Incompleteness of the RUE/NRF inference systems

Maria Paola Bonacina *and* **Jieh Hsiang**
({bonacina,hsiang}@sbcs.sunysb.edu)

Introduction

The *Resolution with Unification and Equality (RUE)* and the *Negative Reflective Function (NRF)* inference rules were proposed in [2] as a generalization of *E-resolution* [3, 1]. These two rules are defined first in a very general form, called the *RUE-NRF inference system in open form*, which is proved to be refutationally complete for first order logic with equality [2]. The RUE-NRF inference system in open form does not represent a mechanical inference system, because it does not specify completely how to compute the inferences. Thus, several refinements intended to turn the RUE-NRF schemes into concrete inference systems are given in [2]. In this note we show that these refinements are not complete.

The RUE-NRF inference rules

Given two clauses $\neg P(t_1 \ldots t_n) \vee A$ and $P(s_1 \ldots s_n) \vee B$, an RUE step consists in generating a resolvent $(A \vee B)\sigma \vee D$:

$$RUE\ \frac{\neg P(t_1 \ldots t_n) \vee A,\ P(s_1 \ldots s_n) \vee B}{(A \vee B)\sigma \vee D}$$

where σ is any substitution and D is a disjunction of inequalities, obtained by applying any number of *Decomposition* steps

$$Decompose\ \frac{f(l_1 \ldots l_n) \neq f(r_1 \ldots r_n)}{l_1 \neq r_1 \vee \ldots \vee l_n \neq r_n}$$

to the inequalities in the disjunction $(t_1 \neq s_1 \vee \ldots \vee t_n \neq s_n)\sigma$. If the two atoms $P(t_1 \ldots t_n)$ and $P(s_1 \ldots s_n)$ unify, the substitution σ is their most general unifier, the set D is empty and RUE reduces to resolution. Similarly, an NRF step

$$NRF\ \frac{t \neq s \vee A}{A\sigma \vee D}$$

applies a substitution σ and any number of Decomposition steps to the inequality $t\sigma \neq s\sigma$ in a given clause. If the two sides of the inequality $s \neq t$ unify, the substitution σ is their most general unifier, the set D is empty and NRF reduces to resolution with $x = x$. The disjunction D is called a *disagreement set* in both RUE and NRF. These inference rules are said to be in *open form* [2], because they do not specify how the disagreement set D and the substitution σ are computed. The *RUE-NRF inference system in open form* [2] is obtained by adding to the RUE and the NRF inference rules a *Factoring* rule, modified in a similar way:

$$Factoring\ \frac{P(t_1 \ldots t_n) \vee P(s_1 \ldots s_n) \vee A}{(P(t_1 \ldots t_n) \vee A)\sigma \vee D}$$

where D is derived by applying any number of Decomposition steps to $(t_1 \neq s_1 \vee \ldots \vee t_n \neq s_n)\sigma$ like in RUE. This set of rules is complete for first order logic with equality with no need of the axioms of equality [2]. Intuitively, this is because the axioms for reflexivity, transitivity and substitutivity (the last ones are also called functional reflexive axioms) are sort of implicitly applied in the RUE-NRF rules. For instance, the basic reflexivity axiom $x = x$ is not needed, because resolution with $x = x$ is a special case of the NRF rule. On the other hand, the authors themselves observe that, in order to implement symmetry, RUE needs to be applied twice, whenever it is applied to a pair of complementary equality literals. Given clauses $s_1 = s_2 \vee A$ and $t_1 \neq t_2 \vee B$, one generally needs to derive two RUE-resolvents: $(A \vee B)\sigma \vee D$, where D is a disagreement set of $s_1 \neq t_1 \vee s_2 \neq t_2$, and $(A \vee B)\sigma' \vee D'$, where D' is a disagreement set of $s_1 \neq t_2 \vee s_2 \neq t_1$.

The RUE-NRF system in open form is complete without the axioms of equality, because it embeds them in a sort of straightforward way. The price to pay for completeness is the very high degree of generality and non-determinism of the inference rules. The RUE-NRF system in open form is too general to be mechanized effectively, as the disagreement set D and the substitution σ can be arbitrary. Consequently, most of the work illustrated in [2] consists in designing criteria to compute D and σ.

The viability criterion is incomplete

The rule which selects the disagreement set determines the number of decomposition steps to be embedded in an RUE-NRF step. All the criteria proposed in [2] are based on a notion of *viable disagreement set*. A disagreement set $D = \{s_i \neq t_i\}_{i=1}^{n}$ is *viable* if and only if it can be partitioned into two sets D_1 and D_2, such that:

1. for every side s of an inequality in D_1 there is a positive equality literal $l = r$ in S such that
   (a) either s and l unify
   (b) or s and l have the same topmost function symbol and there is a viable disagreement set below the pair (s, l)
2. there exists a unifier σ such that $s_i\sigma = t_i\sigma$ for all $s_i \neq t_i \in D_2$.

The viability condition seems to be designed with the purpose of guaranteeing that at least one step can be done in order to erase the generated inequalities. All the pairs (s_i, t_i) in D_2 can be unified and deleted in the same RUE-NRF step where they are generated. For a pair (s_i, t_i) in D_1, it is possible to perform an RUE step with the clause in which the literal $l = r$ occurs. The basic refinement of the RUE-NRF system in open form, proposed in [2], consists in requiring that an RUE-NRF step is performed only if there is a viable disagreement set. Also, it is claimed that if the *topmost viable disagreement set* is selected at each step, the RUE-NRF system is still complete for first order logic with equality without the axioms of equality. Choosing the topmost viable disagreement set means that decomposition is applied, within an RUE-NRF step, until it generates either an inequality $f(l_1 \ldots l_n) \neq g(r_1 \ldots r_n)$ with different topmost function symbols or an inequality $f(l_1 \ldots l_n) \neq f(r_1 \ldots r_n)$, which is viable. The following example contradicts the above claim, showing that if RUE-NRF is restricted to viable disagreement sets, the RUE-NRF system is not complete without the functional reflexive axioms:

Example 1 *Let S be $\{g(f(a)) = a, f(g(x)) \neq x\}$. This set is inconsistent as can be seen by paramodulating $g(f(a)) = a$ into $f(g(x)) \neq x$:*

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

The applied unifier is $\{x \leftarrow f(a)\}$. The possible disagreement sets for an RUE step are $\{x \neq g(f(a)), f(g(x)) \neq a\}$ and $\{x \neq a, f(g(x)) \neq g(f(a))\}$. Neither one is viable, because of the inequalities including the term $f(g(x))$. Since there is no equation whose side starts with the function symbol f, the term $f(g(x))$ does not satisfy the viability requirements. The disagreement set $\{f(g(x)) \neq x\}$ for an NRF step is not viable either, since x occurs in $f(g(x))$, so that the two terms do not unify. No RUE-NRF step with a viable disagreement set can be applied. Thus the restriction to viable disagreement sets is not complete. A topmost viable RUE-NRF refutation can be obtained by adding the functional reflexive axiom $f(y) = f(y)$: RUE applied to $x \neq f(g(x))$ and $f(y) = f(y)$ yields $x \neq f(y) \vee f(g(x)) \neq f(y)$, which is then reduced by NRF to $x \neq f(y) \vee g(x) \neq y$. This can be resolved with $g(f(a)) = a$ to yield the contradiction $f(a) \neq f(a)$.

The RUE-NRF system restricted to viable disagreement sets needs the functional reflexive axioms to handle *collapse equations*, i.e. equations where one side is a variable. The equation $f(g(x)) = x$ which occurs negated in the above example is a collapse equation. The authors of [2] did observe that the filtering effect of the viability criterion is lost, if a collapse equation occurs in S: since every term unifies with a variable, every pair of terms is viable. However, they did not observe that completeness fails to hold in the presence of collapse equations. Intuitively, the equality reasoning in the RUE-NRF system is done by Decomposition, which handles equations in a top down style. This kind of step does not apply to collapse equations, where one side is unknown. The functional reflexive axioms allow to build the term which needs to be assigned to the variable side in order to obtain a refutation. Unfortunately, the functional reflexive axioms make viable any disagreement pair, so that the need for the functional reflexive axioms defeats the purpose of the viability restriction.

The lowest reducible disagreement set heuristic is incomplete

A number of additional variations of the RUE-NRF system are described in [2]. Some of them are not claimed to preserve completeness and are given simply as heuristics. One such heuristic is to choose the *lowest reducible disagreement set*. Let E_S denote the set of all the positive equations which appear, not necessarily as unit clauses, in the given set of clauses S. Decomposition within the RUE-NRF steps is applied until it generates either an inequality $f(l_1 \dots l_n) \neq g(r_1 \dots r_n)$ with different topmost function symbols or an inequality $f(l_1 \dots l_n) \neq f(r_1 \dots r_n)$, such that for some i, $1 \leq i \leq n$, $l_i \neq r_i$ is E_S-irreducible. This heuristic may help preventing the generation of trivially irreducible inequalities, such as $a \neq b$, where a and b are Skolem constants. However, if the RUE-NRF system is restricted to generate only disagreement sets of E_S-reducible inequalities, it becomes trivially incomplete:

Example 2 *Given* $S = \{f(g(a,x)) \neq f(g(x,a)), g(b,y) = g(y,b)\}$*, the inequality* $g(a,x) \neq g(x,a)$*, which leads to the refutation, cannot be generated because neither of its sides is* E_S*-reducible.*

All the other RUE-NRF systems which are claimed or conjectured to be complete in [2] include the restriction to viable disagreement sets and thus are not complete without the functional reflexive axioms. We conjecture that RUE-NRF with the topmost viable disagreement set may be complete for input sets of clauses S whose equational component E_S axiomatizes a *simple* theory. An equational theory is *simple* if any equation $s = t[s]$, i.e. an equation where one side occurs as strict subterm of the other one, is unsatisfiable in the theory. Clearly, a simple theory does not contain collapse equations. However, a restriction to simple theories may be regarded as too strong, casting a shadow on the usefulness of such a result.

References

[1] R.Anderson, Completeness Results for E-resolution, in *Proceedings of the AFIPS 1970 Spring Joint Computer Conference*, AFIPS Press, Reston Va., 653–656, 1970.

[2] V.J.Digricoli, M.C.Harrison, Equality-Based Binary Resolution, *Journal of the ACM*, Vol. 33, No. 2, 254–289, April 1986.

[3] J.B.Morris, E-resolution: An Extension of Resolution to include the Equality Relation, in D.E.Walker, N.M.Norton (eds.), *Proceedings of the IJCAI'69*, 287–294, 1969.