calculus $\mathcal{ERUE}_{fc}$ is Henkin complete.*

Proof. Analogously to Lemma 6 we show that the set of closed Σ-formulas which cannot be refuted by the calculus $\mathcal{ERUE}_{fc}$ (i.e., $\Gamma_\Sigma := \{\Phi \subseteq cwff_o | \Phi_{cl} \not\vdash_{\mathcal{ERUE}_{fc}} \Box\}$) is a saturated abstract consistency class for Henkin models with primitive equality (cf. Def. 1). This entails the assertion by Thm. 1.

The proof is analogous to Lemma 6. Even the abstract consistency properties ∇_e^r and ∇_e^s are proven analogously by employing the generalised paramodulation rule *GPara*, which is by Lemma 9 admissible in $\mathcal{ERUE}_{fc}$.

Lemma 10 (Theorem Equiv.). *$\mathcal{ERUE}_{fc}$ and $\mathcal{ERUE}_f$ are theorem equivalent.*

Proof. Analogous to Lemma 2. The additional or modified rules do not cause any problems.

Question 3 (Theorem Equivalence). The author claims that the calculi $\mathcal{ERUE}$ and $\mathcal{ERUE}_{fc}$ (or $\mathcal{ERUE}_f$) are theorem equivalent. A formal proof will most likely be analogous to questions 1 and 2.

7 Examples

The first (trivial FO) example illustrates the main ideas of $\mathcal{EP}$ and $\mathcal{ERUE}$: $a_\iota \in m_{\iota\to o} \wedge a = b \Rightarrow b \in m$. Sets are encoded as characteristic functions and $\in := \lambda X_\alpha, M_{\alpha\to o}. M\ X$, such that the negated problem normalises to: $\mathcal{C}_1 : [m\ a]^T$, $\mathcal{C}_2 : [a = b]^T$, $\mathcal{C}_3 : [m\ b]^F$. An obvious term-rewriting refutation in $\mathcal{EP}$: $Para(\mathcal{C}_1, \mathcal{C}_2)$, $Triv : \mathcal{C}_4 : [m\ b]^T$; $Res(\mathcal{C}_3, \mathcal{C}_4), Triv : \Box$.[15] A difference-reducing refutation in $\mathcal{ERUE}$: $Res(\mathcal{C}_1, \mathcal{C}_3) : \mathcal{C}_4 : [m\ a = m\ b]^F$; $Dec(\mathcal{C}_4), Triv : \mathcal{C}_5 : [a = b]^F$; $Res(\mathcal{C}_2, \mathcal{C}_5), Triv : \Box$.

We now examine the examples mentioned in Thm. 4 in calculus $\mathcal{EP}$: $\mathbf{E}_2^{Para}$: $[(G\ X_\iota) = p_{\iota\to o}]^T$ (Cantor's theorem) $Func'(\mathbf{E}_2^{Para}), Equiv' : \mathcal{C}_1 : [G\ X\ Y_\iota]^F \vee$

[15] Notation (as already used before): $Res(\mathcal{C}_6, \mathcal{C}_5), Fac$ describes a paramodulation step between $\mathcal{C}_6$ and $\mathcal{C}_5$ followed by factorisation of the resulting clause. $Prim(\mathcal{C}_1|\mathcal{C}_2)$ denotes the parallel application of rule $Prim$ to $\mathcal{C}_j$ and $\mathcal{C}_k$.

$[p\ Y_\iota]^T$, $\mathcal{C}_2 : [G\ X\ Y_\iota]^T \vee [p\ Y_\iota]^F$; $Prim(\mathcal{C}_1|\mathcal{C}_2), Subst : \mathcal{C}_3 : [G'\ X\ Y]^T \vee [p\ Y]^T$, $\mathcal{C}_4 : [G''\ X\ Y]^F \vee [p\ Y]^F$; $Fac(\mathcal{C}_3|\mathcal{C}_4), \mathcal{UNI} : \mathcal{C}_5 : [p\ Y]^T$, $\mathcal{C}_6 : [p\ Y]^F$; $Res(\mathcal{C}_5,\mathcal{C}_6), \mathcal{UNI} : \mathcal{C}_7 : \Box$. $\mathbf{E}_1^{Para}$ and $\mathbf{E}_3^{Para}$ can be proven analogously. The key idea is to employ the positive extensionality rules first. As paramodulation rule is not employed, these proofs are obviously also possible in $\mathcal{ERUE}$.

Example $\mathbf{E}_2^{set}$ focuses on reasoning about sets: $(\{X|\ odd\ X \wedge num\ X\} = \{X|\ \neg\ ev\ X \wedge num\ X\}) \Rightarrow (2^{\{X|\ odd\ X \wedge X>100 \wedge num\ X\}} = 2^{\{X|\ \neg\ ev\ X \wedge X>100 \wedge num\ X\}})$, where the powerset-operator is defined by $\lambda N_{\alpha\to o}.\ \lambda M_{\alpha\to o}.\ \forall X_\alpha.\ \mathbf{M}\ X \Rightarrow \mathbf{N}\ X$. $\mathcal{CNF}(\mathbf{E}_2^{set}), Func, Func' : \mathcal{C}_1 : [(odd\ X \wedge num\ X) = (\neg\ ev\ X \wedge num\ X)]^T$ and $\mathcal{C}_2 : [(\forall X.\ n\ X \Rightarrow ((odd\ X \wedge X > 100) \wedge num\ X)) = (\forall X.\ n\ X \Rightarrow ((\neg\ ev\ X \wedge X > 100) \wedge num\ X))]^F$ where n is a Skolem constant. The reader may check that an application of rule *Para* does not lead to a successful refutation here as the terms in the powerset description do unfortunately not have the *right structure*. Instead of following the term-rewriting idea we have to proceed with difference-reduction and a recursive call to the overall refutation search from within the unification process: $Dec(\mathcal{C}_2), Triv, Func, Dec, Triv : \mathcal{C}_3 : [((odd\ s \wedge s > 100) \wedge num\ s) = ((\neg\ ev\ s \wedge s > 100) \wedge num\ s)]^F$; $Equiv(\mathcal{C}_3), \mathcal{CNF}, Fac, \mathcal{UNI} : \mathcal{C}_4 : [odd\ s]^T \vee [ev\ s]^F$, $\mathcal{C}_5 : [s > 100]^T$, $\mathcal{C}_6 : [num\ s]^T$, $\mathcal{C}_7 : [odd\ s]^F \vee [s > 100]^F \vee [num\ s]^F \vee [ev\ s]^T$; $Equiv'(\mathcal{C}_1), \mathcal{CNF}, Fac, \mathcal{UNI} : \mathcal{C}_8 : [odd\ X]^F \vee [num\ X]^F \vee [ev\ X]^F$, $\mathcal{C}_9 : [odd\ X]^T \vee [num\ X]^F \vee [ev\ X]^T$. The rest of the refutation is a straightforward resolution proof on $\mathcal{C}_4 - \mathcal{C}_9$. It is easy to check that an elegant term-rewriting proof is only possible if we put the succedent of $\mathbf{E}_2^{set}$ in the *right order*: $2^{\{X|\ (odd\ X \wedge num\ X) \wedge X>100\}} = 2^{\{X|\ (\neg\ ev\ X \wedge num\ X) \wedge X>100\}}$. Thus this example nicely illustrates the unavoidable mixed term-reducing and difference-reducing character of extensional higher-order paramodulation.

On the other hand a very interesting goal directed proof is possible within the RUE-resolution approach $\mathcal{ERUE}$ by immediately resolving between $\mathcal{C}_1$ and the unification constraint $\mathcal{C}_2$ and subsequently employing syntactical unification in connection with recursive calls to the overall refutation process (with the extensionality rules) when syntactical unification is blocked.

[Ben99] provides a more detailed discussion of these and additional examples.

8 Conclusion

We presented the two approaches $\mathcal{EP}$ and $\mathcal{ERUE}$ for extensional higher-order paramodulation and RUE-resolution which extend the extensional higher-order resolution approach $\mathcal{ER}$ [BK98a] by a primitive equality treatment. All three approaches avoid the extensionality axioms and employ more goal directed extensionality rules instead. An interesting difference to Huet's original constraint resolution approach [Hue72] is that eager (pre-)unification becomes essential and cannot be generally delayed if an extensionality treatment is required.

Henkin completeness has been proven for the slightly extended (by the additional rule *FlexFlex*) approaches $\mathcal{ER}_f$, $\mathcal{EP}_f$ and $\mathcal{ERUE}_f$. The claim that rule *FlexFlex* is admissible in them has not been proven yet. All three approaches can be implemented in a higher-order set of support approach as presented in [Ben99]. [Ben99] also presents some first ideas how the enormous search space of the introduced approaches can be further restricted in practice, e.g. by introducing redundancy methods.

It has been motivated that some problems cannot be solved in the paramodulation approach $\mathcal{EP}$ by following the term-rewriting idea only, as they unavoidably require the application of the difference-reducing extensionality rules. In contrast to $\mathcal{EP}$ the difference-reducing calculus $\mathcal{ERUE}$ seems to harmonise quite well with the difference-reducing extensionality rules (or axioms), and thus this paper concludes with the question: Can HO adaptations of term-rewriting approaches be as successful as in FO, if one is interested in Henkin completeness and extensionality, e.g., when reasoning about sets, where sets are encoded as characteristic functions? Further work will be to examine this aspect with the help of the Leo-system [BK98b] and to investigate the open questions of this paper.

Acknowledgements. The work reported here was funded by the Deutsche Forschungsgemeinschaft under grant HOTEL. The author is grateful to M. Kohlhase, F. Pfenning, P. Andrews, V. Sorge and S. Autexier for stimulating discussions and support.

References

And71. P. B. Andrews. Resolution in type theory. *JSL*, 36(3):414–432, 1971.
And86. P. B. Andrews. *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof*. Academic Press, 1986.
Bar84. H. P. Barendregt. *The Lambda-Calculus: Its Syntax and Semantics*. North-Holland, 2nd edition 1984.
Ben99. C. Benzmüller. Equality and Extensionality in Automated Higher-Order Theorem Proving. PhD thesis, FB 14, Universität des Saarlandes, 1999.
BK97. C. Benzmüller and M. Kohlhase. Model existence for higher-order logic. Seki-Report SR-97-09, FB 14, Universität des Saarlandes, 1997, submitted to JSL.
BK98a. C. Benzmüller and M. Kohlhase. Extensional higher-order resolution. In Kirchner and Kirchner [KK98], pages 56–72.
BK98b. C. Benzmüller and M. Kohlhase. LEO — a higher-order theorem prover. In Kirchner and Kirchner [KK98], pages 139–144.
Chu40. A. Church. A formulation of the simple theory of types. *JSL*, 5:56–68, 1940.
Dar68. J. L. Darlington. Automatic theorem proving with equality substitutions and mathematical induction. *Machine Intelligence*, 3:113–130, 1968.
Dig79. V. J. Digricoli. Resolution by unification and equality. In W. H. Joyner, editor, *Proc. of the 4th Workshop on Automated Deduction*, Austin, 1979.
Hen50. L. Henkin. Completeness in the theory of types. *JSL*, 15(2):81–91, 1950.
Hue72. G. P. Huet. *Constrained Resolution: A Complete Method for Higher Order Logic*. PhD thesis, Case Western Reserve University, 1972.
KK98. C. Kirchner and H. Kirchner, editors. *Proc. of the 15th Conference on Automated Deduction*, number 1421 in LNAI, Springer, 1998.
Koh94. M. Kohlhase. *A Mechanization of Sorted Higher-Order Logic Based on the Resolution Principle*. PhD thesis, Universität des Saarlandes, 1994.
Mil83. D. Miller. *Proofs in Higher-Order Logic*. PhD thesis, Carnegie-Mellon University, 1983.
RW69. G. A. Robinson and L. Wos. Paramodulation and TP in first order theories with equality. *Machine Intelligence*, 4:135–150, 1969.
SG89. W. Snyder and J. H. Gallier. Higher-Order Unification Revisited: Complete Sets of Transformations. *Journal of Symbolic Computation*, 8:101–140, 1989.
Smu63. Raymond M. Smullyan. A unifying principle for quantification theory. Proc. Nat. Acad Sciences, 49:828–832, 1963.
Wol93. D. Wolfram. *The Clausal Theory of Types*. Cambridge University Press, 1993.