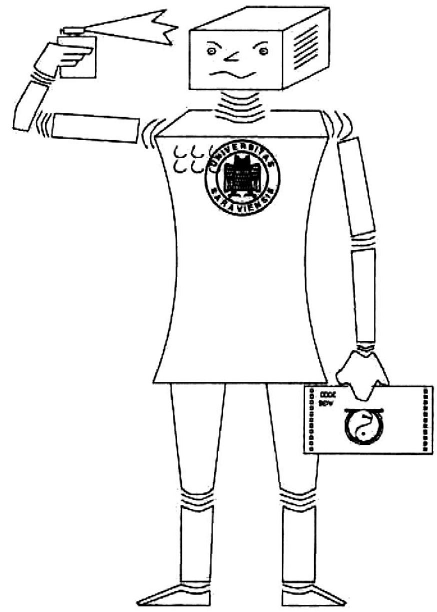
UNIVERSITÄT DES SAARLANDES
FACHBEREICH INFORMATIK
D-66041 SAARBRÜCKEN
GERMANY
WWW: http://www.ags.uni-sb.de/

SEKI Report
ISSN 1437-4447



A remark on higher order RUE-resolution with ERUE

Christoph Benzmüller

chris@ags.uni-sb.de

FR 6.2 Informatik, Universität des Saarlandes

Saarbrücken, Germany

SEKI Report SR-02-05

A remark on higher order RUE-resolution with ERUE

Christoph Benzmüller

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

Abstract

We show that a prominent counterexample for the completeness of first order RUE-resolution does not apply to the higher order RUE-resolution approach ERUE.

Bonacina shows in [BH92] that the first order RUE-NRF resolution approach as introduced in [Dig79, Dig81, DH86] is not complete. The counterexample consists in the following set of first order clauses:

$$\{g(f(a)) = a, f(g(X)) \neq X\}$$

Here X is a variable and f, g are unary function symbols. It is illustrated in [BH92] that this obviously inconsistent clause set cannot be refuted in the first order RUE-resolution approach of Digricoli.

The extensional higher order RUE-resolution variant ERUE has been proposed in [Ben99b, Ben99a] and completeness is analyzed in [Ben99a]. An interesting question is whether the above example is also a counterexample to the completeness of ERUE. The two ERUE refutations presented below illustrate that this is not the case.

We do not present the ERUE calculus here and instead refer to [Ben99b, Ben99a]. In the following we consider $(A \Leftrightarrow B)$ as shorthand for $(A \wedge B) \vee (\neg A \wedge \neg B)$. We furthermore use the $[\ldots]^T$ and $[\ldots]^F$ -notation of [Ben99a] to denote positive and negative literals. Terms are presented in the usual first order style notation, i.e. we write $g(f(a))$ instead of $(g\ (f\ a))$ as done in [Ben99b, Ben99a]. The decomposition rule employed in the refutations below is

$$\frac{\mathcal{C} \vee [h\overline{A^n} = h\overline{V^n}]^F}{\mathcal{C} \vee [A^1 = V^1]^F \vee \ldots \vee [A^n = V^n]^F}\ Dec$$

The reader might be more used to this form of decomposition than to the one employed in [Ben99b, Ben99a]. Compared to the latter the above rule *Dec* also shortens the presentation. The decomposition rule employed in [Ben99b, Ben99a] is more general, i.e. rule *Dec* above is derivable in calculus ERUE.

The first refutation in ERUE presented below (which has been suggested by Chad Brown) employs a flex-rigid unification step (*FlexRig*) in the very beginning. In this key step variable X is bound to an imitation binding that introduces f at head position. The rest of the refutation is then straight forward.

The second refutation shows that there are alternatives to the flex-rigid unification step for variable X at the beginning. The key idea now is to derive the positive reflexivity literal $[f(a) = f(a)]^F$ in clause C_{18}. While positive reflexivity literals cannot be derived in first order RUE-resolution, our example shows that this is (theoretically) possible in ERUE for some symbols and terms occuring in the given clause context, like $f(a)$ in our case.

We now present both ERUE-refutations in detail. f and g are still unary function symbols, while X is a variable. H and Y are freshly introduced variables.

Refutation I

$$
\begin{array}{ll}
 & \mathcal{C}_1 : [g(f(a)) = a]^T \\
 & \mathcal{C}_2 : [f(g(X)) = X]^F \\
FlexRig(\mathcal{C}_2) : & \mathcal{C}_3 : [f(g(X)) = X]^F \vee [X = f(H(X))]^F \\
Solve(\mathcal{C}_3) : & \mathcal{C}_4 : [f(g(X)) = f(H(X))]^F \\
Dec(\mathcal{C}_4) : & \mathcal{C}_5 : [g(X) = H(X)]^F \\
Res(\mathcal{C}_1, \mathcal{C}_5) : & \mathcal{C}_6 : [(g(f(a)) = a) = (g(X) = H(X))]^F \\
Dec(\mathcal{C}_6) : & \mathcal{C}_7 : [g(f(a)) = g(X)]^F \vee [a = H(X)]^F \\
Dec(\mathcal{C}_7) : & \mathcal{C}_8 : [f(a) = X]^F \vee [a = H(X)]^F \\
Solve(\mathcal{C}_8) : & \mathcal{C}_9 : [f(a) = f(a)]^F \vee [a = H(f(a))]^F \\
Triv(\mathcal{C}_9) : & \mathcal{C}_{10} : [a = H(f(a))]^F \\
FlexRig(\mathcal{C}_{10}) : & \mathcal{C}_{11} : [a = H(f(a))]^F \vee [h = \lambda Y \text{.} a]^F \\
Solve(\mathcal{C}_{11}) : & \mathcal{C}_{12} : [a = a]^F \\
Triv(\mathcal{C}_{12}) : & \Box
\end{array}
$$

Refutation II

$$
\begin{array}{ll}
 & \mathcal{C}_1 : [g(f(a)) = a]^T \\
 & \mathcal{C}_2 : [f(g(X)) = X]^F \\
Res(\mathcal{C}_1, \mathcal{C}_2) : & \mathcal{C}_3 : [(g(f(a)) = a) = (f(g(X)) = X)]^F \\
Equiv(\mathcal{C}_3) : & \mathcal{C}_4 : [(g(f(a)) = a) \Leftrightarrow (f(g(X)) = X)]^F \\
n \times Cnf(\mathcal{C}_4) : & \mathcal{C}_5 : [g(f(a)) = a]^T \vee [f(g(X)) = X]^T \\
 & \mathcal{C}_6 : [g(f(a)) = a]^F \vee [f(g(X)) = X]^F \\
Res(\mathcal{C}_6, \mathcal{C}_1) : & \mathcal{C}_7 : [(g(f(a)) = a) = (g(f(a)) = a)]^F \vee [f(g(X)) = X]^F \\
Dec(\mathcal{C}_7) : & \mathcal{C}_8 : [f(a) = f(a)]^F \vee [a = a]^F \vee [f(g(X)) = X]^F \\
Triv(\mathcal{C}_8) : & \mathcal{C}_9 : [f(a) = f(a)]^F \vee [f(g(X)) = X]^F \\
Fac(\mathcal{C}_9) : & \mathcal{C}_{10} : [f(a) = f(a)]^F \vee [(f(a) = f(a)) = (f(g(X)) = X)]^F \\
Triv(\mathcal{C}_{10}) : & \mathcal{C}_{11} : [(f(a) = f(a)) = (f(g(X)) = X)]^F \\
Equiv(\mathcal{C}_{11}) & \mathcal{C}_{12} : [(f(a) = f(a)) \Leftrightarrow (f(g(X)) = X)]^F \\
n \times Cnf(\mathcal{C}_{12}) : & \mathcal{C}_{13} : [f(a) = f(a)]^T \vee [f(g(X)) = X]^T \\
 & \mathcal{C}_{14} : [f(a) = f(a)]^F \vee [f(g(X)) = X]^F \\
Res(\mathcal{C}_{13}, \mathcal{C}_2) : & \mathcal{C}_{15} : [f(a) = f(a)]^T \vee [(f(g(X)) = X) = (f(g(X')) = X')]^F \\
Dec(\mathcal{C}_{15}) : & \mathcal{C}_{16} : [f(a) = f(a)]^T \vee [f(g(X)) = f(g(X'))]^F \vee [X = X']^F \\
Solve(\mathcal{C}_{16}) : & \mathcal{C}_{17} : [f(a) = f(a)]^T \vee [f(g(X')) = f(g(X'))]^F \\
Triv(\mathcal{C}_{17}) : & \mathcal{C}_{18} : [f(a) = f(a)]^T \\
Res(\mathcal{C}_2, \mathcal{C}_{18}) : & \mathcal{C}_{19} : [(f(g(X)) = X) = (f(a) = f(a))]^F \\
Dec(\mathcal{C}_{19}) : & \mathcal{C}_{20} : [f(g(X)) = f(a)]^F \vee [X = f(a)]^F \\
Solve(\mathcal{C}_{20}) : & \mathcal{C}_{21} : [f(g(f(a))) = f(a)]^F \\
Dec(\mathcal{C}_{21}) : & \mathcal{C}_{22} : [g(f(a)) = a]^F \\
Res(\mathcal{C}_{22}, \mathcal{C}_1) : & \mathcal{C}_{23} : [(g(f(a)) = a) = (g(f(a)) = a)]^F \\
Triv(\mathcal{C}_{23}) : & \mathcal{C}_{24} : \Box
\end{array}
$$

The above refutations are admittedly non-trivial. For this particular kind of problems paramodulation therefore seems to be a more appropriate approach. However, we suggest a more thorough analysis to sufficiently clarify this question for the higher order case.

Acknowledgment: I thank Chad Brown, CMU, Pittsburgh, USA, for his comments and contribution.

References

[Ben99a] C. Benzmüller. *Equality and Extensionality in Higher-Order Theorem Proving.* PhD thesis, Department of Computer Science, Saarland University, 1999.

[Ben99b] C. Benzmüller. Extensional higher-order paramodulation and RUE-resolution. In H. Ganzinger, editor, *Proceedings of the 16th International Conference on Automated Deduction (CADE-16)*, number 1632 in LNCS, pages 399–413, Trento, Italy, 1999. Springer.

[BH92] M. P. Bonacina and Jieh Hsiang. Incompleteness of the RUE/NRF inference systems. Newsletter of the Association for Automated Reasoning, No. 20, pages 9–12, May 1992.

[DH86] V. J. Digricoli and M.C. Harrison. Equality-based binary resolution. *Journal of the ACM*, 33(2):254–289, 1986.

[Dig79] V. J. Digricoli. Resolution by unification and equality. In W. H. Joyner, editor, *Proc. of the 4th Workshop on Automated Deduction*, Austin, Texas, USA, 1979.

[Dig81] V. J. Digricoli. The efficacy of rue resolution, experimental results and heuristic theory. In A. Drinan, editor, *Proc. of the 7th International Joint Conference on Artificial Intelligence (IJCAI81)*, pages 539–547, Vancouver, Canada, 1981. Morgan Kaufmann, San Mateo, California, USA.