Publ. Math. Debrecen
51 / 1-2 (1997), 127–132

A note on the completeness of higher order resolution

By I. P. KOSSEY (Uzhgorod)

Abstract. A simple example shows that the completeness theorems of the second and higher order resolution (JENSEN, D.C., PIETRZYKOWSKI, T.) are not valid. The error in the proof is shown and two corrected formulations are proposed which, however, can considerably increase the search space. The practical use of the explained problem is necessary, for example, while proving negative inductive assertions.

In [1] a theorem on the completeness of the second order resolution is stated and (with reference to [2]) proved. In [3] on the completeness of higher order resolution a similar theorem is stated without proof. Unfortunately, both theorems are formulated erroneously.

We reported a short message about this in [4]. Here we describe the problem in detail, give a simple disproving example, show the error of the proof in [1] and propose two corrected formulations with proof.

Consider a set containing one formula:

$$\{\forall q q\}, \tag{1}$$

where q is a variable, $\tau(q) = p$. In other words, q is a nullary predicate variable. The resolution builds no resolvents because the formula

$$\forall q q \tag{2}$$

does not contain negation. On other hand, the set (1) is obviously inconsistent. Indeed, if we omit the quantor, for example, the formulas Q and

Mathematics Subject Classification: 03B15.
Key words and phrases: Resolution method, higher order logic, completeness.
Research was supported by the Hungarian National Foundation for Scientific Research No. T014279.

$\neg Q$ (Q is a nullary predicate constant) are instances of the variable q. The formula (2) asserts semantically that each nullary predicate is true. This obviously is not valid. In some works (for example [5]) the formula (2) is taken as definition of falshood.

We consider some constructions of [1] to show the error in it. For an enumerable set $\mathfrak{X}$ of second order formulas the set $E(\mathfrak{X})$ is defined in [1] as

$$E(\mathfrak{X}) = \{z \mid z \text{ is a substitution of some } X \in \mathfrak{X}\}, \tag{3}$$

and is called an extension of $\mathfrak{X}$.

The proof in [1] supposes that (3) is a semantically complete set of instances of $\mathfrak{X}$. But as we have seen for (1), this is not valid. The resolution essentially uses the semantics of the negation, but the extension (3) does not reflect the semantics of negation. A semantically complete set of instances should contain the negations of all literals which begin with predicate variables or their negation. More precisely consider a literal of the close C, which begins with the predicate variable P (or its negation $\neg P$), and $\tau(P) = (t_1, \ldots, t_n, p)$. Let us call the close $C \circ \mu$ a k-time variation of C on P, where

$$\mu = \{P \to \lambda u_1 \cdots u_n . \underbrace{\neg \ldots \neg} Q u_1 \ldots u_n\}, \tau(Q) = \tau(P)$$

where the (underbraced) negation is repeated k times.

Note that probably in the following explanations it is enough to consider only simple (one-time) variations, if the extensionality ([7]) is accepted. We do not discuss this problem and consider the most general case.

The variation $C \circ \mu$ is an instance of C, therefore $C \circ \mu$ follows from C. Therefore, if we have a consistent set of closes $\mathfrak{X}$, then its expansion $\mathfrak{X}^*$ with variations of $\mathfrak{X}$ is consistent too.

We define now the extension $E^*(\mathfrak{X})$ with variations of the close set $\mathfrak{X}$:

$$\begin{aligned} E^*(\mathfrak{X}) = \{Y \mid Y \in E(\mathfrak{X}) \quad &\text{or } Y \text{ is a variation} \\ &\text{of some element of } E(\mathfrak{X})\}. \end{aligned} \tag{4}$$

Consider again as example the set (1):

$$E(\mathfrak{S}) = \{\forall q q, q\},$$

but $E^*(\mathfrak{S})$ is an infinite set

$$E^*(\mathfrak{S}) = \{\forall q q, q, \neg q, \neg\neg q, \cdots\}.$$

In [1] the inclusion

$$GR(E(\mathfrak{X})) \subseteq E(R(\mathfrak{X})) \tag{5}$$

is proved, where $GR(W)$ is the ground resolution of the set W, $R(W)$ is the general resolution of W.

By the first order resolution method (5) follows the completeness of the resolution [6]. But $E(\mathfrak{X})$ defined in [1] does not contain all semantically true instances of $\mathfrak{X}$. Therefore (5) does not imply the completeness of the resolution. Instead of $E(\mathfrak{X})$ in (5) we should consider $E^*(\mathfrak{X})$:

$$GR(E^*(\mathfrak{X})) \subseteq E^*(R(\mathfrak{X})). \tag{6}$$

The inclusion (6) is not valid in the second order logic. Indeed, for the set (1)

$$E^*(\mathfrak{S}) = \{\forall q q, q, \neg q, \neg\neg q, \cdots\}, \quad R(\mathfrak{S}) = \{\forall q, q\}$$

$\Box \in GR(E^*(\mathfrak{S}))$ but $E^*(R(\mathfrak{S}))$ does not contain $\Box$.

We define

$$\mathfrak{X}^* = \{Y \mid Y \in \mathfrak{X} \text{ or Y is a variation of some element } X \in \mathfrak{X}\}. \tag{7}$$

Now we define a new resolution rule (modified resolution) $R^*(\mathfrak{X})$:

$$R^*(\mathfrak{X}) = R(\mathfrak{X}^*). \tag{8}$$

It is easy to show that

$$GR(E^*(\mathfrak{X})) \subseteq E^*(R^*(\mathfrak{X})) \tag{9}$$

is valid. The proof of (9) repeats the respective proof of (5) in [1], therefore we do not explain it. From (9) the completeness of the resolution follows in a similar way, just as it is proved in [1]. However, in this proof it is important that the modified resolution should be applied not only for the original closes but for all resolvents. This means that variations of the resolvents should be also created. Hence we have

Theorem 1. *The close set S is inconsistent if and only if from S the empty close is derivable with modified resolution.*

To obtain the second formulation, consider the tautology

$$\neg r,\ \neg\neg r, \tag{10}$$

where r is variable of type p. We show how we can derive each variation of each close from (10) with usual resolution:

$$\begin{array}{ll} C = & \boxed{\begin{array}{l} Pe_1 \cdots e_n \\ \neg r. \end{array}}\ , \cdots \\ \neg r,\ \neg & \end{array}$$

The resolvent (the resolved literals are framed) is:

$$C \circ \mu = \neg Qe_1 \cdots e_n, \cdots . \tag{11}$$

(11) is the simple (one-time) variation of C. The two-time variation we obtain in this way:

$$\neg \boxed{\begin{array}{r} Qe_1 \cdots e_n \\ \neg r \end{array}} \begin{array}{l} , \cdots \\ , \neg\neg r \end{array}$$

and so on. In a similar way we obtain the variations, if the close C contains a literal which begins with $\neg P$. From this follows that the double negation in (10) cannot be omitted. Thus, the second formulation is:

Theorem 2. *The set S is inconsistent if and only if from $S \cup \{\neg r, \neg\neg r\}$ the empty close can be inferred with usual resolution.*

All the above statements are valid for higher order resolution too.

We can consider (10) as the axiomatisation of negation in the higher order resolution. Therefore, the axiomatisation of negation is necessary for completeness even if the original set of closes does not contain negation. In [1, 3] it is shown that the axiomatisation of conjunction, disjunction etc. is necessary only if the original set of closes contains these logical functions.

In our example the empty close can be inferred in the following ways. With the Theorem 1:

$$\mathfrak{S}^* = \{\forall q q, q, \neg q, \neg\neg q, \cdots\}.$$

From this $\Box \in R^*(\mathfrak{S})$ follows immediately.

With the tautology (10):

1.	$\forall q q$	given;
2.	q	Q-reduction of 1;
3.	$\neg r, \neg\neg r$	the tautology (10);
4.	$\neg r$	resolvent of 2, 3b, subst.$\{q \to \neg r\}$;
5.	$\Box$	resolvent of $2, 4$.

Both theorems increase the search space of second/higher order resolution. For example, we should add to the induction axioms of [1, 3]

$$\neg P(0),\ P(a(P)),\ P(x)$$

$$\neg P(0),\ \neg P(a(P)+1),\ P(x)$$

their variations

$$\begin{array}{c} P(0),\ \neg P(a(\neg P)),\ \neg P(x) \\ P(0),\ P(a(\neg P)+1),\ \neg P(x) \end{array} \tag{12}$$

to be able to prove assertions of negative kind, for example: the successive Fibonacci numbers have no common divisor (of course, nontrivial, i.e. not 1). To prove this theorem the variations (12) are needed. Otherwise the induction predicate should be renamed (for example, using (10)) to its negation. Obviously the negation means: the successive Fibonacci numbers are relative primes. If the theorem were originally formulated so, the variation or the tautology (10) for our example would not be necessary.

It is interesting to note that in [8] the problem of completeness of the second order Gentzen calculus is not mentioned at all. Probably for the completeness of similar calculi it is necessary to introduce variations.

The level of works [1, 3] is high. How is then such an error possible? We must acknowledge that mathematics is not free from empirism. The natures of errors in mathematical proofs and in complex computer programs seem to be equal. If all mathematical proofs were realisable without error, we should write programs without errors!

References

[1] T. Pietrzykowski, A Complete Mechanization of Second-Order Type Theory, *J. ACM* **2** (1973), 20.

[2] D. C. Jensen, A Two-Valued Type Theory Model Convenient for Resolution Theorem-Proving, CSSR 2055, *Dep. of Applied Analysis and Computer Science, Univ. of Waterloo*, *Waterloo, Ontario, Canada*, 1971.

[3] D. C. Jensen and T. Pietrzykowski, Mechanizing ω-Order Type Theory through Unification, *Theor. Comp. Sci.* **3** (1976).

[4] I. P. Kossey, On the completeness of the higher order resolution method. (L. T. Kusin, ed.), *Collection "Intellectual data banks"*, *Naltschik, USSR.* (in *Russian*)

[5] A. Church, Introduction to the Mathematical Logic, *Princeton, New Jersey* **1** (1956).

[6] J. A. Robinson, Machine-Oriented Logic Based on the Resolution-Principle, *J. ACM* **12** (1965).

[7] G. P. Huet, A Unification Algorithm for Typed λ-Calculus, *Theor. Comp. Sci.* **1** (1975).

[8] Gaisi Takeuti, Proof theory, *North-Holland Publishing Company*, *Amsterdam – London, American Elsevier Publishing Company, Inc. – New York.*

I. P. KOSSEY
UZHGOROD STATE UNIVERSITY
DEPARTMENT OF MATHEMATICS
SEVERNAJA STREET 14
UZHGOROD
UKRAINE

(Received June 19, 1996)